Validating a Proposed Data Mining Approach (SLDM) for Motion Wearable Sensors to Detect the Early Signs of Lameness in Sheep

Submitted for the Degree of Doctor of Philosophy
At the University of Northampton

2019

Zainab Raed Ahmed Al-Rubaye

DEDICATION

To My beloved Iraq

The land where I was brought up, educated, and got my identity

To My Mother

DR. Thikra Kadom Mohamed Hassan

For enduring my absence abroad for PhD study.

For her continuous prayer for me and my success.

To My Father

DR. Raed Ahmed Al-Rubaye

For his wise advice to me for all aspects of life

For keeping me active and determined in all I planned for.

To My Husband

Eng. Alaa Alsaadi

For his patience, support, care, and love at all times and places.

To the soul of my Grandmother

Salima Abdulameer Al-Hassani (Allah mercy her)

She would have kept her prayers for me if she had lived today.

To all my Family members and faithful Friends

For their support and encouragement

INSPIRATIONAL QUOTES

“Whoever follows a path in the pursuit of knowledge. Allah will make easy for him/her a path to Paradise”

Prophet Mohammad (peace upon him)

“No two things have been combined better than knowledge and patience”

Prophet Mohammad (peace upon him)

“Nobody obtains knowledge except the one who lengthens his study”

Imam Ali Bin Abi Talib

“The writing of a man is the symbol of his intellect and the evidence of his merit”

Imam Ali Bin Abi Talib, Ghurar al-Hikam, Page no. 54

DECLARATION

I hereby declare that the entire work conducted in this thesis is original work submitted for the degree of Doctor of Philosophy in Computing, at the University of Northampton, Department of Computer Science and Immersive Technologies. I am the author of this thesis, I certify that I have not previously submitted the material of this work to any institute for obtaining any qualification.

Zainab Raed Ahmed Al-Rubaye

ACKNOWLEDGEMENT

I am grateful to Allah Almighty for enabling me to safely reach the end of my PhD journey and providing me with individuals who have stood by me when I felt weak, giving hope when I felt desperate, and encouraging me when I felt alone during my studies abroad.

My great appreciation to the Ministry of Higher Education and Scientific Research in Iraq for financial support. Excellent thanks to the University of Baghdad for offering a valuable opportunity for me to pursue my PhD in the UK.

My sincere thanks to my supervisor Dr Ali Al-Sherbaz for his continuous pursuit, positive encouragement, and academic guidance through my PhD journey. My appreciation and most profound thanks are also presented to my supervisors Dr Wanda McCormick and Dr Scott Turner for their support, invaluable advice, and insightful feedback.

I would like to thank the University of Northampton for providing a complete academic environment, skilful training workshops, and all the required facilities for the PhD students.

My thanks to Moulton College, which authorised my access to Lodge Farm, the place where I collect my data. Special thanks to the Shepherd 'Tim Perks' who was very helpful in scheduling the observation time at Lodge Farm.

Special thanks to my understanding husband Alaa Al-Saadi, who stands by me all the time and keeps me company in my academic participations.

I'm grateful to my colleagues: Dr Alyaa Al-Barrak for her endless support during the hard times of my study, and Dr Riyadh Abbas for his helpful advice. My extended thanks to Mrs Juliet Dixon-Evans for her organised cooperative assistance.

My heartfelt thanks to my parents, who encouraged me from overseas to successfully work for my qualification. My grateful thanks to all my family member in Iraq who proud of my achievements.

PUBLICATIONS AND AWARDS

Al-Rubaye, Z., Al-Sherbaz, A., McCormick, W. D. and Turner, S. J. (2018) **Sensor Orientation for the Indication of Lameness in Sheep**. Presentation at 3rd Annual *Sensor in Food and Agriculture conference. John Innes Centre*, ***Norwich****, UK, 18-19 July 2018.*

Al-Rubaye, Z., Al-Sherbaz, A., McCormick, W. D. and Turner, S. J. (2018) **Identifying Lameness Movements in Sheep via Sensor Data Analysis**. Poster presented at: *Recent advances in animal welfare science VI, UFAW Animal Welfare Conference, Centre for life,* ***Newcastle****, UK. 28 June 2018.*

Al-Rubaye, Z. (2018) **Sheep Lameness Detection via Wearable Sensor-Based Data Analysis**. Poster presented at: *STEM for Britain exhibition in the engineering section.* T*he House of Commons, the* ***UK parliament****. 12 March 2018.*

Al-Rubaye, Z., Al-Sherbaz, A., McCormick, W., Turner, S. (2018) **Sensor Data Classification for the Indication of Lameness in Sheep**. In *13th EAI Collaborate Computing: Networking, Applications and Worksharing*. ***Edinburgh****. UK, December 11–13, 2017.* Published in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer International Publishing, pp. 309–320.

BBC1 'The ONE Show' broadcast on *16 November 2017* a **Lameness Film** showing the project implication for farmers to identify the early signs of lameness in sheep. Available from: https://www.bbc.co.uk/programmes/b09dy4rm. [Accessed: 4 February 2020].

Al-Rubaye, Z., Al-Sherbaz, A., McCormick, W. D. and Turner, S. J. (2017) **Sheep in Northampton use smart device as a sensor**. Image presented at *the University of Northampton*: Image of Research competition IoR *2017*. **1st place winner** by public vote.

Al-Rubaye, Z., Al-Sherbaz, A., Ghendir, S., McCormick, W. D. and Turner, S. J. (2016) **Lameness detection in sheep via multi-data analysis of a wearable sensor**. Seminar Presentation presented to: *5th Postgraduate Research Symposium,* ***Moulton College****, Northampton, 15 December 2016*.

Al-Sherbaz, A., Al-Rubaye, Z., Ghendir, S., McCormick, W. D. and Turner, S. J. (2016) **The use of multivariable sensor data to early detect lameness in sheep**. Presentation at: *Sensors in Food and Agriculture, Møller Centre, Churchill College, University of* ***Cambridge****, 29-30 November 2016*.

Al-Rubaye, Z., Al-Sherbaz, A., McCormick, W. D. and Turner, S. J. (2016) **Lameness detection in sheep through the analysis of the wireless sensor data**. Presentation presented at *The University of Northampton Graduate School Postgraduate Researcher (PGR) Annual Conference 2016,* ***Northampton****, 14 June 2016*.

Al-Rubaye, Z. (2016) **Lameness detection in sheep through behavioural sensor data analysis**. Poster presented to: *Graduate School 11th Annual Poster Competition, The University of Northampton, 18 May 2016*. **2nd place winner**.

Al-Rubaye, Z., Al-Sherbaz, A., McCormick, W. D. and Turner, S. J. (2016) **The use of multivariable wireless sensor data to early detect lameness in sheep**. Workshop presented to: *School of Science and Technology Annual Research Conference, Newton Building, The University of Northampton,* ***Northampton*** *02 March 2016*.

ABSTRACT

Lameness can be described as painful erratic movements, which relate to a locomotor system and result in the animal deviating from its normal gait or posture. Lameness is considered one of the major health and welfare concerns for the sheep industry in the UK that leads to a substantial economic problem and causes a reduction in overall farm productivity. According to a report in 2013 by ADAS entitled 'Economic Impact of Health and Welfare Issues in Beef, Cattle and Sheep in England', each lame ewe costs £89.80 due to the decline in body condition, lambing percentage, growth rate, and reduced fertility. Thus, early lameness detection eliminates the negative impact of lameness and increase the chance of favourable outcome from treatment. The development of wearable sensor technologies enables the idea of remotely monitoring the changes in animal behaviours or movements which relate to lameness.

The aim of this thesis was to evaluate the feasibility and accessibility of a proposed data mining approach (SLDM) to detect the early signs of lameness in sheep via analysing the retrieved data from a mounted wearable motion sensor within a sheep's neck collar through investigating the most cost effective factors that contribute to lameness detection within the whole data mining process including; sensor sampling rate, segmentation methods, window size, extracted features, feature selection methods, and applicable classification algorithm. Three classes are recognised for sheep while their walking throughout the data collection process (sound, mild, and severe lameness classes). The sheep data were collected using three different sensor applications (Sheep Tracker, SensoDuino, SensorLog) which collect sheep data movements at different sampling rates 10, 5, and 4 Hz. Various sensing data were retrieved in X,Y, and Z dimensions; however, only accelerometer, gyroscope, and orientation readings are considered in the current study. Four sheep datasets are aggregated each of which includes 31, 10, 18, and 7 sheep. The conducted work in this thesis evaluates the performance of ensemble classifiers (Bagging, Boosting, or RusBoosting) using three different validation methods (5-fold, 0.3 hold-out, and proposed one 'Single Sheep Splitting') in comparison to three sampling rates (10, 5, 4 Hz), two segmentation approaches (FNSW and FOSW), three feature selection methods (ReliefF, GA, and RF) and three window sizes (10, 7, 5 $sec.$).

Promising results of lameness prediction accuracies are achieved over most of the combinations (3 sampling rates, two segmentation methods, 3 window sizes, 183 extracted features, 3 feature selection methods, 3 ensemble classification models, and 3 model validation

methods). However, the highest accuracy is revealed by using the `Bagging ensemble classifier 88.92% with F-score of 87.7%, 91.1%, 88.2% for sound walking, mildly walking, and severely walking classes, respectively. The results are obtained using 5-fold cross-validation over a 10 $sec. window$ for sheep data collected at 10 Hz sampling rate using only the accelerometer hardware sensor reading and calculated orientation readings. The number of features selected is 46 optimised by GA using CHAID tree as a fitness function. Conversely, the lowest prediction accuracy of 56.25% with F-score (63.4% sound walking, 51.9% mildly walking, 48.8% severely walking) is recorded when RusBoosting ensemble is applied using 5-fold cross-validation over a 10 $sec. window$ for dataset collected at the 4 Hz. sampling rate.

So, the major research findings recommend that 10 Hz sampling rate is adequate for collect sheep movements, while the best segmentation method is FOSW as 20% of data-points are shared between two successive windows. Whereas, the preferable number of data-points (sheep movements) to be pre-processed is around 100, which is obtained when a 10 $sec. window\ size$ or 7 $sec. window\ size$ is applied. Additionally, the 20 features selected by RF out of 183 features could reveal good accuracy results compared to the whole set of extracted features. Although that GA feature selection method has slower execution time than RF, competitive prediction accuracy could be achieved when the selected features by GA were fed to the classifier. Finally, the acceleration sensor data alone are capable of making the decision about the lame sheep. So no extra hardware sensors like Gyroscope is required for decision making; moreover, the orientation sensor features could be directly derived from Acc which contribute most to lameness detection.

Since the most cost effective factors are identified in this research, the practice in the meanwhile could be applicable for farmers, stakeholders, and manufacturers as no available sensor to detect the lame sheep developed yet. Therefore, the multidisciplinary nature of the conducted research opens diverse paths towards applying further research studies to develop various data mining approaches and practical sensor kits to detect the early signs of sheep's lameness for better farm productivity and sheep industry prosperity in the UK.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D	Two dimensional
3D	Three dimensional
Accu	Accuracy
AdaBoost	Adaptive Boosting
ADAS	the UK's largest independent provider of agricultural and environmental consultancy
AHDB	Agriculture and Horticulture Development Board
ALT-Pedometer	Activity-Lying-Temperature Pedometer
AMS	Automatic Milking System
ANNs	Artificial Neural Networks
AUC	Area Under Curve
BMP	Body Movement Pattern
CART	Classification and Regression Tree
CHAID	Chi-square Automatic Interaction Detectors
CODD	Contagious Ovine Digital Dermatitis
DD	Digital Dermatitis
Defra	Department for Environment, Food & Rural Affairs
Df	Degree of Freedom
DM	Data Mining
DT	Decision Tree
EC	Ensemble Classifier
Emfit	Electromechanical film mat
FFT	Fast Fourier Transform
FNSW	Fixed-size Non-overlapping Sliding Window
FOSW	Fixed-size Overlapping Sliding Window
FR	Footrot
FS	Feature Selection
GA	Genetic Algorithms
GBoost	Gradient Boosting
GPS	Global Positioning System
GRF	Ground Reaction Force
GS	Gait Scoring
GUI	Graphical User Interface

HMM	Hidden Markov Model
HWC	Hoof Weight Crate
Hz	Hertz
IMU	Inertial Measurement Unit
IRT	Infrared Thermography
KDD	Knowledge Discovery in Databases
KNN	K- Nearest Neighbour
LDA	Linear discriminant Analysis
LR	Linear Regression
LS	Locomotion Scoring
ML	Machine Learning
MLP	Multi-Layer Perceptron
NB	Naive Bayesian
NRS	Numerical Rating System
PCA	Principal Component Analysis
PLF	Precision Livestock Farming
PNN	Probabilistic Neural Network
Prec	Precision
QDA	Quadratic Discriminant Analysis
RBF	Radial Basis Function
ReliefF	a name for feature selection method
RF	Random Forest
RusBoosting	Random Under Sampling Boosting
SAS	Statistical Analysis Software
SD	Standard Deviation
SLDM	Sheep Lameness Detection Model
SU	Sole Ulcer
SVM	Support Vector Machine
TPR	True Positive Rate
VAS	Visual Analogue Scale
Vedba	Vectorial of the Dynamic Body Acceleration
WLD	White Line Disease
WSN	Wireless Sensor Network
XGBoost	Extreme Gradient Boosting
Acc	Accelerometer

Acc_Lin	Linear Acceleration
Acc_x	Acceleration readings in the x-axis
Acc_y	Acceleration reading in the y-axis
Acc_z	Acceleration reading in the z-axis
Gyr	Gyroscope
Gyr_x	Gyroscope reading in the x-axis
Gyr_y	Gyroscope reading in the y-axis
Gyr_z	Gyroscope reading in the z-axis
$Orient$	Orientation

1 Chapter One: Introduction

1.1 Problem Description

The current research is a multidisciplinary research study that has been conducted as a collaborative project between the animal welfare department in Moulton College and the Computing Department at the University of Northampton. The current thesis identifies a way of solving a real-world problem (sheep lameness) by utilising sensor technologies for data collection and sophisticated machine learning approaches for data analysis to build a robust model that could adequately predict the early signs of lameness in sheep. The built model could predict a sheep's future status of mildly lame conditions that might be difficult to recognise with the observer's naked eye, sheep as prey species often disguising signs of vulnerability such as limping. The developed approach enriches the field of knowledge that lacks sheep lameness detection studies; furthermore, the application of the proposed system could decrease the prevalence of lameness and enable the shepherd to react quickly to enable better treatment.

1.2 Thesis Outline

This introduction Chapter is started by describing the research problem, then it is followed by sections including lameness definition, welfare and economic implications, and the benefits of early detection of it. The gap in the literature is highlighted followed by stating the aims and objectives of the thesis. A brief structure of the applied methodology is given in a clear flowchart. Finally, this chapter is closed by listing possible research contributions.

The next chapters of the thesis are organised as follows. Chapter Two, investigates the current multidisciplinary research studies in cows and sheep lameness detection approaches and the field of behaviour classification using machine learning techniques. However, the intersection between utilising motion sensors in sheep lameness detection and applying machine learning techniques for lameness prediction in sheep is rarely found. Thus, from this point, the gap in the literature is identified. In Chapter three, an overview of the sensor application used for data collection is given; in addition, sensor deployment and challenges faced in the data collection process are mentioned. It also provides details on developing a data mining approach for the detecting of sheep lameness; including data pre-processing, segmentation, extract walking segments, features extraction and selection, model development, and model validation. Chapter four introduces two parts; the first part demonstrates the Graphical User Interface GUI application that is uniquely designed for the purpose of this study. The second part of Chapter four presents a wide range of intermediate results in addition to the final prediction result accuracies for the built model. These are illustrated with wide-ranging discussions and fair comparisons with other semi-related works to formulate the final recommendations. The thesis is closed with Chapter five, which includes an overall conclusion and inspiring future work ideas to be implemented for enhancing the model and optimising the sensor requirements to save its battery life.

1.3 Lameness in Sheep

1.3.1 Definition and Causes

Lameness is a painful impaired movement disorder, which relates to an animal's locomotory system and causes a deviation from normal gait or posture (Van Nuffel *et al.*, 2015a). The leading causes of lameness in sheep are Footrot (FR) which is a bacterial disease caused by *Dichelobacter nodosus*, Interdigital dermatitis (scald) caused by *Fusobacterium necrophorum*, and Contagious Ovine Digital Dermatitis (CODD) which is caused by the S*pirochaete Treponema* in addition to pre-mentioned causative agents (Gelasakis *et al.*, 2019; (Olechnowicz and Jaśkowski, 2011; Winter, 2008). FR is reported as the most common cause of lameness, resulting in 90% of all sheep lameness cases in the UK (Groenevelt *et al.*, 2015; Scott *et al.*, 2017). Figure 1-1 shows the percentages of lameness caused by FR, scald, and other producers of lameness according to a postal survey conducted by the Royal Veterinary College in 1997 (Defra, 2003).

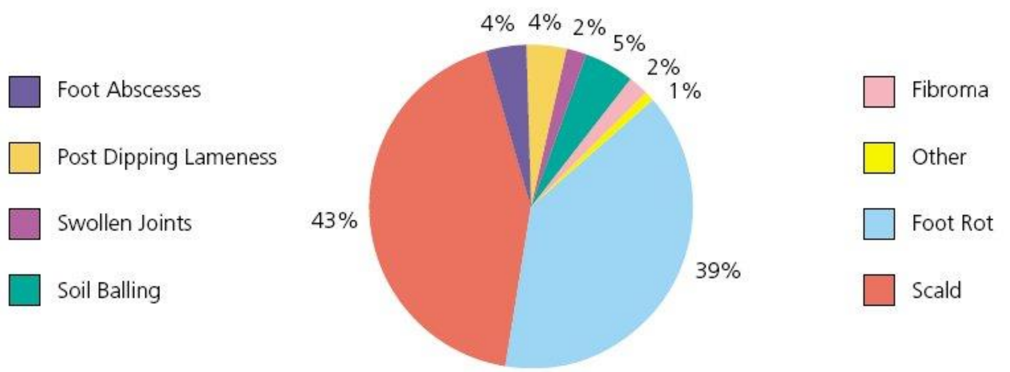


Figure 1-1 Lameness causes based on the Royal Veterinary College survey in 1997 (Defra, 2003).

The infectious nature of FR is commonly increased above 10º C and reaches its peak between April and June and then August to the end of October in the UK (Defra, 2003). So, the UK damp temperature climate changes between mild winter and wet summer provide a perfect environmental condition for FR infectious bacteria to grow and transmit easily and rapidly within the flock. The invasion of the FR bacteria to the horn of the sheep's foot and then its surrounding tissue leads to horn separation in the heel area and could extend beneath the horn, sole, and even the entire hoof in the worst cases, causing different levels of lameness that starts from mild and then develops to moderate and severely lame. As the infected sheep could remain out on pasture for up to twelve days and spread the infectious agents (Defra, 2003; Anzuino *et al.*, 2019), this can have negative implications for the UK sheep industries.

1.3.2 Welfare and Economic Implication

Lameness is considered one of the most significant health and welfare concerns for the sheep industry in the UK, that leads to a substantial economic problem and overall farm productivity decline. Furthermore, the cost of lameness treatment and control consumes a large amount of money in the farm business as described by Lovatt, (2014) see Figure 1-2. According to Nieuwhof and Bishop, (2005), half of the FR cost is spent for preventive measures while the other half is consumed by treatment and lost production; therefore, a reduction in FR incidence could save up to £10 million for UK industries.

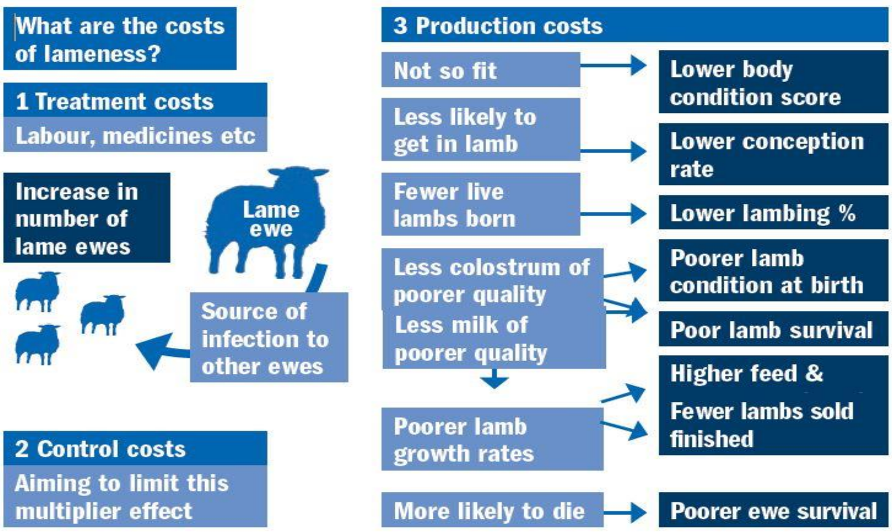


Figure 1-2 The cost of lameness (Lovatt, 2014).

The cost of FR disease to the British sheep industry per year was estimated to be £24 million (Nieuwhof and Bishop, 2005), although statistics from the Agriculture and Horticulture Development Board (AHDB) (an organisation that provides services to beef and lamb levy payers in England) reported that the annual loss from FR alone was around £10 for each ewe in Great Britain (Brian, 2016). Additionally, the latest report from ADAS, '*Economic Impact of Health and Welfare Issues in Beef, Cattle and Sheep in England*', reveals that each lame ewe costs approximately £89.80 because the decline in its performance must be accounted for alongside extra labour and treatment cost (Mary and Wright, 2013). The underlying reasons for the commercial loss in the sheep industry in the UK can be related to various outcomes which are summarised as the decline in the sheep' body condition, lambing percentage, lamb birth weight, growth rate in lambs, wool growth, milk production and poor fertility in the rams (Defra, 2003). This is why lameness is listed as one of the main causes of sheep culling besides infertility and mastitis (Alsaaod *et al.*, 2012; AHDB, 2016).

Consequently, lameness would have an adverse influence on both sheep welfare and farm economy. Preclinical detection of lameness at the farm could increase the level of protection regarding sheep health and the farm's commercial decline and could allow it to be controlled from being spread within the whole flock. Therefore, sheep lameness research studies would be required to assist the farmers in spotting lameness on-farm, as lameness comes at the fifth ranked issue that concerned farmers from 44 farms in a survey for husbandry and health in the UK (Anzuino *et al.*, 2019).

1.3.3 Early Detection Advantages

Since lameness is an endemic disease that cannot be entirely eradicated; however, the early detection of lameness will reduce the disease from spreading very quickly within the flock. A study by Gaudy and Green, (2016); reported on the AHDB website, at the University of Warwick to develop a lameness control plan looking at flocks on three different farms in the UK indicated that the quicker the lame sheep are treated, the less prevalent the disease is and fewer sheep require treatment; these effects are seen primarily if the treatment has been applied within three days of sheep becoming lame. Thus, early lameness detection could actively eliminate the negative impact of lameness and increase the success ratio of treatment by preventing it from being a chronic illness (Alsaaod *et al.*, 2012). Furthermore, the advantages of early lameness detection may be also seen in maximising the total farm income, enhancing the sheep welfare, which leads to improving the entire flock performance and reducing the veterinary, medicine and labour costs (Defra, 2003).

1.4 Research Gap

Reviewing the literature related to lameness detection in sheep yielded inadequate research studies in terms of data collection tools and analysis methods. The existing research studies have used Infrared Thermography (Byrne *et al.*, 2019a), load cells weight platform (Byrne *et al.*, 2019b), and radar sensing (Shrestha *et al.*, 2018) for data collection. However, these tools are costly and require someone to guide the sheep into testing areas such as the load cells or radar sensor place. This is contradicting with the aim of the current research of monitoring sheep in an unattended, not expensive way. Additionally, traditional observations by the trained observer are very time consuming, subjective, and require a lot of effort.

Alternative sensor technologies have emerged to collect sheep motion data in Barwick *et al.*, (2018b); however, the sheep in their experiments were not in real lame conditions; instead, they were simulated by restraining the sheep's front leg using an adhesive bandage. Although Vazquez Diosdado *et al.*, (2018) investigate sheep lameness using motion sensors, the resulted accuracies still need further enhancements. In addition, commercial sensors for detecting sheep lameness have not been developed yet for the benefits of the stakeholders. In contrast to the cattle sector where the IceRobotics company, founded in 2002 and based in Edinburgh (ICEROBOTiCS, 2019), provides a commercial CowAlert application sensing system capable of monitoring cow's health and producing an alert concerning health issues. Recently, a cow lameness alert within CowAlert has been launched by IceRobotics (Chomiak, 2017) to provide daily lameness alerts.

Although this project is not aiming to produce a commercial 'Fitbit' for sheep to monitor their health, including lameness detection, the findings of the current research study would pave the way for other researchers to develop a sensor device which performs monitoring and alarming tasks. This would help the shepherd to identify sheep health issues on the farm and enhance Precision Livestock Farming (PLF) in the sheep sector compared to their cow counterparts. Therefore, the current research could enrich the lack of sheep lameness research studies in term of utilising convenient data collection tools (motion sensors), conducting validated experiments, and developing an original lameness prediction model.

1.5 Research Aims and Objectives

The goal of the conducted thesis was to evaluate the implementation of a data mining approach to detect the early stage of lameness in sheep (mildly lame sheep) by analysing the data being retrieved from a mounted wearable motion sensor within a sheep's neck collar. The validated approach was aimed at being feasible and easy to be accessed by farmers with no extra need for continuous monitoring of the whole flock. Furthermore, the built model was targeted to be economical as data collection, pre-processing, analysis and decision making are all processed into one sensor kit to be mounted in the sheep's collar. Thus, investigating the most cost effective factors contributing to lameness detection is the key focus of this work such as; sensor sampling rate, segmentation method and window size, the most powerful features, the best feature selection methods, and applicable classification algorithm are all experienced to serve the purpose of the research.

The objectives of the current research are as follows:

1- Investigating the lameness detection methods for cows and sheep; including data mining techniques to identify the gap in the literature.
2- Reviewing the data mining classification techniques that have been utilised to classify cows or sheep behaviour to deduce the proper technique for the indication of lameness in sheep.
3- Collecting real-world sheep data from Moulton College Lodge Farm via a wearable sensor device mounted around a sheep's neck at different sampling rates to identify the most convenient sampling rate for identifying sheep lameness.
4- Pre-processing of the sheep sensor raw data in many stages; including data cleaning, missing data manipulation, segmentation, walking segment extraction, features computation, and best features selection.
5- Training the best set of features via various machine learning techniques to determine the most satisfactory prediction accuracy algorithms for sheep lameness detection.
6- Evaluating the trained model using three validation techniques (k-fold, hold-out, Single Sheep Splitting).

1.6 Research Methodology Structure

The sensor-based collected data requires a professional approach to pre-processing, analysing, and decision making in order to classify the sheep status into sound, mildly lame, or severely lame. Figure 1-3 depicts the applied stages of the data mining methodology for lameness detection in sheep in this thesis. Although the full details for each step are explained in Chapter Three, a brief visualise flowchart is presented here as a methodology introductory part within the current Chapter. The proposed approach for Sheep Lameness Detection Model (SLDM) properly provides full data mining steps that could be recommended to future research studies into sheep lameness, as the literature search results are evidently lacking these kinds of studies.

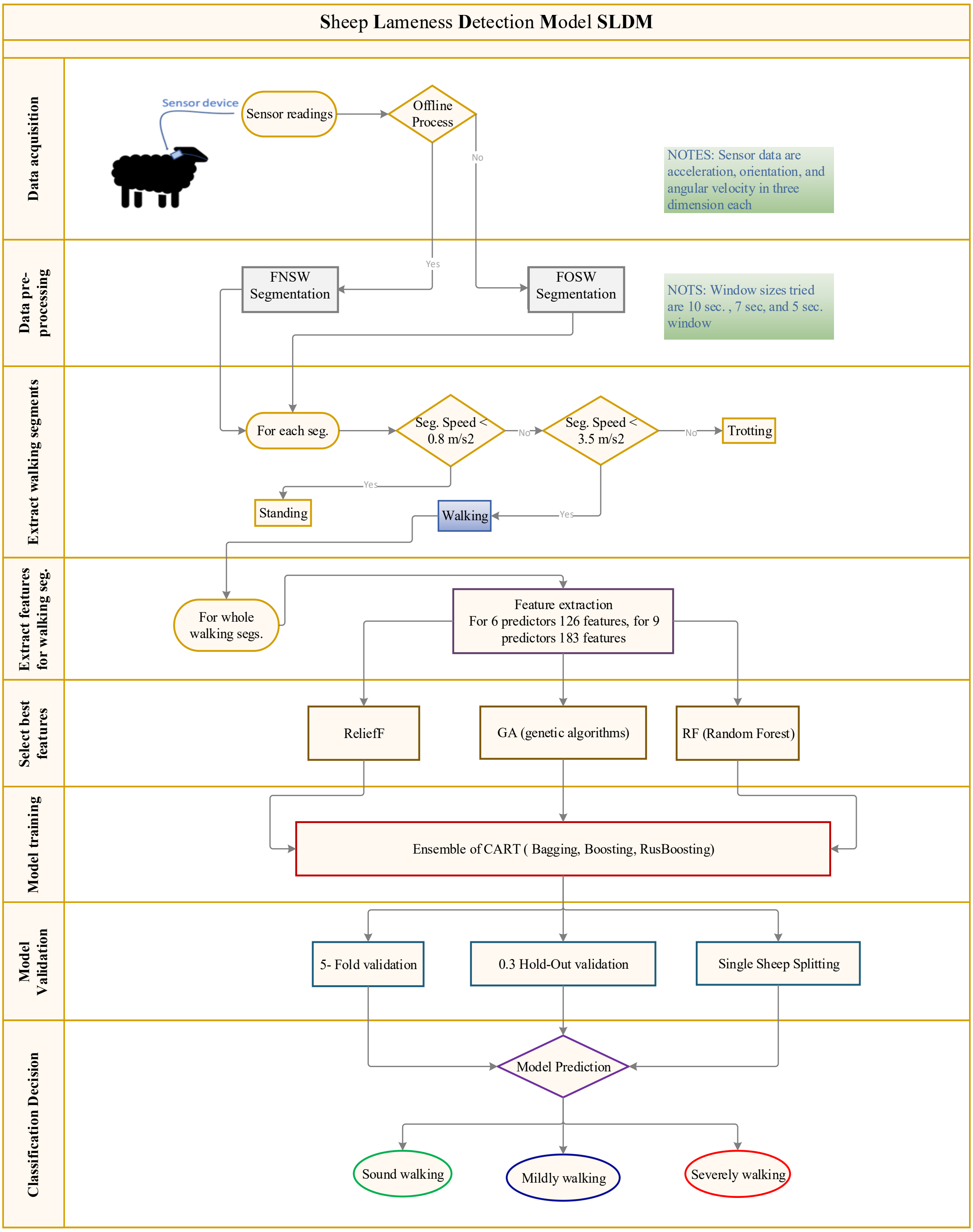


Figure 1-3 The stages of thesis's methodology for developing Sheep Lameness Detection Model (SLDM).

1.7 Possible Research Contributions

No study to date has utilised data mining approaches to classify motion sensor-based data retrieved from a sheep's neck for the purpose of detecting lameness. Instead, some studies have investigated sheep behavioural classification into standing, walking, laying, grazing, ruminating, and other classes for the purpose of grazing at pasture. Thus, the study that has been conducted in this thesis contributes to the field of knowledge as follows:

1- Real-world sheep movement data are collected via a wearable motion sensor at Lodge Farm, Moulton College, Northamptonshire at three different sampling rates 10, 5, 4 Hz. Previous lameness walking movement sheep data could not be found online.

2- Due to the fact that lameness tends to be identified when sheep are walking, only sheep walking data are extracted (aside from standing or trotting movements) to be pre-processed and classified by integrating sheep forward-backwards acceleration (Acc_y) to obtain sheep speed. This process prolongs the sensor battery life as the classification procedure could only work when the sheep are walking. Alternatively, the sensor could be set to sleep mode when the sheep behave differently and not in a walking rhythm.

3- The important features that actively contribute to lameness detection are determined. The Orientation sensor data around the sheep neck (Pitch and Roll angles) are mostly contributing to decision-making as the top-ranked features resulted from three feature selection FS methods are orientation related features.

4- Identification of Acceleration sensor data is able to make a satisfactory decision about a sheep's lameness status without extra energy spend for collecting gyroscope data from the mounted sensor around the sheep neck.

5- Implementation of a genetic algorithm (GA) for feature optimisation and selection reveals competent results compared to other FS techniques such as ReliefF and Random Forest RF.

6- Proposing a method for model validation which is named 'Single Sheep Splitting' that guarantees a proportion of data movement from every single sheep in a dataset to be included in the training and testing set to provide acceptable validation results when compared with to 5-fold and 0.3 hold-out validation methods.

7- A unique user interface application (SLDM) has been designed for the purpose of this thesis. The designed software is enabling the developer to interact, alter the input parameters, and retrain the model as many times as required until acceptable prediction results are achieved.

2 Chapter Two: Multidisciplinary Literature Review and Background

2.1 Introduction

Previous studies in lameness in animals have utilised different types of data collection and data analysis methods, which have been applied in various ways for either animal's illness detection such as lameness or classifying behaviours. Literature studies relating to lameness detection are quite diverse because of the multidisciplinary nature of these research studies. Utilising Computer Science concepts of data mining for knowledge discovery provides a beneficial solution for animal welfare problems such as lameness in sheep. Although the problem of lameness in cattle has been widely addressed and studied, there is a paucity of research to identify sheep lameness in its early stage via using wearable sensors to collect important information that may help to tackle the problem. The divergence in literature could be manifested in Data collection methods, Data analysis techniques, Analysis purpose, and even Target animal. The structure of the reviewed literature is illustrated in Figure 2-1; however, this thesis follows the pathway where boxes with red boundaries appear towards detecting lameness in sheep.

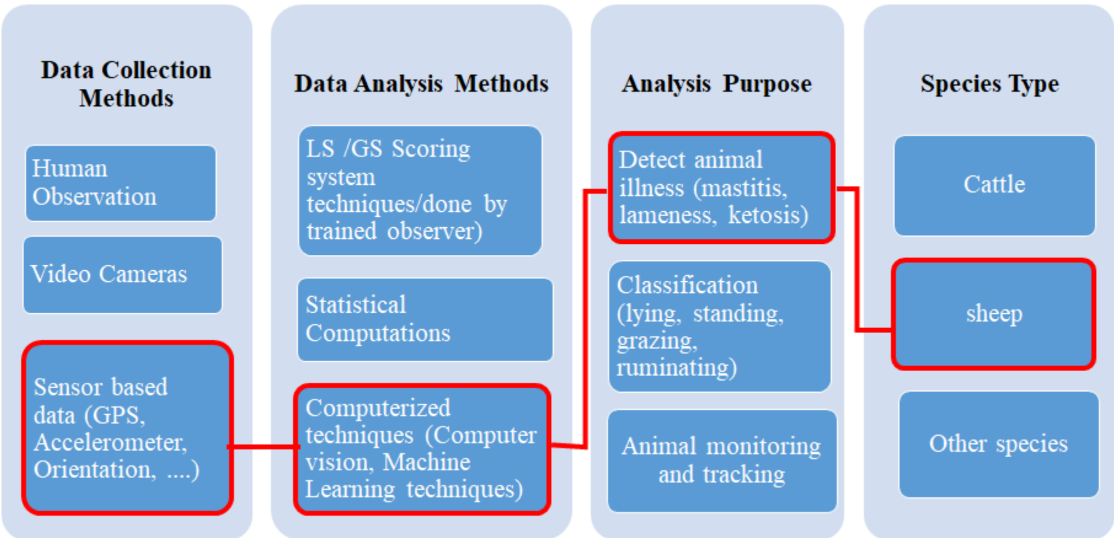


Figure 2-1 Research diversity in data collection, analysis methods, purpose, and target animal.

Originally, lameness has been directly detected using a manual/visual scoring system (Section 2.2.1). However, progress has been made for such lameness detection systems to be worked automatically without human interaction. Therefore, more objective methods for automatic scoring to indicate lameness have been suggested to measure both kinetic (the study of the force in motion) and kinematic (the study of changes in the body's position segments over time (Flower and Weary, 2009). Kinetic can be managed by extracting force information applied to lame limbs and measuring the ground reaction caused by infected hooves, while kinematic principles include assessing specific body changes in respect to time interval using an automatic measurement system (Viazzi *et al.*, 2014; Ramanoon *et al.*, 2018).

2.2 Lameness Detection in Cows

In cattle, the main signs of lameness are identified by Van Nuffel *et al.*, (2015b), who relate the indications for lameness to the changes that are happened in either *animal posture* (back arch posture or body movement pattern), *animal gait* (step overlap, stride duration, and walking) or *animal behaviour* (lying, resting and standing time). In their review, although these changes refer to cows rather than sheep, the collected information could be quite useful to differentiate between lame and non-lame sheep.

Various combinations of automatic kinetic and/or kinematic approaches that have been applied for lameness detection (in respect to cows) are explored starting from Section 2.3. However, lameness detection approaches can be divided in many different ways. One classification could depend on the assessment methods used for lameness detection to be into direct, kinetic, and kinematic approaches (Alsaaod *et al.*, 2019; Ramanoon *et al.*, 2018). Conversely, lameness detection approaches could also be classified according to how the animal's gait or posture-related information is obtained, i.e. according to data collection tools. Hence, the present research planned to follow the latter mentioned classification for lameness detection approaches illustrated in Figure 2-2. The aforementioned approaches are explored in the following sections by mentioning advantages and drawbacks of these approaches in comparison to the motion-based sensor methods which are used to collect data for this research study.

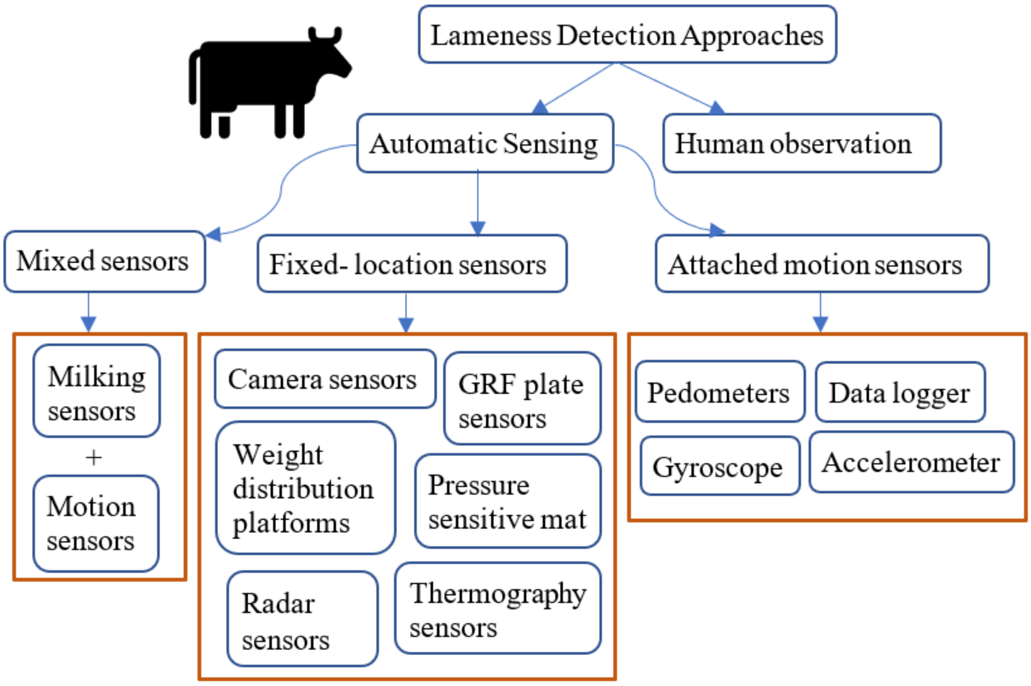


Figure 2-2 Lameness Detection (in cow) according to data collection methods.

2.2.1 Human Observation Approaches

The scale description of lameness varies from mild lame, moderate lame, lame, to severe lame (Helwatkar *et al.*, 2014). Therefore, the assessment of lameness has to be recorded for further analysis and proper treatment action to be taken afterwards. The traditional way of recording the scale of lameness within the flock was done by the trained observer, skilled veterinarian, or agricultural consultant. This scaling method takes enormous effort, is very time-consuming; especially when the whole flock needs to be observed (Wang *et al.*, 2018), and tends to be subjective due to different points of view among observers (Blackie *et al.*, 2011; Van Nuffel *et al.*, 2015b). Basically, two subjective assessment methods for rating the scale of lameness of the individual sheep were used, one is called the Numerical Rating System (NRS), which scales the lameness degree from 0 to 5 points; where '0' represents the non-lame sheep, and '5' represents the severely lame sheep, while the second method is named the Visual Analogue Scale (VAS) that uses a 10 cm line; '0' corresponds to a healthy sheep, with '10' corresponding to a painfully lame sheep (Flower and Weary, 2009; Welsh *et al.*, 1993). Nevertheless, the subjective method for scoring lameness can be implemented with no technical equipments and could suit in-farm assessment; it lacks reliability since it follows the observers' experiences and their biased nature, in addition to the changeable score over time (Flower and Weary, 2009). Therefore, automatic objective kinetic /or kinematic methods have been combined with original direct gait scoring methods; used as a reference standard for system validation, to produce an enhanced automatic lameness detection system.

2.3 Automatic Sensing Approaches

In the last two decades, several sensor technologies have been sought to apply health monitoring due to their reliability and sensitivity compared to the traditional subjective methods (Van Nuffel, *et al*., 2015b). Moreover, the sensors that have been used for remotely monitoring animal's physical behaviour or movement are favourable because of their small size, weight, low cost, and their ability to record behavioural data for an extended period (Vázquez Diosdado *et al.*, 2015; Moreau *et al.*, 2009; Helwatkar *et al.*, 2014). Although the initial monitoring sensing research studies have been applied in the military field, many other global issues are also examined. For example, an indication of a natural disaster, lack of non-sustainable resources alarms, and detection of health monitoring disease, including animals (Helwatkar *et al.*, 2014).

A developed monitoring system for cattle has been intensively reviewed by Rutten *et al.*, (2013) who present sensors which support health management on dairy farms at four levels according to what extent these sensors inform the farmer. The first stage of such a monitoring system is represented by how the data have been collected from animals which means the sensor itself. However, the sensor-based raw data are still hard to understand, unless they are translated into a form that is relevant to cattle gait scoring (Van Nuffel *et al.,* 2015b). The second stage of such system is to use sensor-based information as inputs to the algorithms that provide information related to the individual animal's health (Rutten *et al.*, 2013); for example, such algorithms can identify the changes in the sensor data that are relevant to walking behaviour in order to detect lameness. The third stage is to utilise the output information from the algorithm with a combination of economic information and any other farmers to build a decision support model. The final stage of Rutten *et al.*, (2013) structure represents a final decision regarding the animal's health status as detected by the sensor, which is done either by a farmer or by the automated system itself.

Understandably, the following sections are divergent according to where the sensors are located with respect to the animal body. Thus, fixed location sensor approaches are reviewed in Section 2.3.1.1, while the attached motion sensors are presented in Section 2.3.1.2. Moreover, some methods for lameness detection employ both fixed and/or attached sensors which are explored in Section 2.3.1.3.

2.3.1.1 Fixed-location Sensor Approaches

The types of sensors that have been used for kinetic gait analysis are usually located in a fixed location on a farm; therefore, they are also called non-attached sensors because of their static location away from the animal's body, in contrast to attached sensors that will be discussed in Section 2.3.1.2. Despite the advantage that only a few sensors are required to monitor a herd, the availability of the data collected via this type of sensor is limited because of the short recording time intervals as well as the off-line reaction by farmers (Helwatkar *et al.*, 2014). In real-time applications when the decision needs to be taken, it may not be a useful way to detect lameness. Moreover, these types of non-attached sensors might be impractical in the case of sheep as they left on the farm without shepherd interaction for a more extended period than cows. However, the followings sub-sections will explore some of the research studies that employ static sensors for their data collection process to detect kinetic/kinematic measurements relating to lameness in cattle.

2.3.1.1.1 Camera Sensors

Due to the subjectivity; long time; and effort needed by the observers, surveillance cameras could be an alternative as their features include: continuous recording without human interference besides its objectivity, being less time consuming, and economic (Poursaberi *et al.*, 2010). Cameras have been used as a fixed location sensor that is installed in a specific location on farms to continuously record video footage which is subsequently analysed via computer vision techniques for the sake of lameness detection. The gait characteristics that have been tested and analysed with computer vision techniques in relationship to lameness detection are: back arch curvature (Poursaberi *et al.*, 2010; Viazzi *et al.*, 2013; Viazzi *et al.*, 2014; Van Hertem *et al.*, 2014), body movement pattern BMP (Poursaberi *et al.*, 2011), step overlap (Song *et al.*, 2008; Pluk *et al.*, 2010), hoof release angles (Pluk *et al.*, 2012), variations in the hip joint during walking (Abdul Jabbar *et al.*, 2017), or leg swing (Zhao *et al.*, 2018). Table 2-1 shortly explores the computer vision approaches that have been exploited to detect lameness in cattle.

Table 2-1 Cow lameness detection research studies based on computer vision approaches.

References	Sample size	Data collection tool	Observed features	Analysis methods
Song et al., (2008)	15 lactating cows	AVI video, locomotion scoring	walking cow's hoof locations (step overlap)	Vision Analysis/ validated with manually locomotion scoring system
Pluk et al., (2010)	85 lactating cows	side-view videos, gait scoring	step overlap, body size, and relative step	Computer vision techniques to find the correlation between GS and step overlap
Pluk et al., (2012)	75 lactating Holstein cows	AVI video, pressure-sensitive mat, visual LS	touch and release of hooves angle and range of motion.	Vision Analysis/ gait scoring
Poursaberi et al., (2010)	28 lactating Holstein cows	JPEG images are extracted from AVI video	back posture	Image analysis techniques + statistical analysis (back posture analysis)
Poursaberi et al., (2011)	1200 cows	RGB image	back posture, head position	Image processing/ Body Movement Pattern (BMP) algorithm
Viazzi et al., (2013)	90 cows	Video recordings, visual LS	back arch curvature	BMP algorithm's output classifies into 3 classes by decision tree
Viazzi et al., (2014)	273 cows	2D (side view), 3D (top view) camera, visual LS	back arch, the position of the muzzle	BMP, decision tree learning to classify into a lame and non-lame cow
Van Hertem et al., (2014)	186 cows	3D-camera, LS	back arch, four individual consecutive BMP measurements	Multinomial logistic regression model improves Viazzi et al., (2014) to optimise the classification rate
Abdul Jabbar et al., (2017)	22 Holstein Friesian dairy cows	3D depth video camera, LS	variations in the hip joint during walking, gait symmetry, spine and hind limbs movements	Linear Support Vector Machine (SVM)
(Zhao et al., 2018)	98 cows	Side view video camera, LS	6 features of swing leg motion.	Decision Tree Classifier
(Jiang *et al.*, 2019)	30 dairy cows	video camera	pixel distribution characteristics of each frame image (10 videos for the lame cow, 6 videos for sound cows)	double normal distribution statistical model

Despite the extraction of the features that seemingly relate to gait variables from computer vision techniques that have been investigated by many authors, the implementation of computer vision techniques on the farm is still facing challenges (Hertem *et al.*, 2014). Since surveillance 2D cameras have some limitations due to installation space on the farm that is needed for the side view (Van De Gucht *et al.*, 2017). In addition, the final image might be affected by many factors; for example, lighting conditions and mixed background (Poursaberi *et al.*, 2009; Van De Gucht *et al.*, 2017). On the other hand, the surveillance 3D cameras could solve the problem of 2D continuous changing background and shadows, and overcome the restriction of space required for 2D cameras; however, 3D cameras' field of view is smaller, gait variables to be measured are fewer, and they are more sensitive to the natural light (Viazzi *et al.*, 2014; Vázquez-Arellano *et al.*, 2016; Abdul Jabbar *et al.*, 2017). Ultimately, surveillance cameras might not be practically implemented in the field of detecting lameness in sheep as sheep are left unattended for long periods of time while cows can be monitored at least twice a day when they are milked.

2.3.1.1.2 Ground Force Plate Sensors

Preliminary work on lameness detection in cows was undertaken by Rajkondawar *et al.*, (2002) who propose a walk-through scale system to measure the Ground Reaction Force (GRF) as an indicator for lameness. The GRF system consisted of two parallel force plates each of which with four load cells. When the cow is walking over the parallel plates, the peak of GFR is measured, and seven variables on both left and right hind legs are calculated to identify between a lame and non-lame cow. Then, a later study developed by Tasch and Rajkondawar, (2004) introduces an enhanced algorithm to eliminate the gates causing congestion when a group of cows are walking through the proposed system. More developments are made by Rajkondawar *et al.*, (2006) to their previous model which included adding extra variables to the prior calculated peak of GFR, such as average GFR, stance time, impulse, the area under Fourier transformed curve of GFR which could help to distinguish the lame cow from the non-lame one. Even more, developments have been made to the previous system to measure the GFR in 3 dimensions (Dunthorn *et al.*, 2015) in which sensitivity and specificity are noticeably enhanced.

A further study was conducted by Pastell *et al.*, (2008) who introduced a mat with electromechanical film could be set in any passage within the cow’s walkway. This proposed mat would identify the leg that has a problem by detecting the dynamic of different force-time behaviour where step force is calculated in addition to the stance time. The proposed electromechanical film mat (Emfit) could overcome the drawbacks of the earlier GRF model where the measurements over time were not considered, and the ability to detect the individual lame limb was not addressed. A broader study of gait patterns that used two parallel 3-dimensional force plates to differentiate the gait patterns for sound and lame cows after claw trimming was done by Thorup *et al.*, (2014). Their study reveals that lame cows would display less left-right vertical leg symmetry than healthy cows. Although the study employed a small number of animals, the study potentially provided a base for lameness detection compared with the 1D force plate of (Rajkondawar *et al.*, 2006).

2.3.1.1.3 Weight distribution Platforms Sensors

An initial study that advocated the idea of measuring the weight distribution of the cow's leg as a method for lameness detection was done by Pastell *et al.*, (2006). Strain gauge balances were installed into the milking robot where the weight of each limb was calculated separately using load cells. Several measurements were calculated; for instance, the average and the total weight, each limb's weight variation, the number and the frequency of kicks, and the total time in the milking robot. The primary results of this earliest study illustrate that the limb with foot disorder could be detected (Pastell *et al.*, 2006). Furthermore, Pastell and Kujala, (2007) developed a 4-balance platform on the floor to measure each leg weight separately during milking as a way to automatically detect lameness. The authors' expert model of Probabilistic Neural Network (PNN) is used for a classification task with two layers, one is a radial basis layer and the second is a competitive layer. Most of the cow's legs that have a problem are detected with 1.1% alarm error rate. The results show that there is a change in weight distribution between limbs belonging to a lame cow. The aforementioned system can be used with an Automatic Milking System (AMS) on the farm to help the farmer in decision making for better treatment (Pastell and Kujala, 2007).

Unlike 4- balance platform, a platform outside the automatic milking system was investigated by (Neveux *et al.*, 2006) for weight distribution over four legs of the cow. This platform contained four recording units with two load cells to measure how the cows distribute their body weight over the four legs while standing on different surfaces (rubber and concrete). Their study concludes that the measurements of weight distribution might present useful on-farm techniques for the detection of lameness. A later study by Chapinal *et al.*, (2010), used the 4-balanced sensor for the weight distribution of Neveux *et al.*, (2006) with the combination an IceTag accelerometer attached to the right hind leg. The study shows more associative factors that tend to connect to the occurrence of lameness such as standard deviation SD for rear and front legs weight, walking speed, and daily activity; step counts, laying/standing time and its duration.

In addition to the previous predictors, a promising tool has been indicated for lameness assessments by Chapinal and Tucker, (2012), who utilise the weighing platforms to automatically calculate the frequency of steps for front and rear leg pairs. The Logistic

Regression model used suggests that the steps per minute with the rear legs for the lame cows are more than the non-lame ones. Similarly, a simulation study that used Artificial Neural Networks (ANNs) model to classify lame and healthy cows was done by Gupta *et al.*, (2014) where the four-balanced system was used to measure how the weight was distributed for each leg. Their simulation results show the model could predict the cow's health status with more than 80% accuracy rate depending on the body weight distribution.

By comparing the use of weight distribution platforms for detecting lameness in sheep, it may be considered inappropriate tools to collect movements due to sheep has to be led to the sensing area where the platforms installed while sheep are normally left in farm out of control with less monitoring period than cows.

2.3.1.1.4 Pressure Sensitive Mat Sensor

The first pressure-sensitive walkway called Gaitwise system was developed by (Maertens *et al.*, 2011). Away from human interference, Gaitwise system automatically measured kinematic variables of the cow's gait twice a day after milking. The pressure mat provides spatio-temporal data besides the force information of two complete gait cycles while the cow walks through the sensing area. The data were collected on the farm in real-time and evaluated using Linear Discriminant Analysis (LDA). The results showed that asymmetry of variables in step, length, stance time, step time, and step width which leads to further research on lameness detection in cattle. Nevertheless, the measurement success rate was over 80%; it is mostly associated with cow movements and behaviour such as irregular cow traffic due to external factors.

A follow-up study was undertaken by (Van Nuffel *et al.*, 2013), who tested the asymmetricity of gait variables which were repeatedly produced by the Gaitwise system (Maertens *et al.*, 2011) as a high potential indicator for early lameness detection within cows. In addition to the prementioned variables that are produced by Gaitwise system, their tested results show the fluctuation of stride to stride is also a very sensitive indication prior to lameness. Generally, their promising results could differentiate between the lame and severely cow and defines which leg starts to be lame (Van Nuffel *et al.*, 2013). However, the pressure mat like the Gaitwise system provides detailed sensing information; such a system may be impracticable to adopt due to its high cost and its demand for free space to be installed on-farm.

Therefore, a developed simulation approach has been made by Van De Gucht *et al.*, (2017).

Their proposed simulation study downscales the length of the measurement zone to be no more than 3.28 meters to monitor one complete gait cycle without a huge loss of collected data, whereas the size of the individual sensing must be at least 2.58 *10-3 m^2 to overcome the difficulty of imprints recognition. The idea of reducing the cost and the space needed for the Gaitwise system that was previously developed by Maertens *et al.*, (2011), avoids too much loss in gait variables which relate to lameness detection. The accuracy of lameness detection is not decreased when the LDA is applied to classify the cow into; sound, lame, and severely lame; however, the enhanced Gaitwise system can misclassify some lameness classes (Van De Gucht *et al.*, 2017).

2.3.1.1.5 Infrared Thermography Camera Sensors

Infrared thermography is used as a non-invasive tool to detect foot lesions in cattle which may lead to an indication of lameness when the case of inflammation occurs in a lame limb (Schaefer and Cook, 2013). Here, no gait changes are monitored, the foot temperature alteration is captured instead. Due to the changes in blood flow in vessels, the increased temperature that is emitted from the skin surface might be a sign of a foot problem (Alsaaod *et al.*, 2015). The emitted infrared radiation is measured and displayed in a pictorial form which is called a thermogram where each pixel refers to the measured surface temperature of an object (Turner, 1991).

Applications of thermography to detect foot lesions are clinically reviewed by Alsaaod *et al.*, (2015), who evaluate the performance of those techniques to the benefits of lameness management in cattle. For example, the temperature of the coronary band of an affected foot and healthy one is compared in (Alsaaod and Büscher, 2012) whereas the increased temperature in association with a foot lesion is investigated by Wood *et al.*, (2015); Stokes *et al.*, (2012). Similarly, other foot lesions such as white line disease (WLD), sole ulcer (SU) and digital dermatitis (DD) are also studied with infrared thermography imaging, and the results show a linkage between the changing in temperature and presence of foot lesions (Orman and Endres, 2016).

In the same way, such a high level of lameness is recorded by using thermal imaging techniques in contrast to a subjective lameness scoring method. However, the infrared thermal methods are costly; it is worth being applied when it is compared to consequences as severe lameness stages progress (Renn *et al.*, 2014). Although infrared thermography techniques could be a reliable method used to detect the skin temperature of an affected foot which is linked to lameness, it is strongly affected by environmental changes such as air temperature, humidity, debris and dirt (Alsaaod *et al.*, 2019). Thus, for the same pre-mentioned reasons, infrared thermography might be not suitable to be applied in a flock of sheep due to their mobility away from the farmers' attention.

2.3.1.1.6 Radar Sensing

A recent study by Shrestha *et al.*, (2018) utilises radar micro-Doppler signature data that has been previously used for human detection (Kim *et al.*, 2015) to detect lameness in the horse, cattle, and sheep. Five cows were tested while individually walking through a narrow corridor from both anterior and posterior views. The features of mean and SD for centroid and bandwidth of micro-Doppler signature were measured to be classified via SVM and KNN supervised learning classifiers. The results for cow achieved an accuracy of 85% (Shrestha *et al.*, 2018). However, further comprehensive analysis may be need by expanding feature space from micro-Doppler signature to enhance the classifier performance.

2.3.1.2 Motion-based Attached Sensor Approaches

The emergence in smart sensing technology in the section of animal welfare has started to be a promising, sustainable, and affordable choice for a considerable well-being system for animals. The smart sensing system can be clarified to such physical devices that all connect to a computer system for the purpose of data collection, data pre-processing, information exchange, and data analysis (Jukan *et al.*, 2017). Basically, the motion-based sensor technologies assist the application of automatically monitoring animals to determine either their physiological and/or behavioural changes which may have a significant relationship to a specified illness or even tracking animals to identify their locations on a farm via wearable sensors mounted on their body (Jukan *et al.*, 2017). Although the sensor device itself comes with a low price, concerns would be raised when the whole flock or herd would each need to be equipped with an individual sensor. Therefore, the overall cost of the project may increase toward building a completed monitoring system (Van De Gucht et al., 2017). Whereas the automated methods to control the farm bring many advantages to the farmer in terms of time spent, flock size increasing and sensitivity to detect the lameness (Blackie *et al.*, 2011); it is worth embedding such a monitoring system on a farm.

Mainly, mobile sensors attached to the animal's body either leg or collar may be more reliable than fixed location sensors. Sensor devices such as data logger, pedometer, or accelerometer are attached to different parts of an animal's body; especially leg and neck, for the sake of identifying changes in either *behaviour* or *gait posture* which might have a relationship with lameness detection. The data logger is an electric device which records voltage at a set interval. A voltage of 0 is inactive when the animal stands. Conversely, the voltage is set to be 2.5 when the animal lies down (O'Driscoll *et al.*, 2008). Alternatively, the pedometer is an electronic device (equipped with accelerometer) mostly attached to the leg; it calculates the number of steps and the daily activities (lying, standing, walking) taken by the animal (Arcidiacono *et al.*, 2017). Lately, accelerometer devices used for monitoring various behaviours; especially walking, have been adopted as a wearable device that can be integrated into a computer node of wireless networks. It can be defined as an electronic device that calculates the alternation in acceleration and force of an object and transmits raw data in either one, two or three dimensions (Alsaaod *et al.*, 2019).

The majority of research studies that have collected movement data via motion-based sensors to detect lameness are undertaken in cattle rather than a flock of sheep. However, the following sections will explore the approaches that detect lameness in cattle, referring to the sensor placement which is attached to either the leg, back or neck.

2.3.1.2.1 Leg attached sensors methods

- Pedometers:

The exploration of the usefulness of posture scoring for the locomotion process of cows daily activity via a pedometer attached to a hind limb was investigated first by (O'Callaghan *et al.*, 2003), who revealed that the lame cows have a lower level of daily activity compared to the sound ones. Another study using pedometers for activity measurements was carried out by Mazrier *et al.*, (2006). The average number of steps per hour was calculated as an indicator of lameness. It is noticed that 92% of lame cows decrease their activity several days before the clinical signs appear. However, not all cases of developing lameness could be detected in their study.

In addition to counting the number of steps as an indicator for lameness, the focus on monitoring lying behaviour has also been an interest of many studies for the sake of lameness detection. So, lying-down time, the number of bouts, duration, frequency, and SD of lying bouts have been measured via electronic data logger (Ito *et al.*, 2010; Solano *et al.*, 2016).

Further to the previous measurements, lying-down time around feeding time was explored by Yunta *et al.*, (2012). Most of the previous studies reveal that lame cows have more lying time and longer bouts than non-lame cows. However, examining lying behaviour alone is not optimal, unless combined with other features in order to detect lameness efficiently (Ito *et al.*, 2010) including risk factors associated with the lying behaviour of individual cows like lactation stage or environment (Solano *et al.*, 2016). Similarly, standing time is compared to the previously mentioned lying behaviour features to identify the characteristics of the lame limb of the cows, where the IceTag 3D logger device was attached to both rear legs in a pilot performed by (Kokin *et al.*, 2014). The statistical analysis resulting from the IceTag Analyser shows that the lame cow spent less time standing and had a lower activity rate than a sound cow which agrees with previous research studies.

A different analysis method for lying down behaviour of cows has been implemented by Alsaaod *et al.*, (2012), where ALT-Pedometer (Activity-Lying-Temperature) was attached to the foreleg of the cows. Six features of lying behaviour were extracted to feed SVM classifier where binary classification has been implemented to classify into a lame and non-lame cow with an accuracy of 76%. The results present that deviation from normal behaviour is a better indicator for lameness than justifying a threshold value to differentiate between the behaviour of a lame and non-lame cow (Alsaaod *et al.*, 2012).

A recent study employs the use of a pedometer to detect a lame cow by observing lying time, step count and swapping between standing and lying where the data is sent to a fog node to be analysed. Although the work of Byabazaire *et al.*, (2019) has been focused on reducing the amount of data exchange between the fog node and cloud, the indication for a lame cow is identified in 1-day prior to lameness occurring when Random Forest (RF) algorithm is used, while it is 3-days prior to lameness when the KNN algorithm is implemented. Table 2-2 presents pedometers and data loggers used for the indication of lameness in cows.

Table 2-2 Research studies used pedometer sensors for lameness detection in cows.

References	Sensor type/position	No. cows/	Observed features	Analysis methods
(O'Callaghan et al., 2003)	Pedometer/ hind leg	345	No. of steps/hr., milking time	statistical analysis via SAS software
(Mazrier et al., 2006)	Pedometer/ hind leg	46	No. of steps/ hr.	Computer graph is presented
(Ito et al., 2010)	Data Logger	1319 cows (28 farms)	Lying time, no. bouts, bouts duration, SD of bouts	SAS statistical software
(Yunta et al., 2012)	Data logger/right hind leg	10-15 cows from each 10 farms/ 1-min interval	Lying time around feeding, lying time, no. bouts, bouts duration	SAS statistical software
(Alsaaod et al., 2012)	ALT-pedometer/ foreleg	30 cows	Lying time, no. bouts, bouts duration, max & min bout duration, ambient temperature	SVM, with an RBF kernel
(Kokin et al., 2014)	IceTag3D™ logger /both rear legs	33 dairy cows / 16 Hz	Lying and standing time, no. of lying bouts, step count, motion index	SAS software
(Solano et al., 2016)	HOBO data logger/ hind leg	40 cows	Lying time, no. bouts and frequency, bouts time and SD.	Logistic regression
(Byabazaire et al., 2019)	Long-Range Pedometer (LRP)/ front leg	146 cows	Step count, lying time, swap between lying and standing	Random forest, KNN

➢ Accelerometers:

On the other hand, the accelerometer has been employed in research studies to investigate gait characteristics that refer to the occurrence of lameness such as variance in acceleration. Table 2-3 explores the research studies in this section with brief details. An implementation of wavelet analysis to acceleration measurements that were acquired via 3D accelerometers attached to each cow's leg was carried out by Pastell *et al.*, (2009). A higher asymmetry in the variance of forward acceleration over time is noticed in the hind leg of a cow since it was already lame. Similarly, Chapinal *et al.*, (2011) used four legs' which were attached to accelerometer sensors and one extra 3D accelerometer device was fastened around the torso to detect locomotion changes related to lameness. Their findings report that asymmetry of the variance of overall acceleration in both front and hind legs is increased together with overall gait and asymmetry of steps which are assessed visually.

Other acceleration measurements were investigated like root mean square, maximum, and minimum acceleration via a 3D accelerometer sensor attached to the back of the cow as a tool for lameness detection (Mangweth *et al.*, 2012). The prediction model has a success rate to differentiate between lame and non-lame of 91.7% percentage. However, the accelerometer attached sensor research studies may still be limited due to the equipment used that might affect the normal gait pattern if it is used daily (Mangweth *et al.*, 2012).

Again, lying behaviour is explored; however, an accelerometer is used instead of a pedometer. Thorup *et al.*, (2015) investigate the accelerometers data from an attached sensor to the cow's hind leg to record its activity for the indication of early signs of lameness. Principal Component Analysis (PCA) is used to measure the correlation among 13 leg acceleration's variables such as the number of steps and its frequency, the duration time of lying down, standing, and walking. The analysed results show that early lameness detection seems to be sensitive to walking acceleration and walking duration (Thorup *et al.*, 2015).

Furthermore, it has been found that the walking speed and standing bouts might be the best signs along with other laying behaviour activities even for slightly lameness detection in cattle with a sensitivity of 90.2% according to Beer *et al.*, (2016). In their study, a special 3D accelerometer device called RumiWatch was attached to the hind limbs and nose of cows to investigate the lying behaviour associated with lameness.

Although the aforementioned studies that have utilised motion sensors to investigate behavioural features (lying, standing) in relation to lameness deem these as a good indicator, the changes in gait activities (walking) are more precise (Kokin *et al.*, 2014); moreover, abnormal walking is an advanced symptom of lameness (Haladjian *et al.*, 2018) which priorly could be spotted more than behavioural changes. Therefore, Haladjian *et al.*, (2018) present a different technique for lameness detection via building a model for normal walking stride from sensor data attached to the cows' hind leg. Consequently, the abnormality in the walking pattern is detected as a deviation from the build model of one-class SVM classifier (SVM details in Section 0). However, the abnormality detection based on each cow walking pattern produces an individual measurement for sensitivity and specificity. This approach looks to have a higher energy consumption than a baseline model for a huge herd.

Table 2-3 Studies used leg attached accelerometer sensors for lameness detection in cows.

References	Sensor type/position	No. cows/ Sampling rate	Observed features	Analysis methods
(Pastell *et al.*, 2009)	3D accelerometer/ 4 limbs	11 cows / 25 Hz	The symmetry of variance for forward acceleration	Wavelet Analysis
(Chapinal *et al.*, 2011)	3D accelerometer/ 4 limbs + back	12 + 24 in 2 experiments/ 33.3 Hz	Acceleration symmetry in variance + Step symmetry and walking speed from video rercording	SAS statistical software
(Mangweth *et al.*, 2012)	3D accelerometer/ back	-	Acceleration RMS, Min, Max	Forcast prediction model
(Thorup *et al.*, 2015)	IceTag3D, IceRobotics, 3D accelerometer/ hind leg	348 Holstein cows/ 16 Hz	Duration of: laying, standing, walking, and total acceleration of each	PCA, to measure the association among variables
(Beer *et al.*, 2016)	RumiWatch 3D Accelerometers/ hind limbs + head (noise)	41 lame+12 sound/ 10 Hz	Duration of: lying, standing, eating, ruminating, bouts, stride, walking speed	NCSS8 statistical software, univariable logistic regression
(Haladjian *et al.*, 2018)	3 axes Linear acceleration, 3 axes orientation/ hind left leg	10 cows/ 100 Hz.	Deviation in cows' usual gait (detect abnormal walking patterns)	One-class SVM Classifier

2.3.1.2.2 Neck attached sensors methods

According to the literature in Section 2.3.1.2.1, the gait or locomotion characteristics (lying, standing, and walking behaviour) in relation to lameness have been investigated via a leg attached accelerometer sensor. On the other hand, the accelerometer device could be fitted with a collar around the neck to explore neck activities that relate to lameness as well as locomotion ones. However, neck activity may not give full details on lying, standing, or walking activity (Weigele *et al.*, 2018). Instead, a collar neck accelerometer might be a feasible alternative for lameness detection in commercial farms due to the ease in attaching it and is less likely to cause pressure sores or injuries (Nielsen *et al.*, 2010; Kokin *et al.*, 2014).

Furthermore, Mottram, (2012) refers to the reasons behind preferring a neck mounted sensor rather than a leg-mounted one to the possible feature of being used as a sensor node within wireless sensor networks for monitoring animals; consequently, the information could be transmitted to a base station easier than leg-mounted sensors (Mottram, 2012). Also, Mottram, (2012) clarifies in his patent that the leg-mounted devices are more likely to be dirty because of their close location to mud and the faecal area. In addition, the sensor' readings would be affected by the rotational leg's attached sensor beside its deploying difficulty due to the kicking behaviour of the animal. Also, neck attached sensors would probably cause less disturbance to animals and could be limited to moving or rotation while the animal's scratching or smashing (Andriamandroso *et al.*, 2017).

According to the literature, the majority pay attention to extracting behavioural pattern recognition that may be associated with lameness from accelerometer sensors attached to the leg or back. In contrast, neck activities are explored to investigate feeding behaviour or the estrus cycle in a cow (Vázquez Diosdado *et al.*, 2015; Barker *et al.*, 2018; Khanh *et al.*, 2016), grazing, eating, or ruminating behaviours of cows (Nielsen, 2013; Smith *et al.*, 2016; Rahman *et al.*, 2018; Tamura *et al.*, 2019). However, the neck activities relate to lameness via a neck collar fitted with an accelerometer sensor have been introduced in a few research studies (Table 2-4).

- Accelerometer

Earlier, Martiskainen *et al.*, (2009) utilised the accelerometer sensor within the neck collar of cows to develop a learning SVM classification model to differentiate between eight behavioural categories of the cows: standing and standing up; lying and lying down; normal and lame walking; ruminating and feeding. Lame walking behaviour could be predicted with a sensitivity of 65% and specificity of 66% (Section 0). However, further improvements need to be considered regarding sensor data quality and the high computational time of their selected approach to gain better classification accuracy.

A pilot study was introduced by Mottram and Bell, (2010) show that it is possible to relate neck movement to mobility score in an objective manner. A 3D accelerometer sensor around the neck was used for gathering automatic mobility measurements for a cow while walking a 20-meter path for a couple of minutes. The maximum values for 3 axes measurements were calculated in addition to the number of peaks in forward acceleration which exceeded SD above the Mean by one. Kurtosis is also measured for forward and vertical acceleration as it is a metric used for measuring the weight of collected data in tails of its histogram distribution (the high-frequency data points) (Cox, 2017; McNeese, 2016). Surprisingly, the pilot study showed that the most lame cows move their head less than the least lame ones. However, this may not be the case with sheep due to the different body mass of both animals.

A recent study to differentiate between the moderately lame and the sound cow which used accelerometers attached to different body parts of the cow was undertaken by Weigele *et al.*, (2018). Lying behaviour and locomotor activity are measured by attaching the accelerometer device to the hind leg of the cow, while the neck activities are investigated by attaching the accelerometer to the cow's neck collar. A statistical linear mixed-effect model is used to analyse the gathered data which reveal findings in line with previous studies. The moderately lame cows show a reduction in activity, longer lying time, fewer head activities, while no significant in the upright position of locomotion activity was noticed between moderate and sound cows (Weigele *et al.*, 2018). Thus, more investigation may need to be done to overcome the challenges of distinguishing between the early stages of lameness due to the benefit of early treatment action.

- Ear tags

As a different approach for lameness detection in cattle, an ear tag 3D acceleration sensor is used for data collection, this has been presented by Link *et al.*, (2015). The accelerometer in ear tags could be combined within an official ear tag on a cow to detect lameness, in addition to other behaviours such as heat detection and feeding. The result of such research has been shown in unpublished work (Link *et al.*, 2016), where nineteen features were extracted in the processing stage from the magnitude value of the accelerometer on three axes. The acceleration data were gathered into two datasets referring to each sensor sampling rate 1Hz and 10 Hz denoting lame and non-lame dataset, respectively. The target of the study of Link *et al.*, (2016) was to detect lameness within a '4 day period' or before the '4 day period' (day of detecting lameness +3 days prior). The AUC is performed to evaluate the classifier used; this showed that the AUC value for the 1Hz dataset (AUC=0.88) was higher than 10Hz dataset (AUC= 0.71). Moreover, the best result for both datasets was obtained when the SD, 25% quantile, and kurtosis features were pre-processed.

Table 2-4 Studies used neck attached accelerometer sensors for lameness detection in cows.

References	Sensor type/ position	No. cows	Sampling rate	Observed behaviour	Analysis methods	Model Accuracy
(Martiskainen *et al.*, 2009)	3D Accelerometer / neck collar	30	10 Hz	standing, lying, ruminating, feeding, the normal & lame walking, lying down, standing up	Multi-class SVM	Lame walking Sensitivity = 65%, specificity = 66%
(Mottram and Bell, 2010)	3D Accelerometer / neck collar	20	50Hz	Max for 3 axes, max peak for forward acceleration, kurtosis for forward and vertical acceleration	Correlation	Mobility score correlates with acceleration measurements

(Link *et al.*, 2016)	3D head Acceleration/ ear tags	70	10Hz, or 1Hz	19 features: max, min, range, inter-quantile range, ...	AUC	AUC =0.88 (1Hz), AUC= 0.71 (10 Hz)
(Weigele *et al.*, 2018)	3D Accelerometer / leg and neck + noiseband sensor RumiWatch	17	1 Hz.	Lying behaviour, locomotor activity, and neck activity + feeding and rumination behaviours	linear mixed-effects statistical models in R	Correlation exists

2.3.1.3 Fixed Milking Sensors in Combination with Other Sensors

Since the cow has a daily routine of milking and feeding, the existing milking sensor that is already installed in commercial farms can be utilised to detect lameness where the cow passes through regularly. So, a combination of fixed location sensors (milk sensor) and motion sensors (pedometer, accelerometer) for data collection has been introduced in several research studies.

A validated research study was implemented by (Van Hertem *et al.*, 2013) which draws the attention to utilise the existing sensor data to detect lameness. In their research, the night to daytime behavioural data were measured; such as daily milk production, neck activity ratio, and ruminating time. The measured data were used to build a Logistic Regression model to classify cows into lame and non-lame classes with a performance accuracy of 86%.

In addition to the milk meter sensor, weight scale and pedometer sensors were used by Kamphuis *et al.*, (2013) to measure the animals live weight, activity, and milk yield respectively. The authors enhanced a boosting technique based on the Additive Logistic Regression method in combination with regression tree. Although the prediction performance of the developed algorithm was not high enough, it has been shown that the multivariable sensors (three prementioned sensors) outperform a single sensor (univariable) in lameness detection.

Another study combines data from a concentrate feeder robot in addition to the activity sensor and automatic milking sensor to build a dynamic linear model for lameness detection (de Mol *et al.*, 2013). This model detects the changing activity on a daily basis which could be a useful tool for day to day management. Since the data being collected from an automatic milking robot is too large, Garcia *et al.*,)2014) investigated the Partial least squares discriminant analysis method to distinguish the two classes, lame and non-lame. As it is suitable to be applied to multivariate data points where many variables relate to each other; however, none of them could be a single effective indicator for lameness detection. Naturally, the changes in milk production or feeding behaviour of a lame cow might appear after the lameness has developed; therefore monitoring the changes in gait would be more effective than monitoring the changes in milk production for lameness detection (Haladjian *et al.*, 2018). Table 2-5 briefly summarises milking sensor-based approaches including other sensors for collecting data.

Table 2-5 Cow milking sensor in combination with other sensors for lameness detection.

References	Sensor type	Measured data	Analysis method	Model accuracy
Van Hertem et al., (2013)	Milk meter, neck collar tag, ruminating logger	daily milk production, neck activity ratio, ruminating time	Logistic Regression	86%
Kamphuis et al., (2013)	Weigh scales, pedometers, milk meter	animal live weight, activity (via pedometers), and milk yield	Additive logistic regression + regression tree	80% specificity
de Mol et al., (2013)	Automatic milk sensor AMS, activity sensor	Milk feeding amount activity, and milk production	dynamic linear model	85.5%
Garcia et al.,)2014)	automatic milking system (AMS)	More than 30 data point measurements from a milking robot	Partial least squares discriminant analysis	80%

2.4 Lameness Detection in Sheep

In contrast to the previous sections where lameness detection in cattle is intensively reviewed, this section presents the few available research studies where the lameness in sheep was the objective. Figure 2-3 shows the related existing works for lameness detection in sheep which are quite recent. Thus, the indication of lameness in sheep is challenging, and a quite on-demand research topic, the field of knowledge is low for such research studies. The next sub-sections introduce an overview of the currently available research studies that focus on identifying lameness in sheep (see Table 2-6).

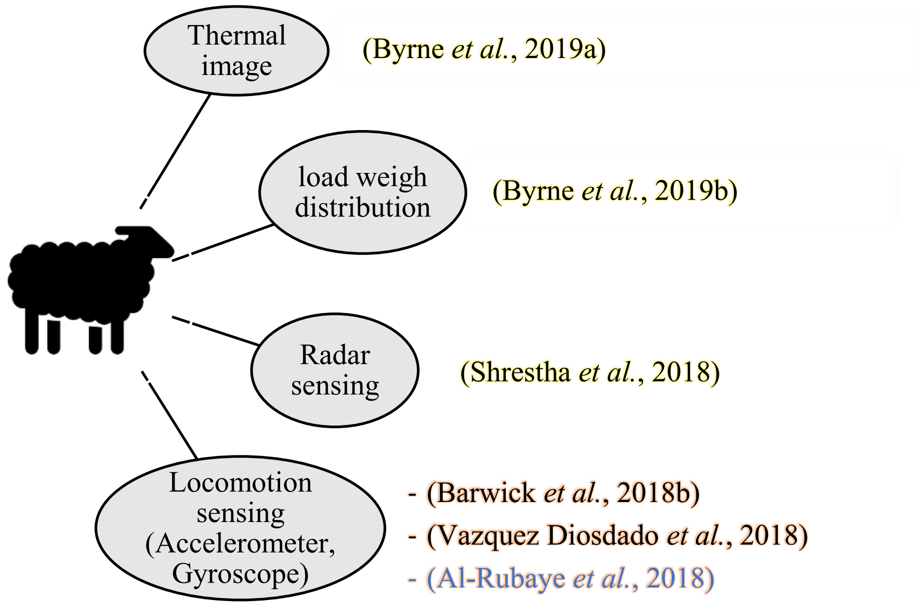


Figure 2-3 Sheep lameness detection approaches applied yet.

2.4.1 Infrared Thermography (IRT)

The feasibility of using thermal imaging for identifying lameness in sheep was investigated Byrne *et al.*, (2019a), who applied three experiments to quantify each hoof's temperature, determine the relationship between the hoof and ambient temperature, and to validate the utilising of IRT to detect infected hooves in sheep. From the experiments, it is noticed that the ambient temperature has no impact on the maximum temperature of the infected hoof (Byrne *et al.*, 2019a). Furthermore, the sensitivity reached 92% when a pre-defined threshold of greater than 9 ^{0}C was defined for the infected hoof (above the mean hoof temperature of five coldest other hooves in the flock). However, the sensitivity declined to 77% when the same threshold was applied with a validation dataset. The IRT may need equipment installation and contract with the aim of the research as early detection of lameness is the objective.

2.4.2 Load Cells Weight Platform

The relationship between the health status of sheep's hooves and the load that each hoof distributes was examined by Byrne *et al.*, (2019b). The ability of a custom hoof weight crate (HWC) is used to measure contralateral load percentage for each pair of hooves (front and back). It is revealed from the applied statistical liner mixed model that the infected hooves (back or front) carry the same load in contrast to the healthy hooves at the same extent where the front hoof carries 60% of weight compared to the back hoof which carries 40% of total weight (Byrne *et al.*, 2019b). The HWC could be useful to detect lame sheep; however, this process may not serve the objective of this research where the unattended way of monitoring is required.

2.4.3 Radar Sensing

As mentioned in Section 2.3.1.1.6, the study of Shrestha *et al.*, (2018) where the radar micro-Doppler signatures are utilised to detect lameness for 5 cows, the sheep are also tested in their study to detect lameness. However, the test was more challenging as sheep are social animals and like to accompany their mates from nose to tail when walking. The hind limb is focused on by the radar signature to extract features for supervised learning classifiers (SVM, KNN). Centre of mass and the intensity of the signature are obtained by calculating mean and SD, respectively (Shrestha *et al.*, 2018). In spite of the results which reveal accuracy around 99% for sheep, the number of sheep included in the experiment were only six, that was divided into three healthy sheep and three lame ones. Moreover, the calculations of sensitivity and specificity were not clearly mentioned in the study.

2.4.4 Locomotion Sensing

Unlike previous approaches, locomotion sensing or rotational movement approaches via accelerometers or gyroscopes respectively has been investigated by only a few researchers. The sensors devices are attached to the sheep's body to acquire data about their movements or neck activities which might relate to lameness detection. Barwick *et al.*, (2018b) attach a 3 axes accelerometer to three different locations on the sheep's body: collar, leg, and ear. The movement's data has been analysed using Quadratic Discriminant Analysis (QDA) where an epoch of 10 seconds is subject to extract the selected movement metric features (3 out 14 using Random Forest) and then analysed. The prediction accuracy yield from the experiment for each deployment is 82%, 35%, and 87% for ear, collar, and leg, respectively. The authors applied the lameness simulation by restricting the sheep's front right leg with an adhesive bandage which might not be identical to a lame sheep's real movements. The accuracy of the model could be affected when the model is tested with a real dataset of sheep movement.

Similarly, Vazquez Diosdado *et al.*, (2018) use 3 axes accelerometers in addition to a 3 axes gyroscope to obtain data from two different locations of sheep ear and collar. An initial result of a classification algorithm for lameness detection is presented. However, the evaluation of sampling frequency (8 Hz, 16 Hz, 32 Hz), window size (3s, 5s, 7s), and sensor position (ear, collar) was conducted by the same research team (Walton *et al.*, 2018) for classifying sheep behaviour into lying, standing, and walking. Many ML techniques are applied to classify sheep into lame or non-lame, the best performance is achieved with Random Forest algorithm with a total accuracy of 68.6% and sensitivity of 78.3% (Vazquez Diosdado *et al.*, 2018).

The current research output (Al-Rubaye *et al.*, 2018) utilises 3D acceleration, 3D orientation, and 3D linear acceleration sensors attached to the sheep's neck. The data are retrieved from sensors at 10 Hz sampling rate. The study aims to determine the best accuracy among various supervised machine learning techniques which can predict the early signs of lameness while the sheep are walking on a flat field. The experimental results show that the DT outperforms other classifiers with an accuracy of 75.46% and a sensitivity of 82.87%, 48.78%, and 87.31% for severely lame, mildly lame, and sound respectively. The experiment also reveals that the orientation sensor data (angles) around the neck are the strongest predictors used to differentiate the three classes of sound, mild, and severe lame.

Table 2-6 Sheep lameness detection research studies.

References	Data collection tool	no. of sheep	Examined location	Observed behaviours/features	Analysis tool	Sensitivity
(Byrne *et al.*, 2019a)	Thermal images, locomotion scoring	9 ewes (30 images)	Front, back hooves	Max, average hoof temperature	SAS statistical analysis	77% (detect infected hoof)
(Byrne *et al.*, 2019b)	Load cells for four individual hoof platforms	20 ewes (lame and sound)	4 hooves	Individual hoof weigh	ASReml statistical package	-
(Shrestha *et al.*, 2018)	radar micro-Doppler signatures	6 sheep (3 sound 3 lame)	Look at hind limbs	Mean of the centroid (centre of mass of micro-Doppler signature), sheep velocity	KNN, SVM	99 %
(Barwick *et al.*, 2018b)	3-axis accelerometer (12 Hz)	10 sheep	Neck, leg, ear	walking, standing, grazing, and lying for (sound and lame)	MatLab, R (QDA)	Lame walking accuracy (ear 98%, collar 83%, leg 96%)
(Vazquez Diosdado *et al.*, 2018)	3-axis accelerometer 3-axis gyroscope	19 sheep	Neck, ear	Lame and non-lame while walking	Microsoft Azure Learning Studio (RF)	78.3%
(Al-Rubaye *et al.*,	3D accelerometer	7 sheep	neck	Sound, lame, and severely lame sheep	MatLab (Decision	Sound 87.31%,

2018)	3D orientation 3D linear accelerometer				tree)	mildly lame 48.78%, Severely lame (82.87%),

2.5 Inertial Measurement Unit (IMU) Sensors for Behavioural Classification

Utilising sensors in livestock farming is widely applicable to dairy cattle and sheep. Recently, monitoring livestock animals on an individual basis might be the main interest of researchers rather than herd/flock management due to its important contribution in developing Precision Livestock Farming (PLF) and farm management applications. Several research studies investigate the use of Inertial Measurement Unit (IMU) sensors for their data collection to retrieve information about standard monitoring system behaviours in both cattle (Smith *et al.*, 2015; Smith *et al.*, 2016; González *et al.*, 2015) and sheep (Haddadi *et al.*, 2011; Walton *et al.*, 2018; Guo *et al.*, 2018). More details are listed in Table 2-7.

IMU can refer to a combination of motion sensors (accelerometers, gyroscope, magnetometer) and location sensors using the GPS that offers an advantage of reading variables from all sensors' type at the same time (Andriamandroso *et al.*, 2016). For example, a magnetometer has been used to detect feeding behaviour in cattle, sheep, and goats by monitoring jaw movements (Mulvenna *et al.*, 2018), while GPS has been utilised to track animals to estimate their distance travelled (Knight *et al.*, 2018), to monitor cow grazing behaviour (James *et al.*, 2016), or to classify different cows' activities (Godsk and Kjærgaard, 2011; de Weerd *et al.*, 2015). Furthermore, GPS or magnetometers might also be combined with accelerometers to derive animal behaviour patterns which help to detect livestock illness and welfare concerns (Bailey *et al.*, 2018). Refer to Table 2-7 to review the research studies using motion sensors/GPS for classification of livestock behaviour.

Table 2-7 IMU sensors for livestock behaviour classification.

References	IMU Sensor type	Animals	Sensor location	Classified/ observed behaviour	Classifier used
(Umstätter et al., 2008)	GPS, tilt sensor (pitch, roll)	10 sheep (2 sites)	neck	Active and inactive	LDA, classification tree, developed DT
(Guo et al., 2009)	GPS, 3-axis accelerometer, 3-axis magnetometer	6 cows	Neck collar	Describe animal movements and transition behaviours	HMM, long-term prediction algorithm
(Moreau et al., 2009)	GPS, tri-axial accelerometer	26 goats	chest belt, dog harness, neck collar	Resting, eating, walking	'Animstat' custom-designed c++ software tool
(Haddadi et al., 2011)	GPS, 3-axis MEMS accelerometer, 3-axis MEMS gyroscope, 3-axis magnetometer	46 sheep	harnesses on sheep	Spatial-temporal(time & distance) patterns associated with social network	K-means clustering
(Mason and Sneddon, 2013)	3-axis accelerometer (in sensor node)	4 ewes	head/ neck	Foraging behaviour (grazing, standing, browsing, … etc.)	PCA to assess the accuracy of assigned behaviour to a given group
(Smith et al., 2015)	GPS, 3-axis MEMS accelerometer (10 Hz), 3-axis magnetometer	10 cows	Neck collar	Grazing, walking, ruminating, chewing, resting, head down, and others	LDA, NB, binary DT, one-vs-one SVM
(Dutta et al., 2015)	3-axis accelerometer 3-axis magnetometer	24 cows	neck	Grazing, Ruminating, Resting, Walking, others	Ensemble
(González et al., 2015)	GPS, 3-axis MEMS accelerometer (4 Hz,10 Hz), 3-axis magnetometer	4 group of 11 steers	below the neck	Foraging, ruminating, travelling, resting, scratching, head shaking, grooming	Decision tree
(Vázquez Diosdado et al., 2015)	GPS, 3-axis accelerometer	6 cows	Neck	Lying, standing, feeding, transitions between standing and lying	DT, k-means, HMM, SVM
(Smith et al., 2016)	GPS, 3-axis MEMS accelerometer (10 Hz), 3-axis magnetometer, pitch & roll	24 cows	Neck collar	Grazing, walking, ruminating, resting, others	SVM, LR, KNN, RF
(Kamminga et al., 2017)	3D accelerometer (200 Hz), 3D gyroscope	4 goats, 2 sheep	Various positions of neck collar	Stationary, foraging, walking, trotting, running	DT, NN, SVM, NB, LDA, KNN, KNN
(Guo et al., 2018)	3-axis accelerometer (20 Hz), 3-axis gyroscope, 3-axis magnetometer, camera	3 lambs	Neck	grazing or non-grazing	Linear Discriminant Analysis (LDA)
(Walton et al., 2018)	3-axis accelerometer (8,16, 32 Hz), 3-axis gyroscope	6 sheep	Ear, neck collar	walking, standing, lying	Random Forest
(Wang *et al.*, 2018)	3-axis accelerometer (1 Hz), GPS	5 cows	Leg tag	feeding, standing, lying, lying down, standing up, normal walking, and active walking	AdaBoost (MBP)

Subsequently, the use of smartphones is suggested to serve the PLF developing process, since smartphones are equipped with relevant high-performance IMU sensors and GPS which simultaneously provides useful information about movements like acceleration, rotational gravity, angular velocity, orientation angles, and location of an object (Debauche *et al.*, 2018; Debauche *et al.*, 2017). For example, IMU has been used to record measurements of cattle ruminating (Andriamandroso *et al.*, 2014), biting (Andriamandroso *et al.*, 2015), and grazing (Andriamandroso *et al.*, 2017).

Furthermore, the built-in sensors in smartphones facilitate the idea of no extra hardware needing to be developed for monitoring purposes, besides the advantages of the ubiquitous features of smartphones (Debauche *et al.*, 2018; Debauche *et al.*, 2017). However, the applications of IMU are not only limited to PLF; instead, various other movable applications are reviewed in (Ahmad *et al.*, 2013), who highlight the most important consideration when IMU sensors are chosen; for instance, package size, data accuracy, response rate, and degree of freedom.

2.6 Data Mining for Analysing Motion-Sensor Data

As described earlier, GPS has been independently used or in combination with other motion sensors to identify different behaviour patterns on the farm which might not directly relate to lameness detection. Instead, it could be used for classifying various physical activities such as standing, lying, grazing, and walking which contributes to developing PLF and animal welfare management.

The sensor-based data either from IMU or GPS is usually called spatial-temporal data which involves spatial properties such as geometry and location, and temporal properties like time interval or timestamp (Rao *et al.*, 2012). The huge amount of collected spatial-temporal data calls for more professional and precise approaches to analyse such relative large datasets since both spatial and temporal dimensions increase the complexity of analysis to discover hidden patterns and trends for these collected data (Rao *et al.*, 2012). Consequently, Data Mining (DM) is emerged to be employed in research studies that aim to build robust computational techniques for analysis of enormous databases of such spatial-temporal datasets.

Due to motion sensors being mounted on animals to collect such spatial-temporal data, the needs for DM analysis techniques is increased. Additionally, the concept of 'reality mining' has explored the idea of cross-collaboration between disciplines to produce more integrated approaches (Krause *et al.*, 2013). Data Mining techniques, which are a convergence of principals refer to many disciplines as mentioned in Figure 2-4 (Jiawei *et al.*, 2012). For example, DM combines statistical principles like sampling, hypothesis, testing, and estimation with Machine Learning (ML) aspects such as searching algorithms, modelling techniques, and learning theories. Although both statistics and ML aim to build a model to describe the input data best, ML are hypothesis-free approaches which could attractively deal with complex data and focus on prediction rather than an inference that is assumed by the traditional statistical approaches to accept or reject (Valletta *et al.*, 2017).

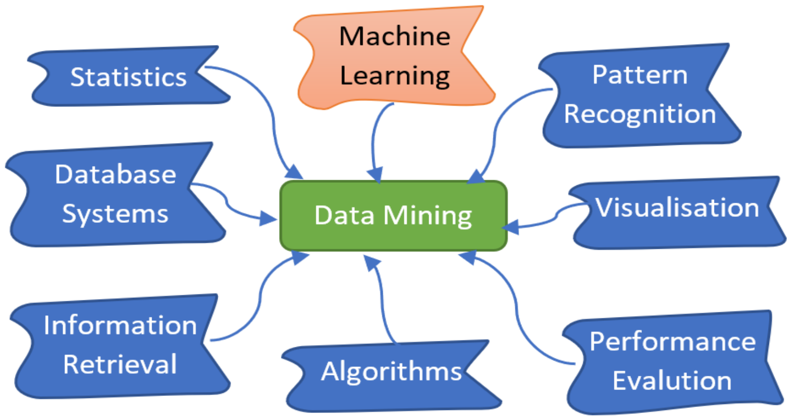


Figure 2-4 Data Mining as a convergence of many domains' principles.

Consequently, the implementation of DM to analyse such spatial-temporal sensor-based data for behavioural classification in both cattle and sheep could play an important role detect some illness concerns such as lameness. However, the number of research studies that have applied DM techniques for behavioural classification may exceed the ones for lameness detection research studies in cattle and would be very limited to lameness detection in sheep.

2.6.1 Data Mining Definition

Data Mining (DM) could be defined as the process of automatically retrieving useful information from a huge data repository, cleaning it; like removing noise, pre-processing it such as extracting related features, analysing it to gain useful insights from data, and finally intelligent decisions could be made for future observations (Aggarwal, 2015; Tan *et al.*, 2006). DM is an integral part of Knowledge Discovery in Databases (KDD) which involve data pre-processing, data mining, and data post-processing (Tan *et al.*, 2005). Therefore, wide variations of real-world problems solving approaches could accumulate under the broad umbrella term of 'Data Mining' (Aggarwal, 2015). As a result, this variation leads to fruitful collaborations between the DM field of computer science disciplines and many other scientific branches to perform interdisciplinary research studies that could solve many real-world problems.

2.6.2 Data Mining Approaches and Tasks

Two tasks for mining data could be applied; one could be predictive tasks which aim to predict the values of an attribute for unseen examples relying on the characteristics of seen examples. Predicting approaches refer to either *classification* if the predicted class has discrete/categorical values or *regression* if the predicted class has continuous/numerical values. On the other hand, descriptive tasks are applied to another purpose for mining data that aims to derive patterns that describe the relationship of data. Descriptive tasks involve either *clustering* which searches for a closely related group of observations that have similar features or *anomaly detection* that detects significant deviations from normal behaviour (Tan *et al.*, 2006; Aggarwal, 2015).

2.7 Machine Learning Background

As an embedded part of mining data, Machine Learning (ML); which is a fast-growing branch of knowledge, could be defined as the process of investigating how the computer machines learn from data and improving the learning performance of ML algorithms based on data in order to recognise important patterns or create a model which consequently makes an intelligent decision depending on the significant extracted pattern or built model (Jiawei *et al.*, 2012).

Basically, two types of learning are followed by ML algorithms divided into supervised and unsupervised to solve the aforementioned predictive and descriptive tasks, respectively. For the predictive task (classification/regression), a model is created depending on the labelled input data; then the created model would be used to predict the class for new unlabelled data. While in descriptive tasks (clustering/anomaly detection) a pattern is derived from unlabelled input data to identify the relationship or the outliers of the input data. In contrast to labelled or unlabelled input data, the output class would be either categorical when the classification algorithm is applied or numerical class when the regression algorithms are applied for model creation (Rokach and Maimon, 2009). Figure 2-5 depicts the main learning algorithms of ML;

however, classification algorithms would be the major focus of this study where the output class is not numeric; instead, the categorical class value would be the target.

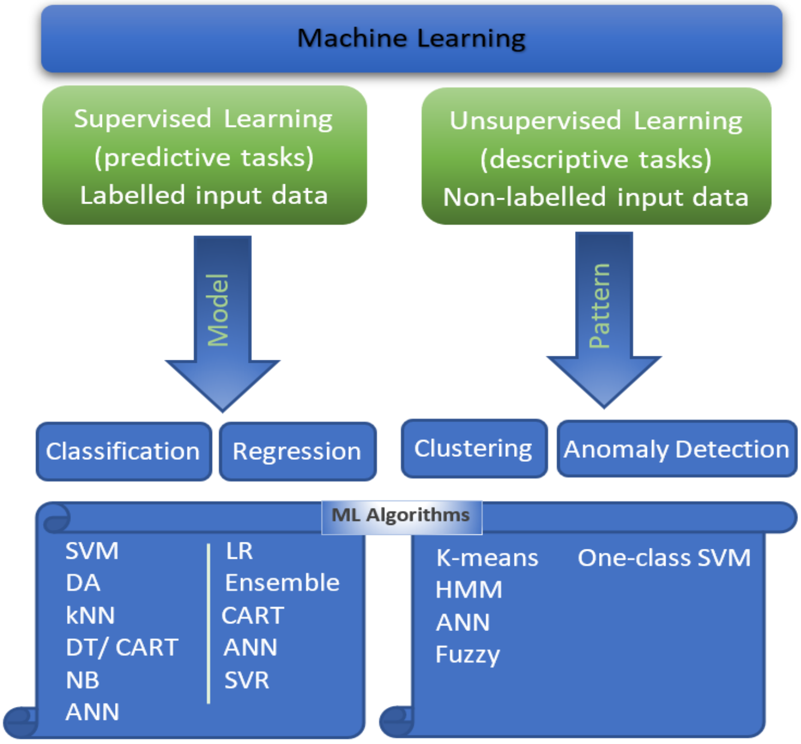


Figure 2-5 Main Machine Learning algorithms refer to predictive and descriptive tasks.

There are several supervised classification algorithms; each follows a different approach of learning; however, the process of choosing the best algorithm to fit all dataset's type is an overwhelming process (MathWorks, 2018a). Although selecting the right algorithm usually depends on trial and error, the final choice of an algorithm to build a specific model would outweigh one benefit against another such as model speed, accuracy, complexity, memory usage, and interpretability (MathWorks, 2018b). The most considerable features for the basic ML algorithms are listed in Table 2-8 (Razavi and Kurfess, 2003; MathWorks, 2018b; Rokach and Maimon, 2009; Mathworks, 2016; Osisanwo et al., 2017). It could be referred to, to meet different requirements.

Table 2-8 Considerable features in selecting the best ML algorithm to specific datasets. (speed of prediction: fast: 0.01 sec, medium: 1 sec., slow: 100 sec.), (memory occupied: small: 1MB, medium: 4MB, large: 100MB), (Accuracy:1-5 from worst to best).

ML Algorithm	LDA	Decision trees	SVM	Naïve Bayes	KNN	Ensemble
Numerical/ categorial Predictors	Numerical (Not categorial)	Both	Both	Both, discrete numerical values, not continuous	Either numeric (Euclidean distance) or categorical (hamming distance)	Both (except subspace ensembles of discriminant analysis classifiers
Binary/multi-class	Binary & multi-class	Binary & multi-class	Combine multiple binary classifiers	Binary & multi- class	Binary & multi-class	Binary & multi-class
Memory occupied	Small (LDA), large (QDA)	Small	Medium to large	Small (simple distributions) medium (kernel distributions or high-dimensional data)	Medium to large	Medium to high depending on the choice of classifier
Speed of prediction	Fast	Fast	Medium (linear), Slow (non-linear). Depend on the number of support vectors	Medium (simple distributions), slow (kernel distributions or high-dimensional data)	Slow (cubic) medium (others)	Medium to slow depending on the choice of classifiers
Fitting speed (training time)	-	Fast	Medium	Depend on data distribution	Depend on data dimention	Medium (boosted tree), slow (bagged tree)
Interpretability	Easy	Easy	Hard (for kernel function)	Easy	Hard	Hard
General Accuracy	-	2	4	1 (depend on dimentions	2	4
Better performance	All classes have a normal distribution, classes are separatable, and large dimensional datasets	Embedded with the hardware system as low memory required	Perform better when the training dataset is balanced (the number of classes is equivalent)	Small datasets with many parameters, when new scenarios need to be predicted where not existed in training phase	Good predictions in low dimensions, multi-classes prediction problem	When predictors behave nonlinearly

2.8 Applications of Machine Learning in Cattle and Sheep behaviour

Apart from lameness detection studies that were discussed earlier (Section 2.2), the classification research studies of cattle and sheep behaviours that might intersect with the current research topic as some monitored behaviours relate to lameness detection; for example, lame walking or neck activity. However, the classification of other behaviours based on analysing data derived from a motion-sensor attached to a farm animal using ML is briefly reviewed in the following sections (it might intersect with aforementioned studies in Table 2-7).

2.8.1 Discriminant Analysis (DA)

DA is a statistical technique dependent on finding a linear combination among predictors to separate the space of continuous measurements of input data into classes according to linear hyperplane decision boundaries (Osisanwo *et al.*, 2017; Kotsiantis, 2007). Linear DA can be extended to quadratic QDA when the decision boundaries are changed to be non-linear. Furthermore, DA could be used either for dimensional reduction of the input features space (Marais *et al.*, 2014) or for behavioural classification in targeted species. Table 2-9 lists the main studies which implement LDA or QDA to identify various animal behaviours.

Table 2-9 Research studies use LDA classifier to investigate livestock behaviour.

References	Sensor type	No/Animal	Sensor location	Classifier used	Classified behaviour
(Umstätter *et al.*, 2008)	GPS, tilt sensor	10 sheep (2 sites)	Neck	LDA, Classification tree, developed DT	Active and inactive
(Watanabe *et al.*, 2008)	3-axis MEMS accelerometer	1 cow	Under the jaw	QDA	eating, ruminating and resting
(Marais *et al.*, 2014)	3-axis accelerometer	5 Sheep	Neck	LDA, QDA	Lying, standing, walking, running and grazing
Van De Gucht *et al.*, (2017)	Pressure mat	45 cows	On the floor	LDA	discriminate among non-lame, mildly lame or severely lame
(Giovanetti *et al.*, 2017)	tri-axial accelerometer	3 sheep	Under the jaw	DA	grazing, ruminating and resting
(le Roux *et al.*, 2017)	GPS, 3-axis accelerometer	5 Sheep, 3 rhinoceros	Neck	LDA	Lying down, standing, walking, running and grazing.
(Radeski and Ilieski, 2017)	3-axis accelerometer	13 sheep (10 ewes, 3 rams)	Left hind leg	DA	Gait (walking, trotting, galloping), posture (standing, lying)
(Guo *et al.*, 2018)	3-axis accelerometer, 3-axis gyroscope, 3-	3 sheep	Neck	LDA	Grazing and non-grazing

	axis magnetometer				
(Barwick *et al.*, 2018a)	3-axis accelerometer	5 sheep (ewe)	Neck, ear, leg	QDA	Grazing, standing, lying, and walking

2.8.2 Decision Trees (DT)

Decision Tree (DT) or Classification and Regression Tree (CART) is a well-known ML classification technique used to predict qualitative or quantitative responses, respectively. DT or CART classifies the training observation according to a multi-stage process, where the input data are recursively divided into sub-groups according to a splitting criterion (Hartley, 2014; James *et al.*, 2013). The qualitative response represents the most commonly occurring class within the sub-group of the training dataset, while the quantitative response refers to the mean response of the training observation of that sub-group (James *et al.*, 2013).

The splitting process follows splitting criteria like entropy or Gini index (to be discussed in Section 3.7.1) to sort each sub-group of (attributes or predictors) according to its class from the set of output classes (Sharma *et al.*, 2013; Wu *et al.*, 2008). Each node in the tree flowchart represents an attribute or feature in an instance (data point/record), every branch is a testing output of splitting criteria, and every leave node is a class label (Razavi and Kurfess, 2003; Sharma *et al.*, 2013).

DT outperform other ML classification techniques (Table 2-8) due to its affordable features such as (Valletta *et al.*, 2017; Tan *et al.*, 2005; James *et al.*, 2013)

A. DT deal with the missing data point.
B. DT does not require a pre-knowledge of data distribution, whether parametric or not.
C. DT could be used to generate classification rules directly.
D. DT is a computationally inexpensive technique to train, evaluate and store.
E. DT is robust to outliers and the presence of noise.
F. It could handle both input data type (numerical and categorical).
G. It can be represented graphically, and no effort needed to be interpreted.

However, the final accuracy of the tree could be affected by irrelevant attributes as might have been chosen accidentally for tree growth of the classification task. Thus, the feature selection process is crucial to select only the most relevant attributes to improve accuracy. Alternatively, a post-pruning process could be performed to reduce tree size and improve accuracy (Tan *et al.*, 2006).

DT has been widely used in the field of livestock behaviour classification into such behaviours that might relate to welfare issues (see Table 2-11). Due to the comprehensive approach of DT, it could simulate a simple human decision-making procedure of 'if-else rules' that test a pre-defined threshold value to make a decision. For example, the 'if-else decision tree' is developed in some literature Table 2-10 to classify the locomotion of a cow. The 'if-else rules' of the decision tree is applied, when the need for data to be processed in the sensor node itself is necessary, as it is a computationally inexpensive rule-based approach.

Table 2-10 Research studies develop rule-based approaches from DT for livestock behaviour.

References	Sensor type	Animal	Sensor location	Analysis approach	Classified behaviour
(de Mol *et al.*, 2009)	2D Accelerometer	6 cows	right hind leg + neck	linear interpolation all measurements of Acc transformed to angles	Lying, standing
(Nielsen *et al.*, 2010)	IceTag3D™ accelerometers	10 cows	hind limbs (2 Acc)	IceTagAnalyzer software	Walking, standing
(de Mol *et al.*, 2011)	3D Accelerometer	3 cows	Right and left hind leg (2 Acc)	Two-step method: tilt sensing (standing), threshold testing (standing-walking)	Lying, standing or walking
(Apinan *et al.*, 2015)	3-axis analog accelerometer (1 Hz)	-	Around leg	Two-step algorithm: the magnitude of each axis (lying), variance of Y-axis (standing, walking-grazing)	Walking-grazing, standing, lying
(Khanh *et al.*, 2016)	3D-accelerometer (Arduino kit)	cow	Neck	Develop 2 thresholds DT depend on mean of a static component of y-axis and Vedba	Lying, standing, feeding
(Khanh *et al.*, 2018)	3-axis accelerometer	cow	Neck	A multi-stage classification tree is embedded into MCU, evaluation	Lying, standing, feeding, drinking

Table 2-11 Research studies use DT technique to investigate livestock behaviour.

References	Sensor type	No/Animal	Sensor location	Classifier used	Classified behaviour
(Nadimi *et al.*, 2008)	2-axis accelerometer	4 cows	Neck	DT	Active and inactive
(Robert *et al.*, 2009)	3-axis accelerometer	15 beef calves	Right rear leg	Classification tree	Lying, standing, walking, the transition between activities
(Vázquez Diosdado *et al.*, 2015)	GPS, 3-axis accelerometer	6 cows	Neck	DT, k-means, HMM, SVM	Lying, standing, feeding, transitions between standing and lying
(González *et al.*, 2015)	GPS, 3-axis MEMS accelerometer, 3-axis magnetometer	4 group each of it 11 steers	below the neck	DT	Foraging, ruminating, travelling, resting, scratching, head shaking, grooming

(Alvarenga et al., 2016)	Tri-axial accelerometer	10 sheep	Under the jaw	DT	Grazing, lying, running, standing and walking
(Tamura *et al.*, 2019)	3-axis accelerometer	38 cows	Neck	Decision tree	Eating, rumination, and lying

2.8.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a non-probabilistic classifier which maps input data features into high-dimensional feature space; where each dimension belongs to a classification feature, of two classes. The two-class datasets are separated by one hyperplane that produces the best accuracy among many other existing hyperplanes in the high dimensional input space. SVM learning can guarantee the best fit function that maximises the margins between two classes (Wu *et al.*, 2008). Fundamentally, SVM is a binary classifier; however, it can be extended to deal with a multi-class problem when one class is considered against all other classes (one-versus-all) or against one other class (one-versus-one) (James *et al.*, 2013).

Good accuracy of the SVM classifier could be obtained when the data points can be separated linearly. However, a transformation into high dimensions data (kernel transform) could be an alternative to quantify the linear decision boundary (Mathworks, 2016). One advantage that SVM classifiers could find is the best classification function, the high accuracy results might be obtained from the SVM learning process. Moreover, overfitting might be prevented when the SVM classifier is applied. Conversely, the SVM is computationally expensive as it requires a large amount of training time, storage space, and extra effort for the result's interpretation is needed among other classifiers (Sharma *et al.*, 2013; Nathan *et al.*, 2012). In addition, SVM might have a limited success rate when it is applied to imbalanced datasets when one class exceeds other classes in the training dataset (Ganganwar, 2012).

The SVM classifiers are implemented in the field of livestock behaviour to classify various behaviours or detect welfare issues. In addition to research studies priorly mentioned in Table 2-7, Table 2-12 briefly explores the other research studies in the field of livestock behaviour that implements SVM for their experiment.

Table 2-12 Research studies use SVM classifier to investigate livestock behaviour.

References	Sensor type	No/Animal	Sensor location	Classifier used	Classified behaviour
(Martiskainen *et al.*, 2009)	Three-dimensional accelerometer	30 cows	Neck	Multi-class SVM	standing, lying, ruminating, feeding, the normal and lame walking, lying down, and standing up
(Alsaaod *et al.*, 2012)	ALT-pedometer	30 cows	Ankles	SVM, RBF-Kernal	Non-lame and lame
(Benaissa *et al.*, 2017)	3-axis accelerometers	16 cows	Leg, neck	SVM, KNN, NB	Lying, standing, feeding
(Haladjian et al., 2018)	3-axis linear acceleration, 3 axis orientation	10 cows	hind left leg	One class-SVM	Normal and abnormal (lame) cow stride
(Mansbridge *et al.*, 2018)	3-axis accelerometer, 3-axis gyroscope	6 sheep	Ear, collar	RF, SVM, KNN, Adaboost	Grazing, ruminating, non-eating

2.8.4 K-Nearest Neighbour (KNN)

KNN is a statistical ML technique where the tested object is classified based on the closest training data point in the space of input features (Sharma *et al.*, 2013). In the KNN method, no explicit training for the input features is required. Instead, the class of the target object is assigned according to the majority class of the much closest data points. The majority votes for noisy examples outweigh when a pre-defined value of *k* is small which may yield a high misclassification error. On the other hand, too many points from other neighbourhood classes may be included when the *k* value is set to be too large (Wu *et al.*, 2008; Sharma *et al.*, 2013). Distance metrics; for example, Euclidean, could be used to calculates the best nearest neighbour (Mathworks, 2016). KNN has been used in animal behaviour classification as it is illustrated briefly in Table 2-13.

Table 2-13 Research studies use KNN classifier to investigate livestock behaviour.

References	Sensor type	No/Animal	Sensor location	Classifiers used	Classified behaviour
(Smith *et al.*, 2016)	GPS, 3-axis accelerometer, 3-axis magnetometer, pitch & roll	24 cows	Neck collar	SVM, LR, KNN, RF	Grazing, walking, ruminating, resting, others
(Benaissa *et al.*, 2017)	3-axis accelerometers	16 cows	Leg, neck	SVM, KNN, NB	Lying, standing, feeding
(Mansbridge *et al.*, 2018)	3-axis accelerometer, 3-axis gyroscope	6 sheep	Ear, collar	RF, SVM, KNN, Adaboost	Grazing, ruminating, non-eating
(Kleanthous *et al.*, 2018), Dataset from	3D accelerometer, 3D gyroscope	4 goats, 2 sheep	Various positions of neck	MLP, RF, KNN, Extreme Gradient Boosting	Grazing, lying, scratching or biting, standing, walking

(Kamminga *et al.*, 2017)			collar	(XGBoost)	

2.8.5 Ensemble Classifier (EC)

Ensemble classifier (EC) is a predictive model where multiple classification models; called weak learners, are combined to produce one optimal model that would increase the predictive quality of a classification task. Although EC has been used to aggregate several DT classifiers to enhance the overall accuracy and reduce the variance of the training dataset (James *et al.*, 2013), EC does not pertain to DT only. When the same classifiers are used in EC, it is called homogenous Ensembles; otherwise, it is named heterogenous Ensemble when a different type of classifiers are used (Smolyakov, 2017).

Three common techniques of ensemble classifier where DT is employed as a weak learner are discussed in the following sub-sections.

2.8.5.1 Bagging or Bootstrap Aggregation

In bagging or bootstrap aggregation, smaller repeated samples; called replicas or bootstrap samples, are generated from the training dataset where DTs are grown on replicas to be all aggregated at the end of the training process. The final prediction of the ensemble classifier is measured by either averaging all the predictions produced from independent trees in regression, or by voting the most commonly occurring class of the predictions in classification (James *et al.*, 2013; Smolyakov, 2017).

2.8.5.2 Random Forest (FR)

Random Forest is identical to Bagging technique where the dataset is divided into replicas; however, some level of differentiation could be achieved in RF. Basically, in Bagging technique, each replica has the same input features (predictors) to be trained by a single DT. Conversely, each replica in RF has a different group of input features with a replacement that is chosen by RF according to a random selection following the same distribution for all trees in the forest (Lutins, 2017b; Breiman, 2001).

As an advantage of the RF, it overcomes the problem of highly correlated trees in EC that might occur when using Bagging, as each tree has the same input parameters rather than in RF where a different group of predictors are used for each tree. Hence, the average variance of different trained models; due to the variations in the input feature, is better than the average variance of similar trained models because of the similarity in the input parameters (James *et al.*, 2013).

2.8.5.3 Boosting

Boosting involves the same idea of Bagging except that, the trees are constructed sequentially depending on the information (e.g. feature importance) from previously grown trees where each tree is trained according to a modified version of the original dataset instead of using bootstrap samples (James *et al.*, 2013). A higher weight is assigned to misclassified examples (to be focused by the next learner), while a lower weight is given to correctly classified examples from the previously trained tree. As a result, the stronger classification in the current stage is re-allocated with a higher weight and so on (Grover, 2017). Therefore, the strong learner is obtained by iteratively adding trees and adjusting the weight of each tree to enhance the accuracy of EC (Mathworks, 2016). Thus, the final prediction of the Boosting ensemble is obtained either through a weighted majority vote in the classification task, or a weighted sum in the regression task (Smolyakov, 2017).

Because the growing of trees in Boosting ensemble takes into account the information from previously built trees, it results in a better performance when compared with RF as a smaller number of trees would be sufficient to achieve the optimal accuracy with a good level of interpretability (James *et al.*, 2013).

Boosting can be applied to various techniques such as:

1. **Ada**ptive **Boost**ing (AdaBoost) where a Decision Stump is used as a weak learner in Boosting ensemble that performs one level of splitting (depth of the tree is one). AdaBoost tries to improve the areas where the base learner fails by working on perfectly fitting every point. However, one drawback could be noticed that AdaBoost is affected by outliers and noisy data as it works to fit every point in training data (Gandhi, 2018).
2. **G**radient **Boost**ing (GBoost) is based on Gradient Descent optimisation problem to find the local optima of a function. It is similar to AdaBoost where each Decision Stump tries to fit every point, but it differs in that for each iteration a decision stump is trained, a loss function is computed to be optimised sequentially until reaching a minimised loss function. The loss represents the difference between the actual and predicted value which are called residuals (Lutins, 2017a; Gandhi, 2018).
3. E**x**treme **G**radient **Boost**ing (XGBoost) follows a similar principle of GBoost; however,

Newton's method is applied to provide a direct route to the minima instead. Generally, XGBoost is faster and higher in performance when compared to GBoost, while GBoost has a wide range of applications (Nielsen, 2016).

Some studies on livestock's behaviour exploit the use of Ensemble Classifier for various behaviours, a brief summary is given in Table 2-14.

Table 2-14 Research studies use Ensemble techniques to investigate livestock behaviour.

References	Sensor type	No/Animal	Sensor location	Classifiers used	Classified behaviour
(Dutta *et al.*, 2015)	3-axis accelerometer, 3-axis magnetometer	24 cows	Neck	Bagging, random subspace, AdaBoost	Grazing, searching, ruminating, resting, scratching
(Wang *et al.*, 2018)	3-axis accelerometer, GPS	5 cows	Leg tag	AdaBoost	feeding, standing, lying, lying down, standing up, normal walking, and active walking
(Mansbridge *et al.*, 2018)	3-axis accelerometer, 3-axis gyroscope	6 sheep	Ear, collar	RF, SVM, KNN, Adaboost	Grazing, ruminating, non-eating
(Kleanthous *et al.*, 2018) The dataset from (Kamminga *et al.*, 2017)	3D accelerometer, 3D gyroscope	4 goats, 2 sheep	Various positions of neck collar	MLP, RF, XGBoot, KNN	Grazing, lying, scratching or biting, standing, walking

2.9 Gap in literature

While the overall goal of current research is to develop a predictive model to indicate the lame status of sheep as early as possible to prevent the disease from being spread all over the flock, monitoring sheep behaviour and gathering useful movement measurements are the first step towards achieving this goal.

Although the number of published studies that validate the application of sensor technology to categorise and quantify sheep behaviour has increased recently (Fogarty *et al.*, 2018), only a few studies (Section 2.4) utilise sensor technology to detect lameness in sheep (Barwick *et al.*, 2018b) in Australia, (Vazquez Diosdado *et al.*, 2018) in Nottingham/UK, and the earlier research output (Al-Rubaye *et al.*, 2018) in Northampton/UK.

Moreover, exploiting ML techniques for the advantage of lameness detection is another desperate shortage in the field of knowledge. Therefore, the idea of fruitful collaborative work to employ the ML principals from computer science for the advantage of animal welfare science (sheep welfare) of the current research study would fill the gap and enrich the field with promising outcomes.

2.10 Chapter Summary

Initially, lameness is detected by trained observers using a scoring system. Alternatively, automatic lameness detection is introduced due to its reliability, speediness, and objectiveness compared to the traditional scoring system. Automatic sensing includes sensors located in a fixed place on the farm (Section 2.3.1.1) or a motion sensor attached to the animal's body (Section 2.3.1.2).

The motion-based sensors are attached to the animal's body to measure the locomotion activity for body, leg, or neck. IMU sensors where readings from different sensors can be obtained at the same time. IMU sensors attached to the animal body are widely used to extract the behavioural status of livestock animals that might relate to lameness; for example, lying, standing, walking behaviours that have shown a relationship with lameness indication in literature.

Due to the myriad of acquired sensor-based data (IMU), ML techniques are investigated for cattle and sheep to develop predictive models to classify various behaviours or lameness detection. However, the ML techniques are widely implemented to classify different livestock behaviours (Section 0) which will contribute to developing PLF and Smart farming in the near future. On the other hand, exploiting ML to detect lameness in sheep (Section 2.9) would lack research studies involving validated experiments, developed models and even data collection tools.

3 Chapter Three: Building a Data Mining Methodology for Sheep Lameness Detection (SLDM)

3.1 Introduction

Sheep lameness detection is not a straightforward task; however, many challenges are addressed, and several requirements need to be met in order to build such an efficient model to detect lameness in sheep; especially in its early stage.

A data mining methodology is constructed to convert the raw data (sheep acceleration movements) into useful information (lameness alarm or indicator) that would contribute to smart farming and PLF to be beneficial in the near future. Basically, each data mining task includes three stages; pre-processing (e.g. cleaning, filtering, feature extraction), developing a learning model (e.g. DT, Ensemble), and post-processing stage (e.g. visualisation, pattern's interpretation) (Tan *et al.*, 2005). However, each step consists of internal sub-steps, which are discussed later in this Chapter. The main stages for developing a Sheep Lameness Detection Model (SLDM) are depicted in Figure 3-1.

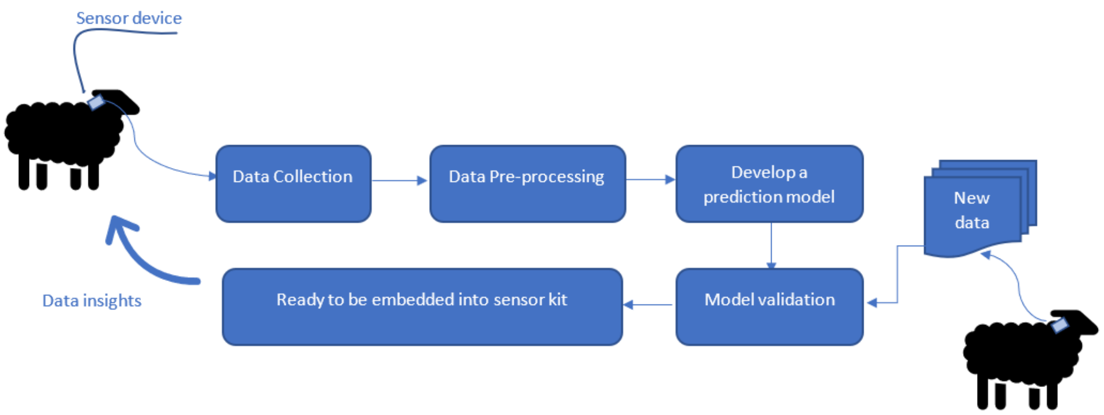


Figure 3-1 Development Stages for Sheep Lameness Detection Data Mining Approach.

The built model could control the spread of disease among the flock and assist the shepherd in spotting the lame sheep without further monitoring's hours. This Chapter includes the practical steps taken towards constructing a data mining methodology that suits sheep data as there is no data mining method which fits all types of data in the world. However, an introductory section about Android-powered sensors; which are used in this research, is presented first in Section 3.2. Then, it is followed by Sections 3.3 and 3.4, which examine how the data are being collected and aggregated from the real-world farm. Afterwards, Section 3.5 explores the data pre-processing stages to prepare the sheep data for the classifier. Section 3.6 explores the feature selection methods that are applied to sheep datasets. Then, the selected features feed the CART decision tree, which is the core classifier in building SLDM, where its characteristics are illustrated in empirical steps in Section 3.7. The validation methods applied to test the developed system are explained in Section 3.8. Finally, the Chapter is closed with a brief summary in Section 3.9.

3.2 Android-powered Sensors

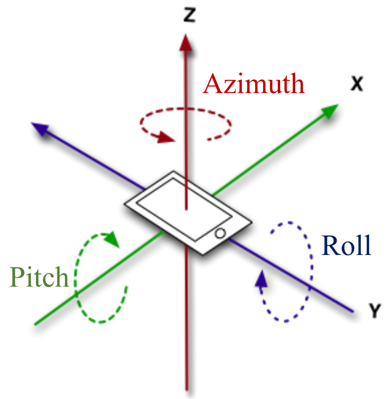
In this research, an Android-powered mobile device is used to serve the purpose of data collection as it has built-in sensors to measure various motion activities and device orientation in high precision and accuracy with three-dimensional measurements (Android Developers, 2019c). Android platforms support three broad categories of sensors, including motion, position, and environmental sensors, refer to Table 3-1. Several sensors are hardware-based sensors which are physical components built into a device such as an accelerometer, gyroscope, and magnetometer, that are capable of deriving their measurements directly from a specific property. On the other hand, software-based sensors or virtual sensors are not physical parts; however, they mimic hardware-based sensors and derive their data from one or more mixed hardware-based sensors such as orientation, linear acceleration, and gravity sensors (Android Developers, 2019c).

Table 3-1 Sensor categories that are supported by Android platforms (* refer to hardware-sensors).

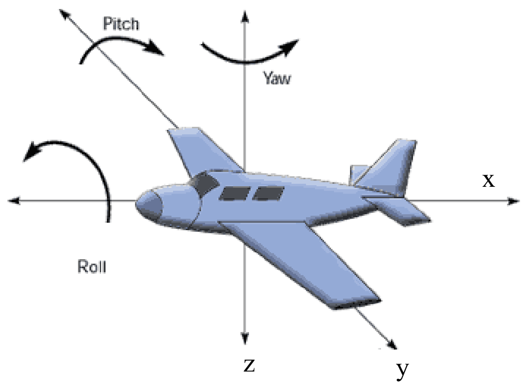
Sensor Category	Sensor include	To measure	Sensor used in research	Unites
Motion	Accelerometer* Gyroscope* Gravity Linear acceleration	acceleration forces and rotational forces	Accelerometer* Gyroscope*	m/s^2 Rad/s
Position	Orientation Magnetometers*	physical position of a device	Orientation (Pitch, Roll, Azimuth)	degree
Environmental	Barometers, Photometers Thermometers	Ambient air temperature, pressure, illumination, and humidity	Not used in research	

3.2.1 Android Coordinate System

The android coordination system is defined relative to the screen when the device is set in its default orientation where the x-axis is horizontal and points to the right, the y-axis is vertical and points upwards, the z-axis is perpendicular and points to the outside of the screen's face (to the sky) Figure 3-2. It is worth mentioning that the orientation system in aviation differs from the Android system (Android Developers, 2019b) as shown in Figure 3-2. Furthermore, the coordination system of an Android device does not swap when the device's screen orientation is changed (Android Developers, 2019b).



(a)



(b)

Figure 3-2 Natural coordinate system of Android device(a) vs default aviation orientation system (b).

3.2.2 Basic Android Sensors Definition

In this research study, the focus has been on three basic types of sensors: accelerometer, gyroscope, and orientation sensors. The first two sensors are hardware-based sensors while the orientation sensor is a software-based one. The following sub-sections explore the definitions of those sensors which would lead to a better understanding of their functioning on three-axis and its application to identify sheep lameness status.

3.2.2.1 Accelerometer (Acc):

Acc measures the object's acceleration beside the force along each axis. Positive acceleration is obtained when the device (in its natural position) moves towards the right, forward, and up for x, y, and z, respectively (Figure 3-2 a). Negative acceleration is obtained doing the opposite. Acceleration includes the gravity (static acceleration component) and linear acceleration (a dynamic component of acceleration without gravity). Equ 3-1 measures acceleration involving both linear acceleration $AccLin$ and gravity.

$$Acc = AccLin + g \qquad \text{Equ 3-1}$$

Where Acc is the acceleration along any axis, $AccLin$ is the linear acceleration excluding gravity, and g is the force of gravity (in case of stability, g= -9.81 m/s^2). For example, the z-acceleration reading of non-moving object (where Acc_z supposed to equal 0) does not equal to zero; however, it is approximately equivalent to $Acc\_z + 9.81$ m/s^2 as [$Acc\_z = 0 - (-9.81)$ = $Acc\_z + 9.81$]. So, the acceleration readings need to be filtered to exclude gravity (Android Developers, 2019d) as a pre-processing step before the classification task is performed refer to Section 3.5.2.3.

3.2.2.2 Gyroscope (Gyr):

Gyr measures the speed of rotation (angular velocity) around each device's axes in radius/second. Gyr follows the same coordinate system of Acc (Figure 3-2 a). The positive rotation is obtained in a counter-clockwise direction. This definition is not the same as the Roll angle used by the orientation sensor (Android Developers, 2019e). Gyr could be used in combination with Acc to produce more accurate motion and direction sensing via integration of Gyr within a 3D space (Ustev, 2015).

3.2.2.3 Orientation ($Orient$):

$Orient$ measures the device's orientation in degrees relative to the earth's magnetic north pole. The orientation angles are calculated by combining accelerometer and magnetometer sensors (Android Developers, 2019a). The orientation includes three angles Azimuth, Pitch, and Roll that measure the degree of rotation about z, x, and y axes, respectively (Figure 3-2 a).

1- Azimuth (degree of rotation about the z-axis): measures the angle between the direction of the device's current compass and the magnetic north. Azimuth is equal to 0˚ when the top edge of the device faces the North, while it equals 180˚ when the top edge faces South. Conversely, Azimuth equals to 90˚ and 270˚ when the top edge of the device faces East and West respectively.
2- Pitch (degree of rotation about the x-axis): measures the angle between the top edge of device in its natural orientation and the ground. In this case, a positive pitch angle is obtained, while the tilt in the opposite direction measures a negative pitch angle. The range of pitch angle is between -180˚ to 180˚.

3- Roll (degree of rotation about the y-axis): measures the positive roll angle when the left edge of the device is in its natural position tilting towards the ground; oppositely, the negative angle is measured when the direction of the right edge tilts towards the ground. The range of roll angle is between -90˚ to 90˚.

3.3 Data Collection Process

A Galaxy S4 Android 5.0 mobile device was chosen to be a prototype sensor tool for collecting movement measurements from sheep due to its own ubiquitous variety of built-in sensors. Thus, no hardware needs to be developed to serve the purpose of data collection. However, the focus of the current research is on the analysis of sensor-based data not on developing a hardware sensor.

Furthermore, utilising IMU sensors in smartphones was previously suggested for monitoring cattle behaviour in the literature (Andriamandroso *et al.*, 2014; Andriamandroso *et al.*, 2015; Andriamandroso *et al.*, 2017; Debauche *et al.*, 2017). This idea supports the area of developing PLF as massive collected data being sent to the cloud (Debauche *et al.*, 2018) either to be gathered with other data sources or to be processed for the sake of decision making to inform the farmer via a phone application. However, the limitation could occur in battery drainage, energy consumption, memory storage, and communication method.

3.3.1 Data Collection Location and Challenges

The data collection experiments were conducted at the University of Northampton/ Moulton College Lodge Farm, Northamptonshire, United Kingdom (52°18'02.7"N 0°51'56.8"W) 52.300755, -0.865783. Therefore, the ethical approval and risk assessment request to visit the Lodge Farm was authorised by the Moulton College research committee in April 2016. See Appendix A for the signed document.

To overcome some challenges when data were being collected (Figure 3-3), special farm clothes and waterproof boots were required to be worn during the data collection stage to meet the security and safety conditions on the Farm. Moreover, it was not a trifling procedure to catch, chase, and deal with sheep while deploying the sensor. So, a special Induction day was given on 14 April 2016 to get familiar with the environment. Importantly, an additional hand was required to help with the data collection. Thus, authorised access to the Farm was obtained for the researcher's husband to enter the farm and help in the data collection process. The last challenge was to keep the sensor in a fixed location around the sheep's neck; therefore, a plastic clip was used to keep the sensor stable on the sheep's neck. Figure 3-3 reflects the data collection process at Lodge Farm in Moulton College.

(a) (b) (c) (d)

Figure 3-3 Data Collection at Lodge Farm (a) custom wear, (b) data collection assistant, (c) sensor fixer, (d) video recording process.

3.3.2 Sensor Deployment at Lodge Farm

In the real-world experiment of data collection, the sensor device was kept in a sport mobile phone case that had a visible plastic cover to resist severe environmental conditions such as rain, muddy soil, or scratches by other sheep in the flock. The wearable collar was attached to the neck where the surface of the device faced the sky, and the upper edge pointed to the sheep's head. This position approximately simulates the natural coordinate system (Figure 3-2) of the

Android device where the orientation reading would be more reliable if the Roll angle is about 0˚ (Android Developers, 2019b). Figure 3-4 illustrates the sensor coordination on the sheep according to the Android system (which differs from the aerospace system). The Orientations of the x, y, and z coordination in this research study were lateral, anterior-posterior, and dorso-ventral, respectively (Figure 3-4).

The sensor was mounted for each sheep for a period of ten to fifteen minutes, which was recommended by other sheep behaviour studies (will explain later) to be adequate to log movement data for a sheep while walking to detect mild lame status. This period would probably be equivalent to the required period for the observer to identify the lame sheep manually.

Video footage was also taken (Figure 3-3 (d)) via Canon or Sony cameras or even by phone camera while the movement of an individual sheep was measured to compare with each sheep's status (sound, mild lame, or severe lame) for the purpose of data labelling in the pre-processing stage. Unlike behaviour classification research, in this research study, the synchronised labelling of sheep behaviour is less important when compared to each sheep's common status of lameness or sound.

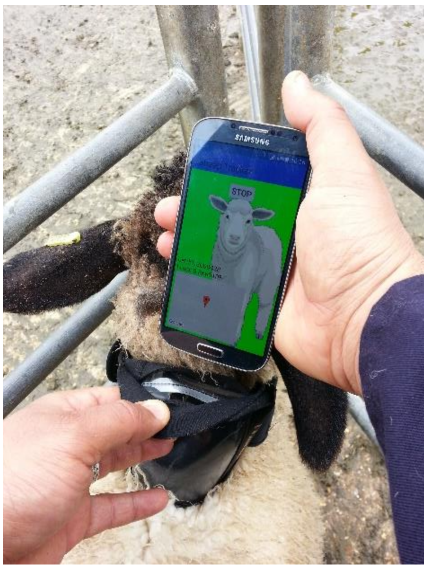
Figure 3-4 Sensor deployment at Lodge Farm with its orientation.

3.3.3 Sensor Reading Applications

Three types of sensor applications were used to retrieve the movement and orientation measurements from sheep in various data collection occasions. In each application, a different setting was applied; such as sampling rate, the type of sensors activated during the experiments, the number of sheep deployed with the sensor, and the dedicated sheep class (sound, mild or severe lame) that were already spotted by the skilful shepherd 'Tim Perks' who has lived and worked in Lodge Farm for more than 30 years. The following sub-sections explore the type of retrieval application used in this research study.

3.3.3.1 Sheep Tracker

Sheep Tacker is a specially designed application (Ghendir, 2016) which serves the purpose of collecting data in three dimensions from Accelerometer, Gyroscope, and Orientation. In addition to Latitude, Longitude, Date and Time. The sampling rate was 5 Hz, which means 5 readings per second were obtained for 3-axes of each sensor. Consequently, nine predictors were utilised for the classification task of 3 axes readings for each Acc, Gyr, and $Orient$ sensors; whereas, the position predictors (latitude and longitude) were neglected in this study as the aim is to detect sheep lameness with as few predictors as possible for the process. The collected data file was automatically stored in the device storage in both Excel and Text file format (Figure 3-5).



(a)

Acc_X	Acc_Y	Acc_Z	Gyr_X	Gyr_Y	Gyr_Z	Orien_Pitch	Orien_Roll	Orien_Yaw	Latitude	Longitude	Date	Time
0.050082	-0.22117	-0.07469	-0.02779	-0.03085	0.026573	-53.02451	-0.05377463	225.03409	52.30089	-0.8658	13-06-2016	17:15:17
-0.11052	-0.26927	-0.19796	-0.01497	0.00336	0.031765	-52.808792	0.06869861	224.32187	52.30089	-0.8658	13-06-2016	17:15:17
-0.02418	-0.27016	-0.15455	-0.0113	-0.03574	0.016493	-52.968174	0.26895723	226.16624	52.30089	-0.8658	13-06-2016	17:15:17
-0.06942	-0.23755	-0.09012	0.001833	-0.02169	0.026267	-53.093227	0.32356656	226.04578	52.30089	-0.8658	13-06-2016	17:15:17
-0.17096	-0.23012	-0.15663	0.03207	-0.04551	0.007941	-53.587345	0.97334677	228.49358	52.30089	-0.8658	13-06-2016	17:15:17
0.058182	-0.32348	-0.08496	-0.03146	-0.06903	0.065668	-53.765694	1.2735475	229.30037	52.30089	-0.8658	13-06-2016	17:15:17
-0.15928	-0.19992	-0.26547	0.040928	-0.02779	0.047647	-53.786777	1.5950792	228.51857	52.30089	-0.8658	13-06-2016	17:15:17
-0.09206	-0.27908	-0.10007	0.015882	-0.07208	0.050091	-54.10323	2.0876245	229.58357	52.30089	-0.8658	13-06-2016	17:15:17
-0.18438	-0.33388	-0.08495	-0.01527	0.050396	0.014661	-53.959427	2.0533495	228.43246	52.30089	-0.8658	13-06-2016	17:15:19
-0.09422	-0.24054	-0.18083	-0.01161	-0.02841	0.01405	-53.952087	2.0122075	227.00894	52.30089	-0.8658	13-06-2016	17:15:19
-0.16684	-0.22498	-0.0477	-0.04276	-0.00275	0.040623	-53.544044	2.3280542	224.57312	52.30089	-0.8658	13-06-2016	17:15:19
-0.10896	-0.21481	-0.11133	-0.03971	-0.05345	0.023213	-53.116642	2.5605192	223.7972	52.30089	-0.8658	13-06-2016	17:15:19
-0.12255	-0.53634	-0.1873	-0.08369	-0.04581	-0.01588	-52.654	2.6319528	224.37631	52.30089	-0.8658	13-06-2016	17:15:19
-0.06229	-0.24824	0.020244	-0.1066	0.003971	0.046731	-51.869354	2.696872	225.78148	52.30089	-0.8658	13-06-2016	17:15:19
-0.1232	-0.2438	-0.15496	-0.01985	-0.00764	0.07758	-51.343388	3.6180077	226.90079	52.30089	-0.8658	13-06-2016	17:15:19
-0.1308	-0.30066	-0.17724	-0.00855	-0.08827	0.030543	-51.432987	4.1012707	229.29439	52.30089	-0.8658	13-06-2016	17:15:19
-0.1502	-0.30938	-0.06954	-0.05895	-0.04123	0.031765	-51.084072	4.7872376	229.34932	52.30089	-0.8658	13-06-2016	17:15:19
-0.1103	-0.27561	-0.18884	-0.01405	-0.02382	0.010996	-50.83191	4.8854227	228.64523	52.30089	-0.8658	13-06-2016	17:15:19
-0.14891	-0.22849	-0.18048	-0.02718	0.069333	0.083994	-50.410435	4.8218355	228.38245	52.30089	-0.8658	13-06-2016	17:15:21
-0.14821	-0.26998	-0.25838	-0.04917	0.01405	0.068417	-50.273785	4.1655383	227.69916	52.30089	-0.8658	13-06-2016	17:15:21
-0.07854	-0.35644	-0.29025	-0.02016	0.139888	0.0562	-49.883865	4.526917	228.5031	52.30089	-0.8658	13-06-2016	17:15:21
0.123511	-0.06423	0.104595	-0.00489	0.406225	-0.10415	-49.882793	1.0673242	224.03394	52.30089	-0.8658	13-06-2016	17:15:21
-2.39456	-0.76196	-0.76487	0.28222	0.677755	-0.14539	-51.431282	-8.0555	218.9253	52.30089	-0.8658	13-06-2016	17:15:21
-0.18719	-0.29782	-0.54693	-0.00275	0.361021	0.249538	-52.455627	-6.4397664	229.64476	52.30089	-0.8658	13-06-2016	17:15:21

(b)

Figure 3-5 Sheep Tracker sensor (a), an example of collected data in Excel file format (b).

The data collected through this type of sensor were gathered from **two attempts** on two different days. The **first trial** was done on June 2016 where 10 sheep were equipped with the sensor (once at a time) for 10-15 minutes to test the device and look at the collected data in its first trial. From the collection of 10 sheep, two sheep were prepared for the next step (one mildly lame, the other sheep was sound) Table 3-2. The remaining 8 sheep's data files were only explored for getting the first impression of how the data would look like and has been excluded from the next steps.

A **second visit** to Lodge Farm was on September 2016 where 23 sheep were attached with the sensor for the approximately same period of the first attempt. Only data from 22 sheep were prepared for the next step due to the 4th sheep getting an empty data file at the end of the experiment. 14 sheep were mild to severely lame, while the other 8 sheep in the tested group were sound. Table 3-2 explains the metadata of both attempts with the Sheep Tracker sensor. At the end of the experiments, the stored files in the mobile storage were transferred to a computer device for pre-processing and analysis.

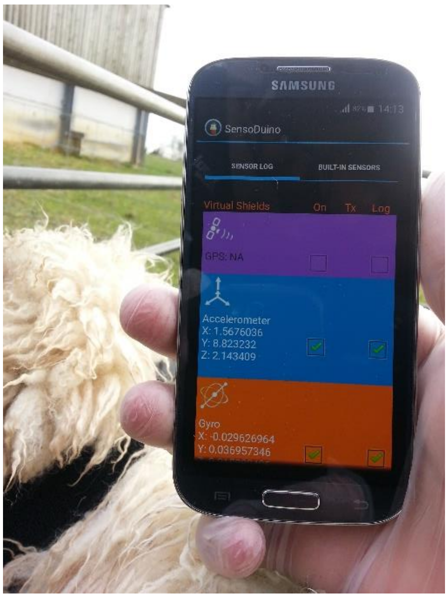
Table 3-2 Metadata for data collected from Sheep Tracker sensor at 5 Hz sampling rate.

Attempt No.	Date	# Sheep in each experiment	Datalog time	Sensor readings of interest	No. of reading records (rows)	# Sheep considered for the next step
1st attempt	13 June 2016	10 sheep (sound & lame)	≈ 10 mins	Acc, Gyr, and Orient	5 Hz ×10 mins ×60 sec.=3000 readings	2 sheep (sound and mildly lame), other files for a prior test only
2nd attempt	23 Sep. 2016	23 sheep (8 sound, 15 mild to severely lame)	5 to 10 mins.	Acc, Gyr, and Orient	5 Hz ×5 or 10 mins ×60 sec.=1500 to 3000 readings	22 sheep (8 sound, 7 mildly lame, 7 severely lame), one empty file
Total obtained sheep and their class			5 Hz			24 sheep (9 sound, 8 mildly lame, 7 severely lame)

3.3.3.2 SensoDuino

SensoDuino is a free Android application which can log and transmit Android built-in sensor readings to the Arduino controller or any other Android device via Bluetooth HC-05 module (Bitar, 2013). Many different motion and environmental Android sensors can be recognised by SensoDuino which includes hardware-based sensors (Accelerometer, Gyroscope, and Magnetometer) as well as a software-based sensor (Orientation, Linear acceleration, Gravity, Rotational vector) besides other sensors such as GPS, Pressure, Humidity, Temperature, Proximity, Light and Audio level sensors. Although several sensors of SensoDuino were activated in the sheep' data collection experiments, the only considered sensor readings for the next pre-processing stage were Accelerometer (Acc), Gyroscope (Gyr), Linear accelerometer ($AccLin$), and Orientation ($Orient$) for each 3-axes.

SensoDuino can be configured to capture data at every 100 milliseconds to 10 minutes according to manufactures limits (Bitar, 2013). However, the sampling rate of SensoDuino in the sheep data collection experiments was set to be at 10 Hz. Furthermore, the 4 Hz sampling rate was also tried for the same group of sheep (Table 3-3). Thus, 10 readings or 4 readings per second were obtained from 3-axes for each activated sensor which resulted in 12 predictors (columns) being considered for the next step. At the end of each sheep's deployment with SensoDuino, the log data were saved automatically in the phone's Stick Card into a text file that can be read in Excel as a delimited comma format as appears in (Figure 3-6).



(a)

```
SensoDuino log file created:
DATE: 2017/9/26
TIME: 14:37:30
------------------------------
Accelerometer, 0, -1.6172832,7.363966,5.3636103, 3
Gyro, 1, 0.092240654, 0.14080444,-0.017715093, 3
Orient, 2, 8.327302,-52.558083,-9.767258, 3
Gravity, 3, -1.3403261,7.7861867,5.8094077, 3
AccLin, 4, -0.27695715,-0.4222207,-0.44579744, 3
RotVec, 5, 0.4471119,0.053982552,-0.04263663, 3
Time, 6, 2,37,30, 3
Accelerometer, 7, -1.2294226,7.8571715,5.5581393, 3
Gyro, 8, 0.026572637, 0.010995574,0.07238753, 3
Orient, 9, 8.740568,-52.95484,-9.785093, 3
Gravity, 10, -1.3499097,7.827285,5.751673, 3
AccLin, 11, 0.120487094,0.029886723,-0.1935339, 3
RotVec, 12, 0.45109758,0.05455546,-0.0457666, 3
Time, 13, 2,37,30, 3
Accelerometer, 14, -1.3509283,7.7344685,5.509058, 3
Gyro, 15, 0.0580322, 0.022907447,-0.0018325958, 3
Orient, 16, 8.956314,-53.14722,-9.926576, 3
Gravity, 17, -1.373285,7.847078,5.7190776, 3
AccLin, 18, 0.022356749,-0.11260939,-0.21001959, 3
RotVec, 19, 0.4525769,0.054827917,-0.04633867, 3
Time, 20, 2,37,31, 3
Accelerometer, 21, -1.3736732,7.833828,5.5940523, 3
Gyro, 22, 0.02321288, -0.029626964,0.047036625, 3
Orient, 23, 8.926149,-53.32324,-9.839288, 3
Gravity, 24, -1.3640951,7.86511,5.6964626, 3
AccLin, 25, -0.009578109,-0.031281948,-0.10241032, 3
RotVec, 26, 0.45423773,0.055024605,-0.046404224, 3
Time, 27, 2,37,31, 3
Accelerometer, 28, -1.364695,7.6105685,5.41748, 3
Gyro, 29, 0.023823744, 0.014049901,0.0125227375, 3
Orient, 30, 8.733725,-53.44793,-9.638855, 3
Gravity, 31, -1.3379325,7.877838,5.685068, 3
AccLin, 32, -0.026762486,-0.2672696,-0.26758814, 3
RotVec, 33, 0.4549593,0.053501204,-0.045224868, 3
Time, 34, 2,37,31, 3
```

(b)

Figure 3-6 SensoDuino sensor (a), an example of collected data in Text format (b).

The sheep movement data via SensoDuino were collected through **three attempts** performed on more than one visit to Lodge Farm at Moulton College.

The **first visit** was on 17 January 2017, where 7 sheep participated in the data collection experiment. The sensor was attached around the sheep's neck for 3-7 minutes (one at a time) to retrieve measurements from Acc, Gyr, $AccLin$, and $Orient$ sensors while the sheep were walking on a flat field at 10 Hz sampling rate. The same group of sheep were mounted again with SensoDuino set to be at 4 Hz sampling, more details in Table 3-3. Choosing two different sampling rates for the same group of sheep would justify the optimal sampling rate for lameness detection and how that could affect the classification rate.

The sheep at the time of the first experiment were manually labelled by the expert shepherd at Lodge Farm into either purple or green colour to refer to severely lame and mildly lame sheep respectively, while the non-labelled sheep represented sound sheep status within the flock, see Figure 3-7. Seven sheep participated in the experiment of the first visit with SensoDuino were 2 severely lame sheep, 2 mildly lame sheep, and 3 sound sheep.

Figure 3-7 Manually labelled sheep by the shepherd at Lodge Farm (purple, green, non-labelled sheep's colour refer to severe, mild, and non-lame sheep respectively)

The **second attempt** for data collection with SensoDuino occurred on 26 September 2017. Like the first attempt, the four basic sensors' readings were considered (Acc, Gyr, $AccLin$, and $Orient$) at a sampling rate of 10 readings per second from each axis. Although the data measurements from Gravity and Rotational vector sensors were also obtained during the experiment, these data were neglected in developing SLDM process. Eighteen sheep were equipped with SensoDuino (one at a time) for a period of approximately 5-10 minutes. Ten out of 18 tested sheep were sound, while the remaining 8 sheep had a different level of lameness range from mildly to severely lame (Table 3-3).

During the experiment, the 3rd sheep was tested twice as the first deployment failed because the sensor's collar was unfastened which led to the discarding of the current readings and redeploying the sensor to record a new reading to be taken into account. In addition, sheep number 12 had two different deployments; one test was executed on a flat field, while the other test was done on a grass field rather than a stable yard. The reason for choosing two different walking environments for the same sheep (12th sheep) was to justify the rate of classification error for the SLDM for the sheep while walking on varied terrain.

The **last attempt** with SensoDuino conducted on 26 October 2017. It was a day to remember as it was not only a visit to collect data; but the BBC 'The One Show' team were filming to record a report talking about the research of early lameness detection that had been conducted at the University of Northampton. The attractive report was prepared by Kevin Duala and broadcast on 16 November 2017.

At the same time of recording the report, the data was being collected at Lodge Farm from two sheep, one was mildly lame, and the other was sound. So, an extra mobile device was needed for this purpose besides the one already being used (Galaxy S4) for the data collection. A Galaxy S2 Android 4.1 was mounted on the sound sheep to collect movement data at the same setting of SensoDuino; 10 readings were retrieved every second. Unlike the Galaxy S4 Android 5.0, the Galaxy S2 device does not support the orientation sensors; therefore, no Orient readings were collected. Moreover, when the file was read afterwards, only Acc readings were obtained along 3 axes (3 predictors) due to a setting error, so there were no Gyr readings either.

At Lodge Farm, two sheep; one severely lame and one sound, were mounted with one sensor each (two devices Galaxy S4 and S2 were used) Table 3-3. While the BBC 'The One Show' team were recording their report, the data from deployed sensors were logged through the whole time which approximately lasted for an hour and a half. Consequently, 2 large data files were obtained for each sheep separately due to the long period of recording time at this attempt of the data collection. Unfortunately, the sound sheep file was damaged, and only the lame sheep file was prepared for the next step. Afterwards, when the lame file was read, it was discovered that there was no orientation data in it; thus only Acc, Gyr, and $AccLin$ (9 predictors instead of 12) were recorded in this experiment.

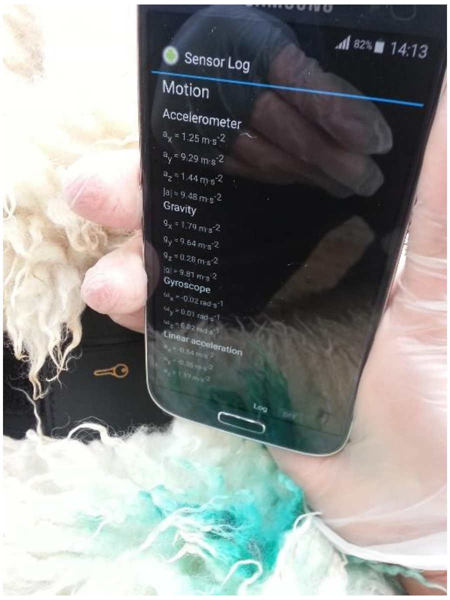
On the same day, another Farm was being visited belonging to Richard Harris near to Lodge Farm at Moulton College, where the BBC team continued to record footage there. Data were collected from two sheep there as well; one mildly lame the other was sound (Table 3-3). After roughly equivalent to 1 hour, the large gathered data file from the lame sheep that was mounted with S2 device only got Acc readings for 3 axes (3 predictors). On the other hand, the sound sheep file, which was mounted with S4 got Acc, Gyr, $AccLin$, and again no $Orient$ readings in it; thus, only 9 predictors were obtained.

Table 3-3 Metadata for data collected from SensoDuino sensor at both 10 Hz and 4 Hz sampling rate.

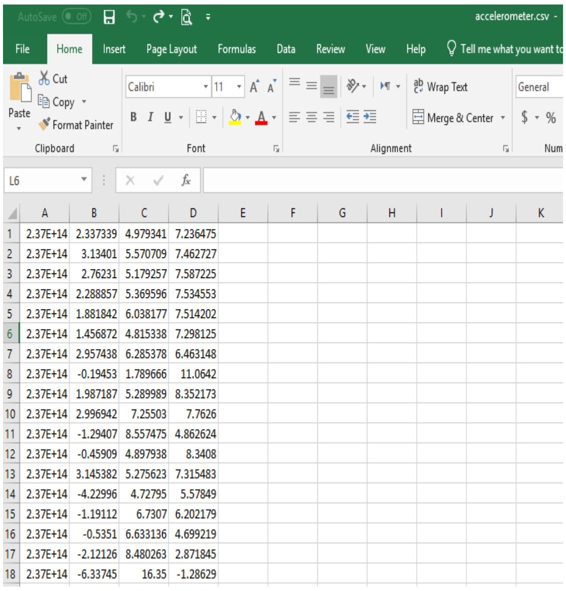
Attempt No.	Date	# Sheep in each experiment	Sampling rate	Sensor readings of interest	No. of reading records (rows)	# Sheep considered for the next step
1st attempt	17 Jan. 2017	7 sheep (3 sound, 2 mildly lame, 2 severely lame)	10 Hz	Acc, Gyr, AccLin, Orient (some missing readings in Gyr)	2000 - 3000 readings	All 7 sheep
			4 Hz	Acc, Gyr, AccLin, Orient (extra readings for Gyr)	500 – 1400 readings	All 7 sheep
2nd attempt	26 Sep. 2017	18 sheep (10 sound, 5 mildly lame, 3 severely lame)	10 Hz	Acc, Gyr, AccLin, Orient	2500- 4200 readings	All 18 sheep
3rd attempt	26 Oct. 2017	4 sheep (1 lame, 1 sound from both Farms)	10 Hz	(Acc, Gyr, AccLin, no Orient) for severely and sound sheep. (Acc only) for mildly lame sheep	Observed for 1-2 hr, (36000- 72000 readings)	3 sheep (1 mildly lame, 1 sound from Richards Farm), (1 severely lame sheep from Lodge Farm)
Total obtained sheep and their class			10 Hz			28 sheep (14 sound, 8 mildly lame, 6 severely lame)
			4 Hz			7 sheep (3 sound, 2 mildly lame, 2 severely lame)

3.3.3.3 Sensor Log

Sensor Log is a free Android application which records sensor data for twelve different sensors at the same time in 3 axes. Although the existing application has been updated (GitHub, 2014) with some new features, the data was collected via the previous release. So, there were 12 output data files that contained each sensor data separately in a CSV format (Figure 3-8). Approximately, the readings were obtained every 5.58 seconds. The available sensors were Accelerometer (Acc), Ambient_temperature, Gravity, Gyroscope (Gyr), Illuminance, Linear_acceleration ($AccLin$), Magnetic_field, Orientation ($Orient$), Pressure, Proximity, Relative_humidity, and Rotation_vector. However, only Acc, Gyr, $AccLin$, and $Orient$ sensors were involved in the lameness detection process.



(a)



	A	B	C	D
1	2.37E+14	2.337339	4.979341	7.236475
2	2.37E+14	3.13401	5.570709	7.462727
3	2.37E+14	2.76231	5.179257	7.587225
4	2.37E+14	2.288857	5.369596	7.534553
5	2.37E+14	1.881842	6.038177	7.514202
6	2.37E+14	1.456872	4.815338	7.298125
7	2.37E+14	2.957438	6.285378	6.463148
8	2.37E+14	-0.19453	1.789666	11.0642
9	2.37E+14	1.987187	5.289989	8.352173
10	2.37E+14	2.996942	7.25503	7.7626
11	2.37E+14	-1.29407	8.557475	4.862624
12	2.37E+14	-0.45909	4.897938	8.3408
13	2.37E+14	3.145382	5.275623	7.315483
14	2.37E+14	-4.22996	4.72795	5.57849
15	2.37E+14	-1.19112	6.7307	6.202179
16	2.37E+14	-0.5351	6.633136	4.699219
17	2.37E+14	-2.12126	8.480263	2.871845
18	2.37E+14	-6.33745	16.35	-1.28629

(b)

Figure 3-8 Sensor Log sensor (a), an example of an Accelerometer CSV file (b).

Concurrently with the data collection trial with SensoDuino on 17 January 2017, the Sensor Log was also operational for the same number of sheep; 2 severely lame, 2 mildly lame, and 3 sound sheep to collect data of Acc, $AccLin$, Gyr, and $Orient$ in 3 axes (see Table 3-4). Nevertheless, the collected data were obtained in separate files (3 columns each in addition to time), the whole data from 4 sensors were manually combined into one file to form 13 columns including a time column.

The variety of sensor applications used with different sampling rates for data collection would be necessary for the sake of comparison of sensors' performance in terms of the most related sensor for the early indication of lameness, its accuracy, and the most suitable sampling rate.

Table 3-4 Metadata for data collected from the Sensor Log sensor at 5.58 Hz sampling rate.

Attempt No.	Date	# Sheep in each experiment	Samplin g rate	Sensor readings of interest	No. of reading records (rows)	# Sheep considered for the next step
Only one attempt	17 Jan. 2017	7 sheep (3 sound, 2 mildly lame, 2 severely lame)	5.58 Hz	Acc, Gyr, AccLin, Orient (some missing readings in Gyr)	200 - 800 readings	All 7 sheep
Total obtained sheep and their class						7 sheep (3 sound, 2 mildly lame, 2 severely lame)

3.4 DataSets Aggregation for Pre-processing Stage

Data were aggregated according to the similarity of their sampling rate, which yielded three final sheep Datasets named as; DataSet1, DataSet2, and DataSet3 and refer to 5 Hz, 10 Hz, and 4 Hz sampling rates respectively. The combination of the final three DataSets with their sub DataSets are listed in Table 3-5.

The sensors involved in the process of lameness detection were Accelerometer (Acc), Gyroscope (Gyr), Linear acceleration ($AccLin$), and Orientation ($Orient$). However, in some data collection trails, some sensors readings were missed. Therefore, the missing sensor readings like $AccLin$ and $Orient$ readings could be calculated from already obtained Acceleration readings (Section 3.5.2.1).

As the aim of the research is to detect the early lameness signs in sheep with less sensor power consumption and a smaller set of attributes, the methodology to reach this aim needs to be tried with a different combination of DataSet' characteristics for optimal calculation and accurate level of the disease indication. Thus, the only Accelerometer hardware sensor readings aimed to be retrieved for the final lameness detection process as the software Orientation sensor (Pitch and Roll) sensors readings could be retrieved from Equ 3-2 and Equ 3-3. Furthermore, the software Linear accelerometer sensor ($AccLin$) could be calculated from Equ 3-5.

Table 3-5 Final Sheep Datasets for the next pre-processing stage (* indicates readings with missing values).

Sheep DataSets		Original Source	Data collection tools	Sample rate	Sensor manipulated	Total No. of Sheep for the next step
DataSet1	a	Table 3-2 (1st + 2nd attempts)	Sheep Tracker	5 Hz	Acc, Gyr, $Orient$	24 sheep (9 sound, 8 mildly lame, 7 severely lame)
	b	Table 3-4	Sensor Log	5.58 Hz	Acc, $Orient$, $AccLin$, Gyr^*	7 sheep (3 sound, 2 mildly lame, 2 severely lame)
DataSet1_all		Table 3-2 + Table 3-4	Sheep Tracker + Sensor Log	≈ 5 Hz	$Acc + Orient$	**31** sheep (12 sound, 10 mildly lame, 9 severely lame)
DataSet2	a	Table 3-3 (1st attempt)	SensoDunio	10 Hz	Acc, $Orient$, $AccLin$, Gyr^*	7 sheep (3 sound, 2 mildly lame, 2 severely lame)
	b	Table 3-3 (2nd attempt)			Acc, Gyr, $Orient$	18 sheep (10 sound, 5 mildly lame, 3 severely lame)
	c	Table 3-3 (3rd attempt) BBC			Acc only	3 sheep (sound, mildly lame, and severely lame)
DataSet2_all		Table 3-3 (1st +2nd +3rd attempts)	SensoDunio	10 Hz	$Acc + Orient$	**28** sheep (14 sound, 8 mildly lame, 6 severely lame)
DataSet3		Table 3-3 (1st attempt 4 Hz)	SensoDunio	4 Hz	Acc, Gyr, $Orient$	7 sheep (3 sound, 2 mildly lame, 2 severely lame)
DataSet3_all					Acc, $Orient$	**7** sheep (3 sound, 2 mildly lame, 2 severely lame)

3.5 Sensor Data Pre-processing

Normally, the real-world data is likely to be imperfect, incomplete, noisy, inconsistent, and redundant. At this stage, the importance of data preparation is essential for the next step of data mining (García *et al.*, 2016). In KDD, the data-pre-processing is considered as a powerful tool to generate more qualitative datasets than the original ones which could significantly enhance the data mining process (Zhang *et al.*, 2003). Although the pre-processed dataset is the final training set, which is manipulated by the classifier, a time-consuming procedure is undertaken to produce this final dataset (Kotsiantis *et al.*, 2006).

There are several methods that have been applied in predictive DM tasks which are reviewed by Alexandropoulos *et al.*, (2019). However, the focus of the following sections will be on the methods that are implemented on Sheep DataSets for the purpose of lameness detection. Figure 3-9 illustrates the steps that are followed to pre-process the Sheep DataSets in order to be classified into its class (sound, mildly lame, or severely lame) according to the classifier employed.

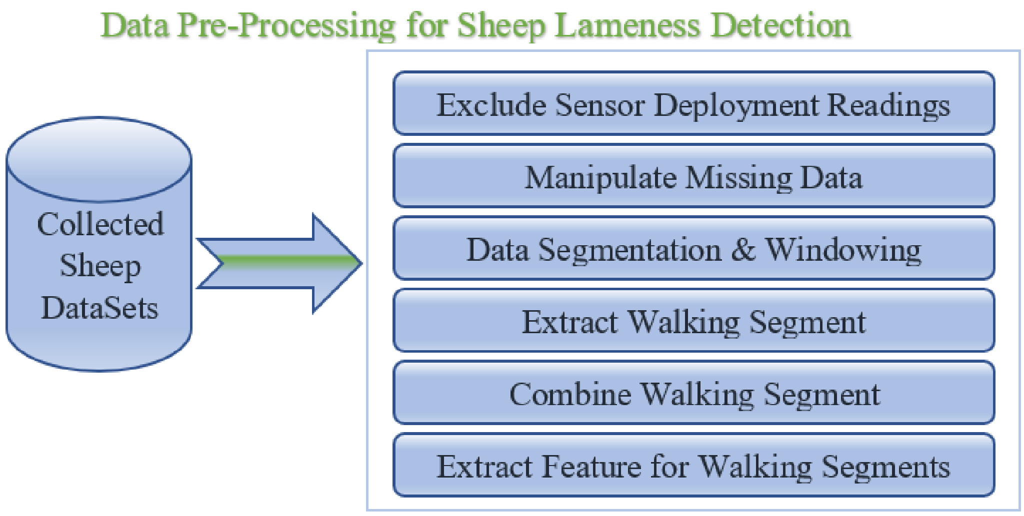


Figure 3-9 Data Pre-Processing Stages of the Sheep DataSet for Lameness Detection.

3.5.1 Noisy Data Manipulation (Exclude Deployment Time Readings)

In the Data collection stage, when the individual sheep was caught for deploying the sensor within the collar around its neck, the recorded sensor readings at deployment time (the time when the sensor was put on and off an individual sheep) were unreliable as the targeted sheep could make abnormal movements. Furthermore, the sheep needed time to settle with the new attached equipment in order to get into its normal walking pattern. The data gathered at deployment time would be noisy and may affect the classification process accuracy. Therefore, a chunk of reading records or instances needed to be removed from the final training data set. The time that was required to be discarded can be specified by the user (shown in the next chapter). The preferred chosen time to discard its sensor readings records was a 3 second period as no more data would be neglected for the next stage.

3.5.2 Missing Data Manipulation

Some sensor data are missed in the data collection process as it is clearly illustrated in Table 3-5. Orientation sensor readings are totally missed in DataSet2_c while Gyroscope readings are partially missed in DataSet1_b, DataSet2_a, and completely missed in DataSet2_c. Thus, the following subsections provide solutions to these raised issues due to the data collection process.

3.5.2.1 Calculating Pitch and Roll from Accelerometer

As mentioned earlier, Accelerometer is a hardware-based sensor whose data is already collected through the data collection step. However, the readings measurements for the software-based sensors like Pitch, Roll could be directly derived from the Accelerometer. In order to save the battery drainage of the sensor attached to the sheep for lameness detection, the retrieving sensor readings could be reduced by calculating the software sensor values from already existing hardware sensors. So, the less use of predictors to indicate the early signs of lameness, the more efficient the process would be.

So, the Pitch and Roll of the mounted sensor on the sheep neck could be calculated by using Equ 3-2 and Equ 3-3 respectively.

$$Pitch = atan2\left(Acc_y, Acc_z\right) * {}^{180}/_{\pi} \qquad \text{Equ 3-2}$$

$$Roll = tan^{-1}\left(\frac{-Acc_x}{\sqrt{Acc\_y^2 + Acc\_z^2}}\right) * {}^{180}/_{\pi} \qquad \text{Equ 3-3}$$

Where Acc_x, Acc_y, and Acc_z refer to the Acceleration sensor readings in a given time slice in the x, y, and z axes. Due to the range of Pitch angle being between [-180,180], the Four-quadrant inverse tangent function (*atan2)* was used to obtain values in the closed interval [-pi, pi] based on Acc_y and Acc_z values (MatLab documentation). The values of Pitch and Roll are measured in degrees, so they multiplied by 180/π.

Pitch and Roll are extracted as features in the feature engineering process to be used by the classifier to differentiate between various sheep behaviours in (Alvarenga *et al.*, 2016). Moreover, the same features are included to classify the behaviours of birds and humans (Collins *et al.*, 2015; van Kuppevelt *et al.*, 2019; Zhang *et al.*, 2014; Davila *et al.*, 2017).

Both equations were applied to DataSet2_c (Table 3-5) as Orientation readings were missed after the data were being collected.

3.5.2.2 Manipulate Gyroscope Missing readings

Missing values in the collected datasets could be managed in various ways (Kotsiantis *et al.*, 2006). Some of the ways are either by replacing it within the most frequent value within the vector of data or substituting it by the average value of the data vector. In both sheep datasets DataSet1_b and DataSet2_a (Table 3-5), the missing values of Gyr were replaced by the average value of the other already retrieved data for each axis (Gyr_x, Gyr_y, and Gyr_z). In contrast to DataSet2_c where the whole Gyroscope readings were missed, so in this case, it may not be applicable to calculate the Gyr readings as Gyroscope is a hardware-based sensor which could not be estimated in case of their readings are lost.

3.5.2.3 Calculating Linear accelerometer from Accelerometer

As mentioned in (Section 3.2.2.1), the built-in Accelerometer sensor readings represent raw acceleration values that include both static and dynamic components of raw acceleration data as follows:

$$Raw_{Acc} = Dynamic_{Acc} + Static_{Acc} \qquad \text{Equ 3-4}$$

Where the $Dynamic_{Acc}$ represents the animals' movements only, while $Static_{Acc}$ belongs to the force of gravity field to the earth (Nathan *et al.*, 2012). So, the already obtained sensor readings of $AccLin$ estimate the dynamic acceleration of the body to which the sensor attached. However, the $Dynamic_{Acc}$ (which is equvilant to $AccLin$ software-based sensor) could be calculated by applying the running mean over a selected window size of a given Acc vector.

Therefore, $Dynamic_{Acc}$ is calculated for each element in Raw_{Acc} by subtracting that element from its running mean value of pre-selected window size w (Gleiss *et al.*, 2011; Qasem *et al.*, 2012; Ladds *et al.*, 2017).

$$Dynamic_{Acc}(i) = Raw_{Acc}(i) - \frac{\sum_{j=1}^{w} Raw_{Acc(i)}}{w} \qquad \text{Equ 3-5}$$

Where i represents the acceleration readings vector of a specific axis, while j refers to each acceleration reading within the selected window of size w to calculate the running mean. This process filters Raw_{Acc} from its gravitational component and returns the dynamic components of Acceleration readings that only relate to the sheep movement.

Example 3.1: Assume that the acceleration readings for the forward-backwards movements $Raw\ Acc_y$= (2, 1, -1, -2, 3, 4), and w= 3 (window size). To calculate the $Dynamic_{Acc\_y}$ component only, the running mean is performed by centring the element in the current position in the w and find the average over that window. When the element in the window does not fill w, then the average is taken for the only included elements in w (MatLab documentation). Then the running mean is subtracted from each element of the $Raw\ Acc_y$

$Raw\ Acc_y$ = 2, 1, -1, -2, 3, 4
$Runing\ mean =$ 1.5, 0.66, -0.66, 0, 1.66, 3.5
$Dynamic\ Acc_y$ = 0.5, 0.33, -0.33, -2, 1.33, 0.5

Static acceleration could be used to estimate the animal's posture, while the dynamic acceleration is used to estimate the changes in the behavioural pattern of animals (Bailey et al., 2018). Thus, the Dynamic components of raw acceleration readings are utilised to calculate the speed of sheep, which is an important criterion to classify sheep behaviour into standing, walking, and trotting. The walking behaviour would also act as an important indicator of the lameness detection process.

3.5.3 Sensor Data Segmentations

The purpose of data segmentation in this research is for the sake of choosing the right segments (walking segments) among standing or trotting segments to be included in the classification process to detect the sheep lameness class. In addition, the sensor battery consumption of data transferring from the sensor node to the base station or where the data needs to be collected is the target of interest. Therefore, if only the walking segments are extracted and included in the classification process, that would be more efficient, and an energy-saving process rather than the case where the data are collected for any sheep posture. Thus, the first step after the data were collected is to identify the segments whose behaviour relates to sheep walking only rather than standing or trotting segments (Section 3.5.4); then the lameness classification process is applied based on the extracted walking segments only.

In general, the segmentation process is crucial in the pre-processing stage of a classification task as it could impact the next stage of feature extraction and even affect the accuracy of the classification (Bersch *et al.*, 2014). Furthermore, the complexity of the chosen segmentation method must be considered beforehand; especially in real-world classification tasks as the higher computational method could cause greater battery drainage when it comes to sensor saving energy issues.

The segmentation of data in the pre-procession stage has two different techniques, either online or offline. In online segmentation techniques, the collected data could start to be segmented before the whole datasets are entirely collected. In contrast, offline segmentation techniques require the whole dataset before starting the segmentation process (Bersch *et al.*, 2014).

Due to the purpose of this research being to detect the early signs of lameness in sheep as soon as possible by analysing real-world data from sheep, so the need for using online segmentation methods would be more applicable, rather than the offline ones as the lameness detection task is a real-world problem which needs its collected data to be segmented, once the data acquired.

Besides the online capability of the online segmentation methods, they could perform well on noisy data, are easy to understand due to its simple computation, and commonly used in health monitoring research studies (Keogh *et al.*, 2001; Bersch *et al.*, 2014).

3.5.3.1 Fixed-size Non-overlapping or Overlapping Sliding Window (FNSW, FOSW)

A sliding window is a common online segmentation method that is used to divide the raw input data into small chunks to be dealt with as input segments to the classifier. When a fixed-size sliding window divides the whole data point equally without interference among the adjacent data points, the technique is then called Fixed-size Non-Overlapping Sliding Window (FNSW). Alternatively, the online segmentation technique is called Fixed-size Overlapping Sliding Window (FOSW) when the sliding window has data overlap with a pre-defined ratio (Bersch *et al.*, 2014).

The size of each segment seg_size is calculated by applying Equ 3-6, while the number of segments for each individual data file (sheep) seg_no are obtained from Equ 3-7 in the case of applying FNSW; otherwise Equ 3-8 is applied to calculate the seg_no for FOSW segmentation techniques.

$seg\_size = sz * sr$ **(FNSW, FOSW)** Equ 3-6

$seg\_no = N/_{seg\_size}$ **(FNSW)** Equ 3-7

$seg\_no = N/_{seg\_size - \ seg\_size \ * \ or/100}$ **(FOSW)** Equ 3-8

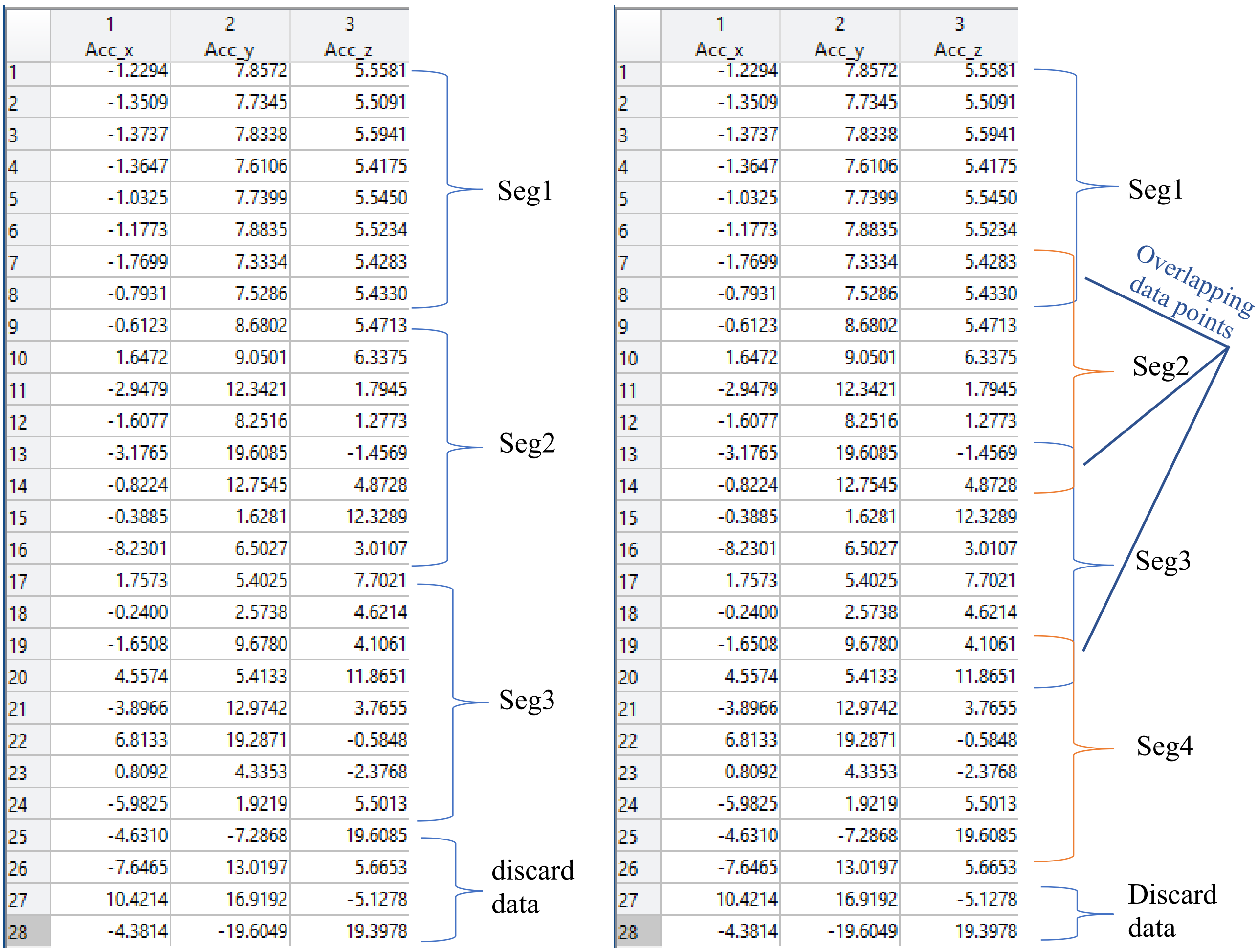
In Equ 3-6, sz represents the length of the pre-selected window in seconds (three options were implemented in the research which are 3, 7, and 10 sec.). Each segment is expected to hold information about the sheep's movement which is tested in the next steps. The shorter window size sz may not be enough to hold the characteristics of an individual movement, and longest sz may conflict two different movements in one segment. Since sheep tend to have less variable movements or transition in behaviour, no need for a long window size (more information), which is normally applied to differentiate complex behaviour rather than simple walking movements (Walton *et al.*, 2018). However, sheep behaviour classification research studies recommend different window sizes sz such as 10 sec. (Alvarenga *et al.*, 2016; Barwick, 2018a), 7 sec. (Walton *et al.*, 2018; Mansbridge *et al.*, 2018; Vazquez Diosdado *et al.*, 2018), and 5 sec. (Marais *et al.*, 2014; le Roux *et al.*, 2017). Thus, three options of window size sz were tested to identify a suitable period for one cycle of sheep movement in the current research.

The other factor in Equ 3-6, sr refers to the sampling rate in Hz of each sensor type used in data acquisition. Due to data aggregation where gathered data produced three groups of DataSets, three sampling rates were obtained 10 Hz, 5 Hz, and 4 Hz for each DataSet, respectively (see Section 3.4).

When the FNSW technique is applied, Equ 3-7 calculates seg_no the number of segments to be obtained from the whole data reading points that belong to each individual sheep within its DataSet. N represents the whole data-points of an individual sheep. The last segment of each individual sheep was discarded each time the number of data points within that segment were less than the seg_size Usually the last data points of each separate sheep file referred to the time when the sensor was taken off from that individual sheep, so no valuable data-points were lost.

Conversely, Equ 3-8 calculates seg_no when FOSW segmentation technique is applied. The N, and seg_no parameters are the same as FNSW; however, or represents the overlapping ratio among the data-points reading of the attached sensor to the sheep neck. The value of or could be identified previously by the user.

Example 3.2: Suppose we have raw Acc sensor data D (X, P), X= $\{x_1, x_2, \ldots, x_i\}$ represents the set of the collected data-points (here N=28), while P= $\{p_1, p_2, \ldots, p_j\}$ represents the set of predictors (3 predictors here Acc_x, Acc_y, Acc_z). Let us assume the sampling rate of sr= 4 Hz, window size sz = 2 sec., and segment overlap ratio or=20%. So, the number of overlap segments ($seg\_size * or /100$) ≅ 2 segments (the fraction segment is rounded to full segment length). Figure 3-10 illustrates the difference between two types of the online segmentation methods used in the research.



	1 Acc_x	2 Acc_y	3 Acc_z
1	-1.2294	7.8572	5.5581
2	-1.3509	7.7345	5.5091
3	-1.3737	7.8338	5.5941
4	-1.3647	7.6106	5.4175
5	-1.0325	7.7399	5.5450
6	-1.1773	7.8835	5.5234
7	-1.7699	7.3334	5.4283
8	-0.7931	7.5286	5.4330
9	-0.6123	8.6802	5.4713
10	1.6472	9.0501	6.3375
11	-2.9479	12.3421	1.7945
12	-1.6077	8.2516	1.2773
13	-3.1765	19.6085	-1.4569
14	-0.8224	12.7545	4.8728
15	-0.3885	1.6281	12.3289
16	-8.2301	6.5027	3.0107
17	1.7573	5.4025	7.7021
18	-0.2400	2.5738	4.6214
19	-1.6508	9.6780	4.1061
20	4.5574	5.4133	11.8651
21	-3.8966	12.9742	3.7655
22	6.8133	19.2871	-0.5848
23	0.8092	4.3353	-2.3768
24	-5.9825	1.9219	5.5013
25	-4.6310	-7.2868	19.6085
26	-7.6465	13.0197	5.6653
27	10.4214	16.9192	-5.1278
28	-4.3814	-19.6049	19.3978

(a) FNSW segmentation

	1 Acc_x	2 Acc_y	3 Acc_z
1	-1.2294	7.8572	5.5581
2	-1.3509	7.7345	5.5091
3	-1.3737	7.8338	5.5941
4	-1.3647	7.6106	5.4175
5	-1.0325	7.7399	5.5450
6	-1.1773	7.8835	5.5234
7	-1.7699	7.3334	5.4283
8	-0.7931	7.5286	5.4330
9	-0.6123	8.6802	5.4713
10	1.6472	9.0501	6.3375
11	-2.9479	12.3421	1.7945
12	-1.6077	8.2516	1.2773
13	-3.1765	19.6085	-1.4569
14	-0.8224	12.7545	4.8728
15	-0.3885	1.6281	12.3289
16	-8.2301	6.5027	3.0107
17	1.7573	5.4025	7.7021
18	-0.2400	2.5738	4.6214
19	-1.6508	9.6780	4.1061
20	4.5574	5.4133	11.8651
21	-3.8966	12.9742	3.7655
22	6.8133	19.2871	-0.5848
23	0.8092	4.3353	-2.3768
24	-5.9825	1.9219	5.5013
25	-4.6310	-7.2868	19.6085
26	-7.6465	13.0197	5.6653
27	10.4214	16.9192	-5.1278
28	-4.3814	-19.6049	19.3978

(b) FOSW segmentation

Figure 3-10 Two types of segmentation techniques for sheep data.

3.5.4 Classify Sensor Data Segments into Three Moving Behaviours

In this research study, gait behaviour of sheep was classified into three classes Standing, Walking, and Trotting by applying threshold limits for the normal walking of sheep. Although all of the sheep during the data collection process were triggered to walk normally on a flat or field area, on many occasions sheep either stood or walked at a faster speed than the normal one, which is then named trotting in this research. Therefore, the research aimed to extract only the walking period to be processed and analysed for the task of lameness classification. Usually, lame sheep are noticed while they walk rather than standing or trotting, as the lame animal is willing to use their affected limbs when walking, in contrast to trotting where they tend to carry their infected limbs (Kim and Breur, 2008).

In sheep gait studies the normal walking speed is identified to be within 1.1-1.3 m/s ranges (Agostinho *et al.*, 2012) or less (Squires *et al.*, 1972); however, this range could be changeable according to the breeds and the environment when the data were collected. In the case of the current research, the sheep in the experiments were encouraged to walk at a slightly faster speed than the normal walking speed of sheep in an open field without monitoring. In the designed software for the purpose of this research study (Section 4.2), the range of normal walking could be pre-defined, which was selected to be between 0.8- 3.5 m/s.

The classification process for the sheep movements was performed by testing the speed of each segment of the sheep file in the targeted DataSet according to the following steps:

Step1: applying Equ 3-5 to the forward-backwards acceleration readings (Acc_y) within the segment to find the $Dynamic\ Acc\_y$ movements of the sheep without gravity interference.

Step2: in order to calculate the velocities corresponding to each dynamic acceleration reading of Acc_y in that segment, the numerical integration with respect to the time between each successive readings (seg_time) is applied using a trapezoidal method. In Equ 3-9, the seg_time is a vector containing time slices starting from 0 to seg_size increasing by $1/\ sr$ (sampling rate). The result of integration $Cum\_velocity$ is a vector equal to Acc_y in size containing commulative velocities corresponding to each sensor reading in that segment. However, Equ 3-10 was applied to obtain the $Pure\_velocities$ vector without its cumulative value.

$$Cum\_velocity = \int_{0}^{seg\_time} Dynamic\ Acc\_y(i) \qquad \text{Equ 3-9}$$

$$Pure\_velocity = \sum_{i=1}^{seg\_size} Cum\_velocity(i+1) - Cum\_velocity(i) \qquad \text{Equ 3-10}$$

Step3: the overall speed of the targeted segment of the sheep data file within the DataSet was calculated according to Equ 3-11 as the speed is the magnitude value of the velocities. The Seg_speed vector is equal in size to the seg_no, where each speed value corresponds to the one segment of the sheep data file.

$$Seg\_speed = \sqrt[2]{\sum_{i}^{n} Pure\_velocity\ (i)^2} \qquad \text{Equ 3-11}$$

Step4: test the speed of each segment within the specified range (upper-speed limit= 3.5 m/s, lower-speed limit = 0.8 m/s). The class of that segment is Standing if the Seg_speed is less than the lower limit. Conversely, the class movement behaviour of the targeted segment is classified Trotting if the Seg_speed exceeds the upper limits. Otherwise, the movement class is considered Walking when the Seg_speed is within the pre-defined limits. The output result is in a table which contains each speed segment with its corresponding class.

The pseudo-code for applying the four steps of the classification process for each sheep movement in a given DataSet is illustrated in Figure 3-11.

```
For each seg in sheep file Do
  For i= 1 to seg_no
      Compute Dynamic Acc_y        // Apply Equ 3-5
      Compute Cum_velocity         // Apply Equ 3-9
      Compute Pure_velocity        // Apply Equ 3-10
      Compute Seg_speed            // Apply Equ 3-11
                                   // Apply segment classification
      IF Seg_speed <= lower-speed limit Then
         Seg_class= 'Standing'
      Elseif Seg_speed <= upper-speed limit Then
            Seg_class= 'walking'
      Else Seg_class= 'Trotting'
      End if
  End
End
```

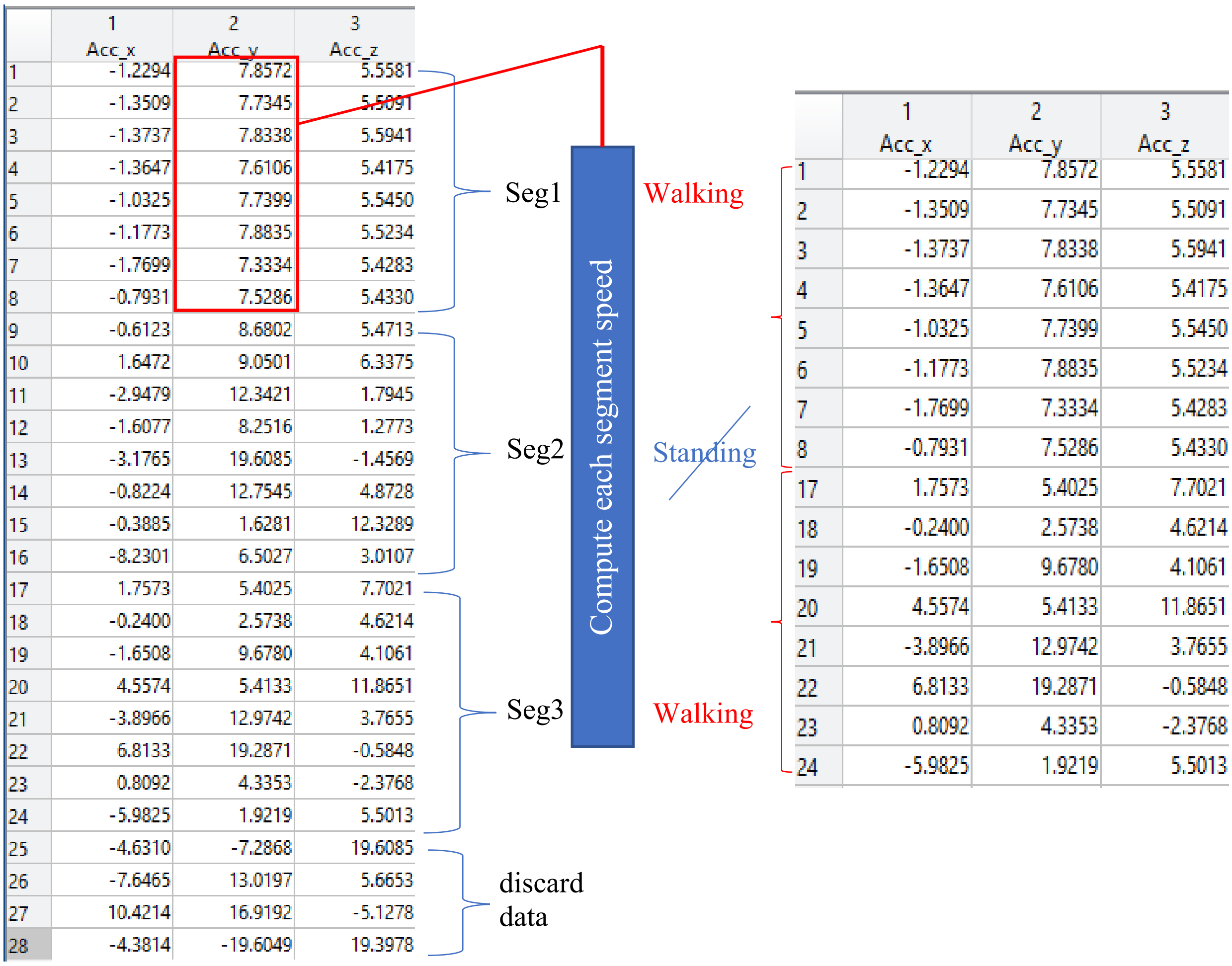
Figure 3-11 The pseudo-code for the classification process of the data sheep file' segments for a given DataSet.

3.5.5 Extract Walking Segments for a Sheep File

As the sheep data file was segmented and classified into three behaviours Standing, Walking and Trotting, the targeted segments for the next step (feature extraction) ought to be performed on the walking segments only. Therefore, the walking segments from each individual sheep file were extracted and the other Standing and Trotting segments were discarded from the next step of feature extraction.

Figure 3-12 illustrates an example for extracting $walk\_seg$ for a single sheep within a DataSet. While the pseudo-code for extracting the walking segments only for the individual sheep is presented in Figure 3-13.

The obtained $walk\_seg$ was added to the sheep file information in addition to the existing ones. An example of a single sheep file within a given DataSet is given with details in Figure 3-14, where each sheep file is in a *Struct* format containing many entries.



	1 Acc_x	2 Acc_y	3 Acc_z
1	-1.2294	7.8572	5.5581
2	-1.3509	7.7345	5.5091
3	-1.3737	7.8338	5.5941
4	-1.3647	7.6106	5.4175
5	-1.0325	7.7399	5.5450
6	-1.1773	7.8835	5.5234
7	-1.7699	7.3334	5.4283
8	-0.7931	7.5286	5.4330
9	-0.6123	8.6802	5.4713
10	1.6472	9.0501	6.3375
11	-2.9479	12.3421	1.7945
12	-1.6077	8.2516	1.2773
13	-3.1765	19.6085	-1.4569
14	-0.8224	12.7545	4.8728
15	-0.3885	1.6281	12.3289
16	-8.2301	6.5027	3.0107
17	1.7573	5.4025	7.7021
18	-0.2400	2.5738	4.6214
19	-1.6508	9.6780	4.1061
20	4.5574	5.4133	11.8651
21	-3.8966	12.9742	3.7655
22	6.8133	19.2871	-0.5848
23	0.8092	4.3353	-2.3768
24	-5.9825	1.9219	5.5013
25	-4.6310	-7.2868	19.6085
26	-7.6465	13.0197	5.6653
27	10.4214	16.9192	-5.1278
28	-4.3814	-19.6049	19.3978

(a) Original sheep (FNSW segmentation)

	1 Acc_x	2 Acc_y	3 Acc_z
1	-1.2294	7.8572	5.5581
2	-1.3509	7.7345	5.5091
3	-1.3737	7.8338	5.5941
4	-1.3647	7.6106	5.4175
5	-1.0325	7.7399	5.5450
6	-1.1773	7.8835	5.5234
7	-1.7699	7.3334	5.4283
8	-0.7931	7.5286	5.4330
17	1.7573	5.4025	7.7021
18	-0.2400	2.5738	4.6214
19	-1.6508	9.6780	4.1061
20	4.5574	5.4133	11.8651
21	-3.8966	12.9742	3.7655
22	6.8133	19.2871	-0.5848
23	0.8092	4.3353	-2.3768
24	-5.9825	1.9219	5.5013

(b) Extracted sheep data with walking segments only

Figure 3-12 Walking segments extraction for the individual sheep data file.

For each individual *sheep* file **Do**

$j = 1$ // counter for the $walk\_seg$ array

For i= 1 to seg_no

IF seg (i). $Class$ == 'Walking' **Then**

$walk\_seg(j)$= seg (i)

$j = j + 1$

End if

End

End

Figure 3-13 The pseudo-code for extraction walking segments for individual sheep.

Name	Value
Acc	table
Gyr	table
Ang	table
AccLin	table
time	table
Main_Class	1x1 cell
seg_data	100x9x30 double
vedba	30x1 double
speed	30x1 double
seg_class	30x1 cell
walk_seg	24x900 double

(a) An example of a single sheep file in a DataSet

1- Acc (X,3), accelerometer sensor readings for 3 axes.
2- Gyr(X,3), Gyroscope sensor readings for 3 axes.
3- Ang(X,3), orientation sensor readings for 3 axes.
4- AccLin(X,3), linear acceleration sensor readings for 3 axes.
5- Time is the sensor time while data were read.
6- Main_class is the observed sheep class.
7- Seg_data is 3D array refers to row, column, page (each page is a segment).
8- Speed is the speed of each segment
9- Vedba is vertical dynamic body acceleration of each segment (will discuss in feature extraction section)
10- Seg_class is the class of each segment after movement classification.
11- Walk_seg is the extracted walking segments only

(b) The details of each field in the sheep struct file

Figure 3-14 An example of an individual sheep file in a given DataSet.

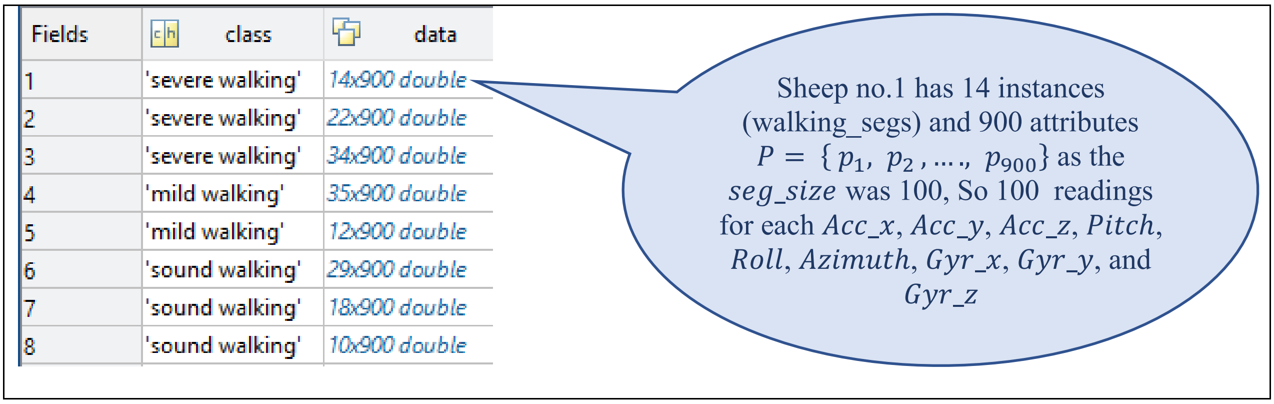
3.5.6 Combine Walking Segments for a DataSet

As mentioned earlier in data aggregation Section 3.4, three final DataSets were obtained, with 31, 28, 7 sheep, respectively (Table 3-5). So, the walking segments $walk\_seg$ for each sheep in that DataSet were combined together to produce a final $raw\_data\_table\ D(X,P)$ which will be ready to use as an input to the chosen classifier to perform the lameness classification task.

In the raw data table $D\,(X,P)$, each row $X = \{\,x_1,\ x_2\,,\dots,\ x_i\}$ represents an instance or example for the classifier to build the prediction model, and each column $P = \{\,p_1,\ p_2\,,\dots,\ p_j\}$ represents a predictor or attribute that the classifier depends on to predict the class of new instance $Y = \{\,y_1,\ y_2\,,\dots,\ y_k\}$, where Y is the class type, and k is the number of classes in a classification problem. As each classification problem could be presented as $D(X,P)\ =\ Y$.

The final $Raw\_data\_table\ D(X,P)$ was obtained by performing two steps:

Step1: in this step, all $walk\_seg$ data of all sheep in the DataSet were combined into one file called $Combine\_data$ (see Figure 3-16) for pseudo-code. The resulting $Combine\_data$ is in a *Struct* format that has entries equal to the number of sheep in that DataSet. Each entry has 2 fields (columns), where the first field refers to the sheep's class and the second field refers to a two-dimensional array of $walk\_seg$ data. The rows of $walk\_seg$ represent walking instances, while the columns represent the set of predictors $P = \{p_1, p_2, \dots, p_j\}$ of that sheep. Figure 3-15 shows $Combine\_data$ in an example of sheep DataSet including 8 sheep.



Fields	class	data
1	'severe walking'	14x900 double
2	'severe walking'	22x900 double
3	'severe walking'	34x900 double
4	'mild walking'	35x900 double
5	'mild walking'	12x900 double
6	'sound walking'	29x900 double
7	'sound walking'	18x900 double
8	'sound walking'	10x900 double

Figure 3-15 An example of a $Combine\_data$ DataSet Sheep file includes 8 sheep, 3 classes, and 147 walking segments (14+22+34+35+12+29+18+10= 147).

Step2: in this step, the data of each individual sheep in a DataSet was combined vertically together to get the Raw_data_table $D(X, P) = Y$. As an example of Figure 3-15, X= $\sum walk\_seg(i)$, where i refers to a sheep number in a DataSet. So, Figure 3-14 An example of an individual sheep file in a given DataSet $X = \{x_1, x_2, \dots, x_{174}\}$ walking segments. $P = seg\_size * no. of\ predictors + Class$, where $seg\_size = 100$ in the presented example, and the number of predictors (sensor readings parameters) = 9 (Acc_x, Acc_y, Acc_z, $Pitch$, $Roll$, $Azimuth$, Gyr_x, Gyr_y, and Gyr_z). Thus, the number of predictors = 100 × 9= 900. In addition, one extra column $Class$ was added for the class type of the current instance, so $P = \{p_1, p_2, \dots, p_{901}\}$. $Y = \{y_1, y_2, y_3\}$ as the class number in the example were three 'severe walking', 'mild walking', and 'sound walking'.

The pseudo-code for both steps including getting $Combine\_data$ and Raw_data_table for a given DataSet is depicted in Figure 3-16. Step1 involves combining walking segments of an individual sheep in a DataSet into one file with its class either 'severe walking', 'mild walking', or 'sound walking'. Step2 includes getting the Raw_data_table ready for next feature extraction, where each row represents a separate instance.

```
Step1
For each individual sheep file in a given DataSet Do
        j = 1                    // counter for new Combined _data
        For i= 1 to sheep_no
             IF sheep (i).Class == 'severe' Then
                 Combine_data (j).Class = 'severe walking'
                 Combine_data (j).data = sheep (i). walk_seg
                 j = j + 1
             Else
                 IF sheep (i).Class == 'mild' Then
                     Combine_data (j).Class = 'mild walking'
                     Combine_data (j).data = sheep (i). walk_seg
                     j = j + 1
                 Elseif
                      IF sheep (i).Class == 'sound' Then
                          Combine_data (j).Class = 'sound walking'
                          Combine_data (j).data = sheep (i). walk_seg
                          j = j + 1
                  End
              End
        End
End
///////////////////////////////////////////////////////////////////////////////////////////////////////////////
Step2
For each individual sheep file in a given DataSet Do
        j = 1                    // counter for new Combined _data
        For i= 1 to sheep_no
              Raw_data.Data(j)= Add Combine_data(i).data         // vertically
              Raw_data.Class(j)= Add Combine_data(i).Class
                                                 // expand to Combine_data(i).data
              j = j + 1
        End
//// Get all variables into table////
        //// Assume first 3 columns for Acc_x, Acc_y, Acc_z
        ////         second 3 columns for Pitch, Roll, Azimuth
        ////         third 3 columns for Gyr_x, Gyr_y, Gyr_z

        Acc= Raw_data (X,P), X= all instances, P=1 to 3 ×seg_size
        Ang= Raw_data (X,P), X= all instances, P=3 ×seg_size +1 To 6×seg_size
        Gyr= Raw_data (X,P), X= all instances, P=6 ×seg_size +1 To 9×seg_size

        Create Raw_data_table= [ Acc, Ang, Gyr, Raw_data.Class]
End
```

Figure 3-16 The pseudo code for combining walking segments for a sheep member of a given DataSet table Raw_data_table ready for the next feature extraction step.

3.5.7 Feature Extraction for Walking Segments

Feature extraction or sometimes called either data transformation, feature engineering, or attribute construction is a very important step in a data mining task as it influences the performance of the final classification task (Su et al., 2014). The feature extraction was implemented over a pre-selected window size (5, 7, 10 sec) in the specially designed software for lameness detection (to be discussed in Section 4.2).

The raw sensor data from the accelerometer, gyroscope, and orientation forms a multi-dimensional DataSet which may need to be optimised to reduce noise and error by extracting a new set of features which are called predictors or attributes to be involved in the classification task. The new set of features tend to be more useful and understandable in terms of structure and accuracy for high dimensional data (Jiawei et al., 2012). For example, the raw data of 10 values could be meaningless to identify the general trend of the current stream of data compared to the average of the tenth values.

According to the studies that have been done in the field of human activity recognition using raw data from an accelerometer (Figo *et al.*, 2010; Bersch *et al.*, 2014), there are many feature extraction techniques to be applied either in the time or frequency domain of the acceleration data stream. Although Figo *et al.*, (2010) survey explores these techniques for human activities, many features have been employed in the field of animal behaviour detection from either an accelerometer or gyroscope sensor. For example, feature extraction in cattle behaviour has been employed in (Rahman *et al.*, 2018; Smith *et al.*, 2015), and in sheep behaviour studies in (Marais *et al.*, 2014; Alvarenga *et al.*, 2016; Kamminga *et al.*, 2017; Barwick *et al.*, 2018a; Walton *et al.*, 2018; Guo *et al.*, 2018; Kleanthous *et al.*, 2018). A combination of features in the aforementioned references was implemented in the current research in addition to extra features which all are listed in Table 3-6.

Twenty-four features were extracted in this research over a pre-selected window (sz) from raw data for each axis of Accelerometer (3 axes), Gyroscope (3 axes), and Orientation (3 angels) sensors of sheep DataSet. As it is illustrated in Figure 3-17, the features were divided into seventeen features from the time domain where basic statistics of each data window were calculated, and seven features from the frequency domain where the signal periodic is described in Fast Fourier Transform FFT.

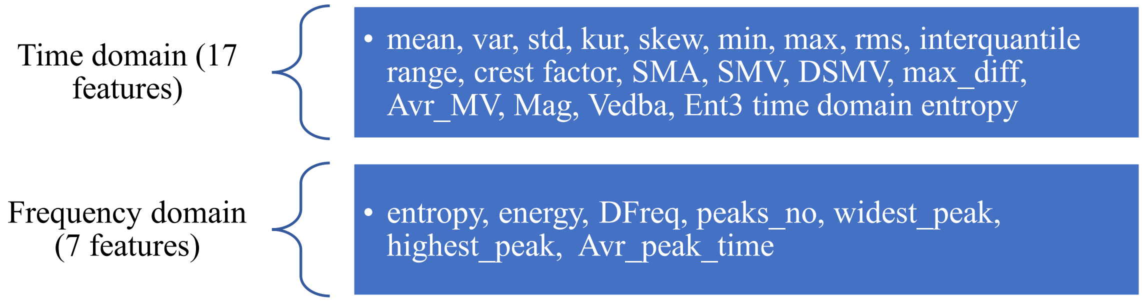


Figure 3-17 Extracted features from raw data of walking sheep.

The names, equations, meanings, and the number of resulting features for each instance (row) in the Raw_data_table excluding the $Class$ column ('severe walking', 'mild walking', or 'sound walking') over a selected window size or segment (sz) for a sheep file in a DataSet are explained in Table 3-6.

The pseudo-code to perform the features extraction for sheep data to be included in the lameness detection classifier is presented in Figure 3-18, the $Raw\_data\_table\ D\ (X, P)$ in each obtained DataSet that already contains only the walking data of sheep with its related other sensor readings has X instances and P predictors by noticing that the last column in P represents the $Class$ of that instance either 'severe walking', 'mild walking', or 'sound walking'. Therefore, the number of P was reduced by applying feature extraction and the best set of P (features) were only considered by the classifier in the next step.

For each $Raw\_data\_table\ D\ (X, P)$ in the feature extraction stage, the output data table was named $Featured\_data\_table\ (X, P`)$. Where X represents the same number of instances in both the Raw and $Featured$ data table. P is the number of predictors in the Raw_data_table $P = seg\_size * np$, where np is the number of predictors that were obtained from sensor readings. For example, if seg_size=100, then $P = \{p_1, p_2, \dots, p_{900}\}$, if seg_size=50, then $P = \{p_1, p_2, \dots, p_{450}\}$. While $P`$ of $Featured\ data\_table$ represents the new calculated features, which equal to 183 features from all axes as explained in Figure 3-18.

Table 3-6 Computed features from time and frequency domain for the sheep walking segments within a sheep DataSet, where the Raw_data_table referes to sensor raw data excluding the last $Class$ column.

Feature name (symbol)	Feature meaning	Feature equations for each Accelerometer, Gyroscope, and Orientation sensor readings	Computed features for each instance in a sheep file $P`$
Mean (μ)	measures the average activity of a selected window. it removes noise, random peaks, smooths data, and kind of axial calibration	$\mu = \frac{1}{sz}\sum_{i=1}^{sz} Raw\_data\_table(i)$	9 (as we have 9 predictors)
Variance (∂)	measures the variability of the data sequence, i.e. the deviation of movement from the mean	$\partial = \frac{1}{sz}\sqrt{\sum_{i=1}^{sz}(Raw\_data\_table(i) - \mu)^2}$	9
Standard deviation (∂^2)	Measure the spread of data within a selected window. It is equal to the square of ∂	$\partial^2 = (\frac{1}{sz}\sqrt{\sum_{i=1}^{sz}(Raw\_data\_table(i) - \mu)^2})^2$	9
Kurtosis (Kur)	Kur is the third standardized moment of each axis per window, measure how the outliers prone to distribute in a selected window	$Kur = \frac{1}{sz}\sum_{i=1}^{sz}\frac{(Raw\_data\_table(i) - \mu)^3}{\partial^3}$	9
Skewness (Skew)	Skew is the fourth standardized moment of each axis per window, measure the degree of data symmetry in a selected window	$skew = \frac{1}{sz}\sum_{i=1}^{sz}\frac{(Raw\_data\_table(i) - \mu)^4}{\partial^4}$	9
Maximum value (Max)	Maximum value within the selected window	Max (selected window of Raw_data_table)	9
Minimum value (Min)	The minimum value within the selected window	Min (selected window of Raw_data_table)	9
Root mean square (Rms)	Measure the energy distribution and randomness of the values within a selected window. It is used in human research to distinguish walking patterns and input to the classifier	$Rms = \sqrt{\frac{1}{sz}\sum_{i=1}^{sz} Raw\_data\_table(i)^2}$	9
Interquartile range (Interq)	Measure the variability of a selected window data	$Interq = Q3 - Q1$, where $Q1$ is the middle of the first half of data, $Q3$ is the middle of the third half of data	9
Crest factor (CF)	Measure the impulsiveness of the selected window, i.e. the sudden movement or behaviour. CF is the ratio of Max value to Rms value of a selected window	$Cf = \frac{Max(Raw\_data\_table)}{\sqrt{\frac{1}{sz}\sum_{i=1}^{sz} Raw\_data\_table(i)^2}}$	9

Signal magnitude area (SMA)	Measure the energy expenditure of walking sheep. Compute absolute integral which represents the area encompassed by the magnitude of acceleration, angular velocity, and angles within the selected window	$SMA_{Acc} = \frac{1}{t} * \int_0^t [\lvert Raw\_data\_table.Acc\_x(t)\rvert + \lvert Raw\_data\_table.Acc\_y(t)\rvert + \lvert Raw\_data\_table.Acc\_z(t)\rvert]\, dt$ An example for SMA for acceleration signal, where $t = seg\_time$	3
Signal vector magnitude (SMV)	Measure the degree of movement intensity of the selected window, also eliminates the inconsistency of sensor orientation	$SMV = \frac{1}{sz} \sum_{i=1}^{sz} sqrt[Raw\_data\_table.Acc\_x(i)^2 + Raw\_data\_table.Acc\_y(i)^2 + Raw\_data\_table.Acc\_z(i)^2]$ An example for SMV for acceleration signal, where $sz = seg\_size$	3
Differential Signal Vector Magnitude (DSMV)	Contribute to dynamic daily activity classification of sheep	$DSMV = \frac{1}{seg\_time} * \int (\lvert \sum_{1}^{seg\_size} SMV' \rvert dt)$ SMV' is the difference between two successive SMV values	3
Maximum difference (Max_diff)	Measure the largest changes between two successive sensor readings for each axis of a selected window	For i= 2 to $sz - 1$ $Max\_diff = Max(Raw\_data\_table(i+1) - Raw\_data\_table(i)$ End	9
Average movement variation (Avr_MV)	Measure the average movement variation along each axis of the selected window	$Avr\_MV = \frac{1}{sz} \sum_{i=1}^{sz-1} \lvert Raw\_data\_table.Acc\_x(i+1) - Raw\_data\_table.Acc\_x(i)\rvert + \sum_{i=1}^{sz-1} \lvert Raw\_data\_table.Acc\_y(i+1) - Raw\_data\_table.Acc\_y(i)\rvert + \sum_{i=1}^{sz-1} \lvert Raw\_data\_table.Acc\_z(i+1) - Raw\_data\_table.Acc\_z(i)\rvert$ An example of Avr_MV for acceleration signal, where $sz = seg\_size$	3

Magnitude (Mag)	Measure the intensity of each sensor reading for 3 axes each within a selected window. Reduce the complexity of sensor orientation	$Mag = [Raw\_data\_table.Acc\_x(i)^2 + Raw\_data\_table.Acc\_y(i)^2 + Raw\_data\_table.Acc\_z(i)^2]^{1/2}$ An example of Mag of Accelerometer sensor readings	3
Vectorial of the dynamic body acceleration (Vedba)	Measure the energy expenditure of a walking speed within a selected window	$Vedba = [Dynamic\_Acc\_x^2 + Dynamic\_Acc\_y^2 + Dynamic\_Acc\_z^2]^{1/2}$ Apply Equ 3-5 first, then calculate $Vedba$ for Accelerometer readings	3
Entropy Time-domain (Ent3)	Measure the impurity of movement data within the selected window.	$Ent3 = \frac{1}{sz}\sum_{i=1}^{sz}(1 + T\_Acc(i)) \times ln(1 + T\_Acc(i))$ $T\_Acc = Raw\_data\_table.Acc\_x + Raw\_data\_table.Acc\_y + Raw\_data\_table.Acc\_z$	3
Entropy Frequency- domain (Ent)	Measure the energy disorder of a selected window. It is used to discriminate the sheep's activities of the same energy.	For each element (i) in the selected window size (sz) **Do** 1- Find the power spectral (PS) of the selected window via discrete Fourier transformation (fft) $PS(i) = \lvert fft(Raw\_data\_table(i))\rvert^2$ 2- Find probability density function of the power spectrum PDF_PS normalised by summation of PS (i.e. normalised by its norm) to be treated as a probability function $PDF\_PS = \frac{PS(i)}{\sum_{i=1}^{sz} PS(i)}$ 3- Find the entropy (Ent) $Ent = \frac{-\sum_{i=1}^{sz} PDF\_PS(i) * Log_2\,(PDF\_PS(i))}{Log_2(sz)}$ **End**	9
Energy (Eng)	Measure the movement complexity of a selected window of walking sheep	$Eng = \frac{1}{sz}\sum_{i=1}^{sz}\lvert fft(Raw\_data\_table)\rvert^2$	9

Dominant frequency (Dfreq)	The 1st coefficient value of the spectral signal which has the largest value within the selected window	$Dfreq = Max(\lvert fft(Raw\_data\_table)\rvert$	9
Number of peaks (nPeak)	Calculates the number of peaks within a selected window	1- Find the absolute values of frequency domain FD for the selected window $FD = \lvert fft(Raw\_data\_table)\rvert$ 2- Find $[Pks, Locs, PW] = findpeaks\ (FD)$ Where, $Pks = \{pk_1,\ pk_2, \dots,\ pk_i\}$, vector of peaks values (local maxima), i represents no. of peaks ($nPeaks$). $Locs$ is the vector of indices at which the Pks happen, and PW is the vector of widths of each found peak in Pks in a selected window.	9
Widest peak (Widest_Peak)	Return the widest peak value in a selected window	$Widest\_Peak = Max(PW)$ Where PW is the vector of peaks' width values obtained from the function $findpeaks\ (FD)$	9
Highest peak (Highest_Peak)	Find the highest peak value in a selected window	$Highest\_Peak = Max\ (Pks)$ Where Pks is the vector of local maxima values obtained from the function $findpeaks\ (FD)$	9
Average peak time (Avr_peak_time)	Measure the average time between successive peaks in second.	$\mu\_\text{Diff}\_peaks\_Locs = \frac{1}{Pks} * \sum_{i=1}^{Pks}(Locs(i+1) - Locs(i))$ Where, $\mu\_\text{Diff}\_peaks\_Locs$ is the mean of differences of peaks' distance, and Pks is the found peaks vector. $Avr\_peak\_time = \mu\_\text{Diff}\_peaks\_Locs * \frac{1}{sr}$ Where sr, is the sample rate of a selected window	9
Total obtained features			**183**

```
For each sheep in Raw_data_table of a given DataSet Do
    Get seg_size = sz * sr                    // sz = window size, and sr= sampling rate
        For i= 1 to  X                        //X= no. of instances
            k = 1                             // counter for the new P`
            For j= 1 to P  Step seg_size      // P=seg_size*np
                All_feature_extract (i,k)  = Compute features (Raw_data_table)
                                              // Apply Table 3-6 for each segment
                k = k + 1
            End
    End
End
Create Featured_data_table= [ All_feature_extract, Raw_data_table.Class]
```

Figure 3-18 The pseudo-code for the feature extraction for each sheep file in a given DataSet.

3.6 Feature Selection (FS) for the Classifier

One of the significant key issues in performing a machine learning task is the feature selection (FS) which could be defined as the process where irrelevant features (features which have no effect on the class) and redundant features (features taking the role of another one) are being removed from the original dataset for the sake of obtaining a smaller optimal set of features (predictors) that would be sufficient to effectively describe the dataset and predict the class (label) of new instances (Alexandropoulos *et al.*, 2019).

The new selected features may be adequate to construct a more accurate and concise classifier that performs well in the classification task (Alexandropoulos *et al.*, 2019; Mwadulo, 2016). Since the FS process improves the interpretation of the generated model as the visualisation of the model formed from the fewer features is more understandable and comprehensible than the original set of features (Mwadulo, 2016).

Moreover, FS avoids the model over-fitting when it highly fits the trained dataset and not performing well on new unseen examples (García *et al.*, 2016; Mwadulo, 2016). Another advantage could be expected when the FS is applied is that the learning process tends to be faster and occupies less memory storage as the search space specified by the features is reduced. However, extra computations may be added to the overall data mining task when the FS is applied as $2^{P`}$ possible combination subset would search for the optimal selection and that would be complicated even if the size of feature search space is not too big (Kotsiantis *et al.*, 2006; Tang *et al.*, 2014).

Generally, FS approaches could be divided into two common ways *filter* and *wrapper*; however, a *hybrid* FS method represents a mixture of two previous ones, while another embedded FS method exists to bridge between the filter and wrapper. The description of fundamental work' principles of FS methods is depicted in Figure 3-19. The new subset of features is named $Selected\_features\_table(X, A)$, where X represents the number of observations (rows) and A represents the optimal subset of features (columns) from the original feature set $P`$.

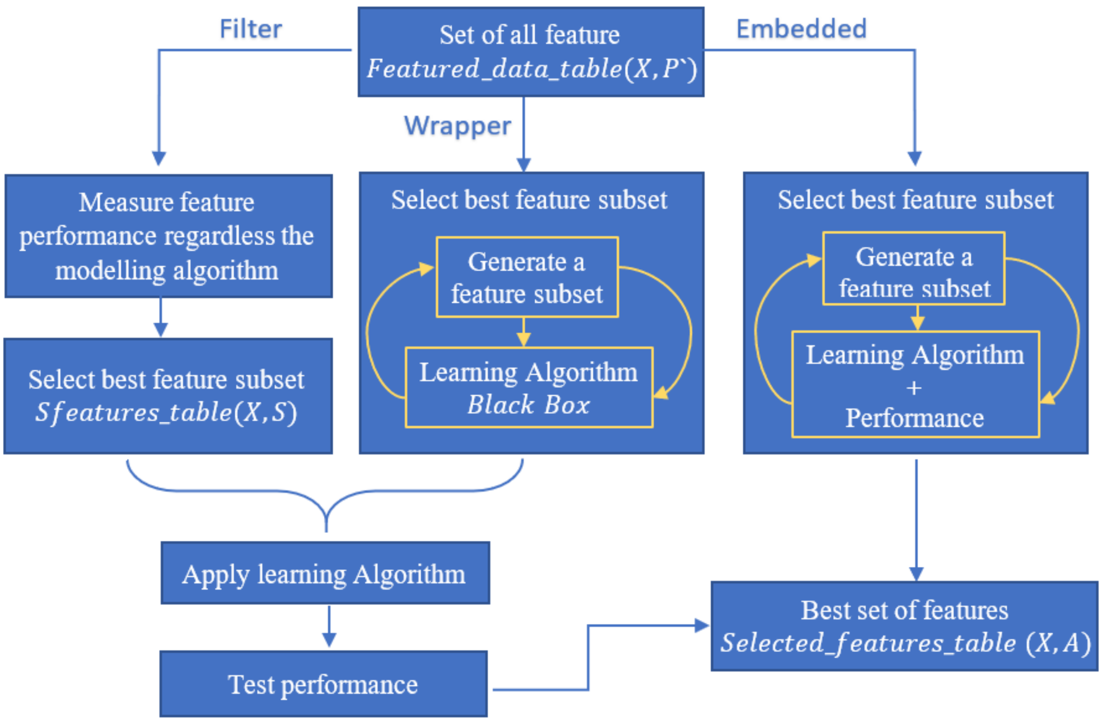


Figure 3-19 Feature selection approaches for a classification task.

The FS algorithm searches the whole space of features to only include the optimal features in the training set that will be used by the classifier in the next stage. Therefore, two basic components may need to be considered in the FS process; an *algorithm* to be proposed to select the best set of features, and an *evaluation function* to measure the integrity of the prior selected features (Kotsiantis *et al.*, 2006; Alexandropoulos *et al.*, 2019). In addition, the search for the best set of features by the selected algorithm would be stopped via a proper *stopping criterion*. The searching process could be implemented by either adding or deleting non-effective features or by meeting a chosen evaluation function (Kotsiantis *et al.*, 2006).

A review including feature selection methods with their application has been presented by Jović *et al.*, (2015), while a specific review on feature selection for classification task only has been explored in (Mwadulo, 2016; Tang *et al.*, 2014).

In the current research, three approaches of FS were applied and compared for the sheep dataset to select the optimal set of features suitable for the lameness classification task. The benefits and drawbacks of each approach are presented in Table 3-7.

Table 3-7 Feature Selection (FS) approaches applied to $Featured\_data\_table$ $(X, P`)$ sheep DataSet for SLDM with a brief description of their searching concepts, general benefits and drawbacks.

	Advantages	Disadvantages	Searching technique	Applied FS in SLDM
Filter	• Faster for searching an optimal subset of features • Independent of the learning algorithm	• General feature subsets obtained. • Lack of interaction with a learning algorithm • Lower classification performance • Does not evaluate feature's redundancy	Filter best features based on either their distance, information, correlation or consistency (regardless of the model used later	RelifF
Wrapper	• Higher classification performance • optimal feature subset obtained • Take into account features dependencies • Take into account the interaction between feature subsets and the classification model	• Slower to find the optimal features • Biased towards the learning algorithm used as an objective function • Computentially intensive • Chances of model overfitting	search the feature space either sequentially (forward, backwards) or apply heuristic search by evaluating a different subset of features to meet an objective function.	GA
Embedded	• lower computational cost than a wrapper • take into account the features dependences and interaction with the classification model • Search locally for the features that offer better classification • combine the comparable	Required algorithms have their own built-in feature selection methods to be applied	Features to be weighted to regularise learning model based on objective function to minimise the fitting error	RF

	efficiency of the filter and the accuracy of wrapper methods • perform FS during the learning time			

3.6.1 ReliefF

The basic idea of ReliefF is to estimate the features ($P`$) by weighting them according to their relevance to each other to distinguish among classes in a dataset. The pseudo-code of ReliefF algorithm is illustrated in Figure 3-20 (ROBNIK-ˇSIKONJA and KONONENKO, 2003). Firstly, in step1, prior weights $W(P`) = 0.0$ are given to the vector of features (attributes) in the dataset.

In step2, the algorithm iteratively selects a random instance R, and searches for its k-nearest neighbour instances (observations) in a given Dataset. The k-nearest neighbour instances are called $Hits$ if they belong to the same class of R, while the k-nearest neighbour instances of different classes to R are called $Misses$. The k-nearest neighbour instances to R are calculated according to Manhattan distance (sum of the absolute differences).

Finally in step3, the quality estimation of all predictors (features) $W(P`)$ are updated by decreasing the quality estimation of predictors that have different values to $Hits$, and increasing the estimation of predictors that have different values to $Misses$ (Kotsiantis *et al.*, 2006). The contribution to updated weights is kept between [0,1] intervals.

As a result, the first top predictors (features) which have the highest weight in a descending sorted vector $W(P`)$ are selected by retrieving their indices to be the best optimal set of features for sheep dataset.

Step1: Set the weight vector W to prior value of 0 for each
$W[1,2, \dots P`] = 0.0$ // vector equal to the no. of predictors

Step2: For each instance in $Featured\_data\_table(X, P`)$ of a given DataSet **Do**

Select a random instance $R = (X_r, P`)$
Find k-nearest Hits instances H_j to R $Hits = \frac{\sum_{j=1}^{k} |R - H_j|/k}{X}$
For each Class $C \neq R.Class$ **Do**
from Class C **Find** k-nearest Misses $M_j(C)$
$Misses = \frac{\sum_{j=1}^{k} |R - M_j(C)|/k}{X}$

Step3: Update W
For j= 1 to $P`$
$W(i,j) = W(i,j) - Hits + \sum_{C \neq R.Class} \left[\frac{P(C)}{1 - P(R.Class)} * Misses \right]$
End
End

Figure 3-20 The pseudo-code for ReliefF feature selection method.

3.6.2 Genetic Algorithm GA

GA is a heuristic optimisation search, based on the principle of 'survival of the fittest' of Darwin (Mwadulo, 2016). GA algorithm deals with a set of solutions called (chromosome or individuals), which represents a set of features to be optimised for the best features set. By mimicking natural evolution, the fitter chromosomes that have a higher probability are to be chosen for the next generation. An evaluation function is used to compute the fitness of each chromosome to be selected, while the selected chromosome (features) follows an application of genetic operators, such as crossover and mutation (Figure 3-21) to improve the selection for the fittest features (Il-Seok Oh *et al.*, 2004).

In recent time, GA has great attention in the field of feature selection because any formula for fitness estimation could be implemented (Too *et al.*, 2019). GA has been exploited for feature selection for various databases and has proven that it could uncover the hidden relationship between the features and Class, assist in the dimensionality reduction process, and improve the performance of the classifier (Smith and Bull, 2005; Babatunde *et al.*, 2014).

A developed GA that employs CHAID decision tree (**Ch**i-square **A**utomatic **I**nteraction

Detectors) as a fitness function for the feature selection of sheep dataset is applied in the current research, as the CHAID decision tree followed a non-parametric procedure where no prior assumption to the underlying data is needed (Miller *et al.*, 2014). The pseudo-code is presented in Figure 3-23 and an explanation for its implementation is explained in the following steps.

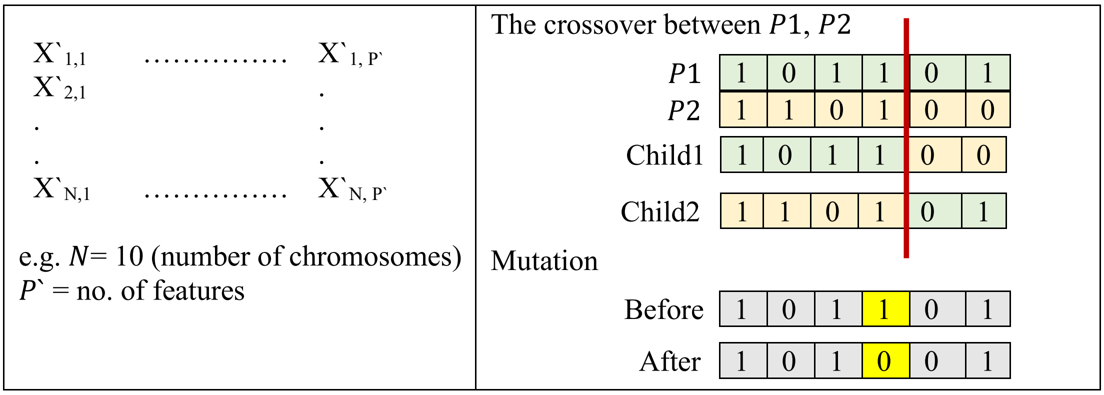


Figure 3-21 GA initialisation (left), GA operation (right).

Step1: Initial population

The feature selection process of a given sheep dataset $Featured_data_table(X, P`)$ that follows the GA algorithm includes: identify the number of selected chromosomes $N$, each of which of size $P$` denotes the number of features. The GA algorithms operate on binary search space as each randomly selected chromosome is a combination of bit strings called genes (features) where '1' denotes the selected features, and '0' refers to the features that are not selected for the evaluation process (fitness calculation). For example, each element in N looks like $X` = \{1,0,1,0,0,0,1,1,1\}$. $Z$ is a one-dimensional vector of size $P$` which initialises with 0; however, at the end of the process the best-selected features were set by '1'.

Step2: Fitness function

For each X`==1, a fitness estimation function is invoked to calculate the fitness of the features whose indices are only equal to '1'. So, not all features (183 features) are involved in the fitness calculation. The output of the fitness function calculation is a probability vector that corresponds to each feature its value is set to '1'.

Any fitness function could be used; however, the CHIAD decision tree is implemented in the current research due to its considerable estimation of the importance of the predictors. CHAID decision tree performs a curvature test analysis that applies a chi-square test Chi^2 between each predictor and its $Class$ vector (response vector) to measure the significance and assess the hypothesis that two variables are unassociated.

Chi^2 is used as a split criterion to construct a CHAID tree by summing the squares of differences between observed O and expected E frequencies of observations in respect to $Class$ vector (Sayad, 2011). Then, the best split predictor variable (best feature) is chosen as it minimizes the significant $p-values$ (< 0.05) of Chi^2 tests between each predictor (feature) and its corresponding $Class$ vector. The following step explains the procedure (MatLab documentation, 2019; Susanti *et al.*, 2017):

1. Since the Chi^2 test measures the difference between two categorical variables, the continuous features type are converted into categorical ones by partitioning it into its quartiles (levels) and a new nominal variable combines each original observation to its partition that occupies Figure 3-22.

j (columns: sound, mild, severe); i (rows: Level1–Level4)

(An examined feature)	sound	mild	severe	Sum
Level1 (1st quartile range)	...	…	…	
Level2 (2nd quartile range)	…	…	…	
Level3 (3rd quartile range)	…	…	…	
Level4 (4th quartile range)	…	…	…	
Sum				Total sum

Figure 3-22 Frequency table for one predictor and corresponding 3 categorical classes in $Class$ of sheep dataset.

2. For each level in the partitioned feature and each jth class in $Class$ (i.e. for each *cell* in Figure 3-22) apply the following:
   - Compute the expected frequency (Equ 3-13).
   - Compute Chi^2 test (Equ 3-12) to examine the significance of the association between each level in the partition feature and $Class$.

$$Chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\sqrt{\frac{(O_{ij}-E_{ij})^2}{E_{ij}}} \qquad \text{Equ 3-12}$$

$$E_{ij} = \frac{O_i * O_j}{O_t} \qquad \text{Equ 3-13}$$

O_i is the total sum of observation in the ith level for all classes, and O_j is the total sum of the observation of jth Class in an examined predictor, and O_i is the total sum. r represents the number of observations, while c denotes the number of classes. Df is the degree of freedom which is computed by multiplying $(the\ number\ of\ observation\ -1)$ by $(the\ number\ of\ classes\ -1)$.

3. Find $p - value$ (Equ 3-14) for each jth level in the partitioned feature. If it is less than 0.05, it means that there is a dependency between the tested variables; otherwise, there is no significant relationship.

$$p - value = \sqrt{\frac{Chi^2}{O_t * \sqrt{Df}}} \quad , \; Df = (r-1)(c-1) \qquad \text{Equ 3-14}$$

4. Select the jth level in the partitioned feature that produced the smallest $p - value$ (the lowest $p - value$, the most significant it is).
5. The best split predictor in each node is used to construct CHAID tree and is chosen according to the predictor that minimises the $p - value$ of Chi^2 between each predictor and the $Class$ (response variable).

Step3: Start generation

- Crossover and Mutation

From N, two parents $P1$ and $P2$ chromosomes according to the Roulette wheel probability selection have been chosen to apply a single-point crossover (Figure 3-21), the position of the crossover point is selected randomly. The two resulting children are merged into one chromosome called $Newp$. Then, a mutation process is performed on the $Newp$ (of double size of parent) where one gene is flipped from ‘1’ to ‘0’ or vice versa (Figure 3-21) for a random selection of genes with the total $Newp$ size in respect to the specified mutation rate.

- Merge population and Select the best chromosome Z

The $NewP$ is merged with X`, then the merged population is sorted and the best N chromosomes are selected for the next generation, while the rest of the chromosomes in merged population are discarded. The top first chromosome in the sorted merged population is selected (Z), and to be updated each iteration until the maximum number of generations T is met.

```
Step1: Initialise the parameter of GA randomly
        N = 10            //No. of chromosome in population
        T = 50            // Maximum number of generations
        CR = 0.7          // Crossover rate
        MR = 0.2          // Mutation rate
        Initialise the population X`(N,P`) with a random number either '0' or '1' of size P`
        Z(1,P`) = 0       // initialise Z with zeros
Step2: Compute Fitness function
    For i= 1 to N Do
        For j= 1 to P` Do
            Compute Fit (X`(i,j) == 1)          // CHAID decission tree
        End
    End
Step3: Start generation
Do until iteration <= T
    For k= 1 to round (CR*N)                    // no. of crossover
        InvFit = 1 − Fit                        // compute the inverse of fitness
        Fit_prop = InvFit(i)/ ∑_1^i InvFit      // compute the inversed fitness probability
        Select two parents P1,P2                // Roulette wheel selection depending on Fit_prop
        NewP = [P1,P2]                          // Apply single point crossover between P1,P2
    End
    For k=1 to size (Newp)
        Apply mutation to NewP                  // Randomly selected genes respect to MR
        Fit = fit(NewP)                         //Compute fitness for NewP
    End
    Mergred_pop = [X`,NewP]                     //Add Newp to current population
    Sort (Merged_pop) according to their highest fitness
    X` = Merged_pop                             // Update X` with the best N chromosomes
    Fit = fit (Merged_pop)                      // Update Fit with the best N values
    Z = Merged_pop(1,P`)                        // Update Z with the 1st best chromosome, where
                                                its gene equal to '1' means the feature is selected
Repeat
Step4: Return the best selected features set correspond to the genes equal to '1' in Z
```

Figure 3-23 The pseudo-code for CHAID Genetic Algorithm for feature selection of sheep dataset.

3.6.3 Random Forest RF for Feature Selection

Although to RF was introduced in Section 2.8.5.2, a brief summary is given here. RF is about bagging numerous decision trees which are called weak learners to obtain a global optimal classifier (learner) that overcomes the overfitting problem when the training accuracy of a model is higher than the accuracy of the same model when testing with unseen data. The vote for the final class is assigned by the majority of votes of all trees in the ensemble (Maxwell *et al.*, 2018). Each tree in the ensemble is trained with a random subset of features while one set is kept for testing the error rate of that tree called out-of-bag (oob) dataset. The overall accuracy of RF is estimated by averaging the oob error over the number of trees in the ensemble to provide an independent estimate for accuracy (Breiman, 2001).

The splitting criterion that is used in RF for feature selection is the curvature test (CHAID) which is introduced in the previous section. Since it is recommended to use the Chi test when there are many levels of unique values of the input feature set like continuous sheep datasets. Whereas CART tries all possible cut points (explained in Section 3.7.1), CHAID tries fewer cut points than CART as the continuous input feature is converted to categorical ones and CHAID test between categories for the best splitting point that minimises the $P - vaule$ of Chi test.

The importance of features is the oob error. The observations in the oob dataset are not used for constructing the tree; instead, they are employed as an internal validation set to estimate oob error. A flowchart for a RF working concept is presented by (Boulesteix *et al.*, 2012) Figure 3-24 shows the steps of how the RF is exploited for feature selection.

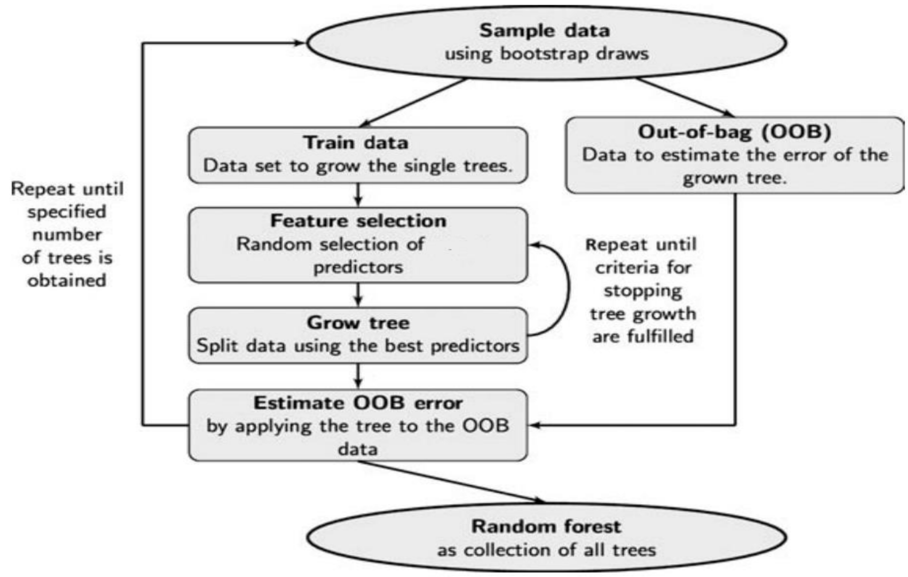


Figure 3-24 RF algorithm framework (Boulesteix *et al.*, 2012).

3.7 Construct A Decision Tree Classifier for Sheep Lameness Detection Model (SLDM)

After the optimal set of features has been reduced and the most important features have been chosen in the FS process, the $Selected\_features\_table(X, A)$ becomes ready to train a classifier to fit a model that has the ability to predict new unseen observations (examples) with a reasonable accuracy ratio.

Although there is no one model fits all types of data, the sheep DataSet2_a (Table 3-5) has been examined (as a raw data only) for more than one classifier in the early research output (Al-Rubaye *et al.*, 2018). The results reveal that the Decision Tree DT classifier outperforms their counterpart classifiers when they have been tested with the unseen dataset. Thus, the main classifier to develop a sheep lameness detection model to classify sheep walking into sound, mildly, and severely is the DT and its ensemble. However, other classifiers which were introduced in Chapter two would be used for comparing their performance with DT, while the basic concepts of how DT works is illustrated in the next section.

3.7.1 CART Decision Tree Characteristics

DT is a hierarchical structural form of a classification model that constructs a tree in a top-down greedy search approach (Reddy and Babu, 2018). DT recursively partitions the input dataset (training set) into a small subset according to their feature space in order to find the decision rules set for a robust predictive classification model (Myles *et al.*, 2004).

The tree structure mainly consists of two components: nodes and branches. The top node is named a root node (decision node), and the internal nodes are either parent or child. All these nodes have branches, while the bottom nodes are called leaf nodes and have no branches as they contain the classification result (the class of a classification problem). Each node, including the root node is selected according to the best attribute (predictor) in the training dataset that meets splitting criteria, while the branches connect the tree nodes and each path represents a decision rule that could be traversed from the root node through the internal node to a leave node as 'if-then' rules (Yan-yan Song and Ying Lu, 2015).

CART (Classification and Regression Trees) is a binary decision tree first introduced by Breiman *et al.*, (1984) that built a predictive model to detect either discrete or continuous targets, CART can deal with both categorical and/ continuous data types for predictors and target class. The obtained predictive model is constructed by partitioning the data set recursively into subsets and evaluating the information gain, before and after splitting to choose the best split that produces a tree with a minimum error rate. Figure 3-25 illustrates how CART is constructed, and the details of the procedures are explained in the following steps (Adnan, 2017; Tan et al., 2006).

Step1: Determining splitting points (cut points)

The $Selected\_features\_table(X, A)$ contains the best features data set which is employed by CART to construct the predictive model for sheep lameness detection (SLDM). So, the $Selected\_features\_table(X, A)$ is upgraded to be named as training dataset $D(X, A)$, where $X = \{x_1, x_2, \dots, x_i\}$ represents the number of observations (instances/ examples), and $A = \{a_1, a_2, \dots, a_j\}$ represents the number of best-selected features (attributes/ predictors) while the last column refers to the $Class$ of each observation (see Figure 3-26). To determine the possible cut points for each attribute (predictor) vector A in D, A is sorted according to its domain value (unique values) in ascending order. Each element of the sorted A between the lower L and upper U boundaries is tested to be chosen as a cut point candidate.

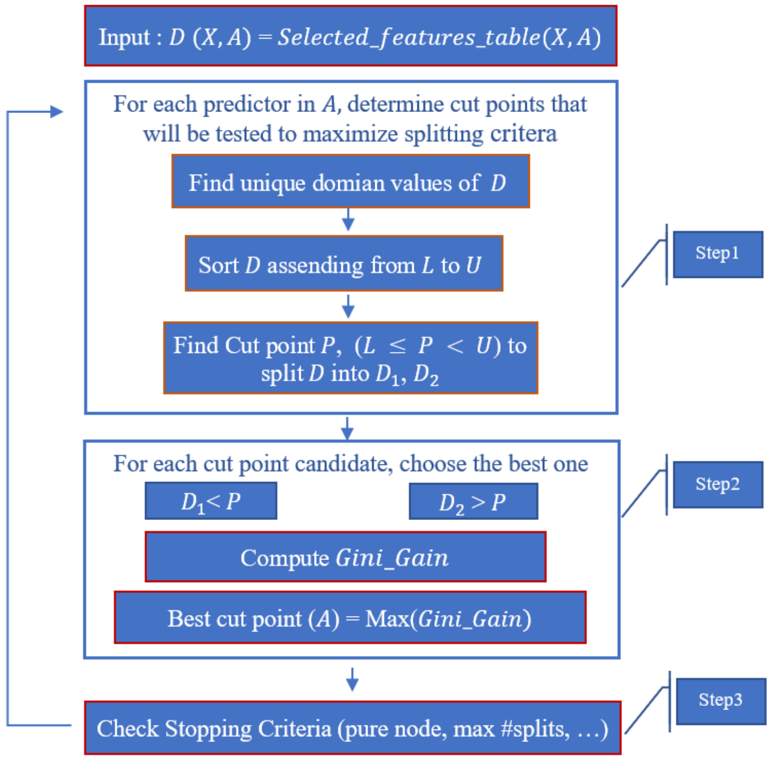


Figure 3-25 CART induction flowchart.

55x10 table

	1 Mean_Acc_x	2 Mean_Acc_y	3 Mean_Acc_z	4 Mean_Pitch	5 Mean_Roll	6 Mean_Azimuth	7 Mean_Gyr_x	8 Mean_Gyr_y	9 Mean_Gyr_z	10 Class
1	0.8515	-0.1054	-1.0412	73.5362	17.2037	202.2657	-0.3298	-0.0424	-0.2072	'mild walking'
2	0.1987	0.1927	0.2539	81.1390	16.8261	128.2018	-0.2795	-0.0268	-0.0198	'mild walking'
3	1.9692	-3.5894	1.6525	68.9420	-44.4189	160.5339	-0.0776	-0.2526	-0.1920	'mild walking'
4	0.1484	-0.0643	0.0360	79.3033	14.2869	231.9066	-0.2895	-0.1107	-0.0987	'mild walking'
5	0.2531	0.2538	0.4113	81.4955	38.9736	257.6466	-0.3551	0.0157	-0.0447	'mild walking'
6	0.1281	-0.0021	0.3213	80.5401	10.6318	189.0921	0.2543	0.0623	0.1444	'mild walking'
7	0.1247	-0.1386	0.5399	83.2682	6.3668	30.6802	-0.0605	-0.0272	-2.9932e-04	'mild walking'
8	0.6458	0.4261	-0.2382	77.0567	13.5367	320.2204	0.2616	-0.0532	0.0570	'mild walking'
9	0.2577	-0.8223	-0.8136	-24.7367	30.8646	276.3762	-0.0729	-0.1445	-0.0442	'severe walking'
10	0.0939	-0.6641	-0.5798	-34.8380	-11.3973	209.2013	-0.1567	0.0203	0.3398	'severe walking'
11	0.3531	-0.5392	-0.2931	-35.0819	-22.2144	91.5866	-0.1145	0.2534	0.0228	'severe walking'
12	0.6687	-1.3830	-0.5966	-32.9846	-7.5708	110.9711	0.0870	0.3033	-0.1219	'severe walking'
13	0.0676	0.3029	-0.7716	-30.3371	-5.4956	68.3981	-0.1238	0.0292	0.0583	'severe walking'
14	-0.0494	-0.1979	-0.0774	-30.5979	-22.0092	296.5524	4.8869e-04	0.1254	0.1935	'severe walking'
15	0.0834	-0.8207	-0.6430	-38.5334	-26.1605	227.4742	0.0553	-0.1906	0.1258	'severe walking'
16	0.8251	-1.4426	-0.4978	-41.0069	-36.0135	111.1087	0.0649	-0.1449	-0.2982	'severe walking'
17	-0.4075	-1.2727	0.5228	-21.3209	-52.4015	101.4113	0.0466	0.0242	-0.1855	'severe walking'
18	0.3148	-1.1270	-1.1100	-25.8440	-37.5833	146.0623	-0.0229	-0.3961	-0.1070	'severe walking'
19	0.2337	-0.1731	-0.7958	-12.4782	-23.8748	142.9905	0.0638	-0.1673	-0.3131	'severe walking'
20	0.3954	-0.4168	-1.7284	-21.2863	-8.1590	231.8840	-0.1838	0.5012	0.0293	'severe walking'
21	0.3565	-0.6496	-0.9314	-22.9028	-19.6995	185.3204	-0.0208	-0.1330	0.0415	'severe walking'

Figure 3-26 An example of sheep training set dataset, where X= 55 # observation, A= 9 # predictors, K= 3 # Classes.

Step2: Choosing the best split cut point

Each candidate cut point splits the dataset into two nodes $D1$ and $D2$ as CART performs binary splitting. For each node, a measurement of impurity is performed (the node is said to be pure if it has observations from the same class). The Gini index ($Gini$) in (Equ 3-15) is a metric to compute the node impurity in CART by summing the squared probabilities of each class in the examined node; where C is the number of classes ($K = 3$ sound, mildly, and severely walking), P_k is the observed fraction of class K over the number of observation for all classes in an examined node. $Gini$ value is between 0 and 1, when $Gini = 0$ that means a pure node contains observations from only one class, in this case it represents a leaf node and no further splitting is required; otherwise, the node is impure and the value of $Gini$ measures the degree of node impurity.

$$Gini\ (D) = 1 - \sum_{K=1}^{C} (P_k)^2 \qquad \text{Equ 3-15}$$

After splitting D into $D1$ and $D2$ and $Gini$ is computed for each new partitioned dataset, the gain in $Gini$ (impurity) is computed between the parent node (node before splitting) and child nodes (nodes after splitting) to find out the best split (cut point). The best split is the one that maximises the impurity gain (ΔI) overall splitting candidates The difference in $Gini$ gain is calculated in Equ 3-16 and the largest difference indicates the better test condition as the best split is the one that maximises the $Gini\_Gain$. In Equ 3-16, $Gini\ (D)$ represents the parent node's $Gini$ before splitting, while $Gini\ (A, D_i)$ represents the child node's $Gini$ after splitting. $N(A)$ is the number of observations in a child node, N is the total number of observations in the parent node. D is the number of partitions which are equal to two as CART is a binary approach that divides each node into two partitions $D1$ and $D2$ as left and right child respectively.

$$Gini\_Gain\ (A, D) = Gini\ (D) - \left(\sum_{i=1}^{D} \frac{N(A)}{N} * \ Gini\ (A, D_i) \right) \qquad \text{Equ 3-16}$$

Step3: Stopping rules

The prementioned process is recursively performed until a stopping criterion is met:

- When a node is pure (the observation of one class is the only observations that exist in that node).

- When the number of observations in a node is less than the minimum parent size that is predefined by the user (default is 10).
- When the number of observations in a node is less than the minimum leaf size that is predefined by the user (default is 1).
- The number of splits reaches the maximum number of splits (default is # observation - 1).

Example 3.3: Consider the dataset in Figure 3-26, where the total no. of observations $X = 55$, no. of classes $K = 3$, and no. of attributes $A = 9$. Then, CART tree is constucted in Figure 3-27, and the calculation for $Gini$ of each node is provided in Table 3-8.

Table 3-8 Gini index of each node of CART presented in Figure 3-27.

# Node	sound	mild	severe	# Observation	$Gini$ # = Gini index * Node Probability
1	18	8	29	55	0.5937
2	2	0	29	31	0.0680
3	16	8	0	24	0.1939
4	1	0	29	30	0.035
5	1	0	0	1	0
6	16	0	0	16	0
7	0	8	0	8	0

$Gini\ \#1 = (1 - ((18/55).\text{^}2 + (8/55).\text{^}2 + (29/55).\text{^}2)) * (55/55)$

$Gini\ \#2 = (1 - ((2/31).\text{^}2 + (0/31).\text{^}2 + (29/31).\text{^}2)) * (31/55)$

$Gini\ \#3 = (1 - ((16/24).\text{^}2 + (8/24).\text{^}2 + (0/24).\text{^}2)) * (24/55)$

… … … …

$Gini\ \#7 = (1 - ((0/8).\text{^}2 + (8/8).\text{^}2 + (0/8).\text{^}2)) * (8/55)$

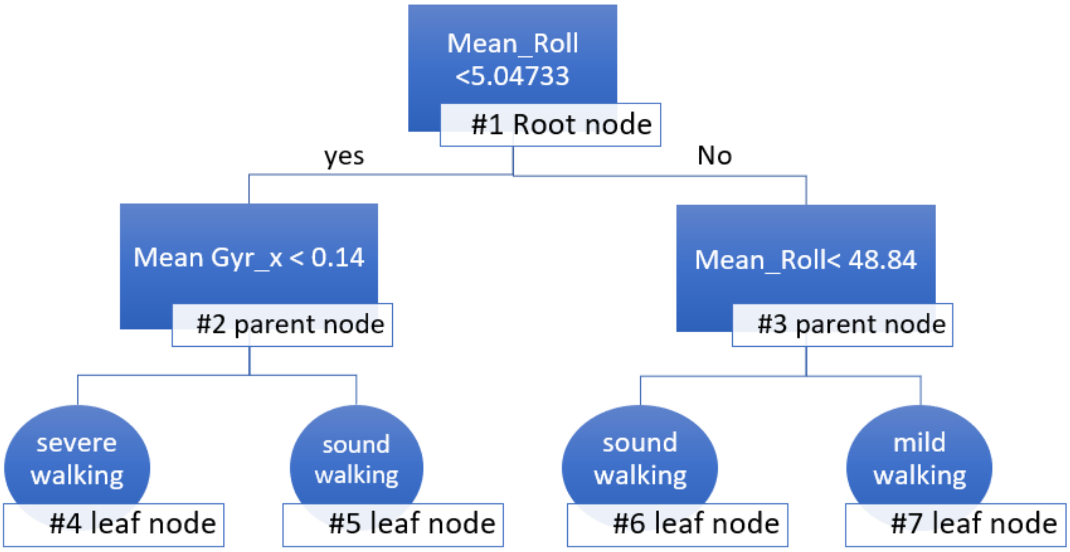


Figure 3-27 CART construction example for a dataset given in Figure 3-26.

3.7.2 The Ensemble of Trees (Bagging, Boosting, and RusBoosting)

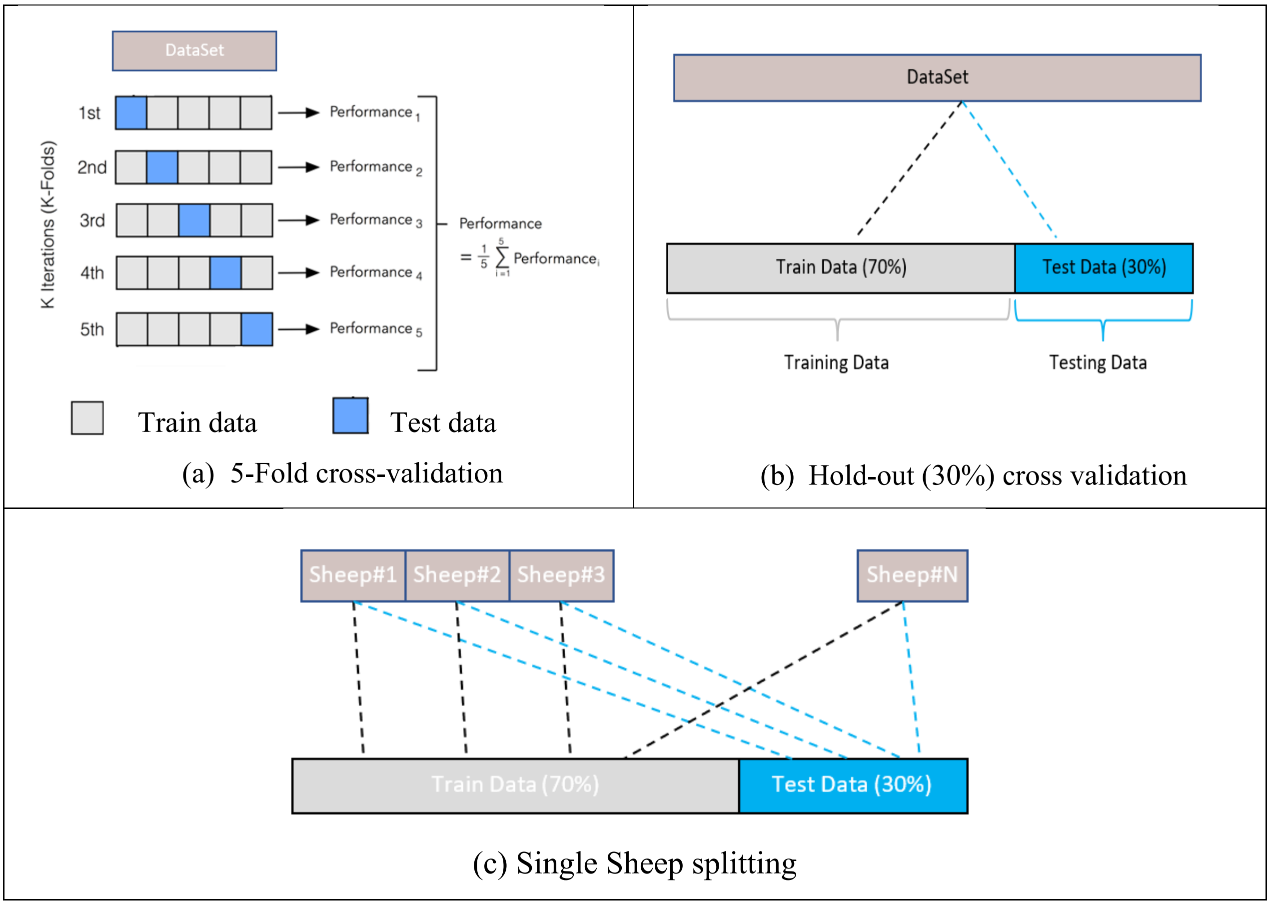
The aggregation of many trees within ensemble techniques increases the level of predictive accuracy of the model; however, the interpretation of the model could be negatively affected (Myles *et al.*, 2004). The number of trees in an ensemble would not cause an overfitting case; in contrast, a sufficient number of trees in an ensemble are developed to reach a settled level of error. So, 100 decision trees are recommended to train an ensemble classifier to reach a satisfying level of performance (James *et al.*, 2013).

Basically, two main techniques are utilised to build an ensemble classifier: Bagging and Boosting. The difference between the two growing techniques would be in the way of growing the trees in the ensemble. In bagging, all trees are constructed once, while in boosting, the growth of the trees happens gradually to increase the model efficiency as the model with smaller trees number would expect to have less execution time. Random Under Sampling Boosting method (RusBoosting) is also used when the dataset has an imbalanced number of classes as it resamples the distribution of classes within the dataset. It constructs the ensemble the same way AdaBoost performs. However, an introduction part for each type was given in 2.8.5. Basically, the ensemble is tested with the sheep dataset file for comparing the performance of a single CART with multiple CARTs in the ensemble.

3.8 SLDM Validation

The performance of the classifier could be evaluated by estimating the number of misclassification records (examples) committed by the classifier on training data; this type of error is called training error or resubstituting error. However, the estimation of the resubstituting error could be optimistic and cause what is called overfitting, when the model fits well with the training example but not with new unseen examples. Thus, the estimation of generalisation error would be rational as it measures the misclassification error on unseen data and has not been employed in building the model (Tan *et al.*, 2005).

To provide an unbiased estimation of the generalisation error, the unseen data that is called test data is tested for the purpose of model validation. In SLDM, the three common methods for evaluating model performance are used (Raschka, 2018); however, one more method is proposed to be applied in the current research that is named Single Sheep Splitting. All validation methods are illustrated in Figure 3-28.



(a) 5-Fold cross-validation

(b) Hold-out (30%) cross validation

(c) Single Sheep splitting

Figure 3-28 Three methods to evaluate the performance of SLDM.

3.8.1 Evaluation Metrics

3.8.1.1 Confusion Matrix

A confusion matrix is a metric that measures the performance of a classifier learner on a set of known class data (Kevin Markham, 2014). It formulates a square matrix with a number of rows and columns equal to the number of classes in a classification problem. For the lameness detection problem, three classes are spotted; sound walking, mildly walking, and severe walking Figure 3-29. The rows represent the actual instances belonging to each $Class$, while the columns refer to predicted instances of these $Class$. The diagonal line represents the overall of True Positive predictions (TP) and True Negative predictions (TN) which mean that the actual classes match the predicted classes. Otherwise, the area above and under the diagonal is called False Negative (FN) and False Positive (FP) (Tan et al., 2006).

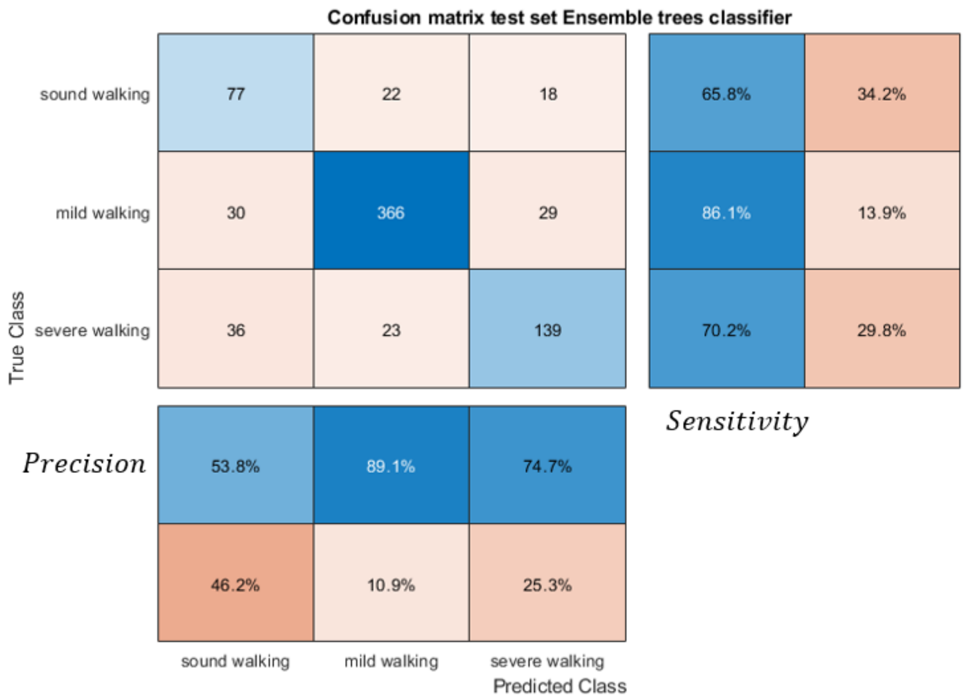


Figure 3-29 Confusion Matrix example for SLDM.

To estimate the accuracy of the classifier's performance and the misclassification error Equ 3-17 and Equ 3-18 could be applied.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad \text{Equ 3-17}$$

$$Missclassification\ rate = (FP + FN)/(TP + TN + FP + FN) \quad \text{Equ 3-18}$$

Where:

TP: the number of examples where the $Class$ of interest is correctly classified by SLDM as it is observed.

FN: refers to the number of examples where the $Class$ of interest is incorrectly classified by SLDM as another $Class$ (ex: mildly walking is misclassified as severely walking by SLDM)

FP: the number of instances (examples) where the $Class$ of interest was incorrectly classified while it is not observed.

TN: the number of instances where the $Class$ of interest was correctly classified as not being observed.

3.8.1.2 Imbalance Dataset Metrics

Although the confusion matrix estimates the accuracy of the model, the imbalance dataset; where the number of classes is unequally distributed, it needs other metrics to evaluate the performance of the classifier. These metrics are calculated from the counts in the confusion matrix such as.

- $Sensitivity$ (Recall) or TPR true positive rate for a positive $Class$ is: the number of correctly classified positive instances divided by the number of positive examples in the data Equ 3-19.
- $Precision$ for a positive $Class$ is: the number of correctly classified positive instances divided by the number of examples predicted by SLDM as positive Equ 3-20.
- $F-score$: is the harmonic mean of $Recall$ and $Precision$ Equ 3-21, the higher F-score the model has, the better the performance is (Tan *et al.*, 2005).

$$Recall = TP/(TP + FN) \quad \text{Equ 3-19}$$

$$Precision = TP/(TP + FP) \quad \text{Equ 3-20}$$

$$F - score = 2 * Recall * Precision/(Recall + Precision) \quad \text{Equ 3-21}$$

3.9 Chapter Summary

Android-powered system sensors were used to collect sheep data from Lodge Farm at Northampton, UK. Three datasets were obtained at 5 Hz, 10Hz, and 4Hz sampling rates as raw data ready to be pre-processed, and then were fed to the developed classifier SLDM. The pre-processing stage includes manipulating noise and missing values, segmentation (FNSW, FOSW) for three various window times 5s, 7s, and 10s as these periods of time are recommended by the researches who conducted sheep behaviour classification and these time slices are adequate to examine the walking sheep for lameness symptoms. Furthermore, data pre-processing involves extracting walking segments and computing features for the walking segment only. 183 features were extracted from the accelerometer, gyroscope, and orientation data; however, for the purpose of eliminating the effect of irrelevant features, three types of the feature selection process are applied: ReliefF, a proposed GA with CHAID fitness function, and RF. Then, the best-selected features would feed the CART decision tree to build a model that classifies sheep lameness status into sound walking, mildly walking, and severe walking. An ensemble of CARTs was applied to overcome overfitting and increase the SLDM accuracy in two ways of the ensemble: bagging and boosting; in addition to the RusBoosting for imbalanced dataset. Finally, the built model was validated with unseen data by using 3 methods of validation: 5-folds cross-validation, hold-out 0.3 of data for testing, and Single Sheep Splitting proposed method. A multi-class confusion matrix was used as a metric to explore the model performance in terms of accuracy, sensitivity, precision, and F-score.

4 Chapter Four: SLDM Implementation, Classification Results and Interpretations

4.1 Introduction

The data mining methodology for sheep lameness detection approach (SLDM) that has been developed for the current research is empirically applied using MatLab programming language as it is a robust tool for data analysis, algorithm development, visualisation and graphics, and numeric computations. MatLab offers a powerful machine learning toolbox, which provides interactive visual environments for investigating data analysis algorithms, evaluating them, and choosing the best algorithm that suits the demand application.

A user-interface for SLDM is designed especially for the current research that allows the user to acquire sensor data, pre-process, and implement the classification algorithm with the option of a user-defined parameter that could be changed during implementation.

The first section of this chapter, 4.2, explores the implementation of SLDM via App Designer in MatLab. Then, it is followed by the results of each task in App Designer with its evaluation. Lastly, a final table for aggregated data is presented in Section 4.3, while the plots for raw data are provided in Section 4.4. Data pre-processing results are presented in 4.5, and a comparison for the best features selection is provided in 4.6. Interesting SLDM train and test results are explored and discussed in Section 4.7. Finally, a comparison of the model's validation methods is given in Section 4.7.3, and the Chapter is closed with a general discussion and comparison with other related sheep lameness prediction studies in Section 4.8.

4.2 User-interface Design for SLDM

The user-interface for SLDM is designed using the rich environments of App Designer in MatLab which provides a Graphical User interface (GUI) that contains visual components to create a design layout view in addition to a code view that is integrated with MatLab editor. App Designer is the programmers' target to build a standalone application that could be executed in any desktop or even web applications where is no need for MatLab's compiler to be installed in a machine; instead, only the executed file is adequate to run the application.

For the purpose of implementing SLDM, a GUI is designed in MatLab to organise the process of building the SLDM in terms of controlling the user-defined parameters, exploring the visual plotting results, and re-executing the implementation with different option parameters. The SLDM interface consists of four main tabs, where each tab performs a different functionality.

4.2.1 Tab1: Get Sensor Data

Three tasks could be conducted in the first tab of SLDM (Figure 4-1). The first task is the sensor ***data acquisition*** from the three sensor application's types that were used in the current research named Sheep Tracker, SensoDuino, and Sensor Log via clicking on the '*Get sensor data*' button. The user can identify the '*discard reading time in sec.*' before the acquisition, which was defined here in 3 seconds. The discard time is the time that is wasted during the deployment of the sensor onto the sheep's neck and taking it off.

The second task is ***data aggregation***. From the whole sheep datasets that were collected (Table 3-5), the aggregation step is required to satisfy that the dataset involves various sheep status ($Class$), and keeps as much as sensor data-points collected. Because the data collected from each sheep has only one class from the group of the classes that are dealt with in this research (severe, mild, sound), the data aggregation is recommended to satisfy the variety of classes to be included in the dataset that is being used to build SLDM. The aggregation process produced four DataSets, which are expressed in Table 4-1.

After the four aggregated DataSets are obtained, the ***summary*** for each DataSet is computed, where the proportion of each $Class$ in a given dataset was calculated from the whole data-points aggregated (Table 4-1). The imbalanced dataset was obtained as the ratio of each $Class$ in the four aggregated datasets are unequal.

The final task to be performed in Tab1 is ***DataSet plotting***. This is required to show how each DataSet is presented. The relationship between the predictors is shown in two forms of plotting: *boxplots* and *matrix of scatter plots*. In *boxplots*, each predictor is grouped in a separate box according to the $Class$ that it belongs to, while the *matrix scatters* plots a matrix of scatter plots grouped every two predictors by their $Class$.

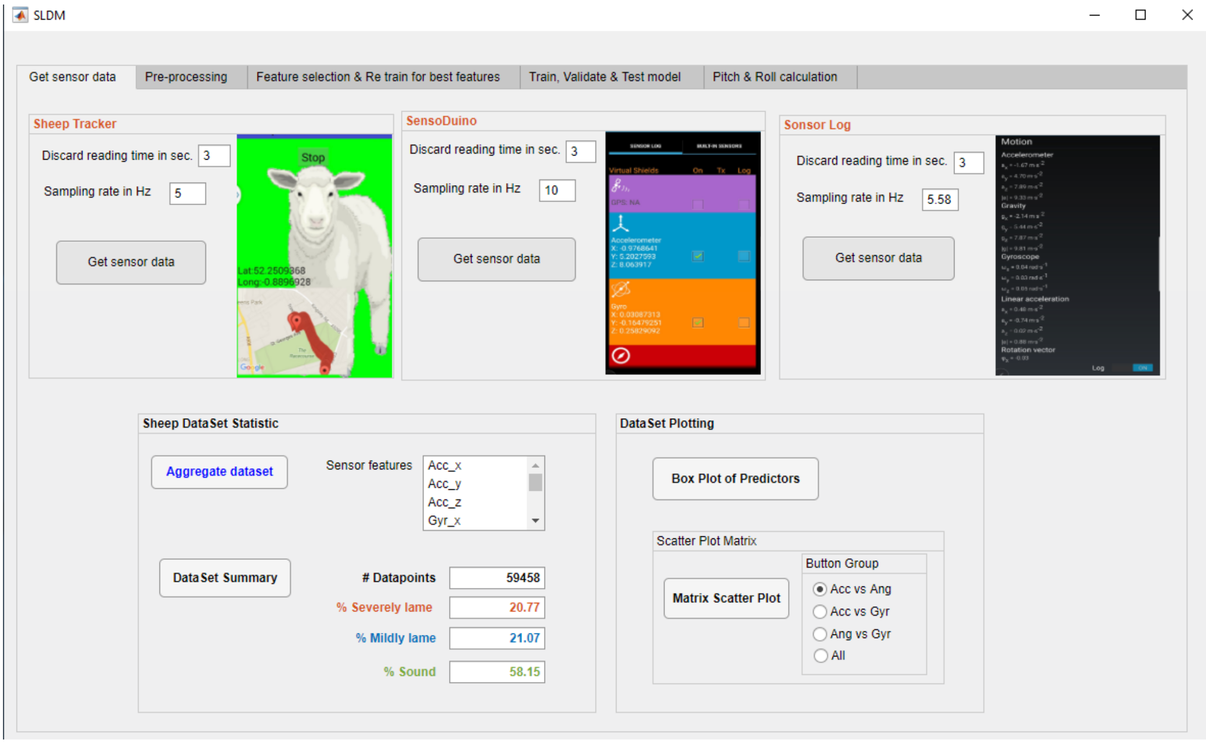


Figure 4-1 The first tab of SLDM.

4.2.2 Tab2: Pre-Processing

Tab2 of the SLDM implements the data pre-processing steps, as shown in Figure 4-2, the pre-processing steps include:

1- ***Segment*** the raw sheep datasets according to the two methods of online and offline segmentation as presented in the methodology Chapter named FNSW and FOSW Section 3.5.3.1. The percentage of '*overlapping*' was set to be 20% overlapping between two consecutive windows in FOSW method. The two segmentation approaches are applied over a chosen window size of 10 sec, 7 sec, and 5 sec. The number of segments (#seg) resulting from each chosen window size for the four aggregated sheep Datasets are presented in a column within the classification tables Table 4-2, Table 4-3, and Table 4-4.

2- ***Classify*** the obtained segments of moving sheep into Standing, Walking and Trotting segments according to pre-defined speed threshold limits that were set to be between 0.8 m/s – 3.5 m/s. The '*running mean window*' that was used to calculate the $Dynamic\ Acc\_y$ to be integrated later on in the speed calculation was set to a number that is equal to 5, 7, or 10 associated with the selected '*window size*' either 5, 7, or 10

$sec.window$. The classification results are presented in4.5.1.1 for each 10, 7, and 5 $sec.window$, respectively. For the visualisation, scatter plots for the segmentation process are depicted in Figure 4-7 for DataSet2_ac, while the plots for DataSet1_all, DataSet2_b, and DataSet3_all are shown in Appendix C. 1, Appendix C. 2, and Appendix C. 3, respectively.

3- ***Combine*** the classified walking segments from each individual sheep in a given dataset into one file called $Combine\_data$, then ***Get*** the raw walking sheep dataset table in one file named Raw_data_table to be prepared for the feature extraction step (refer to Section 3.63.5.7). The summary of each Raw_data_table; which includes the number of segments (instances) belonging to each $Class$ (severely walking, mildly walking, and sound walking) as well as the proportion of each in the four obtained sheep Datasets in (Table 4-1), is presented in Table 4-5, Table 4-6, Table 4-7 for 10, 7, and 5 $sec.window$, respectively.

4- ***Extract features*** that were listed in (Table 3-6) from the four obtained DataSets. The total number of features for the DataSet1_all, DataSet2_b, and DataSet3_all is 183 features, while only 122 features were extracted from DataSet2_ac as it has 6 predictors (no gyroscope readings were included). The type of extracted features is also listed in Table 4-8, which is either time domain or frequency domain as it was mentioned in Figure 3. 17.

5- The feature extraction step is accompanied with ***Computing*** the execution time required for each feature in seconds, and the results are listed in Figure 4-8 for DataSet2_ac, while the results for DataSet1_all, DataSet2_b, and DataSet3_all are presented in Appendix D. 1, Appendix D. 2, and Appendix D. 3, respectively.

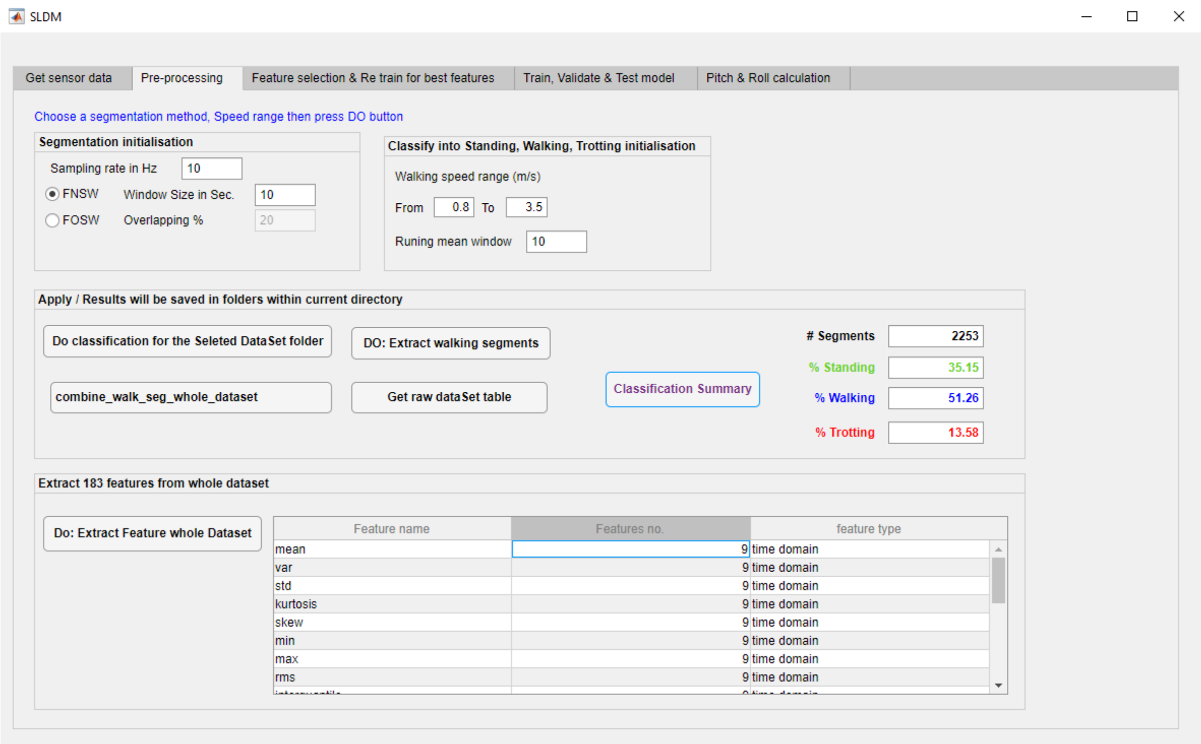


Figure 4-2 The second tab of SLDM.

4.2.3 Tab3: Feature selection & Retrain for Best Features

Two main tasks are conducted in Tab3 of the SLDM Figure 4-3. The first task performs the steps for '***Feature Selection'***, and the second task applies the steps for '***Train for the best no. of features'*** to find out the best number of features that reveal the highest accuracy of lameness prediction.

In ***Feature Selection*** FS task, three approaches RelifF, GA, and RF are implemented (refer to Section 3.6) to figure out the most FS method that reveals the highest prediction percentage of sheep lameness status for the four obtained sheep DataSets. The first tried FS method is Relief F, which accepts user input for k nearest neighbour instances in '*No. nearest neighbour*' field in SLDM. k is identified by 10 in the current execution of SLDM. Alternatively, when GA is applied for FS, the number of best-ranked features is displayed in '*Selected features NO. by GA*' field. Whereas in the third FS method; which is RF, the number of trees to be trained are determined by 100 trees in the current implementation while the accuracy of RF classifier could be retrieved in '*RF Accuracy*' field of SLDM. For the three applied aforementioned FS methods, the execution time is calculated and shown in the '*Execution time in sec.*' field; however, the comparison results over $10, 7, 5\ sec.window$ for both FNSW and FOSW segmentation methods are provided in Table 4-9, Table 4-10, and Table 4-11.

After each implementation of the three methods, the rank of feature's importance is displayed in the '*Features' Rank*' Listbox in Tab3 of SLDM. Due to the number of features being 183 for DataSet1_all, DataSet2_b, and DataSet3_all, whereas 122 features for DataSet2_ac., the results of ranked features from each group of DataSets (4) for FNSW and FOSW (2) and over 10, 7, 5 $sec.window$ (3) produced long tables of ranked features which are given in Appendix E. So, In Appendix E, each of the 4 sheep Datasets has 3 related tables each of which reveals the ranked feature results over $10, 7, 5$ $sec.window$. Each table has 6 columns of 183 or 122 features (rows) that are all conveyed in Appendix E; however, a photo for the first 25 ranked features are presented in Section 4.6.2.

Next, the '***Train for the best no. of features'*** task is performed which is the process to decide how many features could be selected from the ranked list of features that mostly minimise the classification error. Therefore, the performance of a single CART for the ranked features is tested and validated with 5-fold validation method. The validated results are plotted in Appendix F and discussed in Section 4.6.3 for each 4 sheep DataSet for both (FNSW and FOSW) over $10, 7, 5$ $sec.window$. The plottings also show the highest accuracy obtained with its associated number of features.

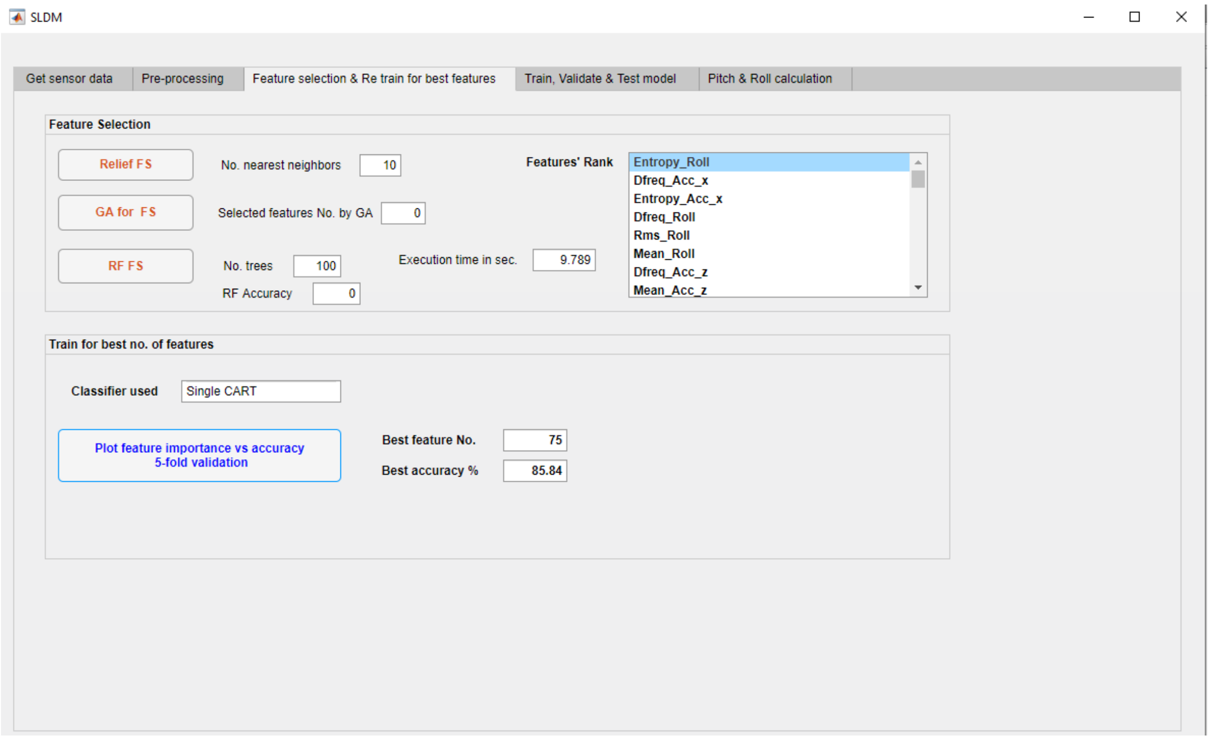


Figure 4-3 The third tab of SLDM.

4.2.4 Tab4: Train, Validate & Test Model

The final Tab in the SLDM interface design is Tab4 Figure 4-4, where two main tasks are implemented. The first task is training the lameness prediction model and '***Apply Ensemble***' of CART either using the Bagging method when '*Bag*' is chosen or using Boosting method when '*AdaBoostM2*' is selected by the user. Another option is available for imbalanced Datasets, which is '*RusBoost*' which stands for Random Under-Sampling Boosting. The '*RusBoost*' effectively classifies the imbalanced dataset as it is under-sampling the majority of the class uniformly and randomly; this method might produce a better classification rate compared to '*AdaBoostM2'* method.

For each tree in the ensemble, the '*MaxNumSplits*' represents the number of maximum splits each CART could perform, which equals to the $(number\ of\ instances - 1)$ by default; however, the user can choose any number of splits to control the depth of each trained tree. In the current research, the number of maximum splits is set to its default. In addition, the number of predictors (features) at each split '*No. of predictors at each splits*' is set to "all" in order to include all predictors for each CART's execution. Finally, the number of features that would be considered in execution is set to the first 20 features from the ranked features tables obtained from the previous Tab3.

The second task of Tab4 is validating the trained model '***Model Validation***' using two common methods either 'K-fold' or 'Holdout'; in addition to, a proposed method that is called 'Single Sheep Splitting' (refer to 3.8). In the current execution, the number of folds needed to validate the trained model is set to be '5', while the percentage of hold-out data is 30%; so, the trained data is 70%, and the built model is tested for the rest of 30% of data that not seen by the model. The 'Single Sheep Splitting' validation proposed method is applied when all features are chosen instead of the first 20 ranked features.

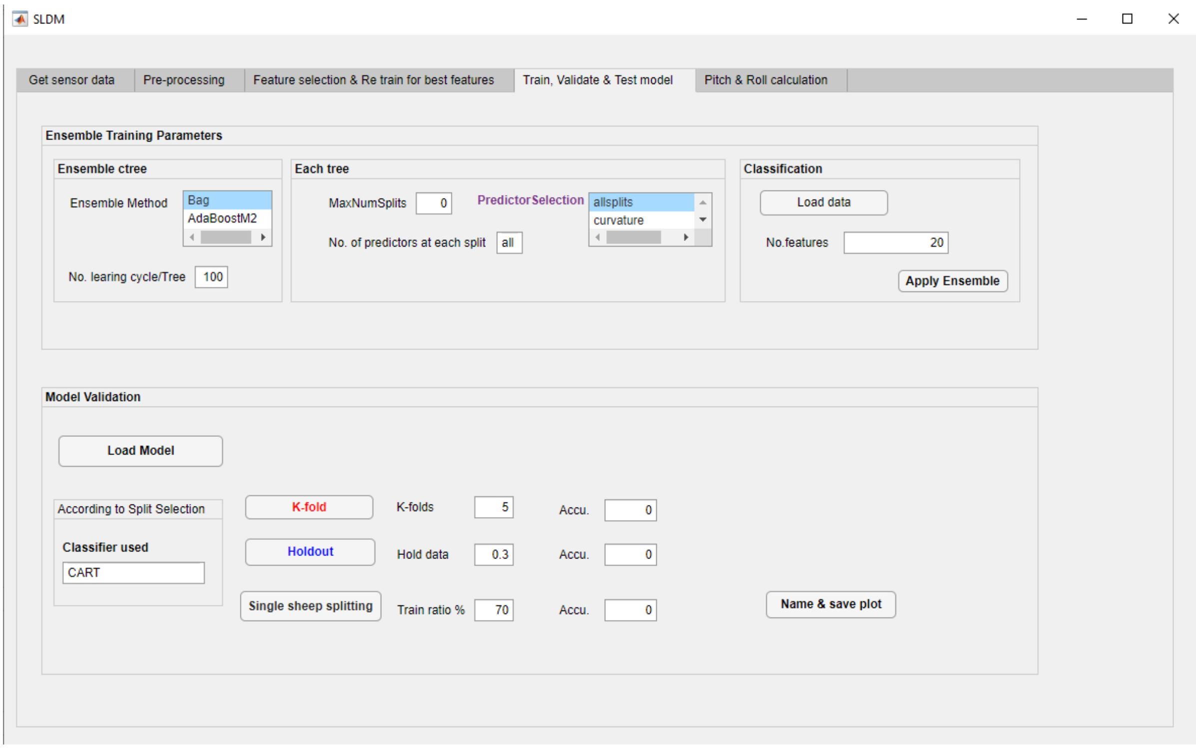


Figure 4-4 The fourth tab of SLDM.

4.3 Data Acquisition and Aggregation Results

The final aggregated sheep DataSets were set to be four sheep DataSets that are expressed in Table 4-1. The DataSet1_all, DataSet2_b, DataSet3_all have 3 acceleration readings, 3 gyroscope readings, and 3 orientation readings in 3 dimensions horizontal, vertical, and orthogonal except that DataSet2_ac has only 3 acceleration readings and 3 orientation readings.

The aggregated datasets show imbalanced datasets as each class's proportion of severe, mild, and sound are not equally distributed. Usually, the imbalanced datasets are obtained from a real-world application like the one tackled in the current research when many obstacles could be faced in collecting an equal number of samples of each class from the sheep flock.

The imbalanced datasets that are collected from a real-world application could affect the classifier performance negatively, where the prediction accuracy of the minority class would underestimate the prediction accuracy of the majority class. This could happen because most of the data mining algorithms assume an equal distribution of all classes in the trained dataset, and the error from each class has the same cost (Ganganwar, 2012).

The process towards building SLDM was implemented for each 4 obtained sheep DataSet as each of them has different characteristics that prevent combining all the collected sheep data into one data file.

Table 4-1 Final aggregated sheep Datasets.

Dataset Name	# predictors	Sensor App used	Sample rate	# Datapoints (# sheep)	Severe ratio%	Mild ratio%	Sound ratio%	Notes
DataSet1_all	3 Acc, 3 Gyr, 3 Orient	Sheep Tracker+ Sensor Log	5 Hz	64384 (31 sheep)	22.34	55.69	21.97	Missing Gyr from Sensor Log are manipulated, then all aggregated
DataSet2_ac	3 Acc, 3 Orient	SensoDuino	10 Hz	124806 (10 sheep)	46.56	30.12	23.32	Gyr discard for DataSet2_a and DataSet2_c, Orient are calculated for DataSet2_c, then all aggregated
DataSet2_b	3 Acc, 3 Gyr, 3 Orient	SensoDuino	10 Hz	59458 (18 sheep)	20.77	21.07	58.15	Same DataSet
DataSet3_all	3 Acc, 3 Gyr, 3 Orient	SensoDuino	4 Hz	5342 (7 sheep)	22.97	23.19	53.84	Redundant Gyr values are manipulated first

4.4 Plotting Raw Data

The plot of the four aggregated raw sheep datasets was depicted in forms of *Matrix of scatter plots* and *Boxplots* in the following Sections (4.4.1 and 4.4.2), respectively. It is shown from the plotting that each group of predictors is strongly correlated, and it would be challenging to have a clear linear separation of predictors in order to indicate the lameness status of sheep. This might be due to the lame sheep having the ability to pretend to walk normally in order to hide their pain in case of uncommon situations that might face the flock; for example, when the observation is in progress, and the flock is being monitored by an observer like the case in the current research, or when the shepherd's dog is used on the farm to gather the flock. Also, the sheep could challenge themselves to walk normally when the flock is being attacked in open fields by wild animals like a wolf or fox.

Due to the decision of identifying the sheep's status directly from the raw data not being a straightforward task, the raw data is usually pre-processed to be trained by the classifier. However, the orientation group (Azimuth, Pitch, and Roll) of the raw sheep data could positively contribute as a good indicator for lameness detection. The hypothesis of the orientation group mostly contributing to lameness detection prediction is confirmed in the earlier current research output (Al-Rubaye *et al.*, 2018). In the previous publication Appendix H, the orientation sensor data (angles) reveal a substantial effect on differentiating among severely lame, mildly lame and sound classes of sheep in spite of using the raw data of DataSet2_a without any pre-processing steps.

4.4.1 Matrix of Scatter Plots for Sheep Raw Data

The matrix of scatter plot in Figure 4-5 refer to DataSet2_ac, which has 6 predictors (Acc_x, Acc_y, Acc_z, Azimuth, Pitch, and Roll). Each cell in the matrix of scatter plot showing a relationship between two predictors; conversely, the diagonal line shows the relationship between each predictor and itself. Due to DataSet2_ac has 6 predictors ($P = 6$), a plot matrix of $P * P$ plots were depicted in Figure 4-5 for DataSet2_ac as an example. However, the matrix of scatter plot for the other three sheep DataSet1_all, DataSet2_b, and DataSet3_all that have 9 predictors ($P = 9$) are presented in Appendix B. 1, Appendix B. 2, and Appendix B. 3Appendix B, respectively.

The plots in Figure 4-5 shows a widespread of each class, which causes class overlapping. The reason for this refers to the raw data having no cleaning step where the walking segments are extracted to be trained by the classifier to only classify the walking sheep status into sound, mildly lame, and severely lame walking. Thus, the pre-processing step is crucial in this case where the walking segments are extracted, and 183 features are computed and fed into the classifier (CART) in order to build the Sheep Lameness Detection Model (SLDM).

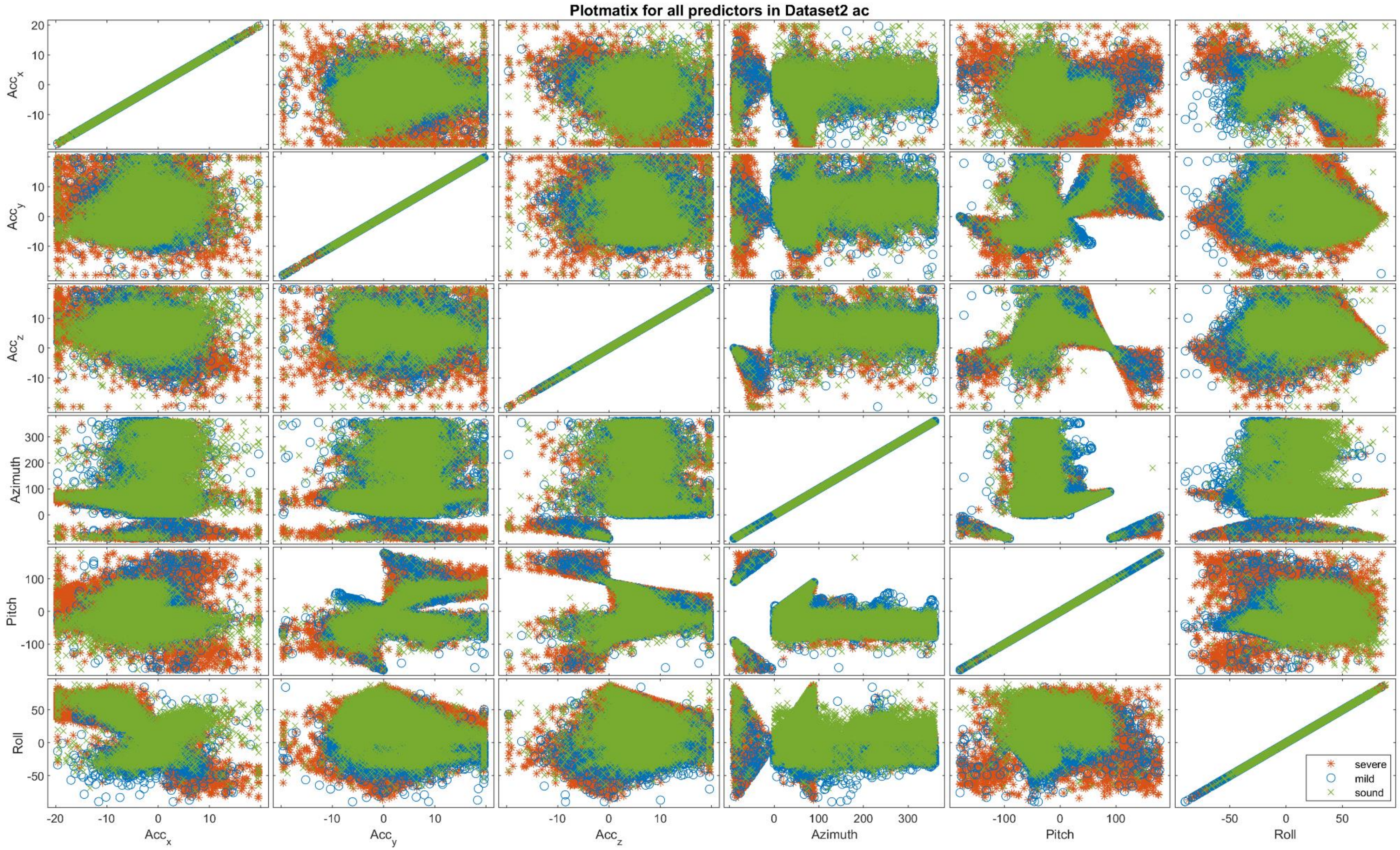


Figure 4-5 Scatter Plot matrix for raw Sheep DataSet2_ac, where *, o, and x represent severe, mild, and sound *Classes* in the DataSet.

4.4.2 Boxplots for Sheep Raw Data

Another form of plotting is a boxplot where each predictor is grouped in a separate box according to its belonging $Class$; either sound, mild, or severe. The boxplots of Figure 4-6 refer to DatSet2_ac that has six predictors (Acc_x, Acc_y, Acc_z, Azimuth, $Pitch$, and $Roll$). The depictions for sheep DataSet1_all, DataSet2_b, and DataSet3_all are illustrated in Appendix B. 4, Appendix B. 5, and Appendix B. 6, respectively.

Again, it would be not a straightforward process to distinguish among the three classes from raw data directly as the centres of each predictor box are convergence among the sound, mild and severe class. So, the importance of applying data pre-processing would be required for better prediction.

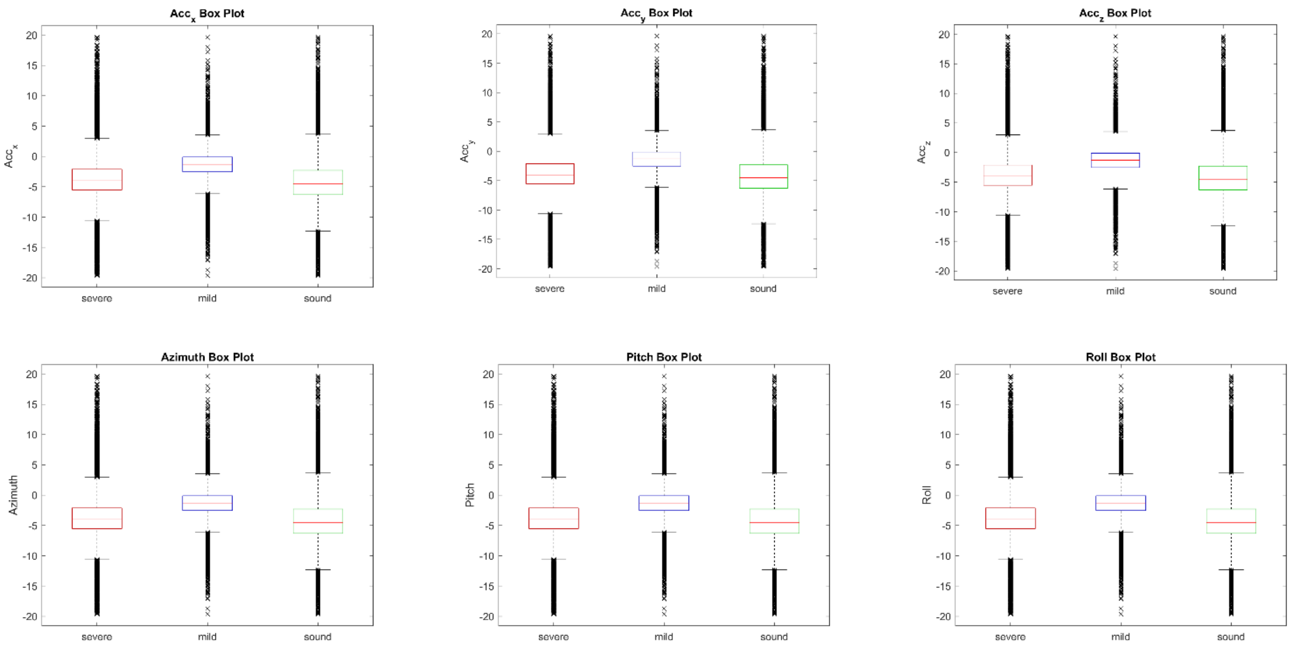


Figure 4-6 Box Plots for each predictor in raw sheep DataSet2_ac.

4.5 Data Pre-Processing Results and Discussion

The four sheep DataSets were pre-processed each for the two segmentation approaches (FNSW and FOSW) that were discussed in Section 3.5.3.1. For each segmentation's approach, three window sizes sz were chosen (10, 7, and 5 sec.) to segment each Dataset according to the selected sz. The classification results for the sheep walking segments are presented in Section 4.5.1, while the walking sheep datasets results are listed in Section 4.5.2. The results of the final step of the pre-processing stage; which is feature extraction and time execution for each feature, is presented in Section 4.5.3.

4.5.1 Sheep Movements Classification Results

The classification of sheep movements firstly checks the sheep's walking speed limits to be between 0.8 to 3.5 m/s. These limits; which were identified in the SLDM interface (pre-processing tab), could be re-defined for many executions; however, the selected walking speed limits in the current research could keep many data-points of a moving sheep even if these speed limits exceed the normal sheep walking that was identified in the literature.

The '*running mean window*' was set to be associated with the selected window size sz; for example, if $sz = 10$ for a Dataset with 10 Hz sampling rate, 100 data-points in the segment were tested for a running mean of each 10 neighbour points in the segment. Alternatively, the running mean was applied separately to neighbour points in the case of $sz = 7$, for the same DataSet as the segment size contains 70 data-points. This choice may guarantee the calculation of speed in relevance to the window size sz.

4.5.1.1 Sheep movements proportion for Standing, Walking, and Trotting Segments

Actually, the proportion of each class depends on the sheep's behaviour during the data collection experiments. In spite of this, each sheep was triggered to walk during the experiment; many times, the sheep tended to stand or start to trot if they felt the presence of the observer.

The percentages of Standing segments, Walking segments, and Trotting segments for each selected window size $10, 7, and\ 5\ sec.$ for the four aggregated sheep datasets DataSet1_all, DataSet2_ac, DataSet2_b, and DataSet3_all are provided in
Table 4-2, Table 4-3, and Table 4-4, respectively.

Table 4-2 Sheep movement classification for two segmentation approaches over 10 $sec.window$ for the four aggregated sheep datasets.

Window size		10 Sec.							
Segmentation type		FNSW				FOSW (20%)			
Segments ratio for Standing (S), Walking (W), and Trotting (T)		# segs	S %	W %	T %	# segs	S %	W %	T %
DataSet1_all	(seg_size= 50), (10 s * 5 Hz)	1273	24.51	52.32	23.17	1563	24.82	52.53	22.65
DataSet2_ac	(seg_size= 100), (10 s * 10 Hz)	1244	45.34	52.25	2.41	1542	44.94	52.66	2.4
DataSet2_b		585	12.99	76.07	10.94	716	12.01	77.93	10.06
DataSet3_all	(seg_size= 40), (10 s * 4 Hz)	130	6.92	43.08	50	156	7.05	41.03	51.92

Table 4-3 Sheep movement classification for two segmentation approaches over 7 $sec.window$ for the four aggregated sheep datasets.

Window size		7 Sec.							
Segmentation type		FNSW				FOSW (20%)			
Segments ratio for Standing (S), Walking (W), and Trotting (T)		# seg	S %	W %	T %	# seg	S %	W %	T %
DataSet1_all	(seg_size= 35), (7 s * 5 Hz)	1826	35.43	50.22	14.35	2253	35.15	51.26	13.58
DataSet2_ac	(seg_size= 70), (7 s * 10 Hz)	1777	64.38	34.89	0.73	2213	64.53	34.57	0.9
DataSet2_b		841	19.02	76.93	4.04	1035	19.23	77.58	3.19
DataSet3_all	(seg_size= 28), (7 s * 4 Hz)	188	11.17	60.64	28.19	228	13.16	58.77	28.07

Table 4-4 Sheep movement classification for two segmentation approaches over 5 $sec.window$ for the four aggregated sheep datasets.

Window size		5 Sec.							
Segmentation type		FNSW				FOSW (20%)			
Segments ratio for Standing (S), Walking (W), and Trotting (T)		# seg	S %	W %	T %	# seg	S %	W %	T %
DataSet1_all	(seg_size= 25), (5 s * 5 Hz)	2559	47.13	46.82	6.06	3172	47.07	47.07	5.86
DataSet2_ac	(seg_size= 50), (5 s * 10 Hz)	2492	73.72	26.04	0.24	3103	73.93	25.72	0.35
DataSet2_b		1180	29.41	69.75	0.85	1459	28.99	69.98	1.03
DataSet3_all	(seg_size= 20), (5 s * 4 Hz)	263	21.29	64.64	14.07	323	21.67	65.63	12.69

For the **DataSet1_all**, the proportion of walking segments in the 10 $sec.window$ for both FNSW and FOSW is 52.32% and 52.53%, which is approximately equal to more than half of the total collected data-points. This ratio of the obtained Walking segments could be considered as a representative ratio from the whole data-points. The aforementioned ratios of the obtained walking segments are a little bit better than the ones obtained from the 7 $sec.window$, where 50.22% and 51.26% are obtained for both segmentation approaches. In contrast to the results of the acceptable walking segments ratio of 10 and 7 $sec.window$, the 5 $sec.window$ size produces lower walking segments ratios of 46.82% and 47.7% for both FNSW and FOSW segmentation approaches.

The data-points of each segment in the $10, 7, 5$ $sec.windows$ are 50, 35, and 25 respectively, which reveals that the smaller number of data-points in a segment could not be considered as a representative segment size for a sensor data in the 5 Hz sampling rate. That means 50 data-points could describe the behaviour of the sheep better than 35 or 25 data-points. The results also reveal that whatever the window size is, the performance of FOSW is better than FNSW segmentation approach because 20% of the data-points are shared between every two successive windows. Although overlapping causes some data-points to be repeated in segments, it produces much better segmentation results than FNSW.

For the **DataSet2_ac** with 6 predictors and 10 Hz sampling rate, the proportion of walking segments over the 10 $sec.window$ exceeds half of the data-points which equal 52.25% and 52.66% for FNSW and FOSW, while the proportion of walking segments in two segmentation approaches is dropped to 34.89% and 37.57% over the 7 $sec.window$ and 26.04% and 25.72% over the 5 $sec.window$ size. The reason could be due to the sheep from attemp3 of SensoDunio (refer to Table 3.3) where the sheep were observed for an extended period of time, approximately more than one hour in an unattended procedure on the farm. So, there was a greater standing period of time than walking because the sheep were not triggered to walk as normal at that time. This is why the standing proportion is much higher the walking proportion, while the limited proportion of trotting segments appears since they tend to trot when the observer gets closer to encourage the sheep to walk.

The data-points of each segment in the $10, 7, 5$ $sec.windows$ are 100, 70, and 50, respectively. The fair walking segments proportion (more than half of the data-points) for the 10 Hz sampling rate is obtained over a $10\ sec.window$ (segment size = 100) which indicates an agreement with the DataSet1_all.

DataSet2_b (9 predictors) generates approximate proportions of walking segments over $10, 7, 5\ sec.windows$ for both FNSW and FOSW as follows (76.07%, 77.93), (76.93, 77.58), and (69.75%, 69.98%). All the obtained segment proportions were over half of the data-points as the sheep in the experiment of data collection were walking for most of the experiment time. The best proportion was obtained over a $10\ sec.window$ with FOSW. Again, an agreement of obtaining the acceptable walking segment or more than that is consistent with DataSet2_ac where the segment size=100.

In contrast with the 3 sheep DataSets, the **DataSet3_all** (4 Hz sampling rate) produces the best proportion of walking segments over the $5\ sec.window$ (64.64%, 65.63%) for two segmentation approaches, and it is followed by the results obtained over the $7\ sec.window$ (60.64%, 58.77). The lowest ratio for walking segments is obtained over the $10\ sec.window$ which are (43.08%, 41.03%). While the data-points of each segment in the $10, 7, 5\ sec.windows$ are 40, 28, and 20 respectively; the best walking proportion is obtained from the smallest segment size that contains 20 data-points. The reason could refer to the sampling rate of the Sensor Log used to collect the sheep data.

In general, the FOSW outperforms the FNSW as some data-points are shared between every two successive windows. In addition, DataSets with 10 Hz and 5 Hz have a walking proportion of over 50% for the window size of $10\ sec.$ and $7\ sec.$ Conversely, the DataSet with 4 Hz sampling rate produces walking segments of more than half of the data-points when $5\ sec.$ and $7\ sec.window$ sizes are applied. As a conclusion, the $7 sec.window$ could suit 10, 5, 4 Hz sensor sheep data.

4.5.1.2 Sheep movements plots for Standing, Walking, and Trotting Segments

The scatter plots in Figure 4-7 depicts Standing, Walking, and Trotting segments each in colours green, blue, and red, respectively for sheep DataSet2_ac over $10,7, 5\ sec.window$ sizes for both FNSW and FOSW segmentation method. In each plot, the x-axis represents the $Speed$ of each segment, while the y-axis refers to the $Vedba$. The $Vedba$ is the vectorial dynamic body acceleration that measures the energy expenditure of a sheep walking within a selected window (refer Table 3-6 in Chapter three for feature extraction). The other plots for sheep DataSet1_all, DataSet2_b, and DataSet3_all are presented in Appendix C. 1, Appendix C. 2, and Appendix C. 3, respectively.

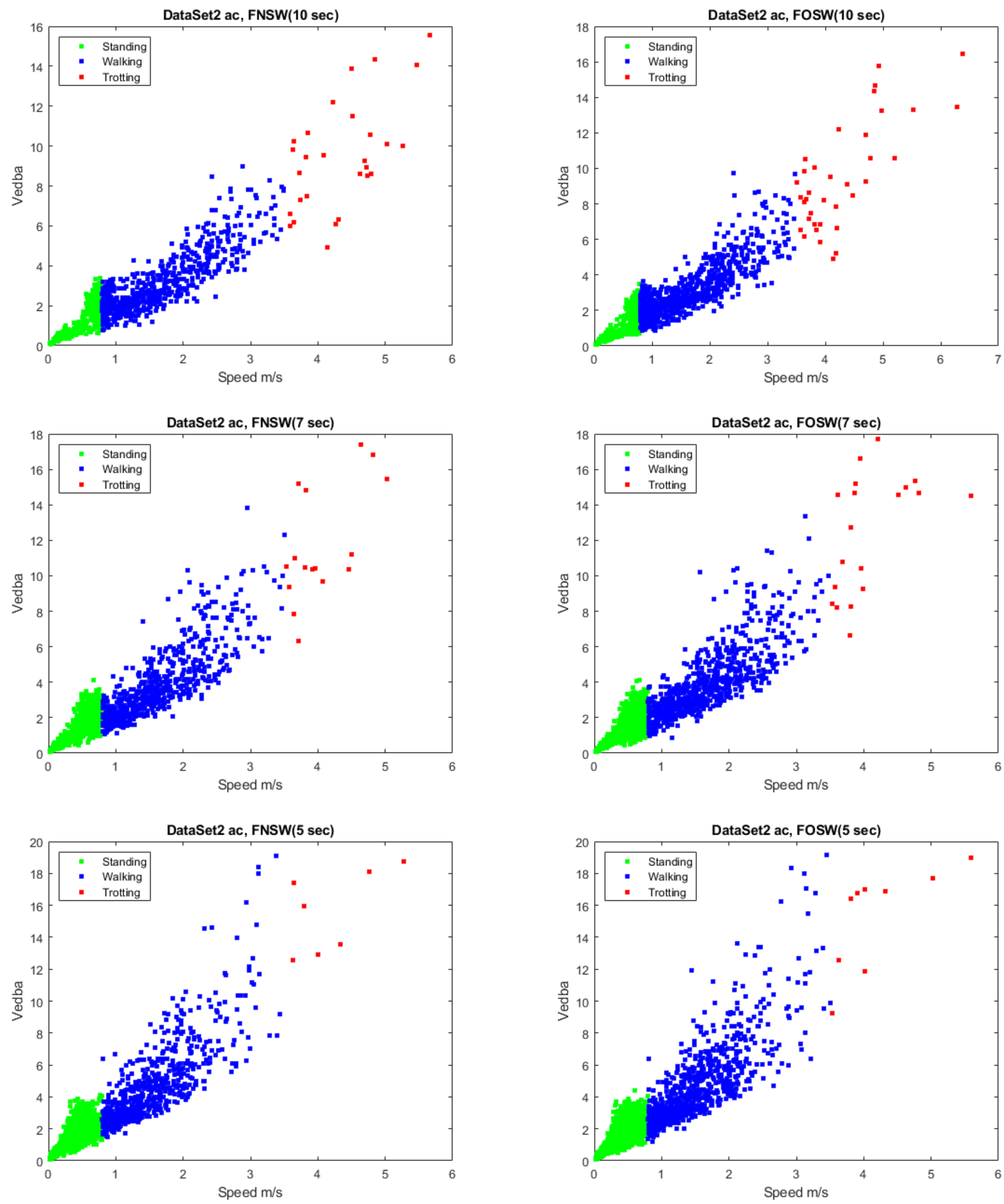


Figure 4-7 Scatter plots of the DataSet2_ac, where movement's classification is done over (10 sec, 7 sec, and 5 $sec.window$) for two segmentation approaches (FNSW and FOSW).

It is shown from the presented plots that the $Speed$ and $Vedba$ could increase together and decrease together as the energy spent for a standing sheep is less than the walking or trotting sheep and vice versa. It also appears from the presented figures that the segments (which represent every single point in the graph) are more scattered when the $Speed$ of the segment is increased, whereas they close together for Standing segments. The reason commonly referred to that minimal energy is spent while standing vs walking or trotting.

4.5.2 Walking Sheep DataSets Results

After the walking segments that were obtained from the previous stage, the walking segments of each sheep in a DataSet is aggregated together into one file. The $Class$ name was updated from 'sound', 'mild', and 'severe' to 'sound walking', 'mild walking', 'severe walking' respectively. Each sheep DataSets with its total walking segments (# total instances), and the percentage of 'sound walking', 'mild walking', and 'severe walking' are each listed in the following tables over 10 $sec.$, 7 $sec.$, and 5 $sec. windows$.

For example, the total instances of DataSet1_all after combining the walking segments of all sheep in the given DataSet are 666 instances. In FNSW, 127, 391, and 148 segments out of 666 refer to sound walking, mild walking, and severe walking in green, blue, and red colour in the respective presented tables. While 156, 490, and 175 segments for sound walking, mild walking and severe walking are obtained from the FOSW approach. The aforementioned numbers are obtained over a 10 $sec. window$. Alternatively, in the 7 $sec. window$, in FNSW, 212, 467, and 238 segments for sound, mild, and severe walking are gathered, while in FOSW 263, 592, and 300 segments for sound, mild, and severe walking are gathered. Lastly for 5 $sec. window$, in FNSW, 333, 507 and 358 for sound, mild, and severe walking segments are obtained, while in FOSW there are 418, 636, 439 segments for sound, mild, and severe walking gathered.

Generally**, DataSet1_all** has the 'mild walking' class, this is the dominant class over the 3 selected window sizes and approximately more than half of instances refer to the 'mild walking' class, while in **DataSet2_ac**, over the $10, 7, 5$ $sec. windows$ 'severe walking' class represents the majority class; however, it occupies two-thirds of total instances. Otherwise, 'sound walking' class is the majority class of **DataSet2_b**, where it exceeds half of the instances over the 3 selected window sizes. Similarly, **DataSet3_all** has the 'sound walking' class as the majority class over 'sound walking' or 'severe walking' classes over

10, 7, 5 $sec.windows$.

Table 4-5 Results of walking sheep Datasets over 10 $sec.window$ used in the feature extraction stage.

Window size		**10 Sec.**							
Segmentation type		**FNSW**				**FOSW (20%)**			
# Segments (instances) for Sound walking (Sw), Mild walking (Mw), and Severe walking (Srw)		# total instances	#Sw (%)	#Mw (%)	#Srw (%)	# total instances	#Sw (%)	#Mw (%)	#Srw (%)
DataSet1_all	(seg_size= 50), (10 s * 5 Hz)	666×451	127	391	148	821×451	156	490	175
			19.07	58.71	22.22		19.00	59.68	21.32
DataSet2_ac	(seg_size= 100), (10 s * 10 Hz)	650×601	204	194	252	812×601	252	241	319
			31.38	29.85	38.77		31.03	29.68	39.28
DataSet2_b		445×901	255	97	93	558×901	325	121	112
			57.30	21.80	20.90		58.42	21.68	20.07
DataSet3_all	(seg_size= 40), (10 s * 4 Hz)	56×361	29	12	15	64×361	34	12	18
			51.76	21.43	26.79		53.13	18.75	28.13

Table 4-6 Results of walking sheep Datasets over 7 $sec.window$ used in the feature extraction stage.

Window size		**7 Sec.**							
Segmentation type		**FNSW**				**FOSW (20%)**			
# Segments (instances) for Sound walking (Sw), Mild walking (Mw), and Severe walking (Srw)		# total instances	#Sw (%)	#Mw (%)	#Srw (%)	# total instances	#Sw (%)	#Mw (%)	#Srw (%)
DataSet1_all	(seg_size= 35), (7 s * 5 Hz)	917× 316	212	467	238	1155×316	263	592	300
			23.12	50.93	25.95		22.77	51.26	25.97
DataSet2_ac	(seg_size= 70), (7 s * 10 Hz)	620× 421	168	157	295	765× 421	217	189	359
			27.10	25.32	47.58		28.37	24.70	46.92
DataSet2_b		647×631	377	140	130	803×631	471	173	159
			58.27	21.64	20.10		58.66	21.54	19.80
DataSet3_all	(seg_size= 28), (7 s * 4 Hz)	114×253	60	27	27	134×253	70	31	33
			52.63	23.68	23.68		52.24	23.13	24.63

Table 4-7 Results of walking sheep Datasets over 5 $sec.window$ used in the feature extraction stage.

Window size		**5 Sec.**							
Segmentation type		**FNSW**				**FOSW (20%)**			
# Segments (instances) for Sound walking (Sw), Mild walking (Mw), and Severe walking (Srw)		# total instances	#Sw (%)	#Mw (%)	#Srw (%)	# total instances	#Sw (%)	#Mw (%)	#Srw (%)
DataSet1_all	(seg_size= 25), (5 s * 5 Hz)	1198×226	333	507	358	1493×226	418	636	439
			27.80	42.32	29.88		28.00	42.60	29.40
DataSet2_ac	(seg_size= 50), (5 s * 10 Hz)	649×301	185	172	292	798×301	232	206	360
			28.51	26.50	44.99		29.07	25.81	45.11
DataSet2_b		823×451	483	182	158	1021×451	597	228	196
			58.69	22.11	19.20		58.47	22.33	19.20
DataSet3_all	(seg_size= 20), (5 s * 4 Hz)	170×181	91	40	39	212×181	112	48	52
			53.53	23.53	22.94		52.83	22.64	24.53

4.5.3 Feature Extraction Results and Time Calculation

For each of the walking sheep dataset results in the previous Section 4.5.2, the extracted features are expressed in Table 4-8. The type and the number of features for 9 and 6 predictor Datasets are provided as well.

Table 4-8 The number and the type of each feature extracted from the four sheep DataSets.

Feature name (Table 3-6)	**# Features** (DataSet1_all, DataSet2_b, DataSet3_all)	**# Features** (Dataset2_ac)	**Feature_type**
Mean (μ)	9	6	Time-domain
Variance (∂)	9	6	Time-domain
Standard deviation (∂^2)	9	6	Time domain
Kurtosis (Kur)	9	6	Time-domain
Skewness (Skew)	9	6	Time domain
Maximum value (Max)	9	6	Time-domain
Minimum value (Min)	9	6	Time-domain
Root mean square (Rms)	9	6	Time-domain
Interquartile range (Interq)	9	6	Time-domain
Crest factor (CF)	9	6	Time-domain
Signal magnitude area (SMA)	3	2	Time-domain
Signal vector magnitude (SMV)	3	2	Time-domain
Differential Signal Vector Magnitude (DSMV)	3	2	Time-domain
Maximum difference (Max_diff)	9	6	Time-domain
Average movement variation (Avr_MV)	3	2	Time-domain
Magnitude (Mag)	3	2	Time-domain
Vector of the dynamic body acceleration (Vedba)	3	2	Time-domain
Entropy Time-domain (Ent3)	3	2	Frequency-domain
Entropy Frequency- domain (Ent)	9	6	Frequency-domain
Energy (Eng)	9	6	Frequency-domain
Dominant Freq	9	6	Frequency-domain
peak analysis (nPeaks, Widest_Peak, Highest_Peak, Avr_peak_time)	36	24	Frequency-domain
Total features	183	122	

The execution time of each feature was calculated over 10 $sec.$, 7 $sec.$, and 5 $sec. window$ and for both segmentation methods FOSW and FNSW in order to compromise the lowest execution time of features with its importance in relation to the sensor energy consumption. The execution time for each extracted feature for DataSet2_ac, which has 6 predictors collected at 10 Hz., is presented in Figure 4-8 as an example among the other sheep DataSet1_all, DataSet2_b, and DataSet3_all that are explored in Appendix D. 1, Appendix D. 2, and Appendix D. 3, respectively.

The overall results reveal that the required execution time in FNSW is less than the execution time of the feature extraction in FOSW. This is due to the number of segments in FOSW being higher than the total number of segments in FNSW. In addition, FNSW is an online segmentation method that could be directly implemented inside the sensor processor before the data is transported to the base station for analysis. The FNSW approach helps in producing a faster alarm system for the shepherd showing the sheep lameness status than the FOSW.

However, the average execution time is less than one second for most of the extracted features for the four Datasets overall window sizes of $10\,sec.$, $7\,sec.$, and $5\,sec.$ and for both segmentation approaches. Exceptionally, the Interquartile range (Interq) feature and Peak analysis features that include four related peak calculation features consume an execution time of over 1 second, as shown in the presented figures (Figure 4-8 and Appendix D). It is worth mentioning that the time for the Peak analysis feature is, in fact, an accumulation of four internal features: ($nPeaks$, $Widest\_Peak$, $Highest\_Peak$, and Avr_peak), so its execution time looks surprisingly high because Peak analysis requires the data within the selected window to be transferred in its frequency domain, where FFT calculations required extra execution time.

For the **DataSet1_all** (5 Hz, 9 predictors), the lowest execution time for features is obtained over the $10\ sec. window$. While the best execution time achieved for the **DataSet2_ac** (10 Hz, 6 predictors) is over both $7\ sec. and\ 5\ sec. windows$ with maximum peak analysis of 3.015 and 3.024 seconds, respectively. The best execution time is slightly increased by 0.5 seconds for the **DataSet2_b** (10 Hz, 9 predictors) over the $10\ sec. window$. Conversely, the best execution time for the **DataSet3_all** (4 Hz, 9 predictors) drops and reaches its shortest time of around 1 second over the $5\ sec. window$.

Generally, the smallest segment size consumes less time to extract the features than the much bigger window size. However, the window size would have to be compatible with the sampling rate of the sensor used to collect sheep data. Due to the small sampling rate of sensor readings (like 4 Hz), valuable information relating to sheep behaviour may be lost in comparison to the 10 Hz sampling rate.

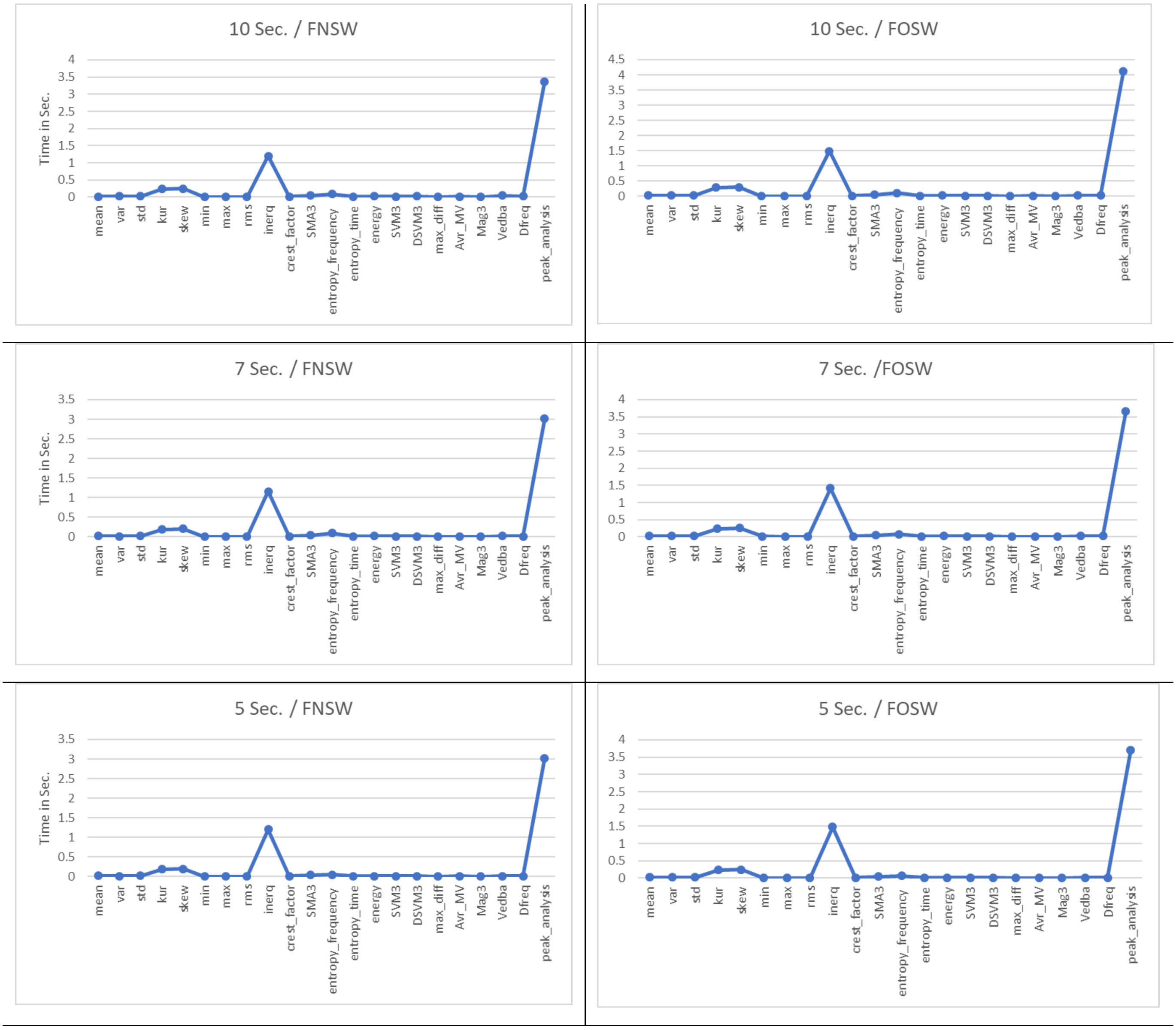


Figure 4-8 Execution time of features for DataSet2_ac (10 Hz).

4.6 Best Feature Selection Comparison Results and Discussion

The process of Feature Selection (FS) was applied to the four aggregated sheep DataSets. For each DataSet, three FS methods were applied where two segmentation methods (FNSW and FOSW) were tested over three selected window sizes $10, 7, and\ 5\ sec. window$. The obtained results are divided into three sections; in Section 4.6.1, the comparison of execution time for each scenario is provided, while in Section 4.6.2, the ranked features obtained from applying three FS methods (ReliefF, GA, RF) for each scenario were presented; however, only the first

25 features are shown as the whole list of ranked features for each FS method is provided in Appendix E. The last Section 4.6.3, explores the performance of a single CART algorithm to test for the best number of features to be considered for the lameness detection in sheep.

4.6.1 Execution Time Calculation for FS

The execution time spent for each aforementioned scenario is presented in this sub-section, where Table 4-9, Table 4-10, and Table 4-11 refer to the execution of each scenario over the $10, 7, and\ 5\ sec. windows$, respectively.

Table 4-9 Time execution comparison for feature selection methods over 10 $sec. window\ size$.

Window size	10 Sec.							
Segmentation type	FNSW				FOSW (20%)			
Execution time for FS methods in **sec.**	# instances	**ReliefF**	**GA**	**RF**	# instances	**ReliefF**	**GA**	**RF**
DataSet1_all ($sz = 50$)	666 × 183 +1	3.921	294.7	168.3	821 × 183 +1	3.791	347.2	179.9
DataSet2_ac ($sz = 100$)	650 × 122 +1	2.1	218.4	77.4	812× 122 +1	2.57	250.2	112.6
DataSet2_b ($sz = 100$)	445 × 183 +1	2.453	257.9	165.7	558 × 183 +1	2.513	277.4	176.4
DataSet3_all ($sz = 40$)	56 × 183 +1	0.628	99.56	67.18	64 × 183 +1	0.679	102.8	70.78

Table 4-10 Time execution comparison for feature selection methods over 7 $sec. window\ size$.

Window size	7 Sec.							
Segmentation type	FNSW				FOSW (20%)			
Execution time for FS methods in **sec.**	# instances	**ReliefF**	**GA**	**RF**	# instances	**ReliefF**	**GA**	**RF**
DataSet1_all ($sz = 35$)	917 × 183 +1	7.234	401.1	181.5	1155 × 183 +1	5.071	409.2	194.4
DataSet2_ac ($sz = 70$)	620× 122 +1	2.41	214.1	77.04	765× 122 +1	2.6	214	79.58
DataSet2_b ($sz = 70$)	445 × 183 +1	3.755	307.8	162.5	558 × 183 +1	3.569	355.6	166.8
DataSet3_all ($sz = 28$)	114 × 183 +1	0.835	131.8	97	134 × 183 +1	2.116	119.2	102.6

Table 4-11 Time execution comparison for feature selection methods over 5 $sec. window\ size$.

Window size	5 Sec.							
Segmentation type	FNSW				FOSW (20%)			
Execution time for FS methods in **sec.**	# instances	**ReliefF**	**GA**	**RF**	# instances	**ReliefF**	**GA**	**RF**
DataSet1_all ($sz = 25$)	1198 × 183 +1	6.471	568.4	181.7	1493 × 183 +1	6.711	613.9	195.5
DataSet2_ac ($sz = 50$)	649 × 122 +1	2.26	193	75.69	798 × 122 +1	2.268	227.3	77.26
DataSet2_b ($sz = 50$)	445 × 183 +1	4.634	406.9	180.7	558 × 183 +1	4.513	560	173.8
DataSet3_all ($sz = 20$)	170 × 183 +1	1.14	153.5	121	212 × 183 +1	1.949	186.1	127.6

The results in the presented three tables show that GA takes the longest time to be performed, then it is followed by RF, while the ReliefF method takes only a few seconds to be performed. However, the test for better lameness detection accuracy from each FS is tested in Section 4.6.3.

Balancing between the time spent and the accuracy of the classifier for lameness detection in sheep could be justified to meet the requirement. In the current research, the focus is on detection of lameness in its early stage, so the execution time spent has less concern than the accuracy of the alarm given for the classification of lameness.

In other cases, when the FS process is needed to be run online and deployed in the sensor kit itself, in this case, the execution time takes more attention than classification accuracy.

4.6.2 Ranked Features Results and Discussion of Feature Selection

The ranked features obtained from applying three FS methods (ReliefF, GA, RF) for each of the four aggregated sheep DataSets are shown in the following sections. For each DataSet, results are depicted in three tables for the first 25 ranked features, while the rest of 183 ranked features (9 predictors for DataSet1_all, DataSet2_b, and DataSet3_all) or 122 ranked features (6 predictors for DataSet2_ac) are all presented in long tables in Appendix E.

4.6.2.1 DataSet1_all (5 Hz) Best Ranked Features

DataSet1_all has 9 predictors with 183 calculated features. The following discussion reflects how the features are ranked according to the 3 FS methods used for the first 25^{th} features. A snapshot for the 25^{th} ranked features for DataSet1_all are presented in Figure 4-9, Figure 4-10, and Figure 4-11 over $10, 7, and\ 5\ sec.window$ sizes, while the whole lists are explored in Appendix E. 1, Appendix E. 2, and Appendix E. 3, respectively.

1. **ReliefF results:** The first 25^{th} ranked feature of ReliefF FS method for the $10\ sec.window$ (Figure 4-9), shows a majority contribution of the $Orient$ related features; especially the $Pitch$ and $Roll$ angles, the angles around $x-axis$ and $y-axis$, respectively. In FNSW (19 features out of the first 25^{th} features are $Pitch$ & $Roll$ related features, 4 features are Gyr_y, Gyr_z related features, and 2 only features are

Acc and Acc_y related features), while in FOSW (18 features out of the first 25th features are $Pitch$ & $Roll$ related features, 5 Gyr related features, and only 2 Acc related feature). However, the first 11th features are exactly the same for both FNSW and FOSW segmentation approaches. For the rest of the features (from 12th – 25th), approximately the same group of features appeared with a slightly different order at the end of the list.

On the other hand, the first 25th ranked feature of ReliefF FS method for the 7 $sec.window$ (Figure 4-10), shows the same group of features for the first 17th ranked features with slightly flipped order between (3rd and 4th) and (14th, 15th) for both FNS and FOSW. Similar to the result of 10 $sec.window$, the major contribution of features related to the Pitch and Roll are significant. Basically, for FNSW (21 features out of the first 25th features are $Pitch$ & $Roll$ related features, 3 features are Gyr related features, and only one feature (Mv_Acc) is related to Acc) , while in FOSW (20 features out of the first 25th features are $Pitch$ & $Roll$ related features, 3 Gyr related features, and 2 Acc related feature).

Finally, the result of the 5 $sec.window$ (Figure 4-11) reveals the same group of features for both FNSW and FOSW segmentation methods with a slight difference in features' order for the first 18th ranked features. Similarly, to the 10 $sec.window$ and 7 $sec.window$, the contribution of $Orient$ related features outperforms Gyr, and Acc related features. In FNSW (18 $Orient$, 5 Gyr, and 2 Acc related features), while in FOSW (21 $Orient$, 3 Gyr, 1 Acc related features). So, the contribution of the $Orient$ group in FNSW gets more prominence than the FNSW as 20% of data-points are overlapped between every two successive windows in the FOSW segmentation method.

2. **GA results:** The first 25th ranked feature of GA for the 10 $sec.window$ (Figure 4-9) has 17 common features between FNSW and FOSW; however, the order of features is not quite as similar as in ReliefF. Most of the features retrieved by GA belong to (Mean, Var, Std, Kur, and Skew) for Acc, Gyr, or $Orient$ groups. The results of FNSW produce (10 Acc, 8 $Orient$, and 7 Gyr related features), while in FOSW (9 Acc, 9 $Orient$, and 7 Gyr related features) are obtained. Due to the arbitrary initialisation of GA generation, the ranked features might not necessarily refer to their importance like

(nearest neighbours in ReliefF, or out-of-bag estimates by permutation in RF).
Likewise, the most prominent features over the $7\ sec.window$ (Figure 4-10) are ($Mean$, Var, Std, Kur, $Skew$, and Min) in both FNSW(9 Gyr, $Orient$, and 7 Acc related features) and FOSW (12 $Orient$, 8 Acc, and 5 Gyr related features). Whereas only 12 shared features between FNSW and FOSW differ from each other in their order.

Finally, the results of GA FS over the $5\ sec.window$ (Figure 4-11) for DataSet1_all (5 Hz) shows not much difference from the 10 and $7\ sec.window$. Although the obtained features vary among ($Mean$, Var, Std, Kur, $Skew$, and Min), the most related features in FNSW are (12 Acc_x & Acc_y, 8 Gyr, and 5 $Pitch$ and $Roll$). On the other hand, the most relevant features for FOSW are (11 $Orient$, 9 Acc, and 5 Gyr). In addition, only 8 features are present in both FNSW and FOSW in a different order with half of them being $Pitch$ and $Roll$ related features.

Surprisingly, the only order which could be noticed is the presence of features ($Mean$, Var, Std, kur, $Skew$, then Min or Max) whether in FNSW or FOSW for GA implementation. As mentioned before, the arbitrary features' presence refers to the random initialisation of the first generation in GA.

3. **RF results:** the three or four most important features for both FNSW and FOSW over $10, 7, 5\ sec.window$ sizes are $Pitch$ and $Roll$ related features. Simple time-domain features ($Mean$, Max, Min, and CF) are the top four features obtained from all scenarios. The importance of features is computed in RF by estimating the model error when a specified feature value is permutated to observe its influence on the model performance. If the permutation process increases the model error, that means the feature whose value is permutated has an influence on the model performance. On the other hands, when no effect has occurred when a feature value is permutated, then there is no significant importance of this feature and its rank decreases in the final list.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Energy_Roll	Energy_Roll	Mean_Acc_x	Mean_Acc_y	Max_Roll	Max_diff_Pitch
2	Rms_Roll	Rms_Roll	Mean_Acc_y	Mean_Azimuth	Mean_Roll	Max_Roll
3	Dfreq_Roll	Dfreq_Roll	Mean_Acc_z	Mean_Pitch	Max_diff_Pitch	Mean_Roll
4	Max_Roll	Max_Roll	Mean_Azimuth	Mean_Gyr_x	Min_Pitch	Max_Gyr_x
5	Mean_Roll	Mean_Roll	Mean_Pitch	Var_Acc_x	Max_Gyr_x	Max_Acc_y
6	Min_Roll	Min_Roll	Mean_Roll	Var_Acc_y	Rms_Roll	Min_Pitch
7	Cf_Roll	Cf_Roll	Mean_Gyr_x	Var_Acc_z	Entropy_Gyr_y	Entropy_Gyr_x
8	Max_Pitch	Max_Pitch	Mean_Gyr_y	Var_Azimuth	Cf_Roll	Skew_Acc_y
9	Entropy_Pitch	Entropy_Pitch	Var_Acc_x	Var_Pitch	Max_diff_Gyr_z	Rms_Pitch
10	Rms_Pitch	Rms_Pitch	Var_Acc_y	Var_Gyr_x	Mean_Acc_z	Max_diff_Gyr_z
11	Max_diff_Pitch	Max_diff_Pitch	Var_Azimuth	Var_Gyr_y	Entropy_Gyr_z	Max_Pitch
12	Energy_Pitch	Min_Pitch	Var_Pitch	Var_Gyr_z	Max_diff_Acc_z	Rms_Roll
13	Dfreq_Pitch	Energy_Pitch	Var_Gyr_x	Std_Acc_x	Cf_Gyr_x	Interq_Gyr_y
14	MV_Gyr	MV_Gyr	Var_Gyr_y	Std_Acc_z	Entropy_Gyr_x	Min_Acc_z
15	Entropy_Roll	Dfreq_Pitch	Var_Gyr_z	Std_Azimuth	Min_Acc_y	Max_diff_Gyr_x
16	Cf_Pitch	Entropy_Roll	Std_Acc_x	Std_Pitch	Max_diff_Acc_y	Min_Gyr_z
17	Max_diff_Gyr_z	Max_diff_Gyr_z	Std_Acc_z	Std_Roll	Var_Gyr_z	Mean_Acc_x
18	Std_Pitch	Cf_Pitch	Std_Pitch	Std_Gyr_x	Min_Gyr_x	Max_diff_Acc_z
19	MV_Acc	Std_Pitch	Std_Roll	Kur_Acc_x	Var_Acc_y	Min_Gyr_x
20	Min_Pitch	DSVM_Gyr	Std_Gyr_y	Kur_Acc_z	Var_Gyr_x	Cf_Roll
21	mag_Ang	MV_Acc	Std_Gyr_z	Kur_Azimuth	Mean_Pitch	Mean_Acc_z
22	Entropy_Gyr_y	Entropy_Gyr_y	Kur_Acc_x	Kur_Pitch	Rms_Gyr_x	Kur_Gyr_z
23	Max_diff_Acc_y	Max_diff_Acc_y	Kur_Acc_y	Kur_Gyr_x	Max_diff_Acc_x	Dfreq_Roll
24	DSVM_Gyr	Mean_Pitch	Kur_Pitch	Kur_Gyr_y	Dfreq_Gyr_x	Mean_Pitch
25	Mean_Pitch	Entropy_Gyr_x	Skew_Acc_y	Skew_Acc_x	Entropy_Acc_y	Var_Acc_y

Figure 4-9 Ranked feature result for DataSet1_all (5 Hz) over 10 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Energy_Roll	Energy_Roll	Mean_Acc_y	Mean_Acc_y	Max_diff_Pitch	Max_diff_Pitch
2	Rms_Roll	Rms_Roll	Mean_Acc_z	Mean_Acc_z	Max_Roll	Mean_Roll
3	Dfreq_Roll	Max_Roll	Mean_Pitch	Mean_Azimuth	Mean_Roll	Max_Roll
4	Max_Roll	Dfreq_Roll	Mean_Roll	Mean_Pitch	Cf_Pitch	Max_diff_Gyr_z
5	Mean_Roll	Mean_Roll	Mean_Gyr_x	Mean_Roll	Cf_Roll	Cf_Roll
6	Min_Roll	Min_Roll	Mean_Gyr_y	Mean_Gyr_x	Rms_Roll	Rms_Pitch
7	Cf_Roll	Cf_Roll	Mean_Gyr_z	Mean_Gyr_y	Mean_Pitch	Min_Gyr_x
8	Entropy_Pitch	Entropy_Pitch	Var_Acc_x	Var_Pitch	Dfreq_Gyr_z	Min_Gyr_z
9	Max_Pitch	Max_Pitch	Var_Acc_z	Var_Roll	Kur_Gyr_x	Rms_Roll
10	Rms_Pitch	Rms_Pitch	Var_Azimuth	Std_Acc_x	Skew_Acc_y	Mean_Pitch
11	Max_diff_Pitch	Max_diff_Pitch	Var_Roll	Std_Acc_y	Var_Gyr_z	Var_Pitch
12	Energy_Pitch	Energy_Pitch	Var_Gyr_z	Std_Azimuth	Kur_Gyr_z	Min_Acc_y
13	Cf_Pitch	Cf_Pitch	Std_Azimuth	Std_Roll	Max_diff_Acc_y	Max_Gyr_x
14	Dfreq_Pitch	Entropy_Roll	Std_Pitch	Std_Gyr_z	Var_Pitch	Max_diff_Acc_z
15	Entropy_Roll	Dfreq_Pitch	Std_Gyr_x	Kur_Acc_y	Max_Acc_y	Cf_Pitch
16	MV_Gyr	MV_Gyr	Std_Gyr_y	Kur_Azimuth	Rms_Pitch	Entropy_Roll
17	Std_Pitch	Std_Pitch	Kur_Acc_y	Kur_Pitch	Max_Gyr_x	Min_Roll
18	Mean_Pitch	Min_Pitch	Kur_Roll	Kur_Roll	Var_Gyr_x	Dfreq_Roll
19	Entropy_Gyr_y	Mean_Pitch	Kur_Gyr_x	Kur_Gyr_z	Max_diff_Acc_z	Dfreq_Acc_y
20	Min_Pitch	Entropy_Gyr_y	Skew_Acc_z	Skew_Acc_y	Kur_Pitch	Min_Acc_z
21	MV_Acc	MV_Acc	Skew_Pitch	Skew_Acc_z	Min_Acc_y	Min_Pitch
22	Entropy_TimeD_Ang	Max_diff_Acc_y	Skew_Roll	Skew_Pitch	Mean_Acc_z	Highest_peak_Pitch
23	SVM_Angle	Entropy_Gyr_x	Skew_Gyr_x	Skew_Gyr_x	Dfreq_Pitch	Var_Gyr_x
24	SMA_Angle	SVM_Angle	Skew_Gyr_z	Min_Acc_y	Max_diff_Gyr_x	Interq_Acc_x
25	Entropy_Gyr_x	Entropy_TimeD_Ang	Min_Acc_z	Min_Pitch	Min_Acc_z	Skew_Pitch

Figure 4-10 Ranked feature result for DataSet1_all (5 Hz) over 7 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Energy_Roll	Energy_Roll	Mean_Acc_x	Mean_Acc_z	Mean_Roll	Mean_Roll
2	Rms_Roll	Rms_Roll	Mean_Acc_y	Mean_Azimuth	Cf_Pitch	Max_diff_Pitch
3	Dfreq_Roll	Max_Roll	Mean_Gyr_x	Mean_Pitch	Var_Gyr_z	Cf_Roll
4	Mean_Roll	Mean_Roll	Mean_Gyr_y	Mean_Roll	Kur_Gyr_x	Cf_Pitch
5	Max_Roll	Dfreq_Roll	Var_Acc_x	Var_Acc_z	Max_Gyr_x	Min_Acc_z
6	Cf_Roll	Min_Roll	Var_Acc_y	Var_Azimuth	Max_Roll	Highest_peak_Pitch
7	Min_Roll	Cf_Roll	Var_Roll	Var_Roll	Max_diff_Gyr_z	Max_diff_Gyr_z
8	Entropy_Pitch	Entropy_Pitch	Var_Gyr_y	Std_Acc_y	Max_diff_Pitch	Rms_Roll
9	Max_Pitch	Max_Pitch	Var_Gyr_z	Std_Azimuth	Entropy_Pitch	Dfreq_Roll
10	Cf_Pitch	Cf_Pitch	Std_Acc_x	Std_Roll	Mean_Acc_x	Var_Acc_z
11	Max_diff_Pitch	Max_diff_Pitch	Std_Acc_y	Std_Gyr_x	Var_Pitch	Interq_Acc_x
12	Rms_Pitch	Entropy_Roll	Std_Roll	Skew_Acc_x	Var_Gyr_x	Max_Roll
13	Entropy_Roll	Rms_Pitch	Std_Gyr_z	Skew_Acc_y	Rms_Roll	Entropy_Pitch
14	Energy_Pitch	Energy_Pitch	Kur_Acc_x	Skew_Acc_z	Min_Gyr_x	Max_Gyr_x
15	Std_Pitch	Std_Pitch	Kur_Acc_y	Skew_Pitch	Max_Pitch	Mean_Gyr_x
16	MV_Gyr	MV_Gyr	Kur_Pitch	Skew_Roll	Mean_Gyr_x	Var_Gyr_x
17	Dfreq_Pitch	Dfreq_Pitch	Kur_Gyr_z	Skew_Gyr_y	Mean_Pitch	Min_Gyr_x
18	Mean_Pitch	Mean_Pitch	Skew_Acc_x	Min_Acc_x	Var_Acc_y	Rms_Pitch
19	Entropy_Gyr_y	Min_Pitch	Skew_Acc_y	Min_Acc_z	Dfreq_Roll	Skew_Acc_y
20	Min_Pitch	Max_diff_Gyr_x	Skew_Pitch	Min_Pitch	Min_Acc_z	Var_Gyr_z
21	Max_diff_Gyr_x	Max_diff_Acc_y	Skew_Gyr_x	Min_Roll	Highest_peak_Pitch	Max_Pitch
22	Max_diff_Acc_y	Entropy_Gyr_y	Skew_Gyr_y	Min_Gyr_x	Highest_peak_Gyr_z	Max_diff_Gyr_y
23	Max_diff_Gyr_z	SVM_Angle	Min_Acc_x	Min_Gyr_y	Cf_Roll	Min_Roll
24	Entropy_Gyr_x	Entropy_TimeD_Ang	Min_Acc_y	Min_Gyr_z	Kur_Gyr_y	Var_Pitch
25	Max_Acc_y	SMA_Angle	Min_Roll	Max_Acc_x	Kur_Pitch	Mean_Pitch

Figure 4-11 Ranked feature result for DataSet1_all (5 Hz) over 5 $sec.\ window$.

4.6.2.2 DataSet2_ac (10 Hz) Best Ranked Features

DataSet2_ac has 122 features where no Gyr readings are included in this DataSet. The achieved results of 25th ranked features for the three FS methods over $10,7,5\ window$ are shown in Figure 4-12, Figure 4-13, and Figure 4-14 and discussed below, while the whole lists are explored in Appendix E. 4, Appendix E. 5, and Appendix E. 6, respectively.

1- **ReliefF results**: the 1st feature for both FNSW and FOSW over $10, 7, and\ 5\ sec.window$ is the $Entropy\_Roll$ feature which measures the energy disorder of Roll angle of a walking sheep (angle around forward-backwards axis y-axis) in a selected window. Due to no Gyr predictors in DataSet2_ac, the contribution of $Orient$ and Acc features are approximately equal; where 12 features are $Orient$ related and 13 features are Acc related, for all six lists (2 segmentation methods × 3 window sizes). The order of features could be similar for the 25th ranked feature with some alteration between near locations (indices) in the list for all 6 scenarios. The reason for the correlated order of features between FNSW and FOSW for $10, 7, and\ 5\ sec.windows$ could refer to the same number of common features between FNSW and FOSW which are 24, 23, and 21 for $10, 7, and\ 5\ sec.windows$, respectively.

2- **GA results**: the randomness of the resulted features by applying GA FS is the same as DataSet1_all results; however, some new features were added to the first 25th ranked feature for all 6 lists. In addition to the ($Mean$, Var, Std, kur, $Skew$, Min or Max), the feature of (Rms, $Interq$, or CF) are ranked within the first 25 features. The common features between FNSW and FOSW are 11, 12, and 10 for $10, 7, and\ 5\ sec. windows$, respectively.

3- **RF results:** for the first 25 ranked features over $10, 7, and\ 5\ sec. window$ for both FNSW and FOSW, Min_Roll feature is the first ranked feature for all scenarios. The number of shared features within the 25th ranked features for both FNSW and FOSW over $10, 7, and\ 5\ sec. window$ are 18, 18, and 19 features. The $Orient$ related features dominate the Acc related features for all 3 window sizes, for both segmentation methods. Where in $10\ sec. window$, 16 and 15 $Orient$ related features with FNSW and FOSW segmentation methods, respectively. Similarly, the dominant features within the first 25 ranked features over $7, and\ 5\ sec. window$ are the Orientation features with 14 features out of 25.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Entropy_Roll	Entropy_Roll	Mean_Acc_x	Mean_Azimuth	Min_Roll	Rms_Azimuth
2	Entropy_Acc_x	Dfreq_Acc_x	Mean_Azimuth	Mean_Pitch	Mean_Roll	Min_Roll
3	Dfreq_Acc_x	Entropy_Acc_x	Mean_Roll	Var_Acc_x	Rms_Azimuth	Mean_Roll
4	Dfreq_Roll	Dfreq_Roll	Var_Acc_x	Var_Roll	Mean_Acc_x	Mean_Acc_x
5	Rms_Roll	Rms_Roll	Var_Acc_z	Std_Acc_x	Entropy_Acc_x	Interq_Azimuth
6	Mean_Roll	Mean_Roll	Var_Pitch	Std_Acc_y	Vedb_Angle	Dfreq_Roll
7	Rms_Acc_x	Mean_Acc_x	Std_Acc_y	Std_Pitch	Mean_Acc_z	Max_Pitch
8	Mean_Acc_x	Rms_Acc_x	Std_Acc_z	Kur_Acc_y	Rms_Pitch	Mean_Acc_z
9	Max_Roll	Dfreq_Acc_z	Std_Azimuth	Kur_Pitch	Max_diff_Acc_y	Entropy_Roll
10	Energy_Roll	Max_Roll	Std_Pitch	Skew_Acc_x	Entropy_Roll	Min_Acc_z
11	Min_Acc_x	Mean_Acc_z	Std_Roll	Skew_Acc_z	Rms_Roll	Rms_Pitch
12	Dfreq_Acc_z	Energy_Roll	Kur_Acc_x	Skew_Azimuth	DSAM_Angle	Rms_Roll
13	Mean_Acc_z	Min_Acc_x	Kur_Pitch	Skew_Pitch	Mean_Acc_y	Max_Acc_y
14	Energy_Acc_z	Energy_Acc_z	Skew_Acc_y	Min_Acc_x	Max_Pitch	Entropy_Acc_x
15	Rms_Acc_z	Rms_Acc_z	Skew_Acc_z	Min_Acc_z	Min_Acc_z	Mean_Acc_y
16	Energy_Acc_x	Mean_Acc_y	Skew_Azimuth	Max_Acc_x	Skew_Roll	Max_Roll
17	Mean_Acc_y	Energy_Acc_x	Skew_Roll	Max_Pitch	Rms_Acc_x	Min_Pitch
18	Cf_Acc_x	Cf_Acc_x	Min_Acc_z	Max_Roll	Cf_Roll	Dfreq_Acc_x
19	Max_Acc_y	Max_Acc_y	Min_Azimuth	Rms_Acc_x	Mean_Azimuth	Mean_Azimuth
20	Max_Pitch	Rms_Pitch	Min_Roll	Rms_Acc_y	Dfreq_Roll	Vedb_Angle
21	Cf_Roll	Max_Pitch	Max_Acc_z	Rms_Pitch	DSVM_Acc	DSAM_Angle
22	Min_Roll	Mean_Pitch	Max_Azimuth	Rms_Roll	Mean_Pitch	Rms_Acc_x
23	Rms_Pitch	Min_Roll	Max_Pitch	Interq_Azimuth	Var_Acc_x	Skew_Roll
24	Mean_Pitch	Cf_Roll	Max_Roll	Interq_Pitch	Var_Roll	Min_Acc_x
25	SVM_Angle	Energy_Pitch	Rms_Acc_x	Cf_Acc_y	Max_diff_Roll	Rms_Acc_z

Figure 4-12 Ranked feature result for DataSet2_ac (10 Hz) over 10 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Entropy_Roll	Entropy_Roll	Mean_Acc_x	Mean_Acc_y	Min_Roll	Min_Roll
2	Dfreq_Acc_x	Dfreq_Acc_x	Mean_Acc_z	Mean_Roll	Mean_Roll	Mean_Roll
3	Entropy_Acc_x	Dfreq_Roll	Mean_Azimuth	Var_Azimuth	Mean_Acc_x	Mean_Acc_z
4	Dfreq_Roll	Entropy_Acc_x	Var_Acc_x	Std_Acc_z	Mean_Acc_z	Entropy_Acc_x
5	Rms_Roll	Rms_Roll	Var_Acc_z	Std_Azimuth	Max_Pitch	Min_Acc_z
6	Mean_Roll	Mean_Roll	Var_Azimuth	Std_Pitch	Entropy_Roll	Max_Azimuth
7	Dfreq_Acc_z	Dfreq_Acc_z	Var_Pitch	Kur_Acc_y	Mean_Azimuth	Mean_Acc_x
8	Mean_Acc_z	Max_Roll	Var_Roll	Kur_Acc_z	Min_Acc_z	Rms_Roll
9	Min_Acc_x	Mean_Acc_x	Std_Pitch	Kur_Pitch	Kur_Acc_x	Max_Pitch
10	Max_Roll	Min_Acc_x	Kur_Acc_y	Kur_Roll	Rms_Roll	Entropy_Roll
11	Mean_Acc_x	Mean_Acc_z	Kur_Acc_z	Skew_Acc_x	Entropy_Acc_x	Rms_Azimuth
12	Rms_Acc_x	Rms_Acc_x	Kur_Azimuth	Skew_Azimuth	Rms_Azimuth	Mean_Azimuth
13	Energy_Roll	Energy_Roll	Skew_Acc_y	Min_Acc_x	Max_Azimuth	Dfreq_Acc_x
14	Rms_Acc_z	Min_Roll	Skew_Azimuth	Min_Acc_z	Dfreq_Acc_x	Mean_Acc_y
15	Energy_Acc_z	Rms_Acc_z	Min_Acc_x	Min_Pitch	Dfreq_Roll	Vedb_Angle
16	Min_Roll	Energy_Acc_z	Min_Pitch	Min_Roll	Mean_Pitch	Kur_Azimuth
17	Cf_Roll	Cf_Acc_x	Max_Acc_x	Max_Acc_x	Min_Acc_x	Dfreq_Roll
18	Cf_Acc_x	Cf_Roll	Max_Acc_z	Max_Acc_z	Skew_Roll	Max_diff_Acc_y
19	Energy_Acc_x	Energy_Acc_x	Max_Pitch	Max_Pitch	Var_Pitch	DSVM_Acc
20	Max_Acc_x	Max_Acc_x	Max_Roll	Max_Roll	Cf_Acc_x	Var_Pitch
21	Max_Pitch	Entropy_Azimuth	Rms_Acc_y	Rms_Azimuth	Rms_Pitch	Var_Roll
22	Entropy_Azimuth	Rms_Pitch	Rms_Acc_z	Rms_Pitch	SMA_Angle	Min_Pitch
23	Rms_Pitch	SVM_Angle	Rms_Roll	Interq_Acc_x	Max_Acc_x	Max_Acc_x
24	Mean_Pitch	Entropy_Acc_z	Interq_Acc_x	Interq_Pitch	Rms_Acc_x	Min_Acc_x
25	Entropy_Acc_z	SMA_Angle	Interq_Acc_y	Cf_Azimuth	DSVM_Acc	Min_Acc_y

Figure 4-13 Ranked feature result for DataSet2_ac (10 Hz) over 7 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Entropy_Roll	Entropy_Roll	Mean_Acc_x	Mean_Acc_y	Min_Roll	Min_Roll
2	Entropy_Acc_x	Entropy_Acc_x	Mean_Acc_y	Mean_Azimuth	Mean_Roll	Mean_Roll
3	Rms_Roll	Dfreq_Roll	Mean_Acc_z	Mean_Pitch	Mean_Acc_x	Mean_Acc_x
4	Dfreq_Roll	Dfreq_Acc_x	Mean_Azimuth	Mean_Roll	Mean_Azimuth	Entropy_Acc_x
5	Dfreq_Acc_x	Rms_Roll	Mean_Pitch	Var_Acc_x	Max_Pitch	Mean_Acc_z
6	Dfreq_Acc_z	Mean_Roll	Mean_Roll	Var_Acc_y	Rms_Azimuth	Mean_Azimuth
7	Mean_Roll	Mean_Acc_x	Var_Acc_x	Var_Acc_z	Mean_Acc_z	Dfreq_Roll
8	Max_Roll	Min_Acc_x	Var_Azimuth	Var_Azimuth	Max_Acc_x	Rms_Azimuth
9	Mean_Acc_x	Max_Roll	Var_Pitch	Var_Roll	Rms_Roll	Var_Acc_x
10	Mean_Acc_z	Dfreq_Acc_z	Std_Acc_x	Std_Roll	Entropy_Roll	Entropy_Roll
11	Min_Acc_x	Mean_Acc_z	Std_Acc_y	Kur_Azimuth	Dfreq_Acc_x	Rms_Roll
12	Energy_Roll	Energy_Roll	Std_Acc_z	Kur_Pitch	Entropy_Acc_x	Max_Azimuth
13	Min_Roll	Min_Roll	Std_Pitch	Skew_Acc_x	Var_Acc_x	Min_Acc_z
14	Rms_Acc_x	Rms_Acc_x	Kur_Acc_x	Skew_Acc_y	Var_Roll	Max_Acc_x
15	Cf_Acc_x	Cf_Acc_x	Kur_Acc_z	Skew_Acc_z	Dfreq_Acc_z	Min_Pitch
16	Cf_Roll	Energy_Acc_z	Kur_Roll	Skew_Pitch	Mean_Pitch	Cf_Azimuth
17	Rms_Acc_z	Rms_Acc_z	Skew_Acc_x	Min_Azimuth	Min_Acc_z	Skew_Acc_y
18	Energy_Acc_z	Max_Acc_x	Skew_Acc_y	Max_Acc_x	Max_diff_Azimuth	Rms_Acc_z
19	Max_Acc_x	Cf_Roll	Skew_Acc_z	Max_Acc_y	Cf_Acc_x	Var_Pitch
20	Energy_Acc_x	Entropy_Azimuth	Skew_Azimuth	Max_Pitch	Cf_Azimuth	Min_Azimuth
21	Entropy_Azimuth	Energy_Acc_x	Min_Acc_x	Max_Roll	Min_Azimuth	Max_Pitch
22	Std_Acc_x	Std_Roll	Min_Acc_z	Rms_Acc_x	Dfreq_Roll	Var_Roll
23	Entropy_Pitch	Entropy_Acc_z	Min_Roll	Rms_Acc_y	Rms_Acc_x	Mean_Acc_y
24	Min_Pitch	Min_Pitch	Max_Acc_x	Rms_Azimuth	Skew_Acc_y	Dfreq_Acc_x
25	Max_Pitch	Rms_Pitch	Max_Acc_y	Rms_Roll	Vedb_Angle	Min_Acc_x

Figure 4-14 Ranked feature result for DataSet2_ac (10 Hz) over 5 $sec.\ window$.

4.6.2.3 DataSet2_b (10 Hz) Best Ranked Features

The obtained results of the best 25 ranked features of DataSet2_b over $10, 7, and\ 5\ sec.window$ are shown in Figure 4-15, Figure 4-16, and Figure 4-17, while the whole lists are presented in Appendix E. 7, Appendix E. 8, and Appendix E. 9, respectively. The ranked features according to the 3 FS methods used for DataSet2_b; which has 9 predictors and 183 features, are discussed in the following:

1- **ReliefF results:** the two features ranked first in both FNSW and FOSW over $10, 7, 5\ sec.window$ are $Mean\_Acc\_x$ and $Mean\_Roll$. Additionally, the high intersection of features (23, 24, and 24) between FNSW and FOSW could be noticed over the $10, 7, 5\ sec.windows$, respectively. The reason might relate to the technique of ReliefF algorithm to search for the 10th neighbour's instances, which shared the same class.

   The majority contribution of features for both FNSW and FOSW over the $10, 7, 5\ \ sec.windows$ are Acc related features. In more detail, the 10 $sec.window$ for FNSW has 13 Acc, 11 $Pitch$ & $Roll$, and only one Gyr related feature, which is $Interq_{Gyr_x}$, while for FOSW (13 Acc related features, 12 $Pitch$ & $Roll$ related features) and no presence of any Gyr related features within the 25 ranked features. Otherwise, for the 7 $sec.window$, (FNSW: 14 Acc, 11 $Pitch$ & $Roll$ related features, FOSW: 13 Acc, 12 $Pitch$ & $Roll$). Lastly, the 5 $sec.window$ has 14 Acc, and 11 $Pitch$ & $Roll$ related features for both FNSW and FOSW.

   Generally, the correlation of features within the 25 highest ranked list is quite high for ReliefF implementation for the two segmentation methods. Furthermore, there is a rare existence of Gyr related features in the 25 highest ranked list, which eliminates the importance of Gyr readings to predict the lameness status of the sheep.

2- **GA results:** no new feature sets are present in the 25 highest ranked list, the same as previous ($Mean$, Var, Std, kur, $Skew$, Min, Max, or Rms) features appeared in the list. In contrast to the high correlation of features present in ReliefF, the shared features between FNSW and FOSW are 11, 7, and 11 for $10, 7, and\ 5\ sec.window$, respectively.

3- **RF results:** the first three dominant features for all scenarios are (Min_Roll,

$Mean\_Acc\_x$, $Mean\_Roll$). The common features for both FNSW and FOSW within the 25 ranked features over $10, 7, and$ 5 $sec. windows$ are 14, 16, and 13 features. The most contributing features over the 10 $sec. window$ are Acc related features (FNSW: 13 Acc, 10 $Pitch$ & $Roll$, and 3 Gyr related features; FOSW: 12 Acc, 9 $Pitch$ & $Roll$, and 4 Gyr related features). Similarly, over the 7 $sec. window$, Acc related features (FNSW: 12 Acc, 12 $Pitch$ & $Roll$, and one Gyr related features; FOSW: 12 Acc, 10 $Pitch$ & $Roll$, and 3 Gyr related feature) contribute the most. Conversely, over the 5 $sec. window$, Pitch & Roll related features contribute most within 25 highest ranked list of features (FNSW: 10 Acc, 12 $Pitch$ & $Roll$, and 3 Gyr related features; FOSW: 10 Acc, 11 $Pitch$ & $Roll$, and 4 Gyr related features).

Generally, within the 25 highest ranked list of features for all scenarios, the features relating to Acc and $Pitch$ & $Roll$ ranked higher when compared to Gyr related features, which reveals that Gyr readings could not have much effect on the model prediction due to less Gyr features being present in the 25 highest ranked features against Acc and $Pith$ & $Roll$ features.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Acc_x	Mean_Roll	Mean_Acc_y	Mean_Acc_x	Mean_Acc_x	Min_Roll
2	Mean_Roll	Mean_Acc_x	Mean_Azimuth	Mean_Gyr_z	Mean_Roll	Mean_Roll
3	Min_Roll	Entropy_Roll	Mean_Pitch	Var_Acc_y	Min_Roll	Mean_Acc_x
4	Dfreq_Acc_z	Dfreq_Roll	Mean_Roll	Var_Acc_z	Mean_Acc_z	Mean_Acc_z
5	Entropy_Roll	Entropy_Acc_x	Mean_Gyr_x	Var_Azimuth	Mean_Pitch	Rms_Roll
6	Cf_Roll	Dfreq_Acc_x	Var_Acc_y	Var_Gyr_x	Vedb_Acc	Rms_Pitch
7	Mean_Acc_z	Cf_Roll	Var_Acc_z	Std_Acc_z	Rms_Pitch	Vedb_Acc
8	Dfreq_Pitch	Dfreq_Acc_z	Std_Acc_x	Std_Gyr_x	Rms_Roll	Rms_Acc_y
9	Vedb_Acc	Mean_Acc_z	Std_Acc_y	Std_Gyr_y	Min_Pitch	Mean_Pitch
10	Dfreq_Roll	Min_Roll	Std_Acc_z	Kur_Acc_x	Kur_Acc_y	Dfreq_Roll
11	Rms_Pitch	Energy_Pitch	Std_Azimuth	Kur_Azimuth	Dfreq_Pitch	Cf_Gyr_x
12	Dfreq_Acc_x	Vedb_Acc	Std_Roll	Kur_Pitch	Mean_Acc_y	Mean_Acc_y
13	Energy_Acc_y	Rms_Pitch	Std_Gyr_x	Kur_Gyr_x	Max_Roll	Kur_Gyr_x
14	Max_Roll	Rms_Roll	Std_Gyr_y	Skew_Acc_x	Entropy_Acc_x	Min_Gyr_y
15	Energy_Pitch	Max_Roll	Kur_Acc_x	Skew_Gyr_y	Cf_Pitch	Entropy_Roll
16	Dfreq_Acc_y	Dfreq_Pitch	Kur_Acc_y	Skew_Gyr_z	Dfreq_Roll	Max_diff_Acc_y
17	Rms_Acc_y	Energy_Acc_y	Kur_Acc_z	Min_Acc_z	Rms_Acc_x	Kur_Acc_y
18	Entropy_Acc_x	Rms_Acc_y	Kur_Pitch	Min_Azimuth	Max_Gyr_y	Interq_Roll
19	Rms_Roll	Dfreq_Acc_y	Kur_Roll	Min_Pitch	Min_Gyr_x	Max_Roll
20	Cf_Acc_x	Energy_Roll	Kur_Gyr_x	Min_Gyr_x	Dfreq_Acc_x	Cf_Acc_y
21	Min_Acc_x	Min_Acc_x	Skew_Roll	Min_Gyr_y	Rms_Acc_y	Max_Acc_y
22	Mean_Pitch	Mean_Pitch	Skew_Gyr_y	Min_Gyr_z	Skew_Gyr_x	Var_Acc_y
23	Interq_Gyr_x	Entropy_Acc_z	Skew_Gyr_z	Max_Acc_y	Kur_Acc_z	Var_Acc_z
24	Entropy_Acc_z	Cf_Acc_x	Min_Acc_y	Max_Acc_z	Max_Acc_x	Max_Gyr_z
25	Energy_Acc_z	Min_Acc_y	Min_Gyr_x	Max_Azimuth	Min_Acc_z	Dfreq_Acc_x

Figure 4-15 Ranked feature result for DataSet2_b (10 Hz) over 10 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Acc_x	Mean_Acc_x	Mean_Acc_z	Mean_Acc_x	Min_Roll	Min_Roll
2	Mean_Roll	Mean_Roll	Mean_Azimuth	Mean_Gyr_x	Mean_Acc_x	Mean_Roll
3	Cf_Roll	Entropy_Roll	Mean_Pitch	Var_Roll	Mean_Roll	Rms_Roll
4	Entropy_Roll	Cf_Roll	Mean_Roll	Var_Gyr_y	Rms_Pitch	Mean_Acc_x
5	Dfreq_Acc_x	Energy_Pitch	Mean_Gyr_y	Std_Acc_z	Rms_Roll	Mean_Pitch
6	Dfreq_Roll	Rms_Pitch	Mean_Gyr_z	Std_Azimuth	Cf_Acc_x	Min_Acc_y
7	Dfreq_Pitch	Dfreq_Pitch	Var_Pitch	Std_Gyr_y	Mean_Acc_y	Mean_Acc_z
8	Min_Roll	Dfreq_Acc_z	Var_Gyr_z	Kur_Acc_y	Mean_Acc_z	Var_Roll
9	Mean_Acc_z	Mean_Acc_z	Std_Acc_x	Kur_Roll	Kur_Acc_y	Rms_Acc_y
10	Dfreq_Acc_z	Dfreq_Roll	Std_Acc_y	Skew_Acc_z	Skew_Gyr_x	Dfreq_Pitch
11	Max_Roll	Min_Roll	Std_Acc_z	Skew_Gyr_x	SVM_Acc	Rms_Pitch
12	Rms_Pitch	Dfreq_Acc_x	Std_Azimuth	Skew_Gyr_y	Mean_Pitch	Mean_Acc_y
13	Entropy_Acc_x	Max_Roll	Std_Pitch	Skew_Gyr_z	Vedb_Acc	Max_Pitch
14	Dfreq_Acc_y	Dfreq_Acc_y	Std_Gyr_y	Min_Acc_x	Dfreq_Pitch	Max_diff_Acc_y
15	Energy_Pitch	Rms_Roll	Kur_Acc_y	Min_Azimuth	Var_Roll	Dfreq_Roll
16	Cf_Acc_x	Entropy_Acc_x	Kur_Acc_z	Min_Roll	Min_Acc_y	Rms_Acc_z
17	Vedb_Acc	Vedb_Acc	Kur_Pitch	Min_Gyr_y	Entropy_TimeD_Acc	Vedb_Acc
18	Rms_Roll	Cf_Acc_x	Kur_Roll	Max_Acc_y	Dfreq_Acc_x	Entropy_Gyr_z
19	Rms_Acc_y	Mean_Pitch	Kur_Gyr_x	Max_Acc_z	Max_Acc_y	Skew_Gyr_x
20	Energy_Acc_y	Rms_Acc_y	Kur_Gyr_y	Max_Azimuth	Rms_Acc_y	Dfreq_Acc_x
21	Min_Acc_x	Max_Acc_x	Skew_Acc_y	Max_Pitch	Min_Pitch	Min_Acc_z
22	Mean_Pitch	Energy_Roll	Skew_Acc_z	Max_Roll	Kur_Pitch	Interq_Roll
23	Entropy_Acc_y	Energy_Acc_y	Skew_Azimuth	Max_Gyr_z	Energy_Pitch	Min_Gyr_y
24	Max_Acc_x	Min_Acc_x	Skew_Pitch	Rms_Acc_x	Cf_Roll	Entropy_Acc_x
25	Rms_Acc_z	Entropy_Acc_y	Skew_Gyr_z	Rms_Acc_y	Dfreq_Roll	Var_Acc_z

Figure 4-16 Ranked feature result for DataSet2_b (10 Hz) over 7 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Acc_x	Mean_Acc_x	Mean_Acc_x	Mean_Acc_x	Min_Roll	Min_Roll
2	Mean_Roll	Mean_Roll	Mean_Acc_y	Mean_Pitch	Mean_Acc_x	Mean_Roll
3	Cf_Roll	Cf_Roll	Mean_Acc_z	Mean_Roll	Mean_Roll	Mean_Acc_x
4	Min_Roll	Dfreq_Acc_z	Mean_Azimuth	Mean_Gyr_y	Rms_Pitch	Mean_Acc_z
5	Dfreq_Roll	Dfreq_Roll	Mean_Roll	Mean_Gyr_z	Mean_Pitch	Mean_Acc_y
6	Entropy_Roll	Min_Roll	Mean_Gyr_y	Var_Acc_x	Mean_Acc_z	Rms_Pitch
7	Dfreq_Pitch	Max_Roll	Var_Roll	Var_Acc_y	Dfreq_Pitch	Max_diff_Acc_y
8	Rms_Pitch	Entropy_Roll	Var_Gyr_y	Var_Azimuth	Rms_Roll	Min_Acc_y
9	Max_Roll	Dfreq_Acc_y	Std_Acc_y	Var_Gyr_x	Dfreq_Acc_x	Mean_Pitch
10	Dfreq_Acc_z	Dfreq_Pitch	Std_Acc_z	Var_Gyr_y	Mean_Acc_y	Rms_Roll
11	Dfreq_Acc_y	Rms_Roll	Std_Gyr_x	Var_Gyr_z	Vedb_Acc	Var_Roll
12	Energy_Pitch	Mean_Acc_z	Kur_Acc_x	Std_Azimuth	Kur_Acc_y	Min_Pitch
13	Mean_Acc_z	Rms_Pitch	Kur_Acc_y	Std_Pitch	Max_diff_Gyr_z	Skew_Gyr_x
14	Dfreq_Acc_x	Dfreq_Acc_x	Kur_Acc_z	Std_Gyr_x	Cf_Pitch	Dfreq_Pitch
15	Rms_Roll	Rms_Acc_y	Kur_Azimuth	Std_Gyr_y	Cf_Roll	Entropy_Roll
16	Cf_Acc_x	Cf_Acc_x	Kur_Gyr_y	Kur_Acc_x	Interq_Roll	Kur_Acc_y
17	Min_Acc_x	Energy_Pitch	Skew_Acc_x	Kur_Acc_z	Rms_Acc_z	Max_Pitch
18	Vedb_Acc	Energy_Acc_y	Skew_Pitch	Kur_Azimuth	Interq_Gyr_z	Rms_Acc_z
19	Rms_Acc_y	Min_Acc_x	Skew_Roll	Kur_Gyr_z	Min_Pitch	Min_Gyr_x
20	Entropy_Acc_y	Max_Acc_x	Skew_Gyr_z	Skew_Acc_x	Var_Acc_x	Max_Roll
21	Max_Acc_x	Entropy_Acc_y	Min_Acc_x	Skew_Acc_y	Var_Roll	Rms_Acc_y
22	Energy_Acc_y	Entropy_Acc_x	Min_Azimuth	Skew_Azimuth	Var_Gyr_x	Min_Acc_z
23	Mean_Pitch	Vedb_Acc	Min_Pitch	Skew_Pitch	Max_Acc_x	Var_Gyr_z
24	Entropy_Acc_x	Rms_Acc_z	Min_Gyr_z	Skew_Roll	Dfreq_Roll	Max_Gyr_z
25	Rms_Acc_z	Energy_Roll	Max_Acc_x	Skew_Gyr_x	Entropy_Acc_y	Entropy_Acc_x

Figure 4-17 Ranked feature result for DataSet2_b (10 Hz) over 5 $sec.\ window$.

4.6.2.4 DataSet3_all (4 Hz) Best Ranked Features

The results for the best 25 ranked features retrieved from 3 FS methods for DataSet3_all over $10, 7, and\ 5\ sec.window$ are presented in Figure 4-18, Figure 4-19, and Figure 4-20, while the whole lists are explored in Appendix E. 10, Appendix E. 11, and Appendix E. 12, respectively. The discussion for the obtained ranked features is provided as follows:

1- **ReliefF results:** The number of shared features between FNSW and FOSW over the $10, 7, and\ 5\ sec.windows$ are 15, 18, and 19, respectively. The reason behind the $5\ sec.window$ having more shared features between FNSW and FOSW refers to the small sampling rate of DataSet3_all, which is 4 Hz. The $Mean\_Roll$ feature is the first feature in the ranked list for both FNSW and FOSW over the $7\ and\ 5\ sec.windows$, and for FNSW over the $10\ sec.window$; surprisingly, $nPeaks\_Gyr\_z$ is the top feature in FNSW over the $10\ sec.window$. The $Orient$ related features are the dominant features over all three $windows$ for both segmentation methods. The implementation of ReliefF over the $10\ sec.window$ produces (FNSW: 7 Acc, 15 $Orient$, and 3 Gyr related features, while in FOSW: 6 Acc, 14 $Orient$, and 5 Gyr related features). Whereas, the $7\ sec.window$ implementation yields (FNSW: 5 Acc, 10 $Orient$, and 9 Gyr related features, while in FOSW: 3 Acc, 13 $Orient$, and 9 Gyr related features). Lastly for the $5\ sec.window$, the results reveal (FNSW: 8 Acc, 11 $Orient$, and 6 Gyr related features, while in FOSW: 6 Acc, 11 $Orient$, and 8 Gyr related features).

2- **GA results:** the implementation of feature selection by GA over $10, 7, and\ 5\ sec.windows$ between FNSW and FOSW produces several common features which equal to 13, 13, and 14, respectively. Approximately, features from Acc, $Orient$ and Gyr are all involved in the list of 25 highest ranked features for all 6 scenarios. So, no group could be considered dominant between the three groups of features. The appearance of features over the $10, 7, and\ 5\ sec.windows$ could be summarised as follows: (FNSW: 5 Acc, 12 $Orient$, 8 Gyr related features; FOSW: 9 Acc, 7 $Orient$, and 9 Gyr related features) over the $10\ sec.window$. while the distribution of features for the $7\ sec.window$ like (FNSW: 9 Acc, 8 $Orient$, 8 Gyr related features; FOSW: 8 Acc, 10 $Orient$, and 7 Gyr related features). Finally, the contribution of features over the $5\ sec.window$ like (FNSW: 9 Acc, 8 $Orient$, 8 Gyr related features; FOSW: 11 Acc, 8 $Orient$, and 6 Gyr related features). As mentioned, the randomness of ranked features from GA refers to the randomness in initialising the first generation of GA. So, the fitness function would not be computed for features set to be ‘0’ instead of ‘1’.

3- **FR results:** the implementation of RF reveals that the maximum number of shared features between FNSW and FOSW is 11 over the 7 $sec.window$, and it is followed by 8 common features over the 5 $sec.window$. What is surprising is that only one shared feature over the 10 $sec.window$ is common between FNSW and FOSW, which is Max_Pitch. The contribution from all features could be relatively changeable; for example, in the 10 $sec.window$ the distribution of features would be as (FNSW: 7 Acc, 12 $Orient$, 6 Gyr related features; FOSW: 11 Acc, 8 $Orient$, and 6 Gyr related features). On the other hand, the contribution of features in the 7 $sec.window$ would be (FNSW: 6 Acc, 11 $Orient$, 8 Gyr related features; FOSW: 4 Acc, 8 $Orient$, and 13 Gyr related features). Lastly, for the 5 $sec.window$, the features appear as (FNSW: 8 Acc, 8 $Orient$, 9 Gyr related features; FOSW: 5 Acc, 12 $Orient$, and 8 Gyr related features).

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	nPeaks_Gyr_z	Mean_Roll	Mean_Acc_x	Mean_Acc_x	Mean_Pitch	Cf_Pitch
2	Var_Pitch	Cf_Pitch	Mean_Pitch	Mean_Azimuth	Skew_Acc_z	Mean_Acc_x
3	Mean_Roll	Mean_Acc_x	Mean_Roll	Mean_Roll	Entropy_TimeD_Gyr	Var_Acc_y
4	Mean_Pitch	Max_Pitch	Mean_Gyr_x	Mean_Gyr_x	Highest_peak_Azimuth	Min_Acc_y
5	Std_Pitch	Std_Pitch	Var_Azimuth	Mean_Gyr_y	Std_Gyr_x	Mean_Roll
6	Entropy_Roll	Var_Pitch	Var_Gyr_x	Mean_Gyr_z	Std_Acc_z	Var_Pitch
7	Mean_Acc_y	Cf_Roll	Var_Gyr_z	Var_Acc_x	Energy_Gyr_z	Var_Gyr_x
8	Dfreq_Roll	Min_Roll	Std_Azimuth	Var_Pitch	Std_Acc_x	Cf_Azimuth
9	Rms_Roll	Min_Acc_y	Std_Pitch	Var_Gyr_z	nPeaks_Acc_z	Min_Acc_z
10	Widest_Peak_Gyr_x	Entropy_Pitch	Std_Roll	Std_Acc_x	Energy_Roll	Mean_Gyr_y
11	Cf_Pitch	Max_diff_Azimuth	Std_Gyr_x	Std_Acc_z	Std_Gyr_z	Mean_Gyr_z
12	Mean_Acc_x	Max_Roll	Std_Gyr_y	Std_Azimuth	Max_Gyr_y	Max_Pitch
13	Max_diff_Azimuth	Interq_Pitch	Std_Gyr_z	Std_Pitch	Entropy_Roll	Interq_Acc_z
14	Max_Pitch	Mean_Acc_y	Kur_Acc_x	Std_Roll	Max_diff_Azimuth	Dfreq_Roll
15	Avr_peak_time_Gyr_z	Mean_Pitch	Kur_Acc_y	Std_Gyr_x	Min_Pitch	Min_Acc_x
16	Cf_Roll	Highest_peak_Gyr_x	Kur_Azimuth	Std_Gyr_y	Skew_Azimuth	Max_Acc_y
17	Min_Roll	Dfreq_Gyr_x	Kur_Pitch	Std_Gyr_z	Std_Azimuth	Min_Gyr_y
18	Skew_Acc_y	Interq_Acc_y	Kur_Roll	Kur_Acc_x	AV_Ang	Highest_peak_Acc_z
19	Energy_Roll	Dfreq_Roll	Kur_Gyr_y	Kur_Acc_z	Kur_Gyr_x	Min_Roll
20	Dfreq_Acc_x	Var_Acc_y	Skew_Acc_y	Kur_Roll	Vedb_Angle	Vedb_Acc
21	Min_Acc_x	Entropy_Gyr_x	Skew_Pitch	Kur_Gyr_x	Max_Pitch	Max_Acc_x
22	Interq_Pitch	Rms_Roll	Skew_Roll	Kur_Gyr_z	Rms_Acc_x	Dfreq_Pitch
23	Highest_peak_Pitch	Std_Acc_y	Skew_Gyr_z	Skew_Acc_x	Energy_Acc_x	Kur_Acc_x
24	Skew_Acc_z	Mean_Gyr_y	Min_Acc_x	Skew_Acc_y	Rms_Roll	Mean_Gyr_x
25	Var_Acc_y	Mean_Gyr_x	Min_Azimuth	Skew_Acc_z	Cf_Acc_z	Max_diff_Acc_y

Figure 4-18 Ranked feature result for DataSet3_all (4 Hz) over 10 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Roll	Mean_Roll	Mean_Acc_x	Mean_Acc_x	Dfreq_Roll	Rms_Roll
2	Rms_Roll	Mean_Acc_x	Mean_Pitch	Mean_Acc_y	Mean_Pitch	Dfreq_Roll
3	Mean_Acc_x	Cf_Roll	Var_Acc_x	Mean_Acc_z	Mean_Acc_x	Max_Pitch
4	Cf_Roll	Rms_Roll	Var_Acc_y	Mean_Pitch	Cf_Pitch	Mean_Roll
5	Dfreq_Roll	Std_Gyr_y	Var_Acc_z	Mean_Roll	Min_Gyr_y	Max_diff_Pitch
6	Energy_Roll	Max_Roll	Var_Gyr_x	Mean_Gyr_x	Min_Acc_x	Min_Gyr_y
7	Mean_Gyr_z	Rms_Gyr_y	Std_Acc_x	Var_Azimuth	Min_Azimuth	Mean_Gyr_z
8	Std_Gyr_y	Mean_Gyr_z	Std_Acc_z	Var_Pitch	Interq_Gyr_y	Highest_peak_Gyr_y
9	Entropy_Roll	Dfreq_Roll	Std_Pitch	Var_Gyr_y	Mean_Roll	Min_Acc_x
10	Rms_Gyr_y	Std_Pitch	Std_Roll	Std_Acc_x	Rms_Roll	Min_Pitch
11	Max_Roll	Max_diff_Gyr_y	Std_Gyr_y	Std_Acc_y	Mean_Gyr_y	Interq_Gyr_y
12	Max_diff_Gyr_y	Min_Roll	Kur_Acc_x	Std_Acc_z	Var_Acc_y	Min_Roll
13	Var_Gyr_y	Var_Gyr_y	Kur_Azimuth	Std_Azimuth	Min_Gyr_z	Mean_Acc_x
14	Min_Roll	Energy_Roll	Kur_Pitch	Std_Pitch	Entropy_Roll	Interq_Gyr_x
15	Energy_Gyr_y	Var_Pitch	Kur_Roll	Std_Roll	Max_diff_Pitch	Max_diff_Gyr_x
16	Min_Acc_x	Entropy_Roll	Kur_Gyr_y	Kur_Acc_x	mag_Ang	Mean_Gyr_y
17	Interq_Gyr_y	Energy_Gyr_y	Skew_Acc_x	Kur_Azimuth	Var_Acc_z	Mean_Acc_z
18	Interq_Gyr_x	Rms_Acc_x	Skew_Azimuth	Kur_Gyr_y	Max_diff_Gyr_y	Var_Pitch
19	Min_Pitch	Min_Acc_y	Skew_Gyr_x	Kur_Gyr_z	Interq_Pitch	Var_Gyr_y
20	Highest_peak_Gyr_y	Dfreq_Gyr_y	Skew_Gyr_y	Skew_Acc_x	Interq_Gyr_x	Max_Acc_y
21	mag_Ang	Interq_Gyr_y	Min_Gyr_x	Skew_Pitch	Max_Acc_z	Entropy_Gyr_x
22	Max_Acc_y	Highest_peak_Gyr_y	Min_Gyr_y	Skew_Gyr_x	Min_Pitch	Cf_Gyr_x
23	Skew_Acc_y	nPeaks_Azimuth	Min_Gyr_z	Min_Roll	Min_Acc_z	Kur_Gyr_x
24	Entropy_Pitch	Max_Pitch	Max_Acc_x	Min_Gyr_x	Cf_Gyr_z	Max_Gyr_y
25	Rms_Acc_x	Cf_Pitch	Max_Azimuth	Min_Gyr_y	Rms_Gyr_y	Dfreq_Gyr_y

Figure 4-19 Ranked feature result for DataSet3_all (4 Hz) over 7 $sec.\ window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Roll	Mean_Roll	Mean_Acc_x	Mean_Acc_y	Dfreq_Roll	Rms_Roll
2	Mean_Acc_x	Cf_Roll	Mean_Acc_y	Mean_Acc_z	Mean_Gyr_z	Mean_Roll
3	Cf_Roll	Mean_Acc_x	Mean_Pitch	Mean_Azimuth	Mean_Roll	Mean_Gyr_z
4	Dfreq_Roll	Max_Roll	Var_Acc_y	Mean_Roll	Mean_Acc_x	Dfreq_Roll
5	Mean_Gyr_z	Entropy_Roll	Var_Acc_z	Mean_Gyr_x	Rms_Roll	Skew_Acc_z
6	Rms_Roll	Rms_Roll	Var_Azimuth	Mean_Gyr_y	Min_Roll	Max_diff_Gyr_y
7	Entropy_Roll	Dfreq_Roll	Var_Pitch	Mean_Gyr_z	Skew_Acc_z	Mean_Acc_x
8	Cf_Acc_x	Min_Roll	Var_Roll	Var_Acc_x	Interq_Gyr_x	Max_Roll
9	Min_Roll	Max_diff_Gyr_y	Var_Gyr_x	Var_Acc_y	Min_Pitch	Max_diff_Gyr_x
10	Dfreq_Acc_x	Cf_Acc_x	Var_Gyr_y	Var_Acc_z	Var_Gyr_y	Var_Acc_y
11	Max_Roll	Mean_Gyr_z	Var_Gyr_z	Var_Pitch	Highest_peak_Gyr_x	Var_Gyr_y
12	Std_Gyr_y	Rms_Gyr_y	Std_Acc_x	Var_Roll	mag_Ang	Min_Roll
13	Rms_Acc_x	Std_Gyr_y	Std_Acc_y	Std_Acc_y	Kur_Gyr_x	Max_diff_Pitch
14	Std_Pitch	Max_Acc_x	Std_Acc_z	Std_Acc_z	Rms_Acc_x	Dfreq_Gyr_x
15	Max_diff_Gyr_y	Max_diff_Pitch	Std_Azimuth	Std_Azimuth	Max_Acc_x	Min_Acc_x
16	Energy_Roll	Energy_Roll	Std_Pitch	Std_Gyr_x	Skew_Acc_y	Max_Azimuth
17	Rms_Gyr_y	Max_Acc_y	Std_Gyr_x	Std_Gyr_y	Entropy_Acc_y	Interq_Roll
18	Min_Acc_z	Rms_Acc_x	Std_Gyr_y	Std_Gyr_z	Min_Gyr_y	Skew_Roll
19	Highest_peak_Acc_x	Highest_peak_Acc_x	Std_Gyr_z	Kur_Acc_x	Interq_Pitch	mag_Acc
20	Max_Acc_x	Energy_Gyr_y	Kur_Acc_z	Kur_Acc_y	Skew_Gyr_x	Entropy_Gyr_x
21	Min_Acc_x	Std_Roll	Kur_Azimuth	Kur_Acc_z	Var_Pitch	Rms_Gyr_y
22	Std_Roll	Var_Gyr_y	Kur_Gyr_x	Kur_Azimuth	Max_Gyr_z	SMA_Angle
23	Entropy_Pitch	Entropy_Pitch	Kur_Gyr_z	Kur_Gyr_z	Widest_Peak_Acc_z	Entropy_Pitch
24	Highest_peak_Gyr_y	Interq_Gyr_y	Skew_Acc_z	Skew_Acc_x	Var_Gyr_x	Entropy_Roll
25	Interq_Gyr_z	Std_Gyr_z	Skew_Azimuth	Skew_Gyr_x	Max_Acc_z	Max_diff_Gyr_z

Figure 4-20 Ranked feature result for DataSet3_all (4 Hz) over 5 $sec.\ window$.

4.6.3 CART Performance Results' Discussion for the Best Number of Features

In this section, the performance of a single CART classifier is tested to identify the best accuracy associated with the number of features that are fed to the classifier. For the sake of saving sensor energy, the lowest number of features with a fair percentage of lameness prediction accuracy would be preferred. Therefore, the four aggregated sheep DataSets are tested for 3 FS methods (ReliefF, GA, and RF), and 2 segmentation approaches (FNSW & FOSW) over 3 window sizes. Thus, the results for each DataSet are provided in 3 tables over 10, 7, and 5 $sec.window$ sizes. Each table explores 6 CART performance (3 FS methods × 2 segmentation methods). The whole set of results for the four sheep DataSets are provided in Appendix F, whereas the following sections discuss the obtained results for each sheep DataSets.

The prediction accuracy of CART for **DataSet1_all** (Appendix F. 1, Appendix F. 2, and Appendix F. 3) for all scenarios 10, 7, and 5 $sec.window$ when FOSW was applied outperforms the prediction accuracy of CART when FNSW was used; this is because of 20% overlapped data-points were shared between every two successive windows in the FOSW segmentation method.

Furthermore, the highest prediction accuracy between the $windows$ is achieved with the 10 $sec.window$ (50 data-points), and it is followed by the 7 $sec.window$ (35 data-points), and the lowest being the 5 $sec.window$ (25 data-points). That means 50 data-points has higher accuracy in predicting the lame walking sheep than the smaller segment sizes.

In addition, the performance of CART for all 3 FS methods is relatively good; however, on many occasions, the RF could produce better accuracy results. It is worth mentioning that the accuracy of CART could be considerable when approximately 20 features are used to feed the classifier.

Similarly, **DataSet2_ac** (Appendix F. 4, Appendix F. 5, and Appendix F. 6) produces a better accuracy of CART in the FOSW segmentation method for the same aforementioned reason. The accuracy of CART is increased for DataSet2_ac. There are no gyroscope readings in the current sheep DataSet, as the presence of a Gyroscope sensor could increase the sensor energy consumption while not having much effect on the accuracy of the prediction.

Also, the prediction accuracies between the $10, 7, and\ 5\ sec. windows$ are significant for all window sizes, being (100 data-points, 70 data-points, and 50 data-points) for each $10, 7, and\ 5\ sec. window$, respectively.

For FS, the performance of CART for ReliefF FS method is higher than RF and GA; however, satisfactory accuracies could be obtained when approximately 20 features feed the classifier.

Although the **DataSet2_b** (Appendix F. 7, Appendix F. 8, and Appendix F. 9) has the same sampling rate of 10 Hz as DataSet2_ac, the performance of CART drops, the reason could refer to the shorter monitoring time for sheep movement during the data collection process compared to the monitoring time in the previous DataSet2_ac.

The prediction accuracies converge for each $window$ when each segment has 100, 70, and 50 data-points. However, the highest accuracy is obtained from applying CART using FOSW over a 10 $sec. window$.

Satisfactory accuracies could be achieved with around 20 features or less for three FS approaches; except that when applying GA with FOSW in the 10 $sec. window$ which might exceed 20 features to reach an acceptable accuracy prediction.

The last tested DataSet is **DataSet3_all** (Appendix F. 10, Appendix F. 11, and Appendix F. 12), which collected sheep data at a 4 Hz sampling rate. Like other sheep DataSets, the performance of CART is higher for FOSW compared to FNSW over $10, 7, and\ 5\ sec. window$ sizes due to the segment's overlapping by 20%.

The three FS methods perform well; however, RF outperforms the other two FS methods for a single CART classifier over the different window sizes used. Also, 20 features are probably enough to reveal satisfactory accuracies for lameness prediction for all scenarios.

The total of 40, 28, and 20 data-points are allocated for the $10, 7, and\ 5\ sec. windows$, respectively, which produce higher prediction accuracies of CART over the 10 $sec. window$. That is followed by less accuracy over the 7 $sec. window$ and the least accuracy over the 5 $sec. window$.

In conclusion, the performance of CART is represented in Table 4-12, where the darker colour refers to the higher accuracy obtained. The table displays the performance for FOSW segmentation method as it achieves better accuracy than FNSW in all scenarios. The number inside the table's cell represents the total number of data-points in each segment. The segment is used to extract one feature to be chosen or not for the classifier.

Table 4-12 the performance of CART relates to colour density (darker is higher), no. inside cells represent the segment size for each sheep Datasets.

For FOSW	10 $sec. window$	7 $sec. window$	5 $sec. window$
DataSet1_all	50	35	25
DataSet2_ac	100	70	50
DataSet2_b	100	70	50
DataSet3_all	40	28	20

4.7 SLDM Ensemble Train, Test, and Validate Results and Discussion

The ensemble of CART decision trees is applied to the sheep datasets using 'Bagging', 'Boosting', or 'RusBoosting' methods. The four aggregated Datasets are trained separately to make a comparison for the best sensor sampling rate to detect the early signs of lameness as each DataSet was collected according to different sensor readings either 5, 10, or 4 Hz for DataSet1, DataSet2, and DataSet3 respectively. Firstly, the training parameters for ensemble classifiers are explained in Section 4.7.1. Then the achieved results for lameness detection are presented and discussed in Section 4.7.2. Finally, the comparison results between the validation methods used in the research are presented and discussed in Section 4.7.3.

4.7.1 CART Training Parameter

1- The maximum number of splits at each run '$MaxNumSplits$' is equal to its default which is equal to the $number\ of\ training\ records - 1$; for example, $MaxNumSplits = 811$ for DataSet1, which has 812 records in the FOSW segmentation method over 10 sec. window. $MaxNumSplits$ is set to its default for 'Bagging', while it is set to be 20 for both 'Boosting', and 'RusBoosting'.

   $MaxNumSplits$ controls the depth of the tree, and it is one of the stopping criteria for the splitting procedure in CART unless other stopping criteria are met like the number of observations (records) in one of the branch nodes being equal or less than the

‘$MinParentSize = 2$’, or the number of observations of one of the leaf nodes is equal or less than the ‘$MinLeafSize = 1$’.

2- The number of predictors to be selected at random for each split ‘$NumVariablesToSample$’ is set to ‘all’ in order to utilise all provided predictors to the classifier CART.
3- The number of ensemble learning cycles ‘$NumLearningCycles$’ is set to 100 trees, where a good classification accuracy might be obtained between [50-500] trees.
4- Split criterion ‘$SplitCriterion$' is set to ‘gdi’ that refer to Gini's diversity index (Equ 3-15), while the algorithm used to select the best split variable (predictor) ‘PredictorSelection’ is set to ‘$allsplits$’ to invoke the standard CART that maximises the $SplitCriterion$ gain of all possible splits of all predictors in the training dataset.
5- Learning rate for shrinking ‘$LearnRate$’ is set to equal ‘0.1’ as a low rate tends to employ a large number of trees for each learning cycle (James *et al.*, 2013). $LearnRate$ is used only with ‘Boosting’ and ‘Rusboost’ classification methods.

4.7.2 Sheep DataSets Implementation Scenarios

Sheep DataSets classification results are tested when ensemble methods Bagging, Boosting, and RusBoosting over the FOSW segmentation method are applied over $10, 7, 5\ sec. window$ sizes. Each pre-mentioned scenario was carried out for all four sheep DataSets with ‘All features’, ’20 features ReliefF’, ‘Best features GA’, and ’20 features RF’. Firstly, all 183 extracted features are fed to the classifier. Then the same implementation is performed with only the first 20 ranked features of Relief FS method over $10, 7, 5\ sec. windows$ are chosen. The third trail of implementation is made when the best number of features are achieved by GA for FOSW overall selected window sizes in Section 4.6.3. The final scenario is applied for the first 20 ranked features obtained from RF implementation over $the\ 10, 7, 5\ sec. windows$.

The confusion matrices for all pre-mentioned scenarios are presented in Appendix G in details, which contains the confusion matrices for DataSet1_all, DataSet2_ac, DataSet2_b, and DataSet3_all. Each DataSet has 9 tables, first 3 tables depict the confusion matrix from 5-fold cross-validation over $10, 7, 5\ sec. windows$, respectively. The second 3 tables show confusion matrix resulted from 0.3 hold-out validation over $10, 7, 5\ sec. windows$, respectively. On the other hand, the tables with other metrics extracted from the confusion matrices (refer to Section 3.8.1) that involve accuracy (**Accu**), True Positive Rate; which is named sensitivity or recall

(**TPR**), precision (**Prec**), and F-score values (**F-score**) are provided in the following sub-sections with its informative discussion for DataSet1_all, DataSet2_ac, DataSet2_b, and DataSet3_all, respectively.

4.7.2.1 DataSet1_all Ensemble Test Results and Discussion

The classification results for DataSet1_all are presented in Figure 4-21. The sheep data of DataSet1_all is collected at a 5 Hz sampling rate, while 9 predictors are obtained (3 for each Acc, Gyr, and $Orient$). The total number of features are 183 extracted features.

5-Fold Validation Results: The best performance of ensemble classifiers over the $10, 7, and$ 5 $sec. wind$ow is achieved over the 10 $sec. window$, when the number of records (data-points) in a segment (segment size) ($seg\_size = 50$); however, the performance of both ensemble methods drops over the 7 and 5 $sec. windows$ since the segment's data-points are decreased to be equal to 35 and 25, respectively.

So, over the 10 $sec. window$, the best performance of Bagging is obtained when 81 features by GA are utilised, the overall accuracy reached 80.39%, while the F-score value percentage of predicting sound walking, mild walking, and severely walking sheep are 61.8%, 88.1%, and 74.3% respectively. Similarly, the best performance by applying the Boosting classifier is achieved when only 81 features of GA implementation are used as overall accuracy 81.49% is obtained with F-score proportion (63.3% sound walking, 89.2% mild walking, and 75.3% severely walking). Although the performance of Bagging outperforms the Boosting method, the best F-score for mild walking sheep class is obtained with Boosting. Thus, Boosting might be the recommended ensemble method for detecting the early sign of lameness as the mildly lame class is the class of target compared to sound or severely walking classes.

0.3 Hold-out Validation Results: In the hold-out validation method, 30% of DataSet1_all is kept aside for testing the classifier after training it with the complement of 70% of DataSet1. The results also reveal that the best performance for both ensemble methods is achieved over the 10 $sec. window$. Bagging's best performance is registered at an overall accuracy of 77.64% when 'all features', '20 features ReliefF', and '20 features RF' are trained. However, the best F-score for the mild walking class; which is the class of interest, when '20 features RF' are selected to feed the classifier is (57.8% sound walking, 87.3% mild walking, 66.7% severely walking).

Single Sheep Splitting Validation Results: For both ensemble methods, the best performance is obtained from the 10 $sec.window$. The Bagging ensemble accuracy when all features are trained is 79.62%, and the F-scores for each sound walking, mild walking, and severely walking classes are 62.7%, 88.1%, 71%, respectively. On the other hand, the Boosting ensemble outperforms Bagging, where the best accuracy achieved is 81.54%, with F-score values of (70% sound walking, 88.6% mild walking, 71.2% severely walking classes).

Generally, for the 5Hz sampling rate of sheep DataSet1_all, the best window size is 10 $sec.$ where the $seg\_size = 50$. The best ensemble method is Boosting when 81 features by GA are used in 5-fold validation, while Bagging method with 20 features RF performs better when 0.3 hold-out validation is used. In addition, Boosting produces higher accuracy of 81.54% than Bagging with Single Sheep Splitting validation.

DataSet1_all/ FOSW/ 5-fold		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	80.76	59	69.7	63.9	90.8	86.4	88.5	72	72.4	72.2	74.89	56.3	63	59.5	85	80.6	82.7	71.3	72.3	71.8	71.13	58.9	61.7	60.3	79.6	75.3	77.4	70.6	73.5	72
	20 featueres ReliefF	78.93	57.1	62.7	59.8	88.8	86	87.4	70.9	71.7	71.3	72.64	57.8	57.8	57.8	82.8	79.4	81.1	65.7	71.6	68.5	67.58	56.9	56.5	56.7	74.8	72.3	73.5	67.2	71.3	69.2
	Best features GA	80.39	59	64.8	61.8	89.6	86.6	88.1	73.7	75	74.3	72.38	54.4	58.4	56.3	84.3	79.1	81.6	64.7	69.5	67	70.26	59.3	60.8	60	75.9	75.2	75.5	72.4	71.8	72.1
	20 features RF	80.02	60.3	64.4	62.3	89.8	86.6	88.2	70.3	73.7	72	76.02	55.5	62.4	58.7	87.3	81.8	84.5	71.7	74.4	73	71.6	59.8	61.9	60.8	78.8	75.2	77	72.4	75.2	73.8
Boosting	All features	80.76	59.6	66.9	63	89.6	87.8	88.7	74.9	72	73.4	73.68	56.3	61.2	58.6	83.8	80.3	82	69	70.2	69.6	69.73	58.1	61.4	59.7	78.9	74.2	76.5	67.4	70.5	68.9
	20 features ReliefF	78.44	48.1	63	54.6	90.4	85.2	87.7	72	69.2	70.6	72.38	54.4	58.4	56.3	84.3	79.1	81.6	64.7	69.5	67	66.98	56.2	57.3	56.7	76.9	72.7	74.7	62.9	67.3	65
	Best features GA	81.49	60.3	66.7	63.3	89.6	88.9	89.2	77.7	73.1	75.3	73.42	54.4	60.3	57.2	82.1	80.2	81.1	73	70.2	71.6	68.92	56.5	60.5	58.4	78.5	73.1	75.7	67	70	68.5
	20 features RF	81.24	62.2	65.1	63.6	90.2	88	89.1	73.1	75.3	74.2	74.37	55.5	62.7	58.9	85.1	80.5	82.7	69.7	70.6	70.1	69.32	56.9	61.7	59.2	77.2	73.7	75.4	69.7	69.4	69.5

DataSet1_all/ FOSW/ Hold out 30%		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	77.64	57.4	71.1	63.5	88.4	82.8	85.5	65.4	66.7	66	75.43	53.2	67.7	59.6	87.6	79.1	83.1	71.1	72.7	71.9	74.5	59.5	66.4	62.8	84.2	77.7	80.8	74.8	76.6	75.7
	20 features ReliefF	77.64	59.6	60.9	60.2	88.4	85	86.7	63.5	70.2	66.7	69.94	45.6	58.1	51.1	84.7	76.1	80.2	62.2	64.4	63.3	70.47	59.5	59.1	59.3	77.4	77.4	77.4	71	71.5	71.2
	Best features GA	76.83	51.1	61.5	55.8	87.8	84.3	86	69.2	66.7	67.9	74.57	54.4	59.7	56.9	87	79.8	83.2	67.8	75.3	71.4	75.39	61.9	63.9	62.9	83.7	80.7	82.2	76.3	78.1	77.2
	20 features RF	77.64	55.3	60.5	57.8	89.1	85.6	87.3	65.4	68	66.7	73.99	51.9	62.1	56.5	87.6	78.3	82.7	66.7	73.2	69.8	73.6	57.9	65.2	61.3	81.1	77	79	77.9	75.6	76.7
Boosting	All features	75.2	53.2	61	56.8	85	84.5	84.7	67.3	61.4	64.2	71.1	53.2	59.2	56	83.1	76.6	79.7	63.3	68.7	65.9	71.59	54.8	62.7	58.5	82.6	76.6	79.5	71.8	71.2	71.5
	20 features ReliefF	75.61	48.9	54.8	51.7	89.1	85.6	87.3	61.5	62.7	62.1	67.63	43	56.7	48.9	81.4	75	78.1	62.2	59.6	60.9	69.57	58.7	60.2	59.4	80	76	77.9	64.9	68.5	66.7
	Best features GA	75.61	46.8	64.7	54.3	85.7	85.7	85.7	73.1	58.5	65	74.28	51.9	64.1	57.4	85.3	78.2	81.6	72.2	73	72.6	73.6	55.6	68.6	61.4	85.3	76.4	80.6	74	72.9	73.4
	20 features RF	78.86	53.2	69.4	60.2	89.8	84.6	87.1	71.2	68.5	69.8	72.83	51.9	65.1	57.8	84.7	78.1	81.3	67.8	67	67.4	73.83	55.6	69.3	61.7	84.2	76.2	80	76.3	73.5	74.9

DataSet1_all/ FOSW/ Single Sheep Splitting	10 Sec. window										7 Sec. window										5 Sec. window									
	Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging (All features)	79.62	61.5	64	62.7	90.1	86.2	88.1	67.9	74.5	71	78.12	62.4	68.8	65.4	85.7	82.1	83.9	77.7	77.7	77.7	71.61	60.6	65	62.7	81.6	75.1	78.2	67.9	72.1	69.9
Boosting (All features)	81.54	67.3	72.9	70	92.1	85.4	88.6	66.1	77.1	71.2	74.52	58.8	63.3	61	84.1	80.1	82.1	70.2	72.5	71.3	70.97	60.6	62	61.3	80.1	75.5	77.7	67.9	72.7	70.2

Figure 4-21 SLDM classification results of DataSet1_all for 5-fold, 0.3 hold-out, and Single Sheep Splitting validation methods for both ensemble (Bagging & Boosting).

4.7.2.2 DataSet2_ac Ensemble Test Results and Discussion

The classification results for DataSet2_ac is provided in Figure 4-22. The sheep data of DataSet2_ac is collected at a 10 Hz sampling rate, while 6 predictors are obtained (3 for each Acc and $Orient$). The total number of features are 126 extracted features.

5-Fold Validation Results: The best performance for both ensemble classifiers is achieved over the 10 $sec.window$, where the $seg\_size = 100$; however, when seg_size is equal to 70 and 50 over the 7 and 5 $sec.windows$, respectively the results also reveal satisfactory prediction accuracies. The best performance for Bagging reaches overall accuracy of 88.92% with F-score value (87.7% sound walking, 91.1% mild walking, 88.2% severely walking classes) when 46 features retrieved by GA are fed into the classifier over a 10 $sec.window$. Whereas the best performance for Boosting is obtained with an overall accuracy of 88.79% and F-score (87.5% sound walking, 91.3% mild walking, 87.9% severely walking classes) when 20 features RF are used. Although the accuracy of Boosting performance (89.04%) when all features are trained is slightly higher than the prediction accuracy of 88.79% when only 20 features FR is used, The Boosting with 20 features RF is more considerable as it uses much fewer features than the full 126 features.

0.3 Hold-out Validation Results: The performance of both classifiers is similar to each other and are relatively high. In Bagging, the best accuracy 88.89% is obtained when 20 features RF are fed the classifier with F-score (86.5% sound walking, 93.2% mild walking, 87.5% severely walking classes). Similarly, the best accuracy in Boosting is achieved when 20 features RF are fed into the classifier with F-score (86.7% sound walking, 91.8% mild walking, 86.3% severely walking classes).

Single Sheep Splitting Validation Results: Both classifiers perform well and have an accuracy of 87.9% over the 10 $sec.window$ while the F-scores for the mild walking class for Bagging and Boosting are 91.6% and 92.2%, respectively. The performance of the ensemble classifiers over the 7 $sec.window$ produce a significant accuracy of 83.76% and 87.18% for both Bagging and Boosting. However, higher accuracy results of Bagging 84.02% and Boosting 84.43% over the 5 $sec.window$ are revealed.

Generally, for the 10 Hz sampling rate of sheep DataSet2_ac, where no Gyr readings are retrieved from the deployed sensor, the best window size is $10\ sec.$ where the $seg\_size = 100$; however, the $seg\_size\ = 70$ also achieves satisfactory prediction results. The best ensemble method is Bagging when 46 features by GA are used in 5-fold validation and Bagging method with 20 features RF when 0.3 hold-out validation is used.

DataSet2_ac/ FOSW/ 5-fold		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	**88.05**	85.3	87.8	86.5	92.1	88.4	90.2	87.1	88	87.5	**85.62**	77.9	85.4	81.5	86.8	82.8	84.8	89.7	87.3	88.5	**86.09**	77.6	90	83.3	89.3	81.4	85.2	89.7	86.8	88.2
	20 features ReliefF	**87.56**	83.3	87.1	85.2	94.2	89.7	91.9	85.9	86.2	86	**84.31**	77.4	84.4	80.7	86.8	81.6	84.1	87.2	85.8	86.5	**85.46**	77.2	86.5	81.6	86.4	84	85.2	90.3	85.8	88
	Best features GA	**88.92**	86.1	89.3	87.7	93.4	88.9	91.1	87.8	88.6	88.2	**85.49**	79.3	85.6	82.3	85.7	82.2	83.9	89.1	87.2	88.1	**84.59**	79.7	81.9	80.8	88.3	81.6	84.8	85.6	88.3	86.9
	20 features RF	**85.96**	82.1	85.9	84	90.9	87.3	89.1	85.3	85	85.1	**85.1**	77	84.8	80.7	86.2	80.3	83.1	89.4	87.9	88.6	**86.34**	78.4	87.5	82.7	88.8	82.4	85.5	90	88	89
Boosting	All features	**89.04**	84.9	90.3	87.5	94.2	88.7	91.4	88.4	88.4	88.4	**86.54**	77.9	88.5	82.9	88.9	83.2	86	90.5	87.4	88.9	**87.09**	78.4	88.8	83.3	87.9	86.6	87.2	92.2	86.5	89.3
	20 features ReliefF	**85.96**	82.1	85.9	84	90.9	87.3	89.1	85.3	85	85.1	**84.58**	78.3	85.9	81.9	87.3	80.9	84	86.9	86	86.4	**84.96**	76.7	84.8	80.5	86.4	84	85.2	89.4	85.6	87.5
	Best features GA	**88.67**	86.5	88.5	87.5	92.9	89.6	91.2	87.1	89.7	88.4	**85.1**	75.6	84.1	79.6	88.9	82.8	85.7	88.9	86.9	87.9	**85.96**	78	87.4	82.4	85.4	83.8	84.6	91.4	86.4	88.8
	20 features RF	**88.79**	85.7	89.3	87.5	93.4	89.3	91.3	87.8	88.1	87.9	**87.45**	82.5	87.7	85	88.4	83.1	85.7	90	89.7	89.8	**85.96**	77.2	85.6	81.2	87.9	84.2	86	90.6	87.2	88.9

DataSet2_ac/ FOSW/ Hold out 30%		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	**87.65**	86.7	86.7	86.7	91.8	93.1	92.4	85.3	84.4	84.8	**87.34**	80	89.7	84.6	86	84.5	85.2	92.5	87.6	90	**87.45**	76.8	91.4	83.5	90.3	81.2	85.5	92.6	89.3	90.9
	20 features ReliefF	**87.24**	81.3	87.1	84.1	94.5	90.8	92.6	86.3	84.5	85.4	**86.03**	80	88.1	83.9	87.7	78.1	82.6	88.8	89.6	89.2	**86.61**	78.3	90	83.7	87.1	83.1	85.1	91.7	86.8	89.2
	Best features GA	**88.07**	85.3	87.7	86.5	93.2	90.7	91.9	86.3	86.3	86.3	**86.9**	83.1	88.5	85.7	87.7	82	84.8	88.8	88.8	88.8	**85.36**	73.9	83.6	78.5	91.9	79.2	85.1	88.9	90.6	89.7
	20 features RF	**88.89**	81.3	92.4	86.5	93.2	93.2	93.2	91.6	83.7	87.5	**87.77**	80	88.1	83.9	87.7	87.7	87.7	92.5	87.6	90	**85.36**	71	89.1	79	90.3	78.9	84.2	91.7	87.6	89.6
Boosting	All features	**87.24**	84	86.3	85.1	91.8	91.8	91.8	86.3	84.5	85.4	**85.15**	78.5	85	81.6	89.5	81	85	86.9	87.7	87.3	**87.09**	73.9	91.1	81.6	88.7	83.3	85.9	95.4	88	91.6
	20 features ReliefF	**83.95**	84	80.8	82.4	90.4	88	89.2	78.9	83.3	81	**85.15**	78.5	86.4	82.3	91.2	80	85.2	86	87.6	86.8	**86.19**	78.3	90	83.7	87.1	78.3	82.5	90.7	89.1	89.9
	Best features GA	**87.65**	85.3	85.3	85.3	91.8	94.4	93.1	86.3	84.5	85.4	**87.77**	81.5	89.8	85.4	93	80.3	86.2	88.8	91.3	90	**84.94**	68.1	85.5	75.8	85.5	82.8	84.1	95.4	85.8	90.3
	20 features RF	**88.07**	86.7	86.7	86.7	91.8	91.8	91.8	86.3	86.3	86.3	**86.46**	80	85.2	82.5	86	86	86	90.7	87.4	89	**85.77**	72.5	87.7	79.4	88.7	80.9	84.6	92.6	87.7	90.1

DataSet2_ac/ FOSW/ Single Sheep Splitting	10 Sec. window										7 Sec. window										5 Sec. window									
	Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging (All features)	**87.9**	87	84.8	85.9	95.9	87.7	91.6	82.5	90.9	86.5	**83.76**	71.6	82.8	76.8	87.9	81	84.3	89	85.8	87.4	**84.02**	76.1	88.5	81.8	88.9	76.7	82.4	86.4	86.4	86.4
Boosting (All features)	**87.9**	83.1	85.3	84.2	95.9	88.8	92.2	85.6	89.2	87.4	**87.18**	82.1	88.7	85.3	87.9	83.6	85.7	89.9	88.3	89.1	**84.43**	74.6	93	82.8	87.3	76.4	81.5	89.1	85.2	87.1

Figure 4-22 SLDM classification results of DataSet2_ac for 5-fold, 0.3 hold-out, and Single Sheep Splitting validation methods for both ensemble (Bagging & Boosting).

4.7.2.3 DataSet2_b Ensemble Test Results and Discussion

The classification results of DataSet2_b is presented in Figure 4-23. Similar to DataSet2_ac, DataSet2_b has data collected at a 10 Hz sampling rate. However, 9 predictors are obtained (3 for each Acc, Gyr, and $Orient$). The total number of features are 183 extracted features.

5-Fold Validation Results: The satisfactory accuracies are varied between the $10, 7, and\ 5\ sec. window$ sizes. However, the best performance is spotted for Bagging over the 5 $sec. window$ with an accuracy of 71.99 % and F-scores (81.3% sound walking, 59.8% mild walking, 46.2% severely walking classes) when 20 features RF are used to train the classifier. Similarly, Boosting achieves a higher accuracy of 70.13% over the 5 $sec. window$ with F-scores (79.9% sound walking, 56.2% mild walking, 43% severely walking classes) when 20 features RF are utilised. Due to the performance of the classifier being affected by the size of the training dataset, the 5 $sec. window$ dataset performs better than the 10 $sec. window$ because of the number of records for DataSet2_b in 5 $sec.$ and 10 $sec. window$ are 1021 and 558 records, respectively.

0.3 Hold-out Validation Results: The best accuracy results are obtained over the $5\ sec. window$; however, some better accuracies could be achieved over the $10\ or\ 7\ sec. window$ sizes. The Bagging ensemble achieves the highest accuracy of 74.51% over the 5 $sec. window$ with F-scores (83.4% sound walking, 63.2% mild walking, 50.6% severely walking classes) when 20 features RF are used for training. On the other hand, Boosting classifier over the 10 $sec. window$ reveals the best accuracy of 73.65% and F-scores (82.2% sound walking, 61.3% mild walking, 55.2% severely walking classes) when 64 features of GA.

Single Sheep Splitting Validation Results: Both ensemble classifiers have a better performance over the $10\ sec. window$. The Bagging ensemble provides an accuracy of 70.52%, while Boosting reveals a higher accuracy of 74.57% and F-scores (83.4% sound walking, 64.6% mild walking, 47.1% severely walking classes).

Generally, Bagging performs better than Boosting for both 5-fold and 0.3 hold-out validation methods when only 20 features RF are used for training over the 5 $sec. window$ size, where $seg\_size\ = 50$ and the number of data-points are 1021.

DataSet2_b/ FOSW/ 5-fold		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	67.38	95.4	67.7	79.2	42.1	72.9	53.4	13.4	50	21.1	68.74	96.2	68.3	79.9	36.4	75	49	22.6	64.3	33.4	69.34	94.8	69.9	80.5	35.5	73.6	47.9	31.1	60.4	41.1
	20 features ReliefF	68.64	88.6	72.2	79.6	49.6	63.8	55.8	31.3	53.8	39.6	66	85.6	70.1	77.1	40.5	61.9	49	35.8	49.6	41.6	67.87	89.1	71.2	79.2	42.5	68.3	52.4	32.7	48.5	39.1
	Best features GA	68.1	93.8	68.7	79.3	46.3	74.7	57.2	17	48.7	25.2	67.75	93.2	68.2	78.8	36.4	73.3	48.6	26.4	57.5	36.2	66.8	86.3	71.3	78.1	43.9	59.9	50.7	34.2	50.8	40.9
	20 features RF	68.46	91.4	70.2	79.4	43	67.5	52.5	29.5	56.9	38.9	69.49	90.7	71.8	80.2	39.3	68.7	50	39.6	57.8	47	71.99	92.1	72.7	81.3	50	74.5	59.8	36.2	64	46.2
Boosting	All features	66.85	91.7	69	78.7	44.6	65.9	53.2	18.8	47.7	27	67.62	91.1	69.6	78.9	38.2	66.7	48.6	30.2	54.5	38.9	68.66	93.8	69.2	79.6	34.2	68.4	45.6	32.1	64.3	42.8
	20 features ReliefF	65.59	86.2	70	77.3	44.6	60.7	51.4	28.6	46.4	35.4	65.01	88.3	67.8	76.7	38.7	63.8	48.2	24.5	46.4	32.1	67.38	91.1	69.5	78.8	36.4	69.2	47.7	31.1	51.7	38.8
	Best features GA	66.67	89.5	69.5	78.2	45.5	66.3	54	23.2	46.4	30.9	66.13	94.4	67.7	78.9	33.5	69	45.1	23.9	50	32.3	64.54	94.3	65.1	77	31.1	62.3	41.5	12.8	59.5	21.1
	20 features RF	67.56	88	70.6	78.3	45.5	63.2	52.9	32.1	54.5	40.4	68.12	92.4	69.4	79.3	38.2	66.7	48.6	28.9	59.7	38.9	70.13	91.8	70.7	79.9	44.7	75.6	56.2	33.7	59.5	43

DataSet2_b/ FOSW/ Hold out 30%		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	67.07	97.9	66.4	79.1	32.4	75	45.3	15.2	62.5	24.5	65.83	91.5	67.2	77.5	30.8	59.3	40.5	27.7	61.9	38.3	71.57	98.9	70	82	42	80.6	55.2	22.4	76.5	34.7
	20 features ReliefF	68.26	87.6	72.6	79.4	43.2	64	51.6	39.4	52	44.8	64.17	83	70.1	76	40.4	52.5	45.7	34	48.5	40	71.57	89.9	73.9	81.1	49.3	81	61.3	41.4	52.2	46.2
	Best features GA	67.66	95.9	70.5	81.3	35.1	61.9	44.8	21.2	50	29.8	66.67	91.5	68.3	78.2	34.6	56.3	42.9	27.7	68.4	39.4	64.71	77.7	72	74.7	47.8	58.9	52.8	44.8	45.6	45.2
	20 features RF	68.86	90.7	71.5	80	43.2	66.7	52.4	33.3	55	41.5	66.67	86.5	70.5	77.7	34.6	62.1	44.4	42.6	52.6	47.1	74.51	93.9	75	83.4	52.2	80	63.2	41.4	64.9	50.6
Boosting	All features	64.67	88.7	67.2	76.5	37.8	60.9	46.6	24.2	50	32.6	67.92	94.3	68.6	79.4	30.8	72.7	43.3	29.8	58.3	39.4	69.28	97.2	67.4	79.6	34.8	82.8	49	24.1	73.7	36.3
	20 features ReliefF	69.46	88.7	72.9	80	45.9	73.9	56.6	39.4	50	44.1	64.58	86.5	69.7	77.2	38.5	51.3	44	27.7	50	35.6	69.28	93.9	69.7	80	39.1	75	51.4	29.3	58.6	39.1
	Best features GA	73.65	90.7	75.2	82.2	51.4	76	61.3	48.5	64	55.2	61.25	87.2	64.1	73.9	34.6	52.9	41.8	12.8	42.9	19.7	66.99	96.1	66.2	78.4	34.8	72.7	47.1	15.5	69.2	25.3
	20 features RF	67.66	88.7	71.1	78.9	45.9	70.8	55.7	30.3	45.5	36.4	65.83	87.9	68.9	77.2	34.6	60	43.9	34	53.3	41.5	72.55	96.1	71.4	81.9	42	78.4	54.7	36.2	75	48.8

DataSet2_b/ FOSW/ Single Sheep Splitting	10 Sec. window										7 Sec. window										5 Sec. window									
	Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging (All features)	70.52	100	68.2	81.1	50	86.4	63.3	5.9	66.7	10.8	62.8	97.9	64.7	77.9	20.4	57.9	30.2	6	30	10	67.94	95.6	66.8	78.6	35.2	73.5	47.6	23	73.7	35.1
Boosting (All features)	74.57	95	74.4	83.4	55.3	77.8	64.6	35.3	70.6	47.1	63.2	91.1	67.9	77.8	25.9	50	34.1	22	42.3	28.9	66.98	94.5	67.3	78.6	32.4	74.2	45.1	24.6	55.6	34.1

Figure 4-23 SLDM classification results of DataSet2_b for 5-fold, 0.3 hold-out, and Single Sheep Splitting validation methods for both ensemble (Bagging & Boosting).

4.7.2.4 DataSet3_all Ensemble Test Results and Discussion

The classification results of DataSet3_all are displayed in Figure 4-24. The sheep data of gathered data at 4 Hz sampling rate, while 9 predictors are obtained (3 for each Acc, Gyr, and $Orient$). The total number of features are 183 extracted features. Due to the smaller number of records 64, 134 and 212 for $10, 7, and\ 5\ sec. window$ sizes, respectively, the implementation of RusBoosting algorithm is considered instead of Boosting. Thus, the majority class is randomly under-sampled while the proportion of the class is kept the same within the dataset.

5-Fold Validation Results: most of the high accuracy results are obtained over the $5\ sec. window$. However, many significant results could be obtained over the $10\ and\ 7\ sec. windows$. The best performance for Bagging ensemble is achieved when 20 features ReliefF are used for training the classifier over a $10\ sec. window$, where the accuracy is 73.44% and F-scores (80.5% sound walking, 72% mild walking, 58.1% severely walking classes). On the other hand, the best performance for the RusBoosting classifier is obtained over the $5\ sec. window$ with an accuracy of 71.23% and F-scores (76.5% sound walking, 70.5% mild walking, 60.8% severely walking classes) when 20 features ReliefF are fed to the ensemble.

0.3 Hold-out Validation Results: The best performance for Bagging ensemble is achieved with an accuracy of 74.6% and F-scores (82.5% sound walking, 60% mild walking, 61.5% severely walking classes) over the $5\ sec. window$ when 38 features by GA are used for training. The best RusBoosting accuracy is 68.25% with F-scores (77.6% sound walking, 57.1% mild walking, 58.1% severely walking classes) also over the $5\ sec. window$ when 20 features RF are used by the classifier.

Single Sheep Splitting Validation Results: Both ensembles perform well over the $5\ sec. window$ when all features are used. In Bagging, the accuracy reached 74.24% and F-scores (82% sound walking, 62.1% mild walking, 64% severely walking classes), while RusBoosting achieved an accuracy of 71.21% and F-scores (78.8% sound walking, 61.1% mild walking, 66.7% severely walking classes)

Generally, the best accuracy results could be obtained for both ensemble classifiers over the $5\ sec. window$ as more data-points could be dealt with (212) compared to the 7 and

10 $sec. windows$ where 134 and 64 data-points are allocated. In 5-fold validation, Bagging outperforms RusBoosting when 20 features RF are employed. Alternatively, in 0.3 hold-out methods, the best performance is achieved when 38 features by GA are used for training the classifiers.

DataSet3_all/ FOSW/ 5-fold		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	73.44	97.1	70.2	81.5	50	75	60	44.4	88.9	59.2	63.43	84.3	67	74.7	41.9	61.9	50	39.4	52	44.8	69.34	90.2	69.2	78.3	47.9	63.9	54.8	44.2	76.7	56.1
	20 features ReliefF	73.44	85.3	76.3	80.5	75	69.2	72	50	69.2	58.1	61.94	75.7	71.6	73.6	48.4	51.7	50	45.5	48.4	46.9	69.34	80.4	73.8	77	64.6	64.6	64.6	50	61.9	55.3
	Best features GA	73.44	88.2	73.2	80	66.7	72.7	69.6	50	75	60	63.43	85.7	70.6	77.4	41.9	54.2	47.3	36.4	48	41.4	66.51	86.6	69.3	77	37.5	52.9	43.9	50	68.4	57.8
	20 features RF	71.88	85.3	74.4	79.5	58.3	70	63.6	55.6	66.7	60.6	67.91	80	70.9	75.2	54.8	73.9	62.9	54.5	56.3	55.4	73.11	88.4	76.2	81.8	60.4	61.7	61	51.9	77.1	62
RusBoost	All features	59.38	61.8	75	67.8	75	52.9	62	44.4	42.4	43.4	64.18	70	79	74.2	61.3	55.9	58.5	54.5	47.4	50.7	69.34	71.4	80	75.5	62.5	54.5	58.2	71.2	64.9	67.9
	20 features ReliefF	62.5	73.5	75.8	74.6	58.3	53.8	56	44.4	44.4	44.4	62.69	70	80.3	74.8	61.3	51.4	55.9	48.5	44.4	46.4	71.23	74.1	79	76.5	77.1	64.9	70.5	59.6	62	60.8
	Best features GA	56.25	55.9	73.1	63.4	58.3	46.7	51.9	55.6	43.5	48.8	65.67	74.3	85.2	79.4	54.8	48.6	51.5	57.6	50	53.5	67.45	69.6	83	75.7	66.7	52.5	58.8	63.5	57.9	60.6
	20 features RF	60.94	64.7	78.6	71	50	60	54.5	61.1	42.3	50	67.16	68.6	81.4	74.5	58.1	62.1	60	72.7	52.2	60.8	66.51	70.5	76	73.1	66.7	56.1	60.9	57.7	58.8	58.2

DataSet3_all/ FOSW/ Hold out 30%		10 Sec. window										7 Sec. window										5 Sec. window									
		Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
			TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging	All features	52.63	90	64.3	75	33.3	25	28.6				75	85.7	78.3	81.8	66.7	66.7	66.7	60	75	66.7	69.84	97	68.1	80	35.7	62.5	45.4	43.8	87.5	58.4
	20 features ReliefF	68.42	100	71.4	83.3	66.7	50	57.2	16.7	100	28.6	65	71.4	83.3	76.9	55.6	50	52.7	60	50	54.5	65.08	93.9	68.9	79.5	35.7	50	41.7	31.3	62.5	41.7
	Best features GA	68.42	90	69.2	78.2	66.7	50	57.2	33.3	100	50	67.5	81	73.9	77.3	66.7	60	63.2	40	57.1	47	74.6	100	70.2	82.5	42.9	100	60	50	80	61.5
	20 features RF	68.42	100	66.7	80	66.7	66.7	66.7	16.7	100	28.6	72.5	85.7	78.3	81.8	55.6	62.5	58.8	60	66.7	63.2	68.25	93.9	70.5	80.5	35.7	50	41.7	43.8	77.8	56
RusBoost	All features	63.16	60	100	75	100	33.3	50	50	75	60	70	81	81	81	66.7	60	63.2	50	55.6	52.7	66.67	75.8	75.8	75.8	64.3	64.3	64.3	50	50	50
	20 features ReliefF	63.16	80	72.7	76.2	66.7	40	50	33.3	66.7	44.4	60	57.1	80	66.6	55.6	45.5	50	70	50	58.3	57.14	69.7	71.9	70.8	35.7	41.7	38.5	50	42.1	45.7
	Best features GA	63.16	70	77.8	73.7	100	42.9	60	33.3	66.7	44.4	57.5	57.1	85.7	68.5	66.7	54.5	60	50	33.3	40	60.32	63.6	70	66.6	57.1	57.1	57.1	56.3	47.4	51.5
	20 features RF	68.42	90	81.8	85.7	66.7	40	50	33.3	66.7	44.4	67.5	76.2	76.2	76.2	55.6	55.6	55.6	60	60	60	68.25	78.8	76.5	77.6	57.1	57.1	57.1	56.3	60	58.1

DataSet3_all/ FOSW/ Single Sheep Splitting	10 Sec. window										7 Sec. window										5 Sec. window									
	Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe			Accu	TPR sound			TPR mild			TPR severe		
		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore		TPR	Prec	Fscore	TPR	Prec	Fscore	TPR	Prec	Fscore
Bagging (All features)	63.64	83.3	62.5	71.4	50	100	66.7	33.3	50	40	61.36	86.4	61.3	71.7	27.3	75	40	45.5	55.6	50	74.24	91.4	74.4	82	60	64.3	62.1	50	88.9	64
RusBoost (All features)	63.64	41.7	83.3	55.6	75	75	75	100	50	66.7	54.55	54.5	66.7	60	54.5	50	52.2	54.5	42.9	48	71.21	74.3	83.9	78.8	73.3	52.4	61.1	62.5	71.4	66.7

Figure 4-24 SLDM classification results of DataSet3_all for 5-fold, 0.3 hold-out, and Single Sheep Splitting validation methods for both ensemble (Bagging & Boosting).

4.7.3 Comparison of SLDM Validation Methods

Three validation approaches are applied, 5-fold, 0.3 hold-out, and a proposed method 'Single Sheep Splitting'. Regarding the 'Single Sheep Splitting', it is only applied when the whole set of features are fed to the classifier. However, when the number of features that were fed to the classifiers was varied, only 5-fold and 0.3 hold-out validation methods were tried; so, the comparison results of the two validation methods, in this case, are explored in Appendix H. While the performance of the ensemble classifiers for all three validation methods when all features fed to the ensemble classifiers are plotted together in Figure 4-25, Figure 4-26, Figure 4-27, and Figure 4-28 over $10, 7, 5\ sec.windows$ for DataSet1_all, DataSet2_ac, DataSet2_b, and DataSet3_all, respectively.

The presented figures show a variance in ensemble performance among validation methods. The highest classification performance is registered when Single Sheep Splitting was applied over $10\ sec.window$ for DataSet1_all (Figure 4-25), DataSet2_b (Figure 4-27), and DataSet3_all (Figure 4-28) with convergence with other validation methods. Conversely, DataSet2_ac (Figure 4-26) has the highest performance when 5-fold validation was applied; however, the other validation methods have satisfactory performance as well.

In General, the performance of 5-fold validation is higher than 0.3 hold-out or Single Sheep Splitting as in the 5-fold method, the data are split into 5 folds, where the model keeps one fold for testing and utilises the other 4 folds for training, and the final accuracy is the average of the repeated process for 5 times. Although 5- fold validation is suitable for small datasets, it may overfit the trained model as the model might be tested for the same data-points that are previously used to build the model.

Instead, the 0.3 hold-out validation method is tested to ensure that 30% of the observations are not seen by the trained classifier with 70% of data to avoid model overfitting. Alternatively, 'Single Sheep Splitting' is proposed for validation, where 30% of each sheep walking movements are kept for testing, while 70% of the sheep movements feed the trained model. This method guarantees that the test data includes movement from each individual sheep rather than a proportion of the whole population's movements. Therefore, the presented performance results show a satisfactory performance for 'Single Sheep Splitting' compared to 5-fold and 0.3 hold-out validation approaches.

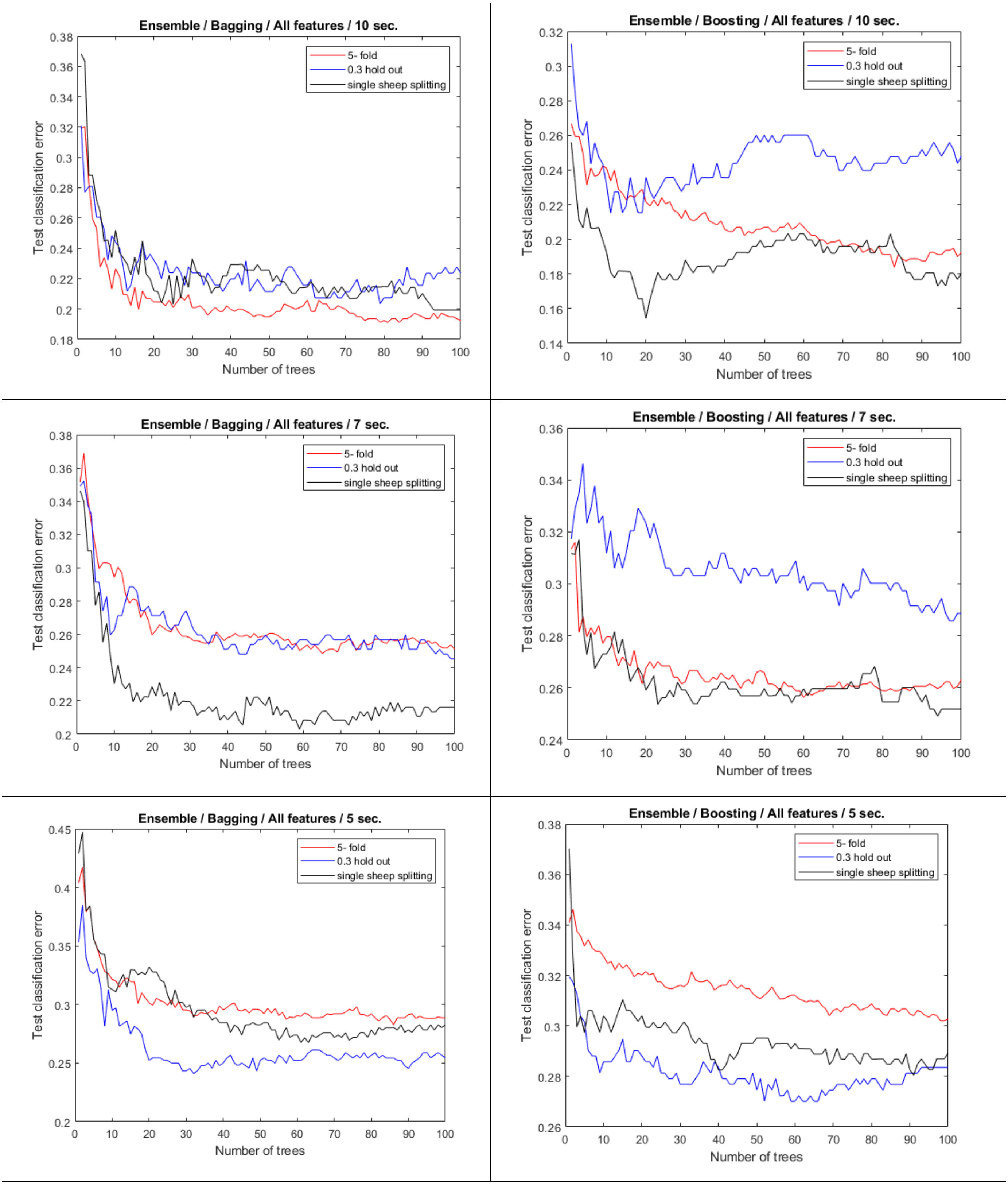


Figure 4-25 Validation techniques comparison for Ensemble (Bag & Boost) classifiers for **DataSet1_all** (all features), FOSW segmentation method over $\mathbf{10, 7, 5}\ sec.window$.

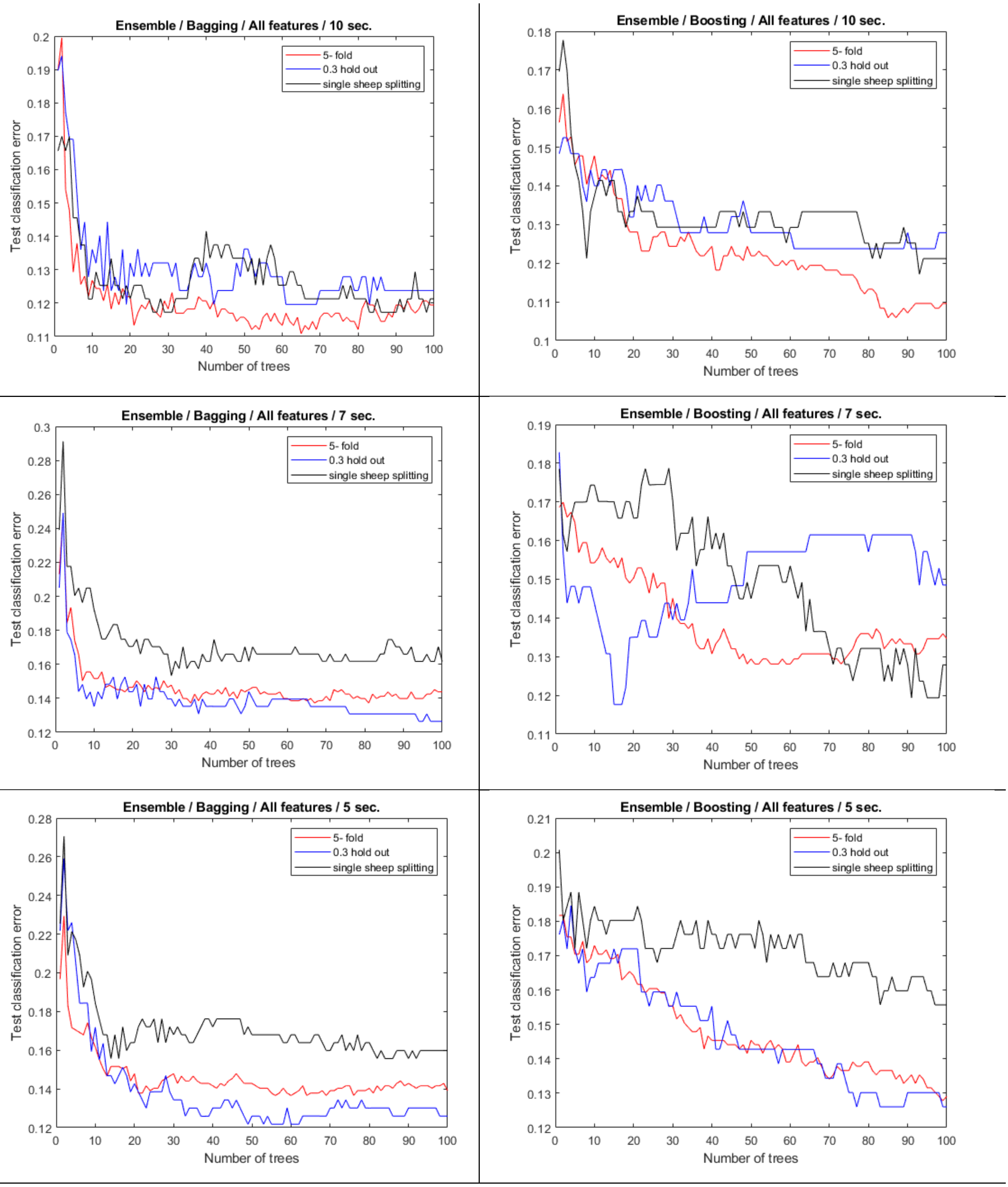


Figure 4-26 Validation techniques comparison for Ensemble (Bag & Boost) classifiers for **DataSet2_ac** (all features), FOSW segmentation method over $\mathbf{10, 7, 5}$ $sec.window$.

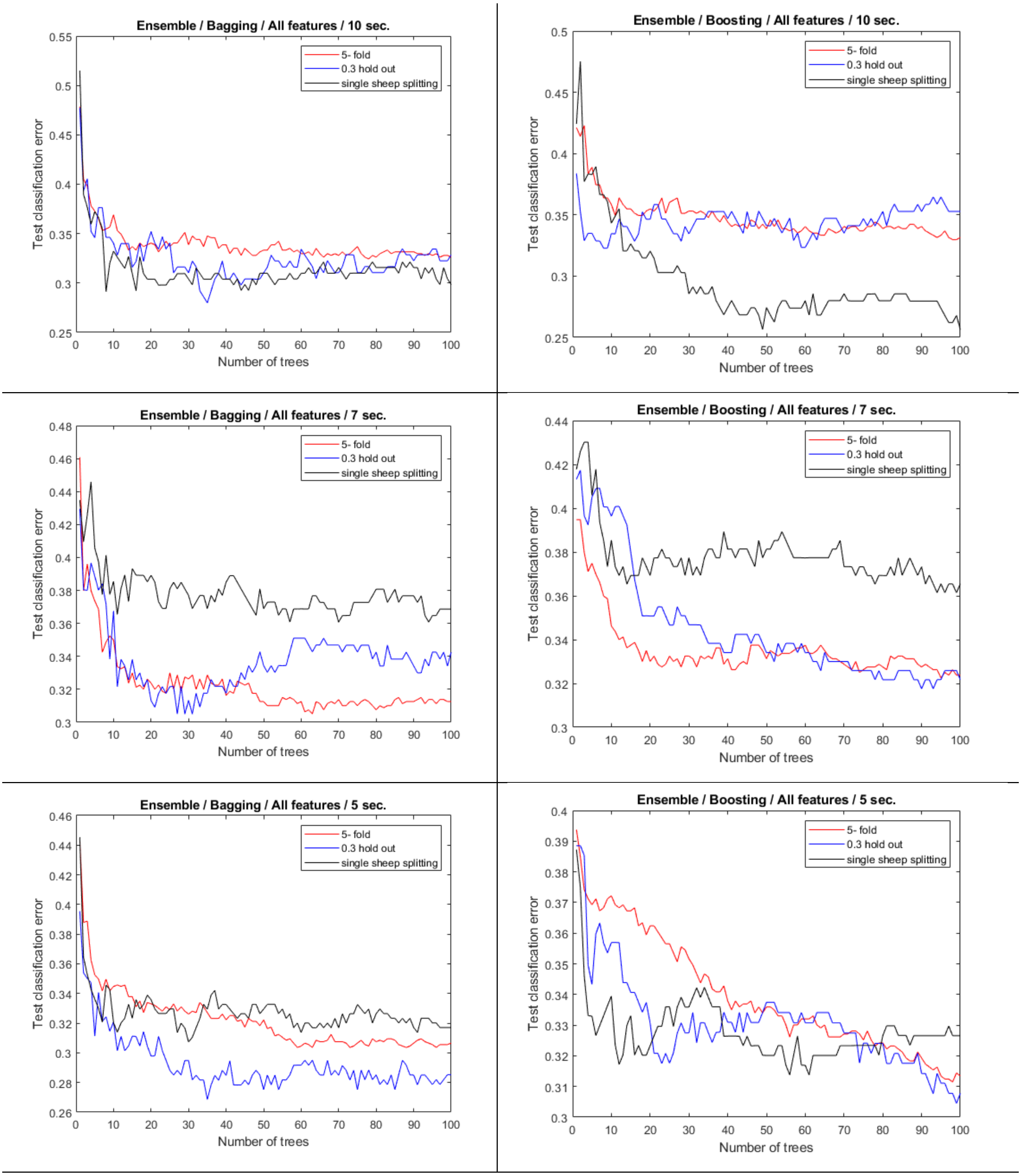


Figure 4-27 Validation techniques comparison for Ensemble (Bag & Boost) classifiers for **DataSet2_b** (all features), FOSW segmentation method over $\mathbf{10, 7, 5}$ $sec. window$.

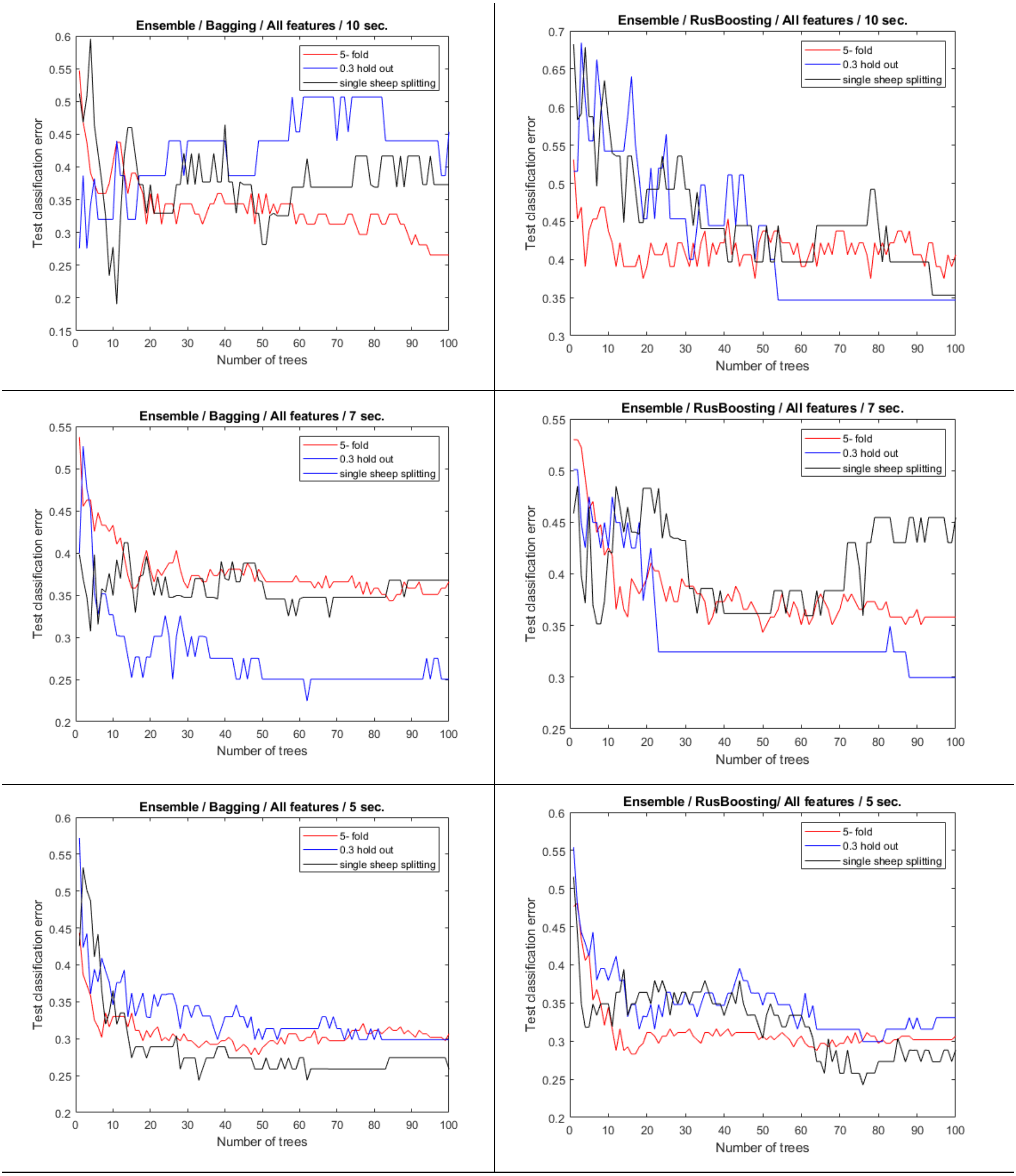


Figure 4-28 Validation techniques comparison for Ensemble (Bag & RusBoosting) classifiers for **DataSet3_all** (all features), FOSW segmentation method over $\mathbf{10, 7, 5}\ sec.window$.

4.8 General Discussion and Comparisons

The proposed approach for SLDM in the current thesis investigates sheep data collection at different sampling rates, pre-processing techniques (2 segmentation method, 3 different window sizes segments), feature extraction, feature selection methods, and 3 classification algorithms for the sake of identifying the most cost effective factors that contribute to detect the early signs of lameness in sheep. To the best of author's knowledge, there is no thorough data mining approach has been proposed yet to assist the classification of sheep lameness problem, and this is also suggested by Vazquez Diosdado *et al.*, (2018). Therefore, this thesis evaluates the effect of several pre-processing methods on the performance of ensemble classifiers (Bagging, Boosting or RusBoosting) for the indication of lameness in sheep when compared with three sampling rates (10, 5, 4 Hz), two segmentation approaches (FNSW and FOSW), three feature selection methods (ReliefF, GA, and RF) and three window sizes (10, 7, 5 sec.). Finally, the ensemble classifiers are evaluated using three different methods (5-fold, 0.3 hold-out, and the proposed one 'Single Sheep Splitting'). Approximately 432 combinations from prementioned options were conducted to identify the best combination.

The validated prediction accuracies from applying the proposed SLDM reveal promising results for most of the combination. The thesis' findings would be beneficial to be recommended in developing a unique sensor with its complete tool kit to detect sheep lameness on-farm basis in the near future; especially in the time of developing of PLF to accelerate the sheep industry in the UK. The best accuracy of 88.92% with F-score of 87.7%, 91.1%, 88.2% for sound walking, mildly walking, and severely walking classes, respectively, is obtained when applying the Bagging ensemble with the 5-fold validation method over the $10\ sec.window$ for a dataset collected at a 10 Hz. sampling rate ($seg\_size = 100$), with only accelerometer hardware sensor being activated, while the orientation sensing data were calculated from accelerometer readings. The number of features that feed the classifier was reduced to 46 features selected by GA feature selection method.

The significant findings achieved in this thesis are compared to what have been examined in a few existing research studies in term of 12 criteria as shown in Table 4-13 to highlight the novelty of the current thesis.

Table 4-13 Comparison of current research with other sheep lameness studies (the blue colour font refers to the setting that the highest accuracy is achieved.

Reference\ Criteria		(Barwick *et al.*, 2018b)	(Vazquez Diosdado et al., 2018)	(Al-Rubaye et al., 2018)	Current work
1	**Sensor Type**	Accelerometer	Accelerometer & Gyroscope	Accelerometer, linear accelerometer, Orientation	Accelerometer, Gyroscope, and Orientation
2	**Sampling rate Hz.**	12 Hz	8, 16, 32 Hz	10 Hz	10, 5, 4 Hz.
3	**#Sheep**	10 sheep (5 tested, 5 companion sheep)	19 sheep	7 sheep	66 sheep
4	**Sensor location**	Neck, leg, ear	Neck, ear	Neck	Neck
5	**Window sizes**	10 sec.	7 sec.	NA	10, 7, 5 sec.
6	**#Extractd features**	14	44 (Mansbridge et al., 2018)	NA	183, or 122
7	**FS algorithm**	RF (Gini index)	RelieF (Mansbridge et al., 2018)	NA	GA, RF and RelieF
8	**#Selected features**	3	10	NA	46, (20 RF, 20 RelieF)
9	**Classifier used**	QDA	RF	Decision tree / single CART	Bagging of CART, Boosting, or RusBoosting
10	**Validation method**	leave-one-out	10-fold	3 unseen sheep data	5-fold, 0.3 hold-out, Single Sheep Splitting
11	**Accuracy, sensitivity(recall), precision**	Collar attached **83**% Accu, 35% recall, 35% precision for lame walking	**68.6**% Accu, 78.3% recall, 67.8% precision	**75.45**% Accu, 82.87% recall, 60.5% precision for severly lame	**88.92**% Accu, 93.4% recall, 88.9% precision for mild walking
12	**# Classes**	sound grazing, sound standing, sound walking, sound lying, and the lame walking	Lame vs non-lame	Sound, mildly lame, and severely lame	Sound walking, mildly walking, severely walking
Notes		Population classification	Individual classification	Population classification	Population classification

Regarding the first criteria ‘Sensor Type’, it is recommended in this thesis that only the accelerometer hardware sensor is adequate to detect lameness in sheep. In contrast to Vazquez Diosdado et al., (2018) who utilise two hardware sensors for the sake of lameness detection. In order to meet the thesis’ aims to determine the most cost effective factor for lameness detection, only one hardware sensor is preferable to develop less energy consumption sensor. Moreover, Barwick et al., (2018b) also utilise the only accelerometer for lameness detection in their research.

Secondly, the **'Sampling frequencies'** affect the classifier performance and sensor energy consumption. In the current research, DataSet2 that collected at 10 Hz sampling rate produces better accuracies compared to DataSets that have been collected at 5 or 4 Hz. Although the sensor energy could be saved when the sampling frequencies are decreased, the continuity of information about behavioural movement might be lost at small sampling rates. In the 10 Hz sensor setting, 10 sensor readings are transmitted every second, while in 4 Hz only 4 readings are retrieved every second. It is recommended that the 10 Hz sampling rate is preferred for the sheep lameness sensor development studies as the prediction accuracies are increased by 8% when 10 Hz is used instead of 5 Hz sampling frequencies. Additionally, Walton *et al.*, (2018) and then used by Vazquez Diosdado *et al.*, (2018) recommend 16 Hz sampling rate among 8 Hz, 16 Hz., and 32 Hz that were tried in their study to identify the most saving energy sampling rate in comparison to the accuracy of lameness detection. Generally, a compromisation has to be made between the cost of a sensor to be developed and the satisfying accuracy of detection within the flock. The more sampling rate reveals more movement information which could waste the sensor energy, while the smaller sampling rates produce satisfactory accuracies as recommended in this thesis. Therefore, investigating the more suitable frequency for sheep movement collection was one of the aims of this thesis to be met.

The third criterion is the **'Number of sheep'** used, which influences the research findings in term of validity and robustness. The selective sample for conducting research is considered representative whenever the number of participants is satisfactory. So, an adequate sample of 66 sheep are participated in the current thesis within a different group of characteristics compared to few numbers of participants sheep of 10, 19, and 7 in (Barwick *et al.*, 2018b; Vazquez Diosdado *et al.*, 2018; Al-Rubaye *et al.*, 2018), respectively.

Although the choice of **'Sensor location'** around the neck is clarified in Section 2.3.1.2.2, it could be summarised as neck mounted sensor is preferable than leg sensor because it is easy to attach, less likely to cause injuries, and less disruption to animals. Furthermore, the neck attached sensor would be a potential for a WSN node for future research studies to enhance PLF in future. However, Barwick *et al.*, (2018b) utilise different sensor location like ear, neck, and leg with prediction accuracy of 82%, 35%, 87%, respectively. In their research, there were no actual lame sheep, instead, the simulation for lameness movement is depended by bending the sheep leg with an adhesive bandage to obtain lame sheep movements like an actual one. In contrast, the current thesis tests an actual lame sheep with different level of lameness (mild and severe) at Lodge farm in Moulton College, which increases the validity of the research compared to other studies.

The fifth criterion to be discussed in relation to other research studies is **'Window size'**. Usually, the more extended window size contains more relevant information that might positively affect the performance of classifiers. The best performance of ensemble classifies is archived over 10 $sec.window$ at 10 Hz. So, 100 data-points is recommended to detect early lameness signs. Moreover, the performance of Bagging and Boosting is relatively significant for most of the selected window sizes 7 $sec.$ and 5 $sec.$ So, the obtained results come in line with (Walton *et al.*, 2018; Vazquez Diosdado *et al.*, 2018), who suggested 7 $sec.window$ for the best sensor energy consumption. As 112 data-points is maintained to detect lameness movement of sheep (7 window size × 16 Hz. = 112 data-points). While Barwick *et al.*, (2018b) used 10 $sec\ window$ size with 12 Hz. sampling rate, which means 120 data-points are dealt to detect lameness at a time. Generally, the recommended data-points (the number of observations) is around 100 which converge with the existing research studies.

The sixth criterion to be investigated in the current thesis was **'Extracted features'**, the number of features to be extracted from the sheep raw movement data. The number of extract feature for the current thesis were either 183 for the DataSets with 9 predictors (raw sensors readings) or 122 extracted features with DataSet2_ac as it has 6 predictors. Although a less number of features were extracted in Mansbridge et al., (2018) and Barwick *et al.*, (2018b), 14 and 44 respectively, the 183 features were extracted here covered a wide range of time domain and frequency domain features that were employed in cattle behaviour studies (Rahman *et al.*, 2018; Smith *et al.*, 2015), sheep behaviour studies in (Marais *et al.*, 2014; Alvarenga *et al.*, 2016; Kamminga *et al.*, 2017; Barwick *et al.*, 2018a; Walton *et al.*, 2018; Guo *et al.*, 2018; Kleanthous *et al.*, 2018), and also human (Figo *et al.*, 2010; Bersch *et al.*, 2014) research studies. It is important to explain that this stage would not be reconsidered to be developed a cost-effective sensor because the best selected features are examined in the current thesis with various 432 scenarios and the recommendation are given for future enhancement studies.

Regarding the seventh criterion in Table 4-13, the performance of three '**Feature Selection'** methods achieves satisfactory results; however, RF could be considered for sheep lameness detection approaches as it has less computational time compared to GA and better accuracy compared to ReliefF. In addition to the low computational cost of RF, it could be feasible to be deployed in a sensor device to detect lameness (Vazquez Diosdado *et al.*, 2018). The RF is also utilised for feature selection in Barwick *et al.*, (2018b); however, the metric used in current research for ranking features depends on minimising p-values of Chi test, in contrast to Barwick *et al.*, (2018b) who identifies the importance of the features according to its minimum Gini values. So, to meet the aims of investigating of the most saving energy factors for the intended sensor to be developed, RF is the best recommended methods for the further research study in the field of lameness detection in sheep.

As a result of the FS step, the final set of **'Selected features'**; the eighth criterion in Table 4-13, are recommended. The set of extracted features were reduced to 20 features that only used by the classifier to detect the lame sheep. Most of the selected features relate to acceleration and orientation group, which were discussed in Section 4.6.2. That means after the sheep was spotted as a walking sheep by testing its speed, the orientation features like $Pitch$ and $Roll$ angels of the rotated neck are most likely to contribute in early lameness detect process. As the head nodding of the sheep could produce a significant indicator for lameness detection, and this was achieved by the thesis findings to fulfil its aims. Furthermore, 46 features that were selected by GA also revealed a competitive accuracy result by the classifier compared to the two other FS methods; however, it has a high computational cost that would be utilised with large datasets collected from different sources on a cloud. On the other hand, the other two competing studies used a smaller number of features 3 by Barwick *et al.*, (2018b), and 10 by Vazquez Diosdado *et al.*, (2018). So, further insights would be needed for future studies to reduce the number of features used by the classifiers,

For the ninth criterion, the **'Ensemble Classifier'** was used as a classification method in this thesis, which revealed significant results that outperform Vazquez Diosdado *et al.*, (2018), who used RF of 8 trees; which is also considered as a form of an ensemble of trees, where each tree was tested separately and the average accuracies from all tress (classifiers) are considered. Whereas the ensemble of 100 trees was applied in SLDM to avoid overfitting and taking into account all possible combination of the collected observation. The ensemble of trees has less computational cost compared to other classification methods (see Table 2-8); so, it would be possible to be implemented into a developed sensor kit for lameness detection in future. Furthermore, various ML classification methods were applied in (Al-Rubaye *et al.*, 2018) to conclude that DT and an ensemble of it could provide the best accuracy results compared to other ML classification methods. For further enhancement in future, more complex ML method would be applied for sake of sheep lameness detection; for example, Deep learning approach; however, it may cost extra sensor energy, memory, and prediction time.

The tenth criterion in Table 4-13 to be compared to other research studies is the **'Validation method'**, which was applied to validate the proposed SLDM. The main idea of validation is to keep part of the observed data named 'test data' to validate the performance of classifier which is trained with the other part of the data named 'train data'. It is worthless to validate the quality of the classifier with the same train data, so an amount of test data are kept for this purpose. The 5-fold cross-validation outperforms the other two methods used in this thesis as the cross-validation method is suitable when the amount of prepared data for pre-processing step were not too huge. For example, DataSet2_ac has 124806 data points, the largest data points among the other sheep datasets (see Table 4-1). These findings come in line with Vázquez-Diosdado *et al.*, (2019) who recommend 10-fold cross-validation with further advice to explore larger datasets. While Barwick *et al.*, (2018b) validate their model using leave-one-out which is one form of 1-fold validation as the processed data point were also not too huge approximately 432,000 data points (5 sheep were observed for 2 hours at 12 Hz. i.e. (5 sheep × 2hr. observation time × 1200 sec. × 12 Hz= 432,000).

The eleventh criterion in Table 4-13 is the **'Evaluation metrics'** including accuracy of the SLDM to detect the lame sheep in addition to sensitivity and precision. Because that SLDM could detect three classes of sheep status (sound, mildly, and severely walking), the overall accuracy of SLDM; which cover the three classes, may not reflect the desired percentage of early signs of lameness detection in sheep as the target was to detect the lame sheep in its early stage (mildly lame walking class). Therefore, the sensitivity calculates the percentage of certain classes (mildly lame for example) in the examined dataset that is correctly classified by the expert shepherd. While the precision refers to the percentage of certain classes that are correctly classified by SLDM. The current thesis' results achieved a detection accuracy of 88.92% (93.4% recall, 88.9% precision for mild walking), while an accuracy of 68.6% (78.3% recall, 67.8% precision) obtained by Vázquez-Diosdado et al., (2019), and 83% Accuracy (35% recall, 35% precision for lame walking) was reached by Barwick *et al.*, (2018b) in collar attached trail in their research. Although the observations for each research study may differ in various setting, the results of this thesis outperform the other research studies which might lack for a full data mining approach like the one proposed in the current thesis (SLDM).

The twelfth criterion in Table 4-13 is the **'Number of Classes'**. Three levels of Classes could be recognised by SLDM in the current thesis (sound, mildly walking, severely walking), while only lame vs non-lame Classes were spotted by Vázquez-Diosdado et al., (2019). If the mildly walking sheep could be observed earlier, that would be beneficial for the farmers to reduce the cost of treatment in laboratories and prevent the other sheep in the flock from being infected. On the other hand, Barwick *et al.*, (2018b) identified the lame walking sheep (with no indication of lameness level) among other Classes of standing, grazing, lying, and sound walking.

Ultimately, it is worth mentioning that sheep research studies in the literature focus on investigating sheep behaviour on pasture for grazing or ruminating research purposes (Marais *et al.*, 2014; Alvarenga *et al.*, 2016; Giovanetti *et al.*, 2017; Guo *et al.*, 2018; Kleanthous *et al.*, 2018; Vázquez-Diosdado *et al.*, 2019; Kleanthous *et al.*, 2019). Furthermore, the field of knowledge in sheep lameness studies that employ a data mining approach in combination with a mounted motion sensor lack for evaluated studies to fill the gap in the literature. Therefore, the importance of applying the recommendation of this thesis would increase the productivity of the sheep industry in farms and positively contribute to PLF. As sheep welfare would be under control when a sensor is mounted on their neck to produce alarms about their health. Furthermore, the practice of this study would help the shepherd to remotely identify the mildly lame sheep in a farm as sheep are more difficult to monitor and more likely to left in fields for grazing with no need for continuous monitoring than cattle which are already have a daily milking routine compared to sheep.

To the best of the author's knowledge, only two recent studies (Barwick *et al.*, 2018b) and (Vazquez Diosdado *et al.*, 2018) utilise motion sensor technology for predicting lameness in sheep in addition to the earlier published work of the current study (Al-Rubaye *et al.*, 2018), which is presented in Appendix I. 11. Therefore, a comparison of the current work with other related studies is listed in Table 4-13 and discussed earlier showing that the best accuracy obtained by the ensemble model (Bagging) 88.92% outperforms prediction accuracies of other current studies according to the pre-listed recommendation.

5 Chapter Five: Conclusion and Future Work

5.1 Introduction

Sensor technologies play a vital role in developing Precision Livestock Farming (PLF) and application of smart farms, as large amounts of information could be collected and analysed to be later used to enhance the overall farm productivity (Shalloo *et al.*, 2018; Bahlo *et al.*, 2019). So, this research utilises motion sensors like Accelerometer, Gyroscope, and Orientation sensors to collect accelerations, angular velocities, and angles readings, respectively, from a mounted sensor around a sheep neck in three-dimensions (vertical, horizontal, and orthogonal) at three different sampling rates (10, 5, 4 Hz). The collected data has been pre-processed, targeted walking data have been extracted, and a classification prediction model has been built to recognise three sheep walking statuses; these are sound walking, the mildly lame walking, and severely lame walking. The built model has been validated to prove its ability to predict new unseen sheep data in the future to predict their class.

Lameness is one of the major concerns in the sheep industry in the UK and is mostly caused by infectious bacteria growing in muddy soil. These bacteria easily transmit to the sheep's foot, causing footrot and results in abnormal walking, and leads in its worst cases to sheep culling if it is not treated early enough. Due to the scale of the estimated annual losses of £10 (Brian, 2016) that reduce farm productivity, the early detection of lame sheep contributes to reducing labour and treatment costs and preventing disease prevalence.

The multidisciplinary nature of the conducted research opens diverse paths for knowledge discovery, since further research studies would be continuously applied to develop various data mining approaches to solve a real-world problem; such as the problem of sheep lameness tackled in this thesis. The application of machine learning has been increasing as the amount of data collected from real-world problems increasing, becoming more sophisticated, and become presented in multi-dimensional space (Maxwell *et al.*, 2018; L'Heureux *et al.*, 2017).

5.2 Summarised Research Findings and Recommendations

The practical implementation toward achieving the aims of this thesis are demonstrated in the significant findings contributing to the field of knowledge. As illustrated in Section 2.8, the fundamental gap identified in literature could be noticed in the limitation in the number of studies that investigate sheep lameness detection utilising the retrieved sensor reading from a sheep neck collar. In contrast to their counterparts in cattle, the literature studies in detecting cattle lameness are satisfactory enough. As a result, commercial sensors to monitor a cow's health are produced by IceRobotics based in Scotland which launched a CowAlert sensor which is commercially available to the stakeholders. In comparison to the cattle industry, the market lacks such an alert system to be developed to monitor sheep health. Although the reason could refer to the paucity in sheep lameness research studies, there are other real-world reasons; cows are more valuable, and most of lameness detection sensors/pedometers are leg mounted sensors, while sheep have skinny legs that make it difficult to attach. Moreover, the most cow developments focus on dairy cattle which are either indoors all the time or come in and at least twice a day, so it is much easier to identify the signs of a lame cow compared to a lame sheep. Therefore, this thesis investigates the whole process of the indication of lameness in sheep starting with collecting sheep data, pre-processing it, features calculation and selection, toward the model building, and validation for future lameness predictions.

In order to demonstrate the requirement for developing a feasible, inexpensive, and handy sensor kit by the shepherd that able to collect sheep movements, analyse collected data, and produce an alarm for abnormal walking segments within sheep movements which might lead to lameness implications in the future, the investigations for the significant factors contributing to decision making are necessarily required in this research. The **first factor** to be investigated is the sample frequency rate which affects the energy consumption of the sensor. High sampling rates cause sensor battery drainage, while fewer frequencies could save more sensor energy (Hounslow *et al.*, 2019). However, a small sampling rate might not be enough to inform of a sheep's status because of the small amount of information retrieved in each sensor reading. Therefore, a compromising solution between sensor energy consumption and the amount of the informative information retrieved was a crucial step to be investigated in this thesis. Empirically, the best classification performance is obtained when the 10 Hz. sampling rate was applied, then followed by 5 Hz and 4 Hz. The findings converge with Walton *et al.*, (2018), who recommend 16 Hz. from tested frequencies of 32, 16, and 8 Hz. Although they recommend

16 Hz., the selected window size used was 7 $sec.window$ (7 window size × 16 Hz. = 112 data-points). While the current thesis recommended 10 Hz. with 10 $sec.window$ (10 window size × 10 Hz. = 100 data-points). So, the number of the data-points that would be dealt by the pre-processing and decision-making steps is the foundation for lameness detection process. Since a compromising could be achieved by either employing higher frequency rate (more sensor energy would be spent) or by increasing the selected window size (more processing time would be required) as the number of data-points to be manipulated equal to the sampling rate multiplied by the selected window size.

The **second factor** is the type of sensor used. As illustrated in Section 3.2.2, two types of Android sensor system are available; a hardware-based sensor and a software-based sensor. The first type required more energy to operate compared to the second type, which is calculated usually from hardware-based sensors. Accelerometer and Gyroscope sensors are hardware-based sensors, while Orientation sensor is a software-based sensor and could be calculated as explained in 3.5.2.1. According to the aim of the research of providing affordable suggestions for a sheep sensor in the market, the less hardware-based sensor is preferable. So, the recommendation is that acceleration sensor readings are adequate to produce satisfactory lameness prediction results, because the highest prediction accuracy of 88.92% is achieved when only an accelerometer sensor was used in addition to the orientation readings that were calculated from the acceleration readings. The suggested recommendations fall in line with other research studies (Kamminga *et al.*, 2018; Kleanthous *et al.*, 2019).

The **third factor** is a combination of segmentation methods (FNSW and FOSW) and segment window size (10, 7, and 5 sec. are applied). The FNSW segmentation method could be implemented in real-time with minimum memory requirements compared to FNSW; however, FOSW produces better prediction results compared to FNSW because of the proportion of overlapped information between every two successive segments. Although the experimental results reveal that the 10 $sec.window$ offers the best prediction accuracies, the 7 $sec.window$ is also competitive, while the 5 $sec.window$ provides fewer accurate predictions for sheep lameness. The 7 $sec.window$ is also recommended by Walton *et al.*, (2018) as a preferable window size from the 3, 5, and 7 $sec.windows$ tried in their study for classifying sheep behaviour into standing, walking, and lying.

The **fourth** examined **factor** is the feature selection algorithm in terms of accuracy and time required to be executed. Three FS methods were tried; which are ReliefF, GA, and RF. The lowest execution time is achieved by applying ReliefF followed by RF, while GA consumes the highest execution time. However, the prediction accuracies obtained when the ensemble classifier was trained by the feature selected by the ReliefF algorithm is quite low compared to ones selected by RF or GA. GA takes a longer time to execute; however, competitive accuracy results are achieved when GA is applied compared to RF. Thus, RF could be preferable to be deployed in a sheep sensor kit for lameness detection for future manufacturing studies for developing a commercial sheep sensor. This opinion agrees with Vazquez Diosdado *et al.*, (2018).

The **final factor** to be examined is the identification of the best machine learning algorithm to classify sheep status into sound walking, mildly lame walking, and severely lame walking. From the fact that no classification algorithm fits all types of data, the test for best performance is applied to raw sheep data gathered at the early stage of the current research. The practical experiment reveals that the decision tree algorithm could suit the raw sheep data with an accuracy of 74.46% compared to other 9 classifiers approached that were applied in (Al-Rubaye *et al.*, 2018) See Appendix I. 11. Further development was implemented as pre-processing steps to the extended sheep datasets in the current thesis were investigated. Bagging and Boosting ensemble classifiers are implemented (100 decision trees were trained) to overcome the problem of overfitting as the final prediction accuracy was the average of the 100th trained classifiers within the ensemble. In addition, the RusBoosting algorithm was tried on the 4 Hz. dataset to overcome the problem of the imbalanced dataset when the number of classes in a dataset is unequal. Regarding the memory requirement for the future sensor kit to be developed for lameness detection, the best recommended setting would be when the applied ML approach occupies less memory within the sensor. As listed in Table 2-8, DT/CART required less memory space than the other approaches; moreover, the memory storage in the future suggested sensor might be increased whenever the complexity of the ML approach is at a higher level. The complexity of classifier and the required sensor memory storage need to be compromised in future studies. Therefore, only 100 trees within the ensemble classifier in the current thesis were recommended as the detection percentage of mildly lame sheep was at a satisfactory level.

The overall findings of the thesis are original as no adequate studies investigate the sheep lameness classification in relation to machine learning implementation. Not only is the predicted model built for lameness detection, a validation study for the whole data mining approach is also conducted to guide researchers for further enhancements in future related studies. Additionally, the current study provides the necessary information required to be considered when manufacturing a sheep sensor kit for monitoring sheep health on-farm and producing health issue alarm.

5.3 Research's Practice and Limitation

The recommendations of this thesis would be applied by future studies to develop a sensor kit for lameness detection in sheep comparable to their counterpart in cattle named CowAlert by IceRobotics. The PLF market lacks a special sensor to monitor sheep health remotely as sheep left grazing in fields for a longer time than cows with no routine milking twice a day. When it comes to the actual practice, the implication of the research findings could be assessed in term of sensor energy consumption, memory space, and the accuracy of lameness detection concerning the sensor price which was targeted in this thesis to be cheap and easily accessible by farmers.

Regarding the energy spent by the sensor to be developed, the target was to prolong the sensor life as much as possible by reducing the sensor power drainage. For example, higher frequency sensors provide more information on sheep movements than lower sampling rates sensors. In this thesis, the collected data from the walking sheep was retrieved at limited low sampling rates (10, 5, 4 Hz.) to keep the battery life longer. Despite the observed time was not too long approximately either for 15 minutes or for 1 hour. Further studies would investigate the effect of the higher frequencies for data collection which is expected to decrease the sensor life and increase the price of the sensor to be manufactured. Although the amount of the collected data at higher frequencies would contain additional information about sheep movement, the most important data would be the one that contributes to decision making. So, in the pre-processing stage, the limitation of low sampling rate could be manipulated by increasing the selected window as the manipulated data-points are equal to the (sampling rates × window size) as discussed in Section 4.8.

Another practice for research findings concern keeping the sensor energy is that setting the sensor to the sleep mood when sheep are not walking as the walking segments are already extracted to be manipulated (Section 3.5.5). This process would save the sensor battery for a longer period. Alternatively, the sensor energy might be spent only for data collection, while the pre-processing stage including; walking segment extraction, feature calculation and selection, training and validating the selected model are all implemented in a common Cloud (Cloud: is a data storage resource on the Internet available to users without direct management), that might require communications cost for data transmission and receiving. However, the work in this thesis was limited to online process which means the data collection and decision making was supposed to conduct in the same sensor kit, while the suggested ideas for future studies would append data communication equipments and offline processing in the Cloud that might be a luxurious solution to the farmers to spot lame sheep in an unattended way.

Additionally, one of the thesis valuable findings that positively affect sensor power consumption recommends that only acceleration hardware sensor is capable to detect the early signs of lameness and it is also used to derive the orientation sensor reading as well. So, no extra hardware gyroscope sensor is required for lameness detection from the mounted sensor on the sheep neck collar. However, more hardware sensors would be examined for the problem of lameness detection in sheep by future research studies.

Regarding the size of memory in the proposed sensor, the smaller size is targeted. So, the amount of the collected data, the number of data-points to be manipulated (sampling rates × window size), and the complexity of the selected ML approach all affect the memory space required. For the current thesis, a recommend data-points were 100 observation (10 Hz, and 10 $sec.window$) which require a small amount of memory in the sensor to be manufactured. Extra data-point would be tried by other future research studies when the price of the sensor is not a matter for the stakeholder. As mentioned, there is no study yet explore the factors affect the sensor design for lameness detection; therefore, more studies in addition to the current thesis research are still needed in the near future. Another factor affects the memory size is the complexity of ML classifier used; so, CART; which is a type of binary DT, was implemented in this thesis as it requires less memory space than other ML methods (Table 2-8), and it is suitable to be embedded into one sensor kit as aimed in this thesis. Alternatively, if an offline practice for the current work would be applied in future, more sophisticated and accurate ML techniques could be practice in Cloud such as Deep learning techniques which keeps learning from the new data that are fed to its learner classifier.

Regarding the accuracy of the proposed validated SLDM, promising lameness detection results are achieved according to the recommended setting in SLDM. So, it would be possible to manufacture an accelerometer sensor kit that collects data at the current recommendation such as 10 Hz for sampling rate, extract walking segments only to be manipulated using FOSW with 20% overlapping with $10\, sec\ window\ size$, 20 selected features by RF, and ensemble classifier. However, further suggestions would be implemented to increase the accuracy of the validated model regardless of the price of the sensor to be manufactured as mention earlier.

Finally, it is important to notice that the accuracy of the proposed SLDM with its current settings could be varied according to alternation in different factors such as sampling frequencies, window sizes, FS methods, and even the labelled class for the collected data. Since the SLDM applied supervised ML techniques (ensemble of CART), which required the class of data to be labelled, unlike unsupervised learning where the class of data no more required. For example, the sheep lameness status (sound, mild, and severe) were primarily labelled for their lameness level by (Tim Perks), the expert shepherd in Lodge Farm. However, the achieved results might be changed if a different shepherd labelled the same sheep for the data collection process. Sheep labelling for sound, mild, or severe is a subjective process; therefore, more objective methods utilising sensor technology are opted to develop PLF. The limitation of employing one expert for data labelling in the current thesis could be overcome in future research studies by employing more than one expert to label the same group of sheep that are allocated for data collection process and reach an agreement among the experts labelling o9f the same group of sheep.

5.4 Future work

The work conducted in this thesis could be improved in terms of hardware and software implementation. So suggested ideas which could improve the conducted research study or be applied to future studies are as follows:

5.4.1 Potential Hardware Improvements

1- Since each sensor mounted into a sheep's neck collar could be a potential sensor node in a Wireless Sensor Network (WSN) that would be utilised for the flock monitoring system, saving the battery life of each node is essentially required. To do so, each sensor node that is mounted on a sheep neck would only work when the sheep is walking and put in sleep mode when no walking behaviour is detected. This process would prolong the sensor life within a WSN.

2- Sheep head movement could be harnessed to produce mechanical energy for self-node battery charging for the sake of gaining longer battery life.

3- As a further approach to saving a sensor battery's life is to deploy solar panels into a sensor kit as an alternative source of energy when the sensor battery is lacking energy.

4- Future sheep studies would combine extra hardware sensors like GPS sensors to track the sheep in the field and monitor their movements in an unattended way.

5- Deploying SLDM as a mobile application requires communication consideration; however, it could be implemented in future for the benefit of shepherds on a farm when an alarm is issued directly to their mobile phones.

5.4.2 Potential Software Improvements

1- Further investigation could be performed to estimate the fitness function for GA optimisation; for example, using KNN instead of CHIAD. That would decrease the time required for execution. Furthermore, the best individual is selected according to the highest fitness value; however, the average fitness value could produce better results.

2- In segmentation, when the total number of segments in each dataset are calculated seg_no some information is lost due to the data-points less than the seg_size being discarded. For example, if a dataset has 149 data-points and $seg\_size = 50$, the total number of segments would be two, each with 50 data-points, while the remaining 49 data-points will be discarded as 49 data-points is less than $seg\_size = 50$. Therefore, to guarantee that no more data will be lost, duplication within seg_size has to be performed if the lost data estimation is more than half of seg_size.

3- Further supervised machine learning algorithms could be implemented to achieve a better prediction performance such as Naïve-Bayesian, ANNs, or Deep learning; however, the interpretation could be a challenge to comprehend compared to CART.

REFERENCES:

Abdul Jabbar, K., Hansen, M.F., Smith, M.L., Smith, L.N. (2017) Early and non-intrusive lameness detection in dairy cows using 3-dimensional video. *Biosystems Engineering*. **153**, 63–69.

Adnan, M.N. (2017) *Decision Tree and Decision Forest Algorithms On Improving Accuracy, Efficiency and KnowledgeDiscovery.pdf*. Charles Sturt University.

Aggarwal, C.C. (2015) *Data Mining*. Cham: Springer International Publishing.

Agostinho, F.S., Rahal, S.C., Araújo, F.A.P., Conceição, R.T., Hussni, C.A., El-Warrak, A.O., Monteiro, F.O.B. (2012) Gait analysis in clinically healthy sheep from three different age groups using a pressure-sensitive walkway. *BMC Veterinary Research*. **8**(1), 87.

AHDB (2016) Managing cull ewes. *Lamb Briefing*. [online]. Available from: http://beefandlamb.ahdb.org.uk/wp-content/uploads/2016/08/Cull-ewes-310816.pdf [Accessed September 28, 2019].

Ahmad, N., Ghazilla, R.A.R., Khairi, N.M., Kasi, V. (2013) Reviews on Various Inertial Measurement Unit (IMU) Sensor Applications. *International Journal of Signal Processing Systems*. **1**(2), 256–262.

Al-Rubaye, Z., Al-Sherbaz, A., McCormick, W., Turner, S. (2018) Sensor Data Classification for the Indication of Lameness in Sheep. In *Collaborate Computing: Networking, Applications and Worksharing*. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer International Publishing, pp. 309–320.

Alexandropoulos, S.-A.N., Kotsiantis, S.B., Vrahatis, M.N. (2019) Data preprocessing in predictive data mining. *The Knowledge Engineering Review*. **34**(May), e1.

Alsaaod, M., Büscher, W. (2012) Detection of hoof lesions using digital infrared thermography in dairy cows. *Journal of Dairy Science*. **95**(2), 735–742.

Alsaaod, M., Fadul, M., Steiner, A. (2019) Automatic lameness detection in cattle. *The Veterinary Journal*.

Alsaaod, M., Römer, C., Kleinmanns, J., Hendriksen, K., Rose-Meierhöfer, S., Plümer, L.,

References

Büscher, W. (2012) Electronic detection of lameness in dairy cows through measuring pedometric activity and lying behavior. *Applied Animal Behaviour Science*. **142**(3–4), 134–141.

Alsaaod, M., Schaefer, A.L., Büscher, W., Steiner, A. (2015) The role of infrared thermography as a non-invasive tool for the detection of lameness in cattle. *Sensors (Switzerland)*. **15**(6), 14513–14525.

Alvarenga, F.A.P., Borges, I., Palkovič, L., Rodina, J., Oddy, V.H., Dobos, R.C. (2016) Using a three-axis accelerometer to identify and classify sheep behaviour at pasture. *Applied Animal Behaviour Science*. **181**, 91–99.

Andriamandroso, A., Lebeau, F., Bindelle, J. (2014) Accurate monitoring of the rumination behaviour of cattle using IMU signals from a mobile device. In A. Hopkins, R. P. Collins, M. D. Fraser, V. R. King, D. C. Lloyed, J. M. Moorby, & P. R. H. Robson, eds. *Grassland Science in Europe, 19 (EGF at 50: the Future of European Grasslands)*. Llandysul, Wales: Gomer Press Ltd., pp. 631–634.

Andriamandroso, A., Lebeau, F., Bindelle, J. (2015) Changes in biting characteristics recorded using the inertial measurement unit of a smartphone reflect differences in sward attributes. In *7th Conference on Precision Livestock Farming'15*. Milan, Italy, pp. 283–289.

Andriamandroso, A.L.H., Bindelle, J., Mercatoris, B., Lebeau, F. (2016) A review on the use of sensors to monitor cattle jaw movements and behavior when grazing. *Biotechnologie Agronomie Societe Et Environnement*. **20**(1), 273–286.

Andriamandroso, A.L.H., Lebeau, F., Beckers, Y., Froidmont, E., Dufrasne, I., Heinesch, B., Dumortier, P., Blanchy, G., Blaise, Y., Bindelle, J. (2017) Development of an open-source algorithm based on inertial measurement units (IMU) of a smartphone to detect cattle grass intake and ruminating behaviors. *Computers and Electronics in Agriculture*. **139**, 126–137.

Android Developers (2019a) Compute the device's orientation. [online]. Available from: https://developer.android.com/guide/topics/sensors/sensors_position#sensors-pos-orient [Accessed April 17, 2019].

Android Developers (2019b) Sensor Coordinate System. [online]. Available from: https://developer.android.com/guide/topics/sensors/sensors_overview.html#sensors-coords

[Accessed April 17, 2019].

Android Developers (2019c) Sensor Overview. [online]. Available from: https://developer.android.com/guide/topics/sensors/sensors_overview [Accessed April 10, 2019].

Android Developers (2019d) Use the acceleration. [online]. Available from: https://developer.android.com/guide/topics/sensors/sensors_motion.html#sensors-motion-accel [Accessed April 17, 2019].

Android Developers (2019e) Use the gyroscope. [online]. Available from: https://developer.android.com/guide/topics/sensors/sensors_motion.html#sensors-motion-gyro [Accessed April 17, 2019].

Anzuino, K., Knowles, T.G., Lee, M.R.F., Grogono-Thomas, R. (2019) Survey of husbandry and health on UK commercial dairy goat farms. *Veterinary Record*. **185**(9), 267–267.

Apinan, A., Rattanawong, T., Kuankid, S. (2015) Classification of the cattle behaviors by using magnitude and variance of accelerometer signal. *Agricultural Engineering International: CIGR Journal*. **17**(4), 415–420.

Arcidiacono, C., Porto, S.M.C., Mancino, M., Cascone, G. (2017) A threshold-based algorithm for the development of inertial sensor-based systems to perform real-time cow step counting in free-stall barns. *Biosystems Engineering*. **153**, 99–109.

Babatunde, O., Armstrong, L., Leng, J., Dean, D. (2014) A Genetic Algorithm-Based Feature Selection. *International Journal of Electronics Communications and Computer Engineering*. **5**(4), 899–905.

Bahlo, C., Dahlhaus, P., Thompson, H., Trotter, M. (2019) The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. *Computers and Electronics in Agriculture*. **156**(November 2018), 459–466.

Bailey, D.W., Trotter, M.G., Knight, C.W., Thomas, M.G. (2018) Use of GPS tracking collars and accelerometers for rangeland livestock production research1. *Translational Animal Science*. **2**(1), 81–88.

Barker, Z.E., Vázquez Diosdado, J.A., Codling, E.A., Bell, N.J., Hodges, H.R., Croft, D.P.,

References

Amory, J.R. (2018) Use of novel sensors combining local positioning and acceleration to measure feeding behavior differences associated with lameness in dairy cattle. *Journal of Dairy Science*. **101**(7), 6310–6321.

Barwick, J., Lamb, David, Dobos, R., Schneider, D., Welch, M., Trotter, M. (2018b) Predicting Lameness in Sheep Activity Using Tri-Axial Acceleration Signals. *Animals*. **8**(1), 12.

Barwick, J., Lamb, David W., Dobos, R., Welch, M., Trotter, M. (2018a) Categorising sheep activity using a tri-axial accelerometer. *Computers and Electronics in Agriculture*. **145**(September 2017), 289–297.

Beer, G., Alsaaod, M., Starke, A., Schuepbach-Regula, G., Müller, H., Kohler, P., Steiner, A. (2016) Use of extended characteristics of locomotion and feeding behavior for automated identification of lame dairy cows. *PLoS ONE*. **11**(5), 1–18.

Benaissa, S., Tuyttens, F.A.M., Plets, D., de Pessemier, T., Trogh, J., Tanghe, E., Martens, L., Vandaele, L., Van Nuffel, A., Joseph, W., Sonck, B. (2017) On the use of on-cow accelerometers for the classification of behaviours in dairy barns. *Research in Veterinary Science*. (April).

Bersch, S., Azzi, D., Khusainov, R., Achumba, I., Ries, J. (2014) Sensor Data Acquisition and Processing Parameters for Human Activity Classification. *Sensors*. **14**(3), 4239–4270.

Bitar, H. (2013) SensoDuino: Turn Your Android Phone into a Wireless Sensors Hub for Arduino. [online]. Available from: http://www.instructables.com/id/SensoDuino-Turn-Your-Android-Phone-into-a-Wireless/ [Accessed April 17, 2019].

Blackie, N., Bleach, E., Amory, J., Scaife, J. (2011) Impact of lameness on gait characteristics and lying behaviour of zero grazed dairy cattle in early lactation. *Applied Animal Behaviour Science*. **129**(2–4), 67–73.

Boulesteix, A.-L., Janitza, S., Kruppa, J., König, I.R. (2012) Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. **2**(6), 493–507.

Breiman, L. (2001) Random Forests. *Machine Learning*. **45**(1), 5–32.

References

Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A. (1984) *Classification and Regression Trees*. Taylor & Francis.

Brian, K. (2016) *Reducing Lameness for Better Returns*.

Byabazaire, J., Olariu, C., Taneja, M., Davy, A. (2019) Lameness Detection as a Service: Application of Machine Learning to an Internet of Cattle. In *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, pp. 1–6.

Byrne, Daire T., Berry, D.P., Esmonde, H., McGovern, F., Creighton, P., McHugh, N. (2019a) Infrared thermography as a tool to detect hoof lesions in sheep. *Translational Animal Science*. **3**(1), 577–588.

Byrne, D.T., Esmonde, H., Berry, D.P., McGovern, F., Creighton, P., McHugh, N. (2019b) Sheep lameness detection from individual hoof load. *Computers and Electronics in Agriculture*. **158**(2019), 241–248.

Chapinal, N., de Passillé, A.M., Pastell, M., Hänninen, L., Munksgaard, L., Rushen, J. (2011) Measurement of acceleration while walking as an automated method for gait assessment in dairy cattle. *Journal of Dairy Science*. **94**(6), 2895–2901.

Chapinal, N., de Passillé, A.M., Rushen, J., Wagner, S. (2010) Automated methods for detecting lameness and measuring analgesia in dairy cattle. *Journal of Dairy Science*. **93**(5), 2007–2013.

Chapinal, N., Tucker, C.B. (2012) Validation of an automated method to count steps while cows stand on a weighing platform and its application as a measure to detect lameness. *Journal of Dairy Science*. **95**(11), 6523–6528.

Chomiak, V. (2017) CowAlert now has daily lameness alerts. *Dairy Tech.* [online]. Available from: https://dairy-tech.uk/cowalert-now-daily-lameness-alerts/ [Accessed September 22, 2019].

Collins, P.M., Green, J.A., Warwick-Evans, V., Dodd, S., Shaw, P.J.A., Arnould, J.P.Y., Halsey, L.G. (2015) Interpreting behaviors from accelerometry: a method combining simplicity and objectivity. *Ecology and Evolution*. **5**(20), 4642–4654.

Cox, V. (2017) *Translating Statistics to Make Decisions*. Berkeley, CA: Apress.

References

Davila, J., Cretu, A.-M., Zaremba, M. (2017) Wearable Sensor Data Classification for Human Activity Recognition Based on an Iterative Learning Framework. *Sensors*. **17**(6), 1287.

Debauche, O., Mahmoudi, S., Andriamandroso, A.L.H., Manneback, P., Bindelle, J., Lebeau, F. (2018) Cloud services integration for farm animals' behavior studies based on smartphones as activity sensors. *Journal of Ambient Intelligence and Humanized Computing*. **0**(0), 0.

Debauche, O., Mahmoudi, S., Andriamandroso, A.L.H., Manneback, P., Bindelle, J., Lebeau, F. (2017) Web-based cattle behavior service for researchers based on the smartphone inertial central. *Procedia Computer Science*. **110**(July), 110–116.

Defra (2003) Lameness in Sheep. [online]. Available from: http://adlib.everysite.co.uk/adlib/defra/content.aspx?id=000IL3890W.18B4NRK0LZK8LC [Accessed March 14, 2017].

Dunthorn, J., Dyer, R.M., Neerchal, N.K., McHenry, J.S., Rajkondawar, P.G., Steingraber, G., Tasch, U. (2015) Predictive models of lameness in dairy cows achieve high sensitivity and specificity with force measurements in three dimensions. *Journal of Dairy Research*. **82**(04), 391–399.

Dutta, R., Smith, D., Rawnsley, R., Bishop-Hurley, G., Hills, J., Timms, G., Henry, D. (2015) Dynamic cattle behavioural classification using supervised ensemble classifiers. *Computers and Electronics in Agriculture*. **111**, 18–28.

Figo, D., Diniz, P.C., Ferreira, D.R., Cardoso, M.P. (2010) Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*. **14**(7), 645–662.

Flower, F.C., Weary, D.M. (2009) Gait assessment in dairy cattle. *Animal : an international journal of animal bioscience*. **3**(1), 87–95.

Fogarty, E.S., Swain, D.L., Cronin, G., Trotter, M. (2018) Autonomous on-animal sensors in sheep research: A systematic review. *Computers and Electronics in Agriculture*. **150**(May), 245–256.

Gandhi, R. (2018) Boosting Algorithms: AdaBoost, Gradient Boosting and XGBoost. [online]. Available from: https://hackernoon.com/boosting-algorithms-adaboost-gradient-boosting-and-xgboost-f74991cad38c [Accessed April 3, 2019].

Ganganwar, V. (2012) An overview of classification algorithms for imbalanced datasets. *International Journal of Emerging Technology and Advanced Engineering*. **2**(4), 42–47.

Garcia, E., Klaas, I., Amigo, J.M., Bro, R., Enevoldsen, C. (2014) Lameness detection challenges in automated milking systems addressed with partial least squares discriminant analysis. *Journal of dairy science*. **97**(12), 7476–86.

García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J.M., Herrera, F. (2016) Big data preprocessing: methods and prospects. *Big Data Analytics*. **1**(1), 9.

Gaudy, J., Green, L. (2016) Focus on lameness How to reduce lameness in sheep. [online]. Available from: http://beefandlamb.ahdb.org.uk/wp/wp-content/uploads/2016/07/BRP-Focus-on-lameness.pdf [Accessed September 28, 2019].

Gelasakis, Kalogianni, Bossis (2019) Aetiology, Risk Factors, Diagnosis and Control of Foot-Related Lameness in Dairy Sheep. *Animals*. **9**(8), 509.

Ghendir, S. (2016) Sheep Tracker.

Giovanetti, V., Decandia, M., Molle, G., Acciaro, M., Mameli, M., Cabiddu, A., Cossu, R., Serra, M.G., Manca, C., Rassu, S.P.G., Dimauro, C. (2017) Automatic classification system for grazing, ruminating and resting behaviour of dairy sheep using a tri-axial accelerometer. *Livestock Science*. **196**(December 2016), 42–48.

GitHub (2014) Sensor Log.

Gleiss, A.C., Wilson, R.P., Shepard, E.L.C. (2011) Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*. **2**(1), 23–33.

Godsk, T., Kjærgaard, M.B. (2011) High Classification Rates for Continuous Cow Activity Recognition Using Low-Cost GPS Positioning Sensors and Standard Machine Learning Techniques. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 174–188.

González, L.A., Bishop-Hurley, G.J., Handcock, R.N., Crossman, C. (2015) Behavioral classification of data from collars containing motion sensors in grazing cattle. *Computers and Electronics in Agriculture*. **110**, 91–102.

Groenevelt, M., Anzuino, K., Smith, S., Lee, M.R.F., Grogono-Thomas, R. (2015) A case report of lameness in two dairy goat herds; a suspected combination of nutritional factors concurrent with treponeme infection. *BMC Research Notes*. **8**(1), 791.

Grover, P. (2017) Gradient Boosting from scratch. [online]. Available from: https://medium.com/mlreview/gradient-boosting-from-scratch-1e317ae4587d [Accessed April 5, 2019].

Van De Gucht, T., Saeys, W., Van Weyenberg, S., Lauwers, L., Mertens, K., Vandaele, L., Vangeyte, J., Van Nuffel, A. (2017) Automatic cow lameness detection with a pressure mat: Effects of mat length and sensor resolution. *Computers and Electronics in Agriculture*. **134**, 172–180.

Guo, L., Welch, M., Dobos, R., Kwan, P., Wang, W. (2018) Comparison of grazing behaviour of sheep on pasture with different sward surface heights using an inertial measurement unit sensor. *Computers and Electronics in Agriculture*. **150**(February), 394–401.

Guo, Y., Poulton, G., Corke, P., Bishop-Hurley, G.J., Wark, T., Swain, D.L. (2009) Using accelerometer, high sample rate GPS and magnetometer data to develop a cattle movement and behaviour model. *Ecological Modelling*. **220**(17), 2068–2075.

Gupta, R.K., Lathwal, S.S., Mohanty, T.K. (2014) Detection of Lameness of Cow based on Body Weight using Artificial Neural Network. . (July 2011), 337–341.

Haddadi, H., King, A.J., Wills, A.P., Fay, D., Lowe, J., Morton, A.J., Hailes, S., Wilson, A.M. (2011) Determining association networks in social animals: Choosing spatial-temporal criteria and sampling rates. *Behavioral Ecology and Sociobiology*. **65**(8), 1659–1668.

Haladjian, J., Haug, J., Nüske, S., Bruegge, B. (2018) A Wearable Sensor System for Lameness Detection in Dairy Cattle. *Multimodal Technologies and Interaction*. **2**(2), 27.

Hartley, A.S. (2014) *EXPLORING STATISTICAL CLASSIFICATION , GIS ANALYSIS AND MAPPING OF HERBIVORE BEHAVIOURS USING ACCELEROMETERS AND HIGH ACCURACY GPS 1 Abstract*. Liverpool John Moores.

References

Helwatkar, A., Riordan, D., Walsh, J. (2014) Sensor Technology For Animal Health Monitoring. In *8th Internacional Conference on Sensing Technology*. pp. 2–4.

Van Hertem, T., Maltz, E., Antler, A., Romanini, C.E.B., Viazzi, S., Bahr, C., Schlageter-Tello, A., Lokhorst, C., Berckmans, D., Halachmi, I. (2013) Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. *Journal of Dairy Science*. **96**(7), 4286–4298.

Van Hertem, T., Steensels, M., Viazzi, S., Romanini, E.C.B., Bahr, C., Berckmans, D. (2014) Improving a computer vision lameness detection system by adding behaviour and performance measures. In *Proceedings International Conference of Agricultural Engineering*. Zurich, pp. 1–8.

Van Hertem, T., Viazzi, S., Steensels, M., Maltz, E., Antler, A., Alchanatis, V., Schlageter-Tello, A.A., Lokhorst, K., Romanini, E.C.B., Bahr, C., Berckmans, D., Halachmi, I. (2014) Automatic lameness detection based on consecutive 3D-video recordings. *Biosystems Engineering*. **119**, 108–116.

Hounslow, J.L., Brewster, L.R., Lear, K.O., Guttridge, T.L., Daly, R., Whitney, N.M., Gleiss, A.C. (2019) Assessing the effects of sampling frequency on behavioural classification of accelerometer data. *Journal of Experimental Marine Biology and Ecology*. **512**(January), 22–30.

ICEROBOTiCS (2019) Sensors. [online]. Available from: https://www.icerobotics.com/researchers/#sensors [Accessed September 22, 2019].

Il-Seok Oh, Jin-Seon Lee, Byung-Ro Moon (2004) Hybrid genetic algorithms for feature selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **26**(11), 1424–1437.

Ito, K., von Keyserlingk, M.A.G., LeBlanc, S.J., Weary, D.M. (2010) Lying behavior as an indicator of lameness in dairy cows. *Journal of Dairy Science*. **93**(8), 3553–3560.

James, G., Witten, D., Hastie, T., Tibshirani, R. (2013) *An Introduction to Statistical Learning*. New York, NY: Springer New York.

James, W.P.J., Brewer, P., Rose, M.T., Williams, M.L., Mac Parthaláin, N. (2016) A novel behavioral model of the pasture-based dairy cow from GPS data using data mining and machine learning techniques. *Journal of Dairy Science*. **99**(3), 2063–2075.

Jiang, B., Song, H., He, D. (2019) Lameness detection of dairy cows based on a double normal background statistical model. *Computers and Electronics in Agriculture*. **158**(July 2018), 140–149.

Jiawei, H., Kamber, M., Pei, J. (2012) *Data Mining: Concepts and Techniques*. 3rd Editio. Elsevier.

Jovic, A., Brkic, K., Bogunovic, N. (2015) A review of feature selection methods with applications. In *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE, pp. 1200–1205.

Jukan, A., Masip-Bruin, X., Amla, N. (2017) Smart Computing and Sensing Technologies for Animal Welfare: A Systematic Review. *ACM Computing Surveys*. **50**(1), 1–27.

Kamminga, J.W., Bisby, H.C., Le, D. V., Meratnia, N., Havinga, P.J.M. (2017) Generic online animal activity recognition on collar tags. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers on - UbiComp '17*. New York, New York, USA: ACM Press, pp. 597–606.

Kamminga, J.W., Le, D. V., Meijers, J.P., Bisby, H., Meratnia, N., Havinga, P.J.M. (2018) Robust Sensor-Orientation-Independent Feature Selection for Animal Activity Recognition on Collar Tags. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. **2**(1), 1–27.

Kamphuis, C., Frank, E., Burke, J.K., Verkerk, G.A., Jago, J.G. (2013) Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological characteristics of dairy cows to detect lameness. *Journal of dairy science*. **96**(11), 7043–53.

Keogh, E., Chu, S., Hart, D., Pazzani, M. (2001) An online algorithm for segmenting time series. In *Proceedings 2001 IEEE International Conference on Data Mining*. San Jose, CA, USA, USA: IEEE Comput. Soc, pp. 289–296.

Kevin Markham (2014) DataSchool/Simple guide to confusion matrix terminology. [online].

Available from: http://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/ [Accessed April 23, 2017].

Khanh, P.C.P., Dinh Chinh, N., Cham, T.T., Vui, P.T., Tan, T.D. (2016) Classification of cow behavior using 3-D OF accelerometer and decision tree algorithm. In *BME-HUST 2016 - 3rd International Conference on Biomedical Engineering*. Hanoi, Vietnam: IEEE, pp. 45–50.

Khanh, P.C.P., Long, T.T., Dinh Chinh, N., Duc-Tan, T. (2018) Performance evaluation of a multi-stage classification for cow behavior. In *2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom)*. Ho Chi Minh City, Vietnam: IEEE, pp. 121–125.

Kim, J., Breur, G.J. (2008) Temporospatial and kinetic characteristics of sheep walking on a pressure sensing walkway. *Canadian Journal of Veterinary Research*. **72**(1), 50–55.

Kim, Y., Ha, S., Kwon, J. (2015) Human detection using doppler radar based on physical characteristics of targets. *IEEE Geoscience and Remote Sensing Letters*. **12**(2), 289–293.

Kleanthous, N., Hussain, A., Mason, A., Sneddon, J. (2019) Data Science Approaches for the Analysis of Animal Behaviours. In D.-S. Huang, M. M. Gromiha, K. Han, & A. Hussain, eds. *Intelligent Computing Methodologies*. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 411–422.

Kleanthous, N., Hussain, A., Mason, A., Sneddon, J., Shaw, A., Fergus, P., Chalmers, C., Al-Jumeily, D. (2018) Machine Learning Techniques for Classification of Livestock Behavior. In *Neural Information Processing*. Springer International Publishing, pp. 304–315.

Knight, C.W., Bailey, D.W., Faulkner, D. (2018) Low-Cost Global Positioning System Tracking Collars for Use on Cattle. *Rangeland Ecology and Management*. **71**(4), 506–508.

Kokin, E., Praks, J., Veermäe, I., Poikalainen, V., Vallas, M. (2014) IceTag3D^{TM} accelerometric device in cattle lameness detection. *Agronomy Research*. **12**(1), 223–230.

Kotsiantis, S.B., Kanellopoulos, D., Pintelas, P.E. (2006) Data Preprocessing for Supervised Leaning. *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE*. **1**(2), 111–117.

Krause, J., Krause, S., Arlinghaus, R., Psorakis, I., Roberts, S., Rutz, C. (2013) Reality mining of animal social systems. *Trends in Ecology & Evolution*. **28**(9), 541–551.

van Kuppevelt, D., Heywood, J., Hamer, M., Sabia, S., Fitzsimons, E., van Hees, V. (2019) Segmenting accelerometer data from daily life with unsupervised machine learning M. S. Buchowski, ed. *PLOS ONE*. **14**(1), e0208692.

L'Heureux, A., Grolinger, K., Elyamany, H.F., Capretz, M.A.M. (2017) Machine Learning With Big Data: Challenges and Approaches. *IEEE Access*. **5**, 7776–7797.

Ladds, M.A., Thompson, A.P., Kadar, J.-P., J Slip, D., P Hocking, D., G Harcourt, R. (2017) Super machine learning: improving accuracy and reducing variance of behaviour classification from accelerometry. *Animal Biotelemetry*. **5**(1), 8.

Link, Y.C., Büttner, K., Karsten, S. (2015) 3D-Head acceleration used for lameness detection in dairy cows. In *66th EAAP Annual Meeting 2015*. Warsaw, Poland.

Link, Y.C., Salau, J., Karsten, S., Krieter, J. (2016) Using classifiers based on 3D-head acceleration for lameness detection in dairy cows.

Lovatt, F. (2014) Causes , control and costs of lameness in sheep. *Veterinary Ireland Journal*. **5**(4), 185–188.

Lutins, E. (2017a) Boosting in Machine Learning and the Implementation of XGBoost in Python. [online]. Available from: https://towardsdatascience.com/boosting-in-machine-learning-and-the-implementation-of-xgboost-in-python-fb5365e9f2a0 [Accessed April 3, 2019].

Lutins, E. (2017b) Ensemble Methods in Machine Learning: What are They and Why Use Them? [online]. Available from: https://towardsdatascience.com/ensemble-methods-in-machine-learning-what-are-they-and-why-use-them-68ec3f9fef5f [Accessed April 3, 2019].

Maertens, W., Vangeyte, J., Baert, J., Jantuan, A., Mertens, K.C., De Campeneere, S., Pluk, A., Opsomer, G., Van Weyenberg, S., Van Nuffel, A. (2011) Development of a real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: The GAITWISE system. *Biosystems Engineering*. **110**(1), 29–39.

Mangweth, G., Schramel, J.P., Peham, C., Gasser, C., Tichy, A., Altenhofer, C., Weber, A., Kofler, J. (2012) Lameness detection in cows by accelerometric measurement of motion at walk. *Berliner und Munchener tierarztliche Wochenschrift*. **125**(9–10), 386–396.

References

Mansbridge, N., Mitsch, J., Bollard, N., Ellis, K., Miguel-Pacheco, G., Dottorini, T., Kaler, J. (2018) Feature Selection and Comparison of Machine Learning Algorithms in Classification of Grazing and Rumination Behaviour in Sheep. *Sensors*. **18**(10), 3532.

Marais, J., Le Roux, S.P., Wolhuter, R., Niesler, T. (2014) Automatic classification of sheep behaviour using 3-axis accelerometer data. In *twenty-fifth annual symposium of the Pattern Recognition Association of South Africa (PRASA)*. pp. 978–0.

Martiskainen, P., Järvinen, M., Skön, J.-P., Tiirikainen, J., Kolehmainen, M., Mononen, J. (2009) Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied Animal Behaviour Science*. **119**(1–2), 32–38.

Mary, V., Wright, N. (2013) Economic Impact of Health and Welfare Issues in Beef Cattle and Sheep in England. [online]. Available from: http://www.eblex.org.uk/wp/wp-content/uploads/2013/04/Economic-Impact-of-Health-Welfare-Final-Rpt-170413.pdf [Accessed September 28, 2019].

Mason, A., Sneddon, J. (2013) Automated monitoring of foraging behaviour in free ranging sheep grazing a biodiverse pasture. In *2013 Seventh International Conference on Sensing Technology (ICST)*. IEEE, pp. 46–51.

Mathworks (2016) Applying Supervised Learning. *What is Machine Learning*.

MathWorks (2018a) Machine Learning in MATLAB. [online]. Available from: https://ch.mathworks.com/help/stats/machine-learning-in-matlab.html.

MathWorks (2018b) Supervised Learning Workflow and Algorithms. [online]. Available from: https://ch.mathworks.com/help/stats/supervised-learning-machine-learning-workflow-and-algorithms.html#bswluh9.

Maxwell, A.E., Warner, T.A., Fang, F. (2018) Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing*. **39**(9), 2784–2817.

Mazrier, H., Tal, S., Aizinbud, E., Bargai, U. (2006) A field investigation of the use of the pedometer for the early detection of lameness in cattle. *Canadian Veterinary Journal*. **47**(9), 883–886.

References

McNeese, B. (2016) *Are Skewness and Kurtosis Useful Statistics?*

Miller, B., Fridline, M., Liu, P.-Y., Marino, D. (2014) Use of CHAID Decision Trees to Formulate Pathways for the Early Detection of Metabolic Syndrome in Young Adults. *Computational and Mathematical Methods in Medicine*. **2014**((Fridline M., mmf@uakron.edu) Department of Statistics, College of Arts and Sciences, University of Akron, Akron, OH 44325-1913, United States PG-), 1–7.

de Mol, P.M., Verhoeven, P.H.F.., Hogewerf, P.H., Ipema, A.H. (2011) Automated behaviour monitoring in dairy cows. In *5th European Conference on Precision Livestock Farming*. Prague, Czech Republic, pp. 70–80.

de Mol, R.M., André, G., Bleumer, E.J.B., van der Werf, J.T.N., de Haas, Y., van Reenen, C.G. (2013) Applicability of day-to-day variation in behavior for the automated detection of lameness in dairy cows. *Journal of Dairy Science*. **96**(6), 3703–3712.

de Mol, R.M., Bleumer, E.J.B., Hogeewerf, P.H., Ipema, A.H. (2009) Recording of dairy cow behaviour with wireless accelerometers. In C. Lokhorst & P. W. G. G. Koerkamp, eds. *Precision Livestock Farming '09*. The Netherlands: Wageningen Academic Publishers, pp. 349–356.

Moreau, M., Siebert, S., Buerkert, A., Schlecht, E. (2009) Use of a tri-axial accelerometer for automated recording and classification of goats' grazing behaviour. *Applied Animal Behaviour Science*. **119**(3–4), 158–170.

Mottram, T. (2012) Method and system for measuring the mobility of an animal.

Mottram, T.T., Bell, N.J. (2010) A novel method of monitoring mobility in dairy cows. In *The First North American Conference on Precision Dairy Management*. Toronto, Canada, pp. 182–183.

Mulvenna, C.C., Wilson, R.P., Marks, N.J., Maule, A.G., Scantlebury, D.M. (2018) The ability of magnetic field sensors to monitor feeding in three domestic herbivores. *PeerJ*. **6**, e5489.

Mwadulo, M.W. (2016) A Review on Feature Selection Methods For Classification Tasks. *International Journal of Computer Applications Technology and Research*. **5**(6), 395–402.

Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., Brown, S.D. (2004) An introduction to decision tree modeling. *Journal of Chemometrics*. **18**(6), 275–285.

Nadimi, E.S., Søgaard, H.T., Bak, T. (2008) ZigBee-based wireless sensor networks for classifying the behaviour of a herd of animals using classification trees. *Biosystems Engineering*. **100**(2), 167–176.

Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., Getz, W.M. (2012) Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *Journal of Experimental Biology*. **215**(6), 986–996.

Neveux, S., Weary, D.M., Rushen, J., von Keyserlingk, M.A.G., de Passillé, A.M. (2006) Hoof Discomfort Changes How Dairy Cattle Distribute Their Body Weight. *Journal of Dairy Science*. **89**(7), 2503–2509.

Nielsen, D. (2016) *Tree Boosting With XGBoost*.

Nielsen, L.R., Pedersen, A.R., Herskin, M.S., Munksgaard, L. (2010) Quantifying walking and standing behaviour of dairy cows using a moving average based on output from an accelerometer. *Applied Animal Behaviour Science*. **127**(1–2), 12–19.

Nielsen, P.P. (2013) Automatic registration of grazing behaviour in dairy cows using 3D activity loggers. *Applied Animal Behaviour Science*. **148**(3–4), 179–184.

Nieuwhof, G.J., Bishop, S.C. (2005) Costs of the major endemic diseases of sheep in Great Britain and the potential benefits of reduction in disease impact. *Animal Science*. **81**(1), 23–29.

Van Nuffel, A., Vangeyte, J., Mertens, K.C., Pluym, L., De Campeneere, S., Saeys, W., Opsomer, G., Van Weyenberg, S. (2013) Exploration of measurement variation of gait variables for early lameness detection in cattle using the GAITWISE. *Livestock Science*. **156**(1–3), 88–95.

Van Nuffel, A., Zwertvaegher, I., Pluym, L., Van Weyenberg, S., Thorup, V., Pastell, M., Sonck, B., Saeys, W. (2015a) Lameness Detection in Dairy Cows: Part 1. How to Distinguish between Non-Lame and Lame Cows Based on Differences in Locomotion or Behavior. *Animals*. **5**(3), 838–860.

Van Nuffel, A., Zwertvaegher, I., Van Weyenberg, S., Pastell, M., Thorup, V., Bahr, C., Sonck, B., Saeys, W. (2015b) Lameness Detection in Dairy Cows: Part 2. Use of Sensors to Automatically Register Changes in Locomotion or Behavior. *Animals*. **5**(3), 861–885.

O'Callaghan, K.A., Cripps, P.J., Downham, D.Y., Murray, R.D. (2003) Subjective and objective assessment of pain and discomfort due to lameness in dairy cattle. *Animal Welfare*. **12**(4), 605–610.

O'Driscoll, K., Boyle, L., Hanlon, A. (2008) A brief note on the validation of a system for recording lying behaviour in dairy cows. *Applied Animal Behaviour Science*. **111**(1–2), 195–200.

Olechnowicz, J., Jaśkowski, J.M. (2011) Lameness in small ruminants. *Medycyna Weterynaryjna*. **67**(11), 715–719.

Orman, A., Endres, M.I. (2016) Use of thermal imaging for identification of foot lesions in dairy cattle. *Acta Agriculturae Scandinavica, Section A — Animal Science*. **66**(1), 1–7.

Osisanwo, F.., Akinsola, J.E.., Awodele, O., Hinmikaiye, J.O., Olakanmi, O., Akinjobi, J. (2017) Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology*. **48**(3), 128–138.

Pastell, M., Kujala, M., Aisla, A.M., Hautala, M., Poikalainen, V., Praks, J., Veermäe, I., Ahokas, J. (2008) Detecting cow's lameness using force sensors. *Computers and Electronics in Agriculture*. **64**(1), 34–38.

Pastell, M., Takko, H., Gröhn, H., Hautala, M., Poikalainen, V., Praks, J., Veermäe, I., Kujala, M., Ahokas, J. (2006) Assessing cows' welfare: Weighing the cow in a milking robot. *Biosystems Engineering*. **93**(1), 81–87.

Pastell, M., Tiusanen, J., Hakojärvi, M., Hänninen, L. (2009) A wireless accelerometer system with wavelet analysis for assessing lameness in cattle. *Biosystems Engineering*. **104**(4), 545–551.

Pastell, M.E., Kujala, M. (2007) A probabilistic neural network model for lameness detection. *Journal of Dairy Science*. **90**(5), 2283–2292.

Pluk, A., Bahr, C., Leroy, T., Poursaberi, A., Song, X., Vranken, E., Maertens, W., Van Nuffel, A., Berckmans, D. (2010) Evaluation of Step Overlap as an Automatic Measure in Dairy Cow

Locomotion. *Transactions of the ASABE*. **53**(4), 1305–1312.

Pluk, A., Bahr, C., Poursaberi, A., Maertens, W., van Nuffel, A., Berckmans, D. (2012) Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques. *Journal of Dairy Science*. **95**(4), 1738–1748.

Pluk, A, Bahr, C., Poursaberi, A., Maertens, W., van Nuffel, A., Berckmans, D. (2012) Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques. *Journal of Dairy Science*. **95**(4), 1738–1748.

Poursaberi, A., Bahr, C., Pluk, A., Berckmans, D., Veermae, I., Kokin, E., Pokalainen, V. (2011) Online lameness detection in dairy cattle using Body Movement Pattern (BMP). *International Conference on Intelligent Systems Design and Applications*. (1), 732–736.

Poursaberi, A., Bahr, C., Pluk, A., Van Nuffel, A., Berckmans, D. (2010) Real-time automatic lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques. *Computers and Electronics in Agriculture*. **74**(1), 110–119.

Poursaberi, A., Pluk, A., Bahr, C., Martens, W., Veermäe, I., Kokin, E., Praks, J., Poikalainen, V., Pastell, M., Ahokas, J., Van Nuffel, A., Vangeyte, J., Sonck, B., Berckmans, D. (2009) Image based separation of dairy cows for automatic lameness detection with a real time vision system. *American Society of Agricultural and Biological Engineers*.

Qasem, L., Cardew, A., Wilson, A., Griffiths, I., Halsey, L.G., Shepard, E.L.C., Gleiss, A.C., Wilson, R. (2012) Tri-Axial Dynamic Acceleration as a Proxy for Animal Energy Expenditure ; Should We Be Summing Values or Calculating the Vector ? *PLoS ONE*. **7**(2).

Radeski, M., Ilieski, V. (2017) Gait and posture discrimination in sheep using a tri-axial accelerometer. *animal*. **11**(07), 1249–1257.

Rahman, A., Smith, D.V., Little, B., Ingham, A.B., Greenwood, P.L., Bishop-Hurley, G.J. (2018) Cattle behaviour classification from collar, halter, and ear tag sensors. *Information Processing in Agriculture*. **5**(1), 124–133.

Rajkondawar, P.G., Liu, M., Dyer, R.M., Neerchal, N.K., Tasch, U., Lefcourt, A.M., Erez, B., Varner, M.A. (2006) Comparison of Models to Identify Lame Cows Based on Gait and Lesion Scores, and Limb Movement Variables. *Journal of Dairy Science*. **89**(11), 4267–4275.

References

Rajkondawar, P.G., Tasch, U., Lefcourt, A.. P., Erez, B., Dyer, R.M., Varner, M.A. (2002) A system for identifying lameness in dairy cattle. *Applied Engineering in Agriculture*. **18**(1), 87–96.

Ramanoon, S.Z., Sadiq, M.B., Mansor, R., Syed-Hussain, S.S., Mossadeq, W.M.S. (2018) The Impact of Lameness on Dairy Cattle Welfare: Growing Need for Objective Methods of Detecting Lame Cows and Assessment of Associated Pain. In *Animal Welfare*. InTech, p. 64.

Rao, K.V., Govardhan, A., Rao, K.V.C. (2012) Spatiotemporal Data Mining: Issues, Tasks And Applications. *International Journal of Computer Science & Engineering Survey*. **3**(1), 39–52.

Raschka, S. (2018) Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. *arXiv*. (January), 1–13.

Razavi, H.A., Kurfess, T.R. (2003) Detection of Wheel and Workpiece Contact/Release in Reciprocating Surface Grinding. *Journal of Manufacturing Science and Engineering*. **125**(2), 394.

Reddy, R.V.K., Babu, U.R. (2018) A Review on Classification Techniques in Machine Learning. *International Journal of Advance Research in Science and Engineering*. **7**(30), 40–47.

Renn, N., Onyango, J., Mccormick, W. (2014) Digital infrared thermal imaging and manual lameness scoring as a means for lameness detection in cattle. *VETERINARY CLINICAL SCIENCE*. **2**(2), 16–23.

Robert, B., White, B.J., Renter, D.G., Larson, R.L. (2009) Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle. *Computers and Electronics in Agriculture*. **67**(1–2), 80–84.

ROBNIK-ˇSIKONJA, M., KONONENKO, I. (2003) Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning*. **53**(1–2), 23–69.

Rokach, L., Maimon, O. (2009) Supervised Learning. In *Data Mining and Knowledge Discovery Handbook*. Boston, MA: Springer US, pp. 133–147.

le Roux, S.P., Marias, J., Wolhuter, R., Niesler, T. (2017) Animal-borne behaviour classification for sheep (Dohne Merino) and Rhinoceros (Ceratotherium simum and Diceros bicornis). *Animal Biotelemetry*. **5**(1), 25.

Rutten, C.J., Velthuis, A.G.J., Steeneveld, W., Hogeveen, H. (2013) Invited review: Sensors to support health management on dairy farms. *Journal of Dairy Science*. **96**(4), 1928–1952.

Sayad, S. (2011) *Real Time Data Mining*. Self-Help Publishers.

Schaefer, A.L., Cook, N.J. (2013) Heat generation and the role of infrared thermography in pathological conditions. In F. Luzi, M. Mitchell, L. N. Costa, & V. Redaelli, eds. *Thermography: Current Status and Advances in Livestock Animals and in Veterinary Medicine*. Brescia, Italy, pp. 69–78.

Scott, P.D., DipECBHM, CertCHP, DSHP, FRCVS (2017) *Lameness Control in Sheep*.

Shalloo, L., O' Donovan, M., Leso, L., Werner, J., Ruelle, E., Geoghegan, A., Delaby, L., O'Leary, N. (2018) Review: Grass-based dairy systems, data and precision technologies. *animal*. **12**(s2), s262–s271.

Sharma, Seema, Agrawal, J., Agarwal, S., Sharma, Sanjeev (2013) Machine learning techniques for data mining: A survey. *2013 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2013*. (I).

Shrestha, A., Loukas, C., Le Kernec, J., Fioranelli, F., Busin, V., Jonsson, N., King, G., Tomlinson, M., Viora, L., Voute, L. (2018) Animal Lameness Detection With Radar Sensing. *IEEE Geoscience and Remote Sensing Letters*. **15**(8), 1189–1193.

Smith, D., Dutta, R., Hellicar, A., Bishop-Hurley, G., Rawnsley, R., Henry, D., Hills, J., Timms, G. (2015) Bag of Class Posteriors, a new multivariate time series classifier applied to animal behaviour identification. *Expert Systems with Applications*. **42**(7), 3774–3784.

Smith, D., Rahman, A., Bishop-Hurley, G.J., Hills, J., Shahriar, S., Henry, D., Rawnsley, R. (2016) Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems. *Computers and Electronics in Agriculture*. **131**, 40–50.

Smith, M.G., Bull, L. (2005) Genetic programming with a genetic algorithm for feature construction and selection. *Genetic Programming and Evolvable Machines*. **6**(3), 265–281.

References

Smolyakov, V. (2017) Ensemble Learning to Improve Machine Learning Results. [online]. Available from: https://blog.statsbot.co/ensemble-learning-d1dcd548e936 [Accessed April 4, 2019].

Solano, L., Barkema, H.W., Pajor, E.A., Mason, S., LeBlanc, S.J., Nash, C.G.R., Haley, D.B., Pellerin, D., Rushen, J., de Passillé, A.M., Vasseur, E., Orsel, K. (2016) Associations between lying behavior and lameness in Canadian Holstein-Friesian cows housed in freestall barns. *Journal of Dairy Science*. **99**(3), 2086–2101.

Song, X., Leroy, T., Vranken, E., Maertens, W., Sonck, B., Berckmans, D. (2008) Automatic detection of lameness in dairy cattle-Vision-based trackway analysis in cow's locomotion. *Computers and Electronics in Agriculture*. **64**(1), 39–44.

Squires, V.R., Wilson, A.D., Daws, G.T. (1972) Comparisons of the Walking Activity of Some Australian Sheep. *Proc. Aust. Soc. Anim. Prod*. **9**, 376–380.

Stokes, J.E., Leach, K.A., Main, D.C.J., Whay, H.R. (2012) An investigation into the use of infrared thermography (IRT) as a rapid diagnostic tool for foot lesions in dairy cattle. *Veterinary Journal*. **193**(3), 674–678.

Su, X., Tong, H., Ji, P. (2014) Activity recognition with smartphone sensors. *Tsinghua Science and Technology*. **19**(3), 235–249.

Susanti, Y., Zukhronah, E., Pratiwi, H., Respatiwulan, Sri Sulistijowati, H. (2017) Analysis of Chi-square Automatic Interaction Detection (CHAID) and Classification and Regression Tree (CRT) for Classification of Corn Production. *Journal of Physics: Conference Series*. **909**(1), 012041.

Tamura, T., Okubo, Y., Deguchi, Y., Koshikawa, S., Takahashi, M., Chida, Y., Okada, K. (2019) Dairy cattle behavior classifications based on decision tree learning using 3-axis neck-mounted accelerometers. *Animal Science Journal*, 1–8.

Tan, P.-N., Steinbach, M., Kumar, V. (2006) Classification : Basic Concepts , Decision Trees , and. In *Introduction to Data Mining*. pp. 145–205.

Tan, P.-N., Steinbach, M., Kumar, V. (2005) *Introduction to Data Mining*. 1st Editio. Pearson Education India.

References

Tang, J., Alelyani, S., Liu, H. (2014) Feature Selection for Classification: A Review. In C. C. Aggarwal, ed. *Data Classification Algorithms and Applications*. CRC Press, 2014, pp. 37–64.

Tasch, U., Rajkondawar, P.G. (2004) The development of a SoftSeparator^{TM} for a lameness diagnostic system. *Computers and Electronics in Agriculture*. **44**(3), 239–245.

Thorup, V.M., Munksgaard, L., Robert, P.-E., Erhard, H.W., Thomsen, P.T., Friggens, N.C. (2015) Lameness detection via leg-mounted accelerometers on dairy cows on four commercial farms. *animal*. **9**(10), 1704–1712.

Thorup, V.M., do Nascimento, O.F., Skjøth, F., Voigt, M., Rasmussen, M.D., Bennedsgaard, T.W., Ingvartsen, K.L. (2014) Short communication: Changes in gait symmetry in healthy and lame dairy cows based on 3-dimensional ground reaction force curves following claw trimming. *Journal of dairy science*. **97**(12), 7679–7684.

Too, J., Abdullah, A., Mohd Saad, N., Tee, W. (2019) EMG Feature Selection and Classification Using a Pbest-Guide Binary Particle Swarm Optimization. *Computation*. **7**(1), 12.

Turner, T.A. (1991) Thermography as an Aid to the Clinical Lameness Evaluation. *Veterinary Clinics of North America: Equine Practice*. **7**(2), 311–338.

Umstätter, C., Waterhouse, A., Holland, J.P. (2008) An automated sensor-based method of simple behavioural classification of sheep in extensive systems. *Computers and Electronics in Agriculture*. **64**(1), 19–26.

Ustev, Y.E. (2015) *User, Device, Orientation and Position Independent Human Activity Recognition on Smart Phones*. Bogazici University.

Valletta, J.J., Torney, C., Kings, M., Thornton, A., Madden, J. (2017) Applications of machine learning in animal behaviour studies. *Animal Behaviour*. **124**(December), 203–220.

Vázquez-Arellano, M., Griepentrog, H.W., Reiser, D., Paraforos, D.S. (2016) 3-D imaging systems for agricultural applications - a review. *Sensors (Switzerland)*. **16**(5), 618.

Vázquez-Diosdado, J.A., Paul, V., Ellis, K.A., Coates, D., Loomba, R., Kaler, J. (2019) A Combined Offline and Online Algorithm for Real-Time and Long-Term Classification of Sheep Behaviour: Novel Approach for Precision Livestock Farming. *Sensors*. **19**(14), 3201.

References

Vázquez Diosdado, J.A., Barker, Z.E., Hodges, H.R., Amory, J.R., Croft, D.P., Bell, N.J., Codling, E.A. (2015) Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. *Animal Biotelemetry*. **3**(1), 15.

Vazquez Diosdado, J.A., CHRISTY, C., MITSCH, J., ELLIS, K., KALER, J. (2018) SMART DATA FOR VETERINARY EPIDEMIOLOGY: COMPARING VARIOUS MACHINE LEARNING ALGORITHMS FOR DETECTION OF LAMENESS IN SHEEP. In *SOCIETY FOR VETERINARY EPIDEMIOLOGY AND PREVENTIVE MEDICINE*. University of copenhagen, pp. 94–101.

Viazzi, S., Bahr, C., Van Hertem, T., Schlageter-Tello, A., Romanini, C.E.B., Halachmi, I., Lokhorst, C., Berckmans, D. (2014) Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. *Computers and Electronics in Agriculture*. **100**, 139–147.

Viazzi, S., Bahr, C., Schlageter-Tello, a, Van Hertem, T., Romanini, C.E.B., Pluk, a, Halachmi, I., Lokhorst, C., Berckmans, D. (2013) Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle. *Journal of dairy science*. **96**(1), 257–66.

Walton, E., Casey, C., Mitch, J., Vázquez-Diosdado, J., Yan, J., Dottorini, T., Ellis, K., Winterlich, A., J, K. (2018) Evaluation of sampling frequency , window size and sensor position for classification of sheep behaviour. *Royal Society Open Science*. **5**(2), 171442.

Wang, J., He, Z., Zheng, G., Gao, S., Zhao, K. (2018) Development and validation of an ensemble classifier for real-time recognition of cow behavior patterns from accelerometer data and location data J. J. Loor, ed. *PLOS ONE*. **13**(9), e0203546.

Watanabe, N., Sakanoue, S., Kawamura, K., Kozakai, T. (2008) Development of an automatic classification system for eating, ruminating and resting behavior of cattle using an accelerometer. *Grassland Science*. **54**(4), 231–237.

de Weerd, N., van Langevelde, F., van Oeveren, H., Nolet, B.A., Kölzsch, A., Prins, H.H.T., de Boer, W.F. (2015) Deriving Animal Behaviour from High-Frequency GPS: Tracking Cows in Open and Forested Habitat D. A. Driscoll, ed. *PLOS ONE*. **10**(6), e0129030.

Weigele, H.C., Gygax, L., Steiner, A., Wechsler, B., Burla, J.-B. (2018) Moderate lameness leads to marked behavioral changes in dairy cows. *Journal of Dairy Science*. **101**(3), 2370–2382.

Welsh, E.M., Gettinby, G., Nolan, A.M. (1993) Comparison of a visual analog scale and a numerical rating scale for assessment of lameness, using sheep as a model. *American Journal of Veterinary Research*. **54**(6), 976–983.

Winter, A.C. (2008) Lameness in sheep. *Small Ruminant Research*. **76**(1–2), 149–153.

Wood, S., Lin, Y., Knowles, T.G., Main, D.C.J. (2015) Infrared thermometry for lesion monitoring in cattle lameness. *Veterinary Record*. **176**(12), 308–308.

Wu, X., Kumar, V., Ross, Q.J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A., Liu, B., Yu, P.S., Zhou, Z.H., Steinbach, M., Hand, D.J., Steinberg, D. (2008) *Top 10 algorithms in data mining*.

Yan-yan Song, Ying Lu (2015) Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*. **27**(2), 130–135.

Yunta, C., Guasch, I., Bach, A. (2012) Short communication: Lying behavior of lactating dairy cows is influenced by lameness especially around feeding time. *Journal of Dairy Science*. **95**(11), 6546–6549.

Zhang, S., Mccullagh, P., Callaghan, V. (2014) An Efficient Feature Selection Method for Activity Classification. In *2014 International Conference on Intelligent Environments*. IEEE, pp. 16–22.

Zhang, S., Zhang, C., Yang, Q. (2003) Data preparation for data mining. *Applied Artificial Intelligence*. **17**(5–6), 375–381.

Zhao, K., Bewley, J.M., He, D., Jin, X. (2018) Automatic lameness detection in dairy cattle based on leg swing analysis with an image processing technique. *Computers and Electronics in Agriculture*. **148**(December 2017), 226–236.

Appendix A. Research Ethical Approval

MOULTON COLLEGE - ETHICAL APPROVAL REQUEST & RISK ASSESSMENT

Approved by Chair of Moulton College Research Committee: 27/4/16.

Date of Assessment:	Review date:
Tasks covered by this assessment	
Location (s)	University of Northampton / Moulton College / Sheep Unit
Name of assessor:	Signature:
Title of project	The use of multivariable wireless sensor data to early detect lameness in sheep
Principal investigator	Zainab Al-Rubaye
Project first supervisor	Ali Al-Sherbaz
Project second supervisor	Wanda McCormick
External supervisor	
Director of studies	Scott Turner
University / Area	University of Northampton / Computer Science and Impressive Technologies
Location of work	Moulton College
Additional contact details	Zainab.al-rubaye@northampton.ac.uk

ETHICS & RA 2014

MOULTON COLLEGE - ETHICAL APPROVAL REQUEST & RISK ASSESSMENT

Moulton College
NORTHAMPTONSHIRE

Brief description of project or activity:

Newly developed sensor technology utilises the idea of automatically monitoring objects; animals, for example, can be monitored to determine the physiological and behavioural indicators, which are subsequently used as inputs to data analysis algorithms. Automated methods to monitor the farm bring many advantages to the farmer in terms of time saving, increasing flock size and sensitivity to detect conditions such as lameness. The type of data collected from the sensor used for recording animal's behaviour depend on the sensor's features and functionality. The sensor that will be used to conduct this research is immensely accurate and sensitive. It provides 3-aixs acceleration, 3-aixs angular velocity, 3-aixs angles (Roll, Pitch, and Heading), longitude, latitude and time of reading which can be set up according to the demanded accuracy.

Brief outline of aims and objectives of research	Lameness has a negative influence on both sheep welfare and farm economy. Therefore, preclinical detection of lameness on farm will increase the level of protection regarding sheep health and farm commerce decline. The main causes of lameness are interdigital dermatitis and foot rot. The latter is a bacterial disease that can be easily transmitted from one sheep to another via pasture. This study will help to remotely record spatial- temporal data related to the sheep behaviour without human interference and to indicate the early signs of lameness by taking advantages of smart wearable sensor technology. This will help the shepherd to prevent the sheep from reaching a severe lameness stage when it becomes difficult to tackle and obviously that will gradually affect the annual farm productivity. This research is developing an automated model to early detect lameness in sheep by relying on the analysis of data that

MOULTON COLLEGE - ETHICAL APPROVAL REQUEST & RISK ASSESSMENT

	will be retrieved from the mounted sensor on their neck collar according to smart predictive data mining techniques.
Brief description of methods (include species and number of animals used if appropriate)	Firstly, few sample of (sound, mildly lame, moderately lame, lame and severely lame) sheep; five sheep for example, will be monitored for a period; one hour for instance, via a sensor that will be mounted on sheep neck within a collar. Then, the mounted sensor will be removed to collect the behavioural data that are stored in sensor memory as an Excel sheet file for later analysis. According to intelligence predictive data mining techniques, the collected data will be analysed to distinguish among the pre-mentioned five lameness scales. Furthermore, a video footage will be taken at the same time of sensor's recording for the later validation process.

ETHICAL CONSIDERATIONS	**Precautions**
Stress caused to sheep during capture & handling	The sheep will be rounded up by the shepherd using standard practices. The researcher will be trained in methods for handling sheep to minimise stress and will be aware of stress indicators. Should an individual animal show excessive signs of stress during handling, the process will be stopped and attempted one further time once the animal has calmed down. Should it still be deemed to be causing excessive stress then the individual will no longer be included in the study.
Injury / discomfort caused by sensor	The sensor will be held on the animal using a standard collar or leg band that is used

	already by industry for that species so no issues are expected. The weight of the sensor will be confirmed that it should not cause any discomfort for the animal for the data collection period. The sheep will monitored during the data collection and the researcher will be able to intervene to remove the sensor should any unforeseen situation arise.

Notes on discussion by panel/ additional precautions to be put in place

MOULTON COLLEGE - ETHICAL APPROVAL REQUEST & RISK ASSESSMENT

Moulton College NORTHAMPTONSHIRE

RISK ASSESSMENT

Persons at Risk : Researcher		**Can the risk activity be eliminated or prevented immediately?**	**Do the College's existing preventative or protective measures reduce the risk to insignificant or low?**
Hazard Injury from sheep (e.g kicking) Transmission of zoonotic infection (e.g. *Salmonella, E.coli, Toxoplasma*)	Nature of risk When working with livestock there is always the risk of injury. As long as good hygiene and biosecurity is practiced the risk should be minimal.	No	Yes

Existing controls including documented policies, training etc.

- Farm induction was scheduled in 14 Apr. 2016 at 1 pm. The researcher was trained in farm practices and handling sheep and had time to 'work-shadow' with the shepherd before commencing data collection to ensure a safe working knowledge of practices.
- PPE will be worn at all times when on the farm (overalls and steel-toe capped wellies) and this will be appropriately cleaned / laundered between visits.
- Biosecurity protocols will be followed (e.g use of foot dips) and good personal hygiene will be practiced following farm visits.
- Should the researcher become pregnant during the course of the study she will inform their supervisor and cease all contact with the sheep.

Additional controls put in place:

ETHICS & RA 2014

Moulton College NORTHAMPTONSHIRE

MOULTON COLLEGE - ETHICAL APPROVAL REQUEST & RISK ASSESSMENT

ETHICAL CONSIDERATIONS & RISK ASSESSMENT Reviewed and authorised by:	Signature	Date
Director of HE:	[signature]	27/4/16.
First project supervisor: Ali Al-Sherbaz	A Sherbaz	26/4/2016
Second project supervisor: Wanda McCormick	[signature]	27/4/16
External panel member: JAMES LITTLEMORE	[signature]	27/4/16

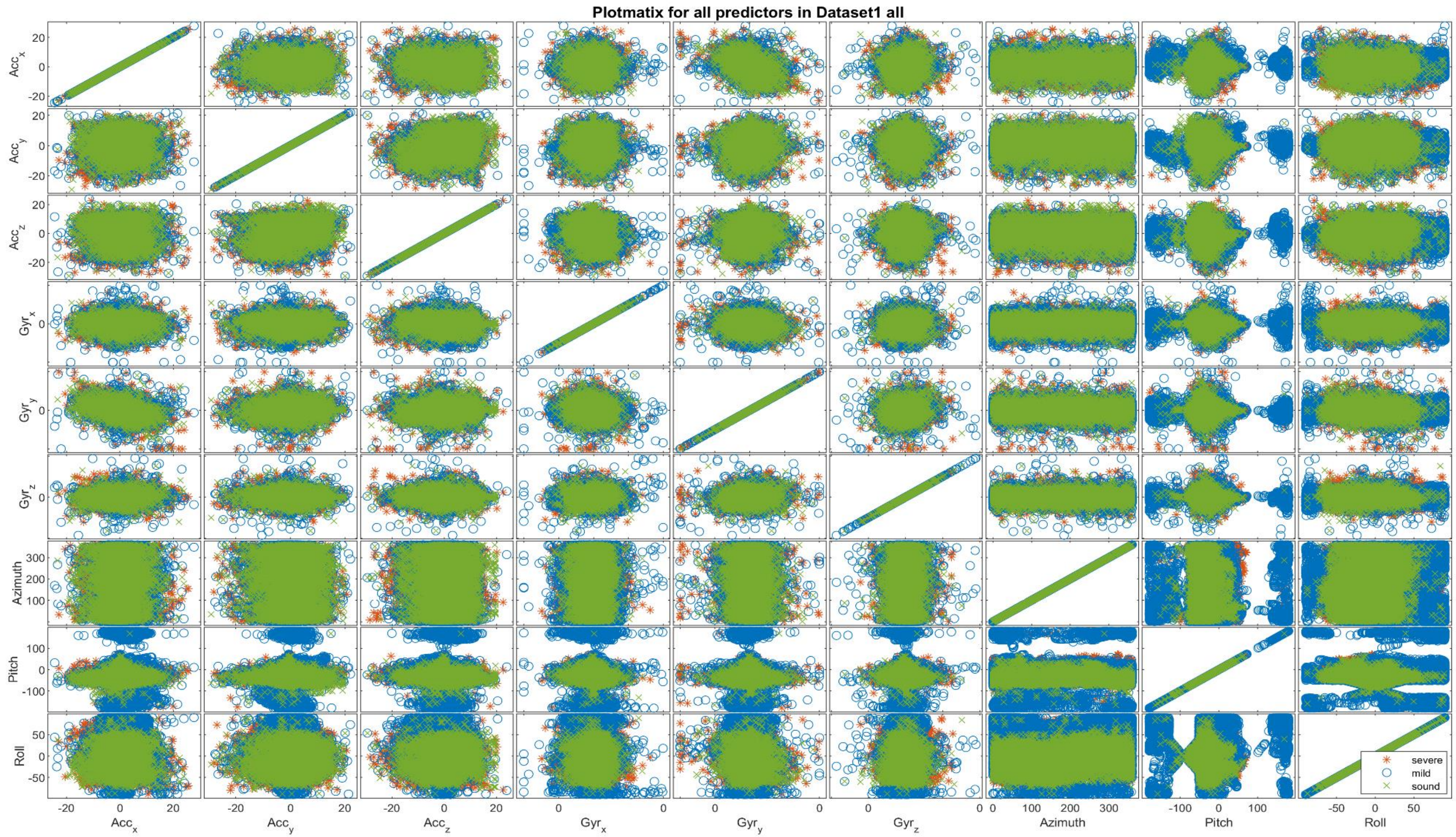
Risk Assessment: Assess potential severity of risk (indicate chosen number) then assess likelihood of risk occurring (indicate chosen number).

Severity x Likelihood = Risk (indicate the resulting number in the risk section).

POTENTIAL SEVERITY?		LIKELIHOOD OF OCCURENCE?		RISK	
Catastrophic (death, widespread illness)	4	Probable (imminent, shortly)	4	9>16	HIGH
Critical (severe injury/damage	3	Reasonably probable (will occur in time)	3	4>8	MEDIUM
Marginal (injury,/damage, not severe)	2	Remote (may occur in time)	2	1>3	LOW
Negligible (minor first aid)	1	Extremely remote (unlikely to occur)	1		

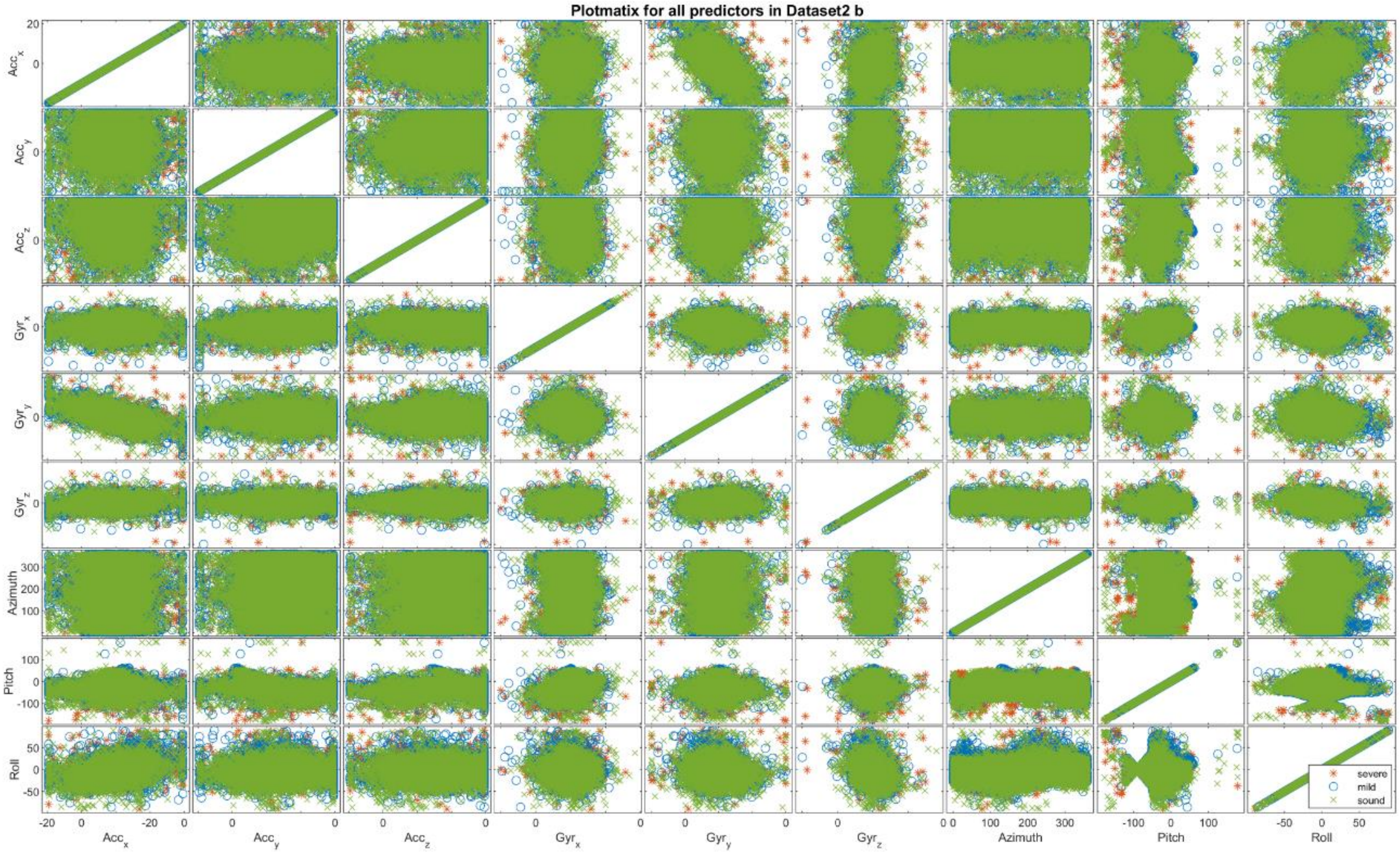
ETHICS & RA 2014

Appendix B. **Results of Sheep Raw Data Plotting**



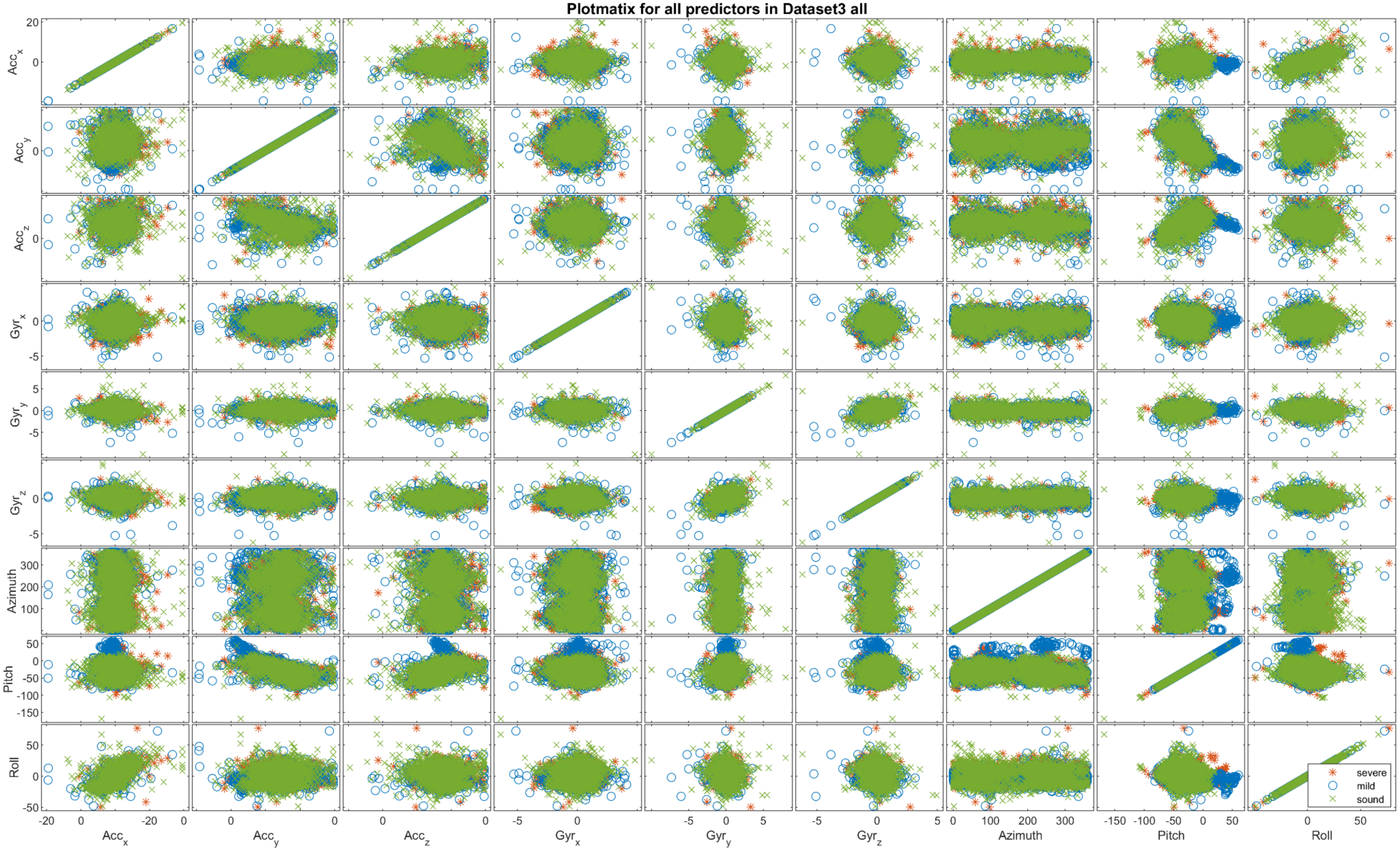
Appendix B. 1 Scatter Plot matrix for raw Sheep DataSet1_all, where *, o, and x represent severe, mild, and sound $Classes$ in the DataSet.

Appendix B. Results of Sheep Raw Data Plotting



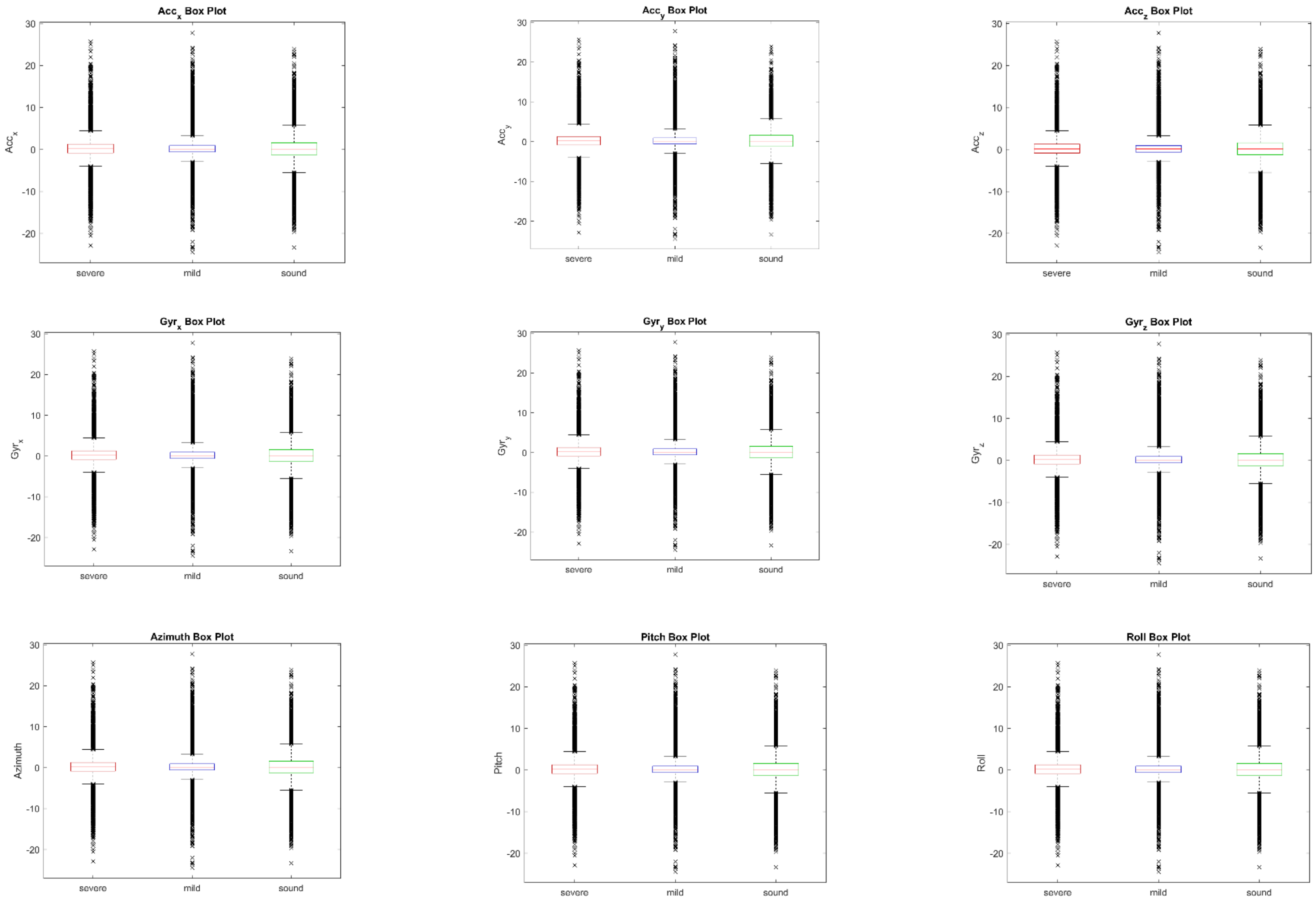
Appendix B. 2 Scatter Plot matrix for raw Sheep DataSet2_b, where *, o, and x represent severe, mild, and sound *Classes* in the DataSet.

Appendix B. Results of Sheep Raw Data Plotting

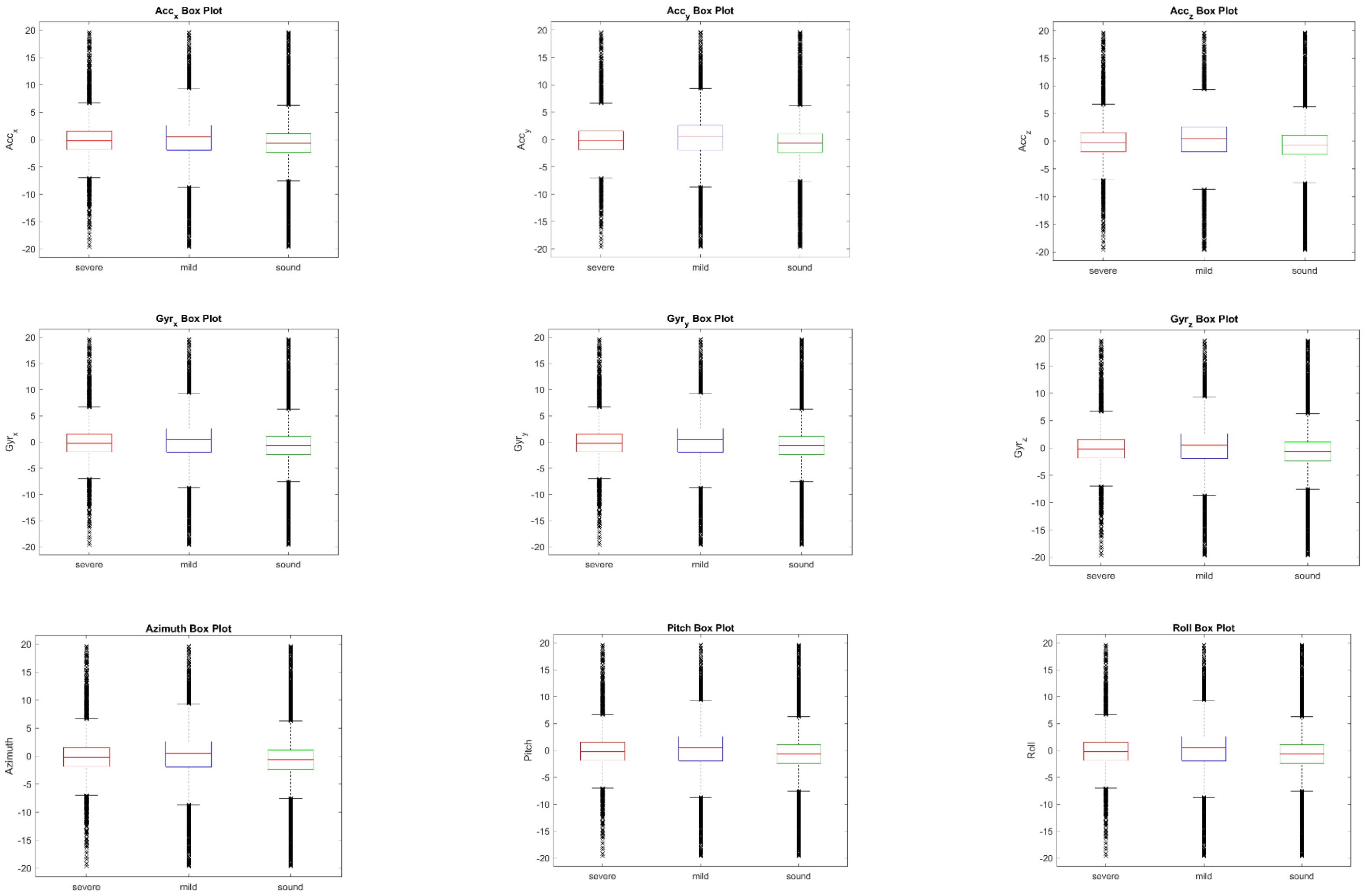


Appendix B. 3 Scatter Plot matrix for raw Sheep DataSet3_all, where *, o, and x represent severe, mild, and sound *Classes* in the DataSet.

Appendix B. Results of Sheep Raw Data Plotting

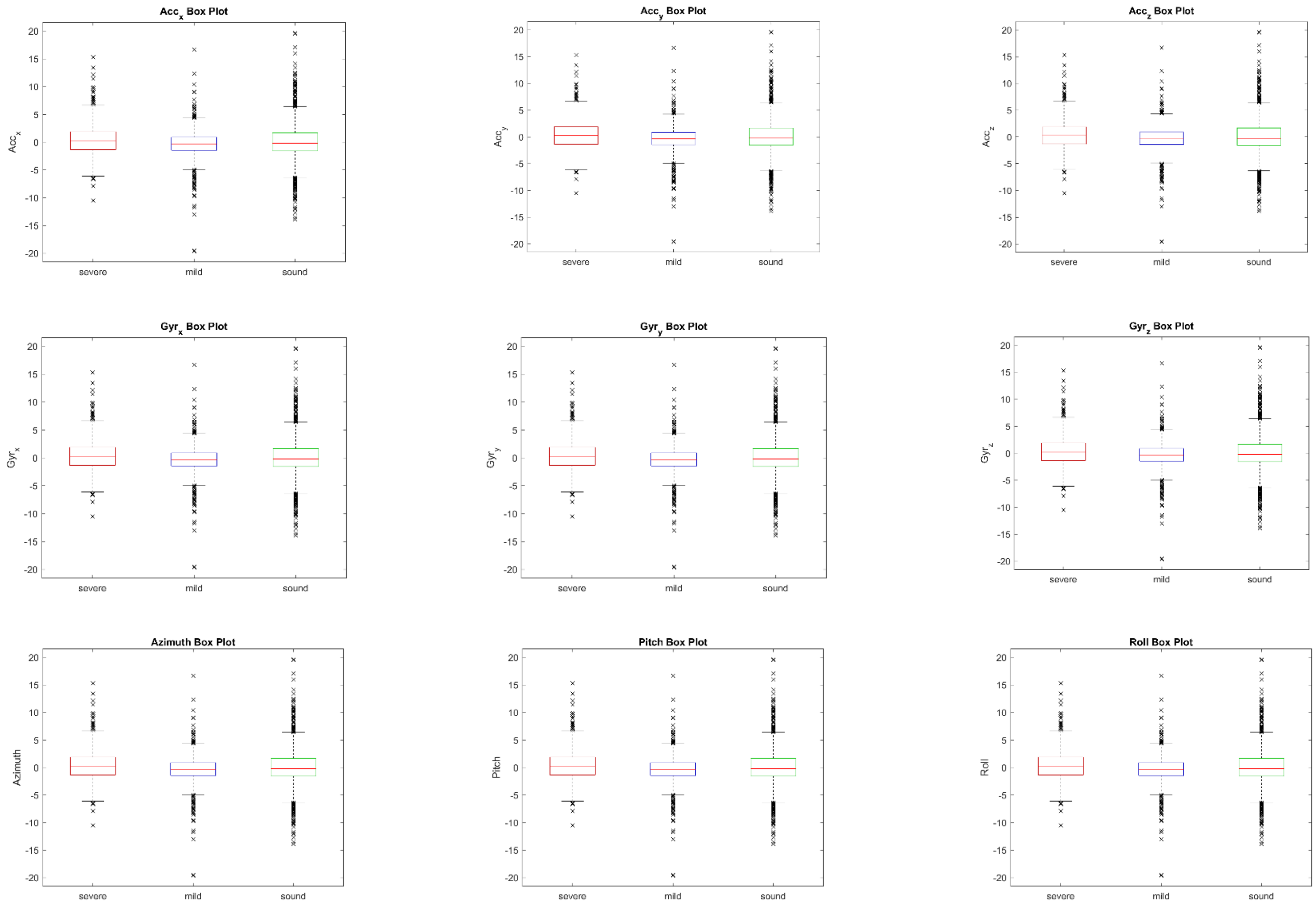


Appendix B. 4 Box Plots for each predictor in raw sheep DataSet1_all.



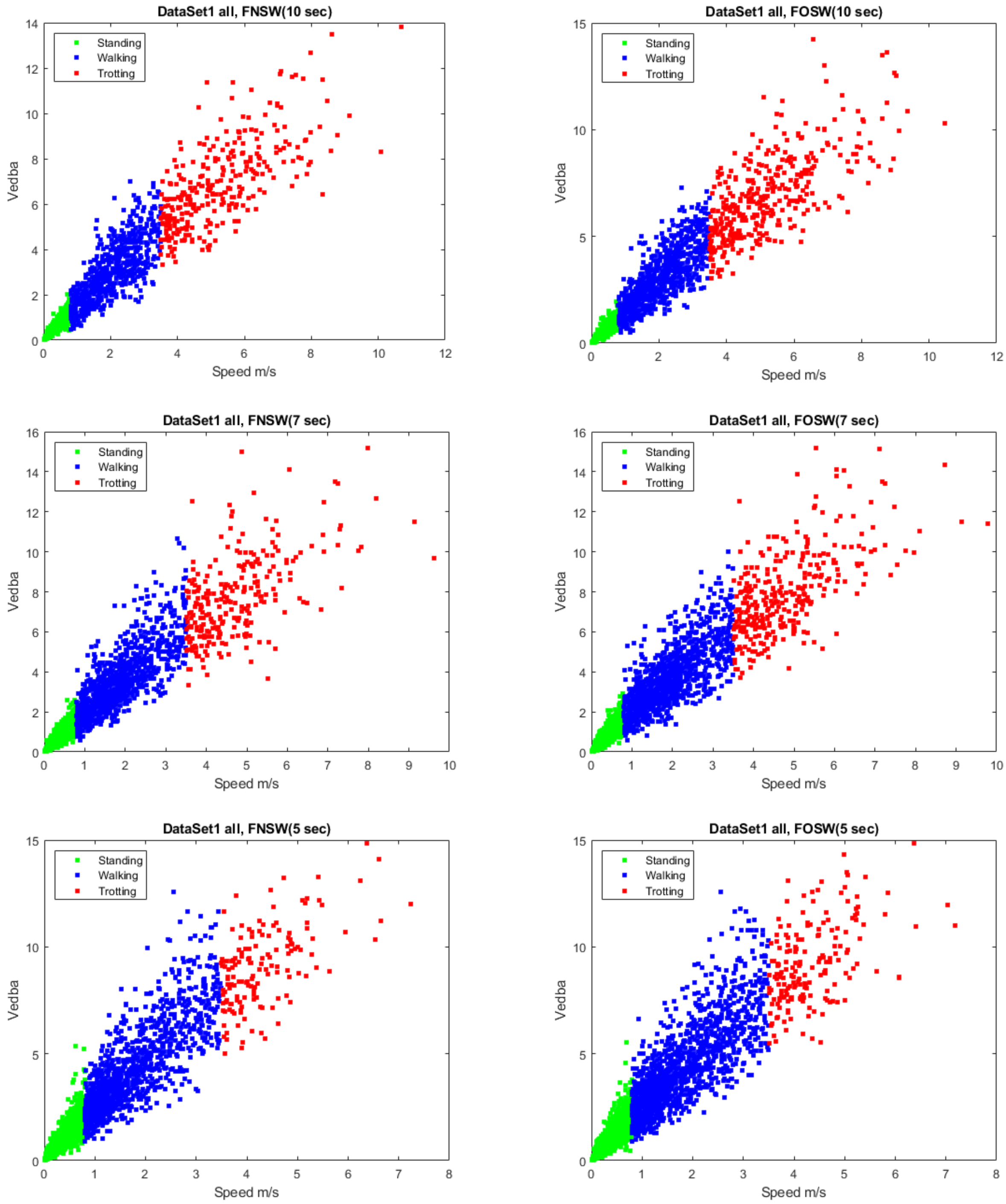
Appendix B. 5 Box Plots for each predictor in raw sheep DataSet2_b.

Appendix B. Results of Sheep Raw Data Plotting

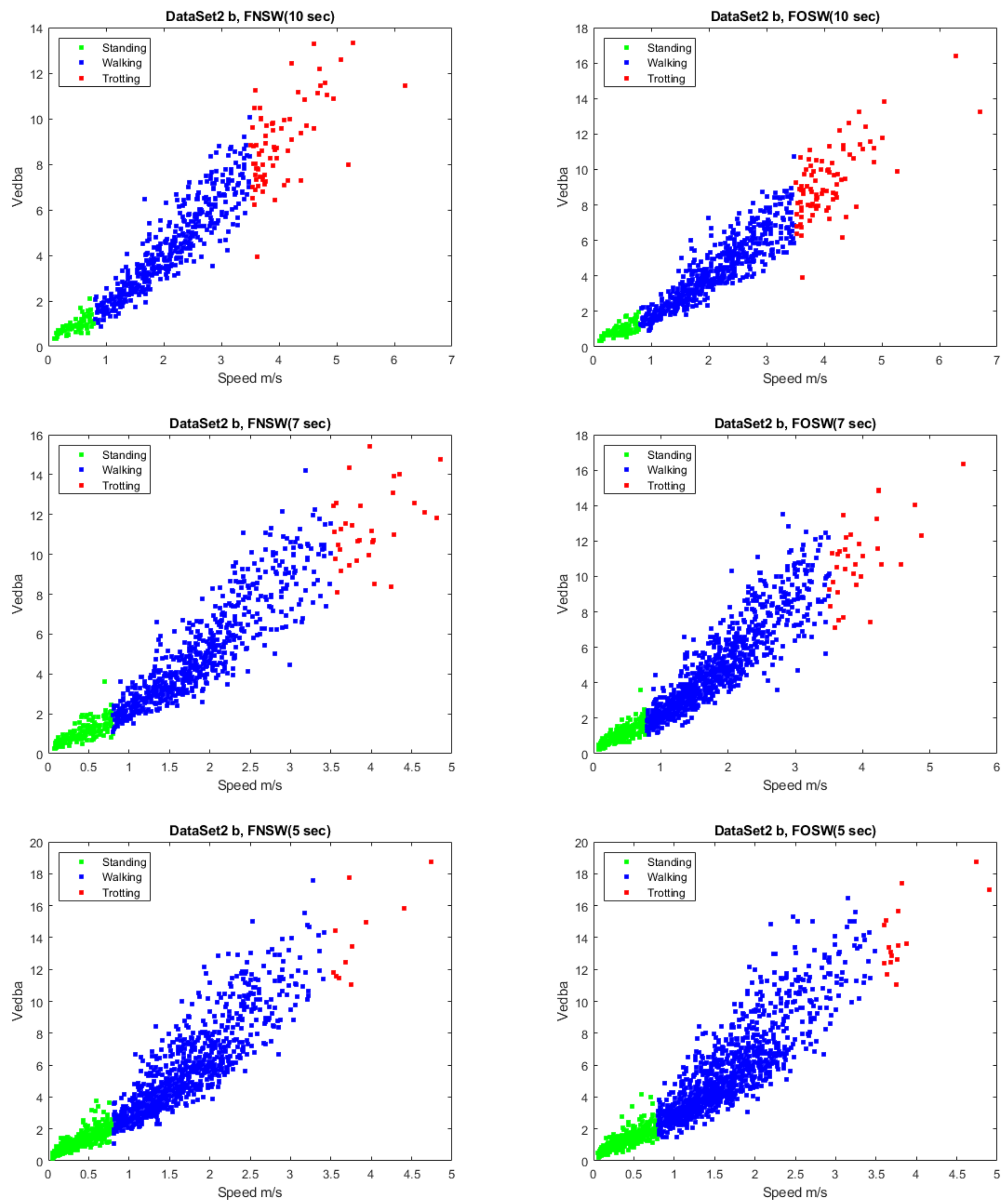


Appendix B. 6 Box Plots for each predictor in raw sheep DataSet3_all.

Appendix C. Sheep movements plots for Standing, Walking, and Trotting Segments

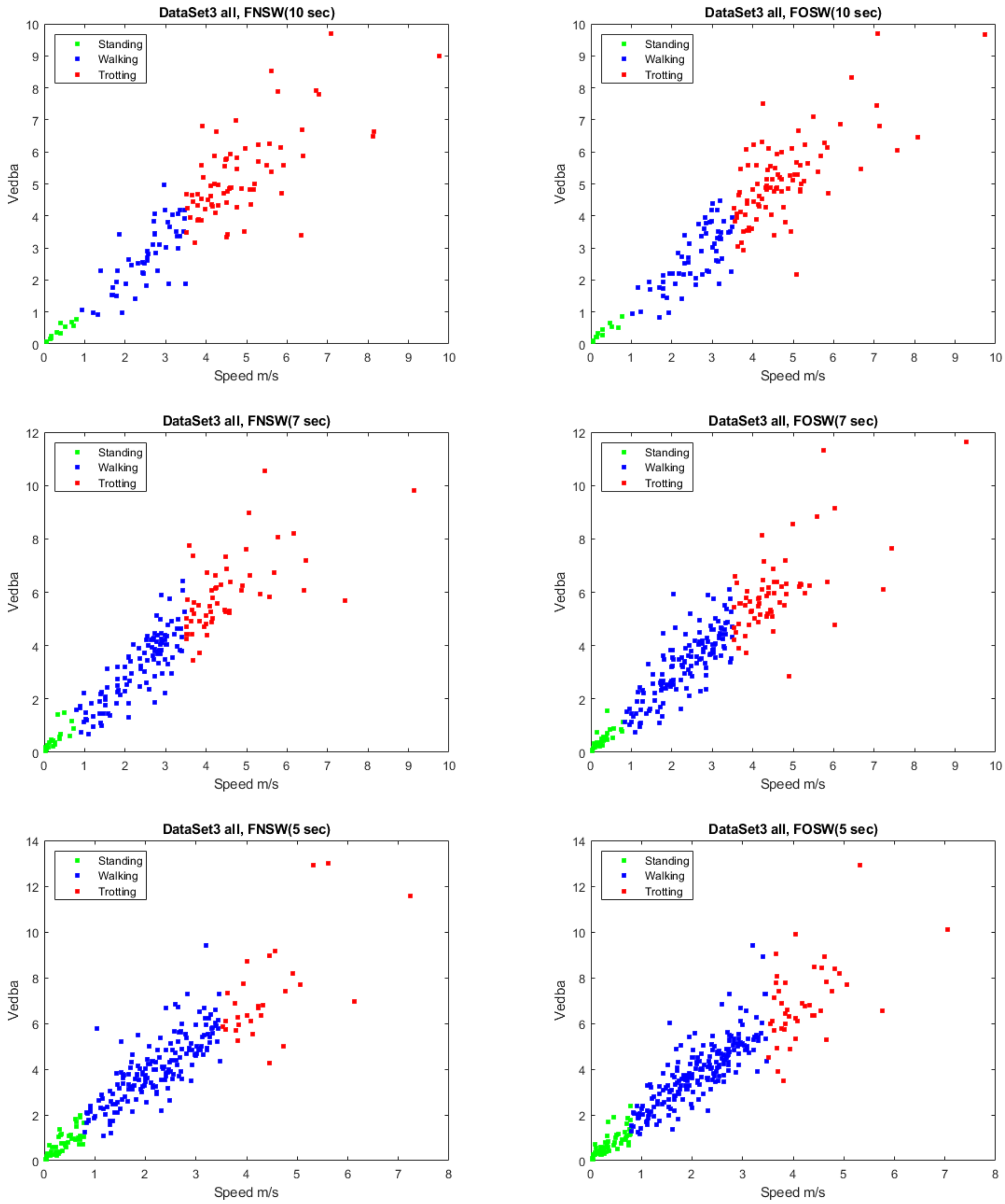


Appendix C. 1 Scatter plots of the DataSet1_all, where movement's classification is done over (10 $sec.$, 7 $sec.$, and 5 $sec. window$) for two segmentation approaches (FNSW and FOSW).



Appendix C. 2 Scatter plots of the DataSet2_b, where movement's classification is done over (10 sec, 7 sec, and 5 $sec. window$) for two segmentation approaches (FNSW and FOSW).

Appendix C. Sheep movements plots for Standing, Walking, and Trotting Segments

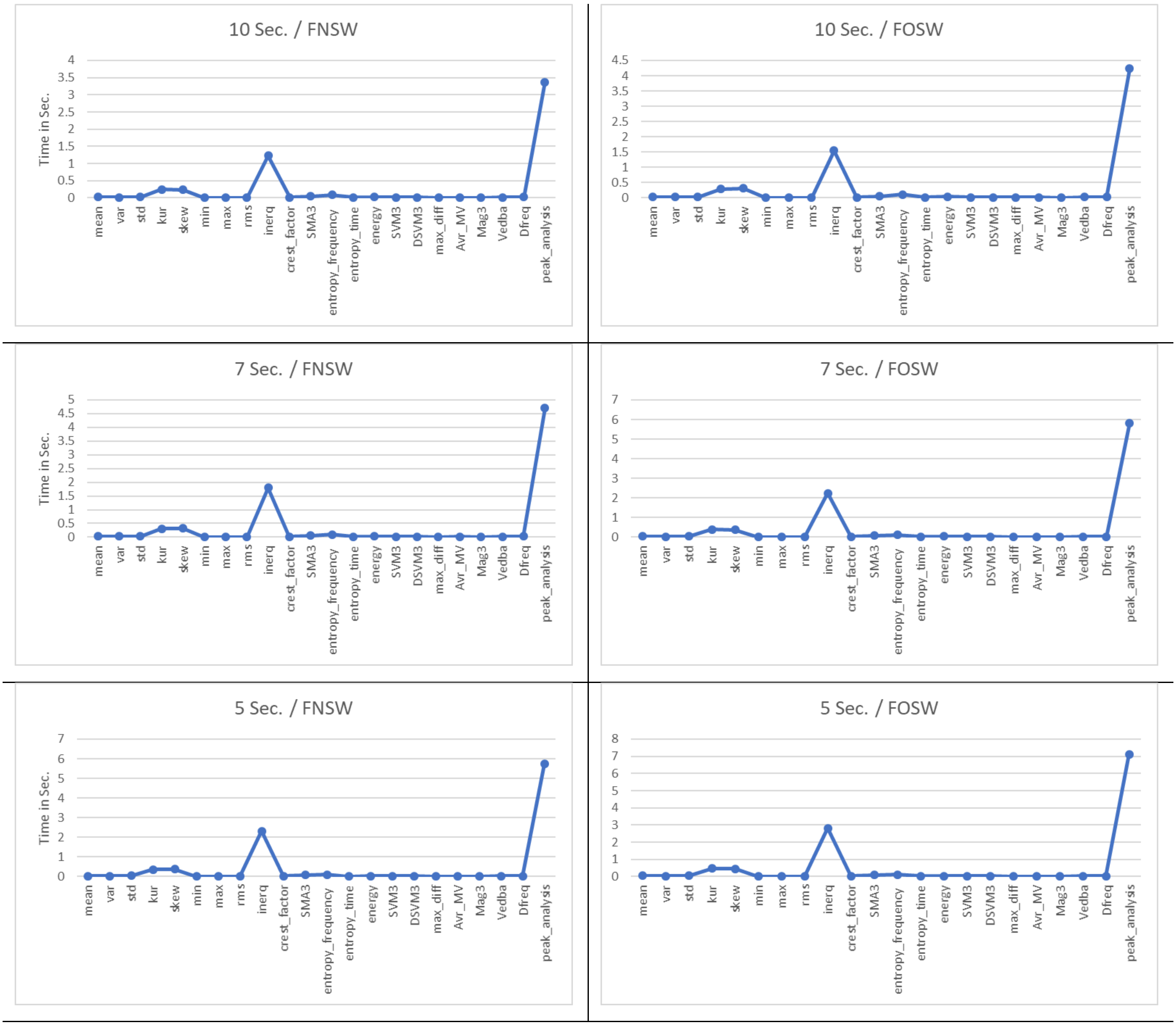


Appendix C. 3 Scatter plots of the DataSet3_all, where movement's classification is done over (10 sec, 7 sec, and 5 $sec.window$) for two segmentation approaches (FNSW and FOSW).

Appendix D. **Time Calculation for the Extracted Features**

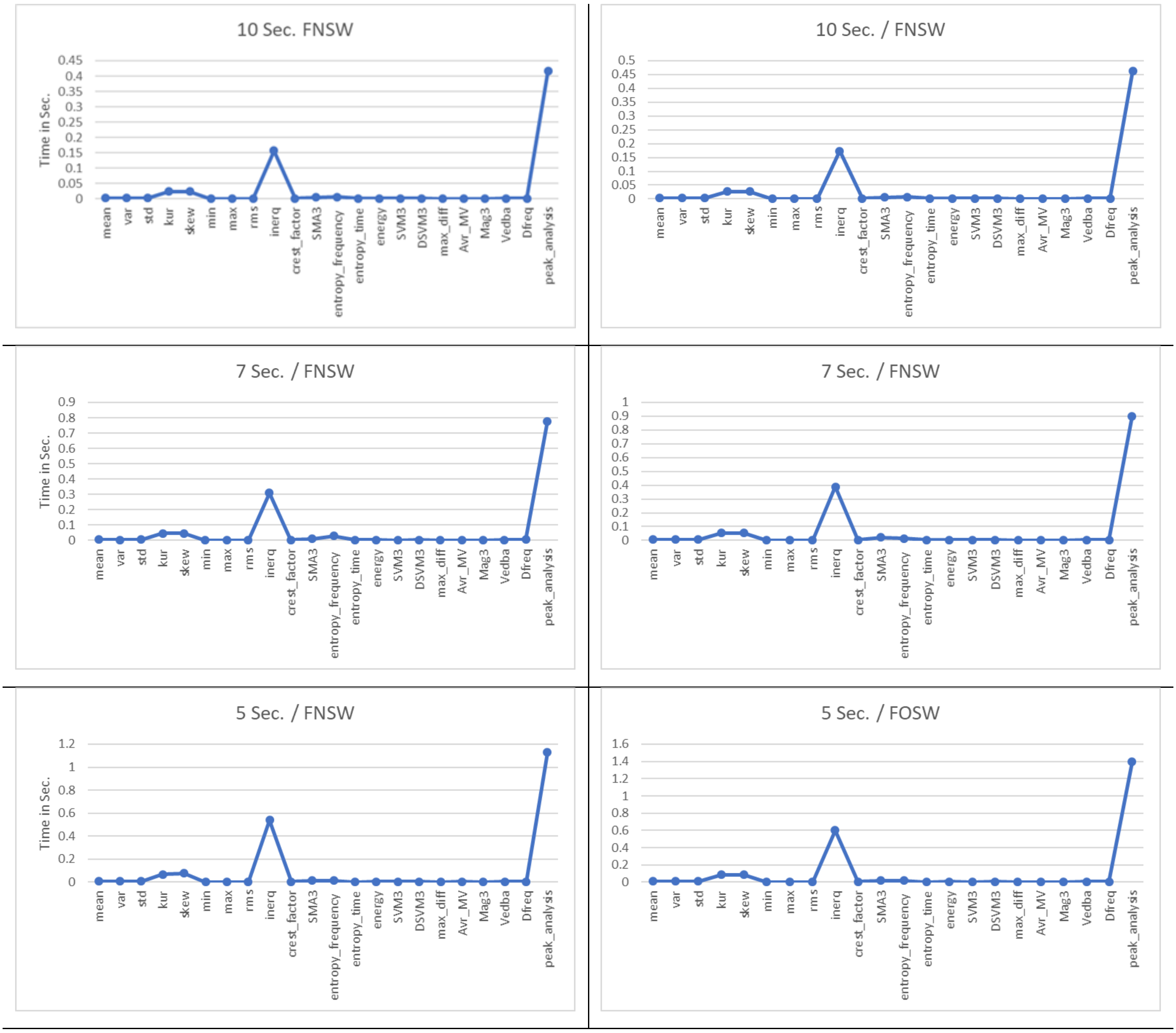
Appendix D. 1 Execution time of features for DataSet1_all (5 Hz).

Appendix D. Time Calculation for the Extracted Features



Appendix D. 2 Execution time of features for DataSet2_b (10 Hz).

Appendix D. Time Calculation for the Extracted Features



Appendix D. 3 Execution time of features for DataSet3_all (4 Hz).

Appendix E. **Ranked Features Tables for Sheep DataSets**

Appendix E. 1 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet1_all** over 10 $sec.window$.

	ReliefF		**GA**		**RF**	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Energy_Roll	Energy_Roll	Mean_Acc_x	Mean_Acc_y	Max_Roll	Max_diff_Pitch
2	Rms_Roll	Rms_Roll	Mean_Acc_y	Mean_Azimuth	Mean_Roll	Max_Roll
3	Dfreq_Roll	Dfreq_Roll	Mean_Acc_z	Mean_Pitch	Max_diff_Pitch	Mean_Roll
4	Max_Roll	Max_Roll	Mean_Azimuth	Mean_Gyr_x	Min_Pitch	Max_Gyr_x
5	Mean_Roll	Mean_Roll	Mean_Pitch	Var_Acc_x	Max_Gyr_x	Max_Acc_y
6	Min_Roll	Min_Roll	Mean_Roll	Var_Acc_y	Rms_Roll	Min_Pitch
7	Cf_Roll	Cf_Roll	Mean_Gyr_x	Var_Acc_z	Entropy_Gyr_y	Entropy_Gyr_x
8	Max_Pitch	Max_Pitch	Mean_Gyr_y	Var_Azimuth	Cf_Roll	Skew_Acc_y
9	Entropy_Pitch	Entropy_Pitch	Var_Acc_x	Var_Pitch	Max_diff_Gyr_z	Rms_Pitch
10	Rms_Pitch	Rms_Pitch	Var_Acc_y	Var_Gyr_x	Mean_Acc_z	Max_diff_Gyr_z
11	Max_diff_Pitch	Max_diff_Pitch	Var_Azimuth	Var_Gyr_y	Entropy_Gyr_z	Max_Pitch
12	Energy_Pitch	Min_Pitch	Var_Pitch	Var_Gyr_z	Max_diff_Acc_z	Rms_Roll
13	Dfreq_Pitch	Energy_Pitch	Var_Gyr_x	Std_Acc_x	Cf_Gyr_x	Interq_Gyr_y
14	MV_Gyr	MV_Gyr	Var_Gyr_y	Std_Acc_z	Entropy_Gyr_x	Min_Acc_z
15	Entropy_Roll	Dfreq_Pitch	Var_Gyr_z	Std_Azimuth	Min_Acc_y	Max_diff_Gyr_x
16	Cf_Pitch	Entropy_Roll	Std_Acc_x	Std_Pitch	Max_diff_Acc_y	Min_Gyr_z
17	Max_diff_Gyr_z	Max_diff_Gyr_z	Std_Acc_z	Std_Roll	Var_Gyr_z	Mean_Acc_x
18	Std_Pitch	Cf_Pitch	Std_Pitch	Std_Gyr_x	Min_Gyr_x	Max_diff_Acc_z
19	MV_Acc	Std_Pitch	Std_Roll	Kur_Acc_x	Var_Acc_y	Min_Gyr_x
20	Min_Pitch	DSVM_Gyr	Std_Gyr_y	Kur_Acc_z	Var_Gyr_x	Cf_Roll
21	mag_Ang	MV_Acc	Std_Gyr_z	Kur_Azimuth	Mean_Pitch	Mean_Acc_z
22	Entropy_Gyr_y	Entropy_Gyr_y	Kur_Acc_x	Kur_Pitch	Rms_Gyr_x	Kur_Gyr_z

23	Max_diff_Acc_y	Max_diff_Acc_y	Kur_Acc_y	Kur_Gyr_x	Max_diff_Acc_x	Dfreq_Roll
24	DSVM_Gyr	Mean_Pitch	Kur_Pitch	Kur_Gyr_y	Dfreq_Gyr_x	Mean_Pitch
25	Mean_Pitch	Entropy_Gyr_x	Skew_Acc_y	Skew_Acc_x	Entropy_Acc_y	Var_Acc_y
26	Entropy_Gyr_x	Entropy_Gyr_z	Skew_Pitch	Skew_Azimuth	Skew_Acc_y	Dfreq_Pitch
27	SVM_Angle	SVM_Angle	Min_Acc_z	Skew_Pitch	Interq_Gyr_y	Cf_Gyr_x
28	SMA_Angle	SMA_Angle	Min_Azimuth	Skew_Gyr_x	Dfreq_Acc_z	Min_Roll
29	Entropy_TimeD_Ang	Entropy_TimeD_Ang	Min_Pitch	Min_Acc_y	Max_diff_Gyr_x	Var_Gyr_z
30	Max_Azimuth	Min_Gyr_z	Min_Gyr_z	Min_Acc_z	Rms_Pitch	Max_diff_Acc_x
31	Entropy_Acc_y	Entropy_Acc_y	Max_Acc_x	Min_Pitch	Dfreq_Roll	Var_Pitch
32	Min_Gyr_z	Std_Gyr_z	Max_Acc_y	Min_Gyr_x	Min_Roll	Mean_Acc_y
33	Dfreq_Azimuth	Max_Acc_y	Max_Pitch	Min_Gyr_z	Min_Gyr_y	Kur_Acc_z
34	Mean_Azimuth	Rms_Gyr_z	Rms_Acc_y	Max_Acc_x	Cf_Pitch	Var_Gyr_x
35	Vedb_Angle	Entropy_Acc_x	Rms_Roll	Max_Acc_y	Max_Gyr_z	Dfreq_Gyr_z
36	Std_Gyr_z	Vedb_Angle	Interq_Acc_y	Max_Acc_z	Dfreq_Acc_y	Var_Acc_x
37	Entropy_Gyr_z	Entropy_Acc_z	Interq_Acc_z	Max_Azimuth	Highest_peak_Acc_y	Min_Gyr_y
38	Std_Acc_y	Std_Acc_y	Interq_Pitch	Max_Roll	Kur_Gyr_y	Skew_Acc_z
39	Rms_Gyr_z	Var_Pitch	Interq_Gyr_x	Max_Gyr_x	Dfreq_Gyr_z	Max_diff_Acc_y
40	Entropy_Acc_x	Max_diff_Gyr_x	Interq_Gyr_y	Max_Gyr_y	Widest_Peak_Acc_x	Highest_peak_Pitch
41	Var_Acc_y	Max_Azimuth	Interq_Gyr_z	Max_Gyr_z	Var_Pitch	Entropy_Gyr_y
42	Rms_Azimuth	Max_diff_Acc_z	Cf_Acc_x	Rms_Roll	Dfreq_Pitch	Kur_Acc_y
43	Entropy_Acc_z	AV_Ang	Cf_Roll	Rms_Gyr_x	Max_diff_Gyr_y	Interq_Acc_x
44	Dfreq_Gyr_y	Var_Acc_y	SMA_Acc	Interq_Acc_x	Mean_Acc_y	Skew_Gyr_z
45	Entropy_Azimuth	Interq_Gyr_y	SMA_Angle	Interq_Acc_z	Max_Acc_y	Kur_Gyr_y
46	Max_Acc_y	Max_diff_Gyr_y	SMA_Gyr	Interq_Azimuth	Skew_Roll	Max_Acc_z
47	Energy_Azimuth	Max_Acc_z	Entropy_Acc_y	Interq_Roll	Rms_Acc_y	Interq_Acc_z
48	Min_Azimuth	DSAM_Angle	Entropy_Gyr_x	Interq_Gyr_x	Var_Acc_x	Entropy_Acc_y
49	Highest_peak_Gyr_y	Highest_peak_Pitch	Entropy_Gyr_y	Interq_Gyr_y	DSVM_Gyr	DSVM_Gyr
50	Max_diff_Gyr_y	mag_Ang	Entropy_Gyr_z	Interq_Gyr_z	Kur_Gyr_z	MV_Gyr
51	Max_Gyr_z	Mean_Azimuth	Entropy_TimeD_Acc	Cf_Acc_y	Max_Pitch	Interq_Gyr_x

52	Highest_peak_Pitch	Dfreq_Azimuth	Energy_Acc_x	Cf_Pitch	Rms_Acc_x	Entropy_Roll
53	Var_Pitch	Max_Gyr_z	Energy_Acc_y	Cf_Roll	Interq_Azimuth	Min_Acc_y
54	Highest_peak_Acc_x	Min_Acc_x	Energy_Pitch	Cf_Gyr_x	Skew_Acc_x	Entropy_Azimuth
55	Max_diff_Gyr_x	SMA_Acc	Energy_Roll	Cf_Gyr_y	Interq_Roll	Skew_Gyr_y
56	Max_diff_Acc_z	Rms_Azimuth	Energy_Gyr_z	SMA_Acc	Entropy_Pitch	MV_Acc
57	Interq_Gyr_y	SVM_Acc	SVM_Gyr	Entropy_Acc_x	Rms_Gyr_z	Dfreq_Acc_x
58	Std_Azimuth	Highest_peak_Acc_x	Max_diff_Acc_x	Entropy_Acc_y	Cf_Acc_z	Kur_Roll
59	Interq_Azimuth	Dfreq_Gyr_y	Max_diff_Pitch	Entropy_Roll	Min_Acc_x	Entropy_Acc_x
60	Interq_Gyr_z	Entropy_Azimuth	Max_diff_Gyr_z	Entropy_Gyr_z	Var_Roll	Widest_Peak_Acc_z
61	Std_Acc_x	Highest_peak_Gyr_y	AV_Ang	Entropy_TimeD_Gyr	Widest_Peak_Pitch	Entropy_Acc_z
62	Rms_Acc_z	Min_Acc_y	mag_Gyr	Energy_Acc_z	Cf_Azimuth	Dfreq_Gyr_y
63	AV_Ang	Var_Gyr_z	Dfreq_Acc_z	Energy_Azimuth	nPeaks_Gyr_z	Rms_Acc_x
64	SMA_Acc	Cf_Acc_z	Dfreq_Azimuth	Energy_Roll	Var_Acc_z	DSAM_Angle
65	SVM_Acc	Energy_Azimuth	Dfreq_Pitch	Energy_Gyr_y	MV_Acc	Kur_Gyr_x
66	DSAM_Angle	Std_Gyr_y	Dfreq_Roll	Energy_Gyr_z	nPeaks_Azimuth	Var_Roll
67	Max_Acc_x	Energy_Gyr_z	Dfreq_Gyr_y	SVM_Acc	Skew_Acc_z	Dfreq_Gyr_x
68	Var_Gyr_z	Entropy_TimeD_Acc	Dfreq_Gyr_z	SVM_Angle	DSVM_Acc	Rms_Gyr_y
69	Min_Acc_x	Rms_Gyr_y	Highest_peak_Acc_x	SVM_Gyr	AV_Ang	Rms_Acc_y
70	Highest_peak_Acc_y	Rms_Acc_y	Widest_Peak_Acc_y	DSVM_Acc	Entropy_Roll	Cf_Acc_y
71	Min_Acc_y	Rms_Acc_z	Highest_peak_Acc_y	Max_diff_Acc_y	Min_Acc_z	Avr_peak_time_Acc_z
72	Rms_Acc_y	Max_Acc_x	Avr_peak_time_Acc_y	Max_diff_Acc_z	Skew_Gyr_z	Max_diff_Gyr_y
73	Entropy_TimeD_Acc	Cf_Acc_y	nPeaks_Acc_z	Max_diff_Pitch	Kur_Roll	Rms_Gyr_z
74	Rms_Gyr_y	Std_Acc_z	Widest_Peak_Acc_z	Max_diff_Roll	Interq_Acc_x	AV_Ang
75	Energy_Gyr_z	Interq_Gyr_z	Highest_peak_Acc_z	Max_diff_Gyr_y	Max_diff_Roll	Vedb_Acc
76	Highest_peak_Azimuth	Interq_Pitch	Avr_peak_time_Acc_z	MV_Gyr	Interq_Gyr_z	Dfreq_Acc_z
77	Std_Gyr_y	Min_Gyr_y	Avr_peak_time_Azimuth	mag_Acc	Highest_peak_Acc_z	Entropy_Gyr_z
78	Min_Gyr_y	Min_Azimuth	Widest_Peak_Pitch	mag_Gyr	Mean_Acc_x	Max_diff_Roll
79	Std_Acc_z	Highest_peak_Acc_y	Highest_peak_Pitch	Vedb_Gyr	MV_Gyr	Var_Acc_z
80	Max_Acc_z	Interq_Azimuth	Avr_peak_time_Pitch	Dfreq_Acc_x	Highest_peak_Pitch	Interq_Roll

81	Interq_Acc_x	Interq_Acc_x	nPeaks_Roll	Dfreq_Acc_y	SVM_Acc	Rms_Gyr_x
82	Cf_Gyr_y	Max_Gyr_x	Widest_Peak_Roll	Dfreq_Acc_z	Max_Acc_x	Skew_Pitch
83	DSVM_Acc	Max_diff_Acc_x	Avr_peak_time_Gyr_x	Dfreq_Pitch	Widest_Peak_Roll	Cf_Acc_z
84	Skew_Roll	Skew_Roll	nPeaks_Gyr_y	Dfreq_Roll	Entropy_TimeD_Gyr	SMA_Gyr
85	Cf_Acc_x	Min_Gyr_x	Widest_Peak_Gyr_y	Dfreq_Gyr_y	Mean_Gyr_y	Var_Gyr_y
86	Max_Gyr_x	DSVM_Acc	Widest_Peak_Gyr_z	Dfreq_Gyr_z	Vedb_Acc	Avr_peak_time_Azimuth
87	Interq_Pitch	Std_Acc_x	Avr_peak_time_Gyr_z	nPeaks_Acc_x	Avr_peak_time_Acc_y	Skew_Azimuth
88	Var_Azimuth	Cf_Gyr_y		Widest_Peak_Acc_x	Highest_peak_Azimuth	Cf_Gyr_z
89	Max_diff_Acc_x	Max_diff_Azimuth		Highest_peak_Acc_x	Var_Gyr_y	Rms_Azimuth
90	Skew_Acc_x	Rms_Acc_x		Avr_peak_time_Acc_x	Cf_Gyr_z	Skew_Acc_x
91	Rms_Acc_x	Interq_Acc_z		nPeaks_Acc_y	Max_Acc_z	Interq_Acc_y
92	Min_Acc_z	Std_Azimuth		Widest_Peak_Acc_y	Avr_peak_time_Gyr_z	Interq_Gyr_z
93	Max_diff_Azimuth	Skew_Acc_x		Highest_peak_Acc_y	Min_Gyr_z	Mean_Gyr_z
94	Highest_peak_Gyr_z	Skew_Acc_y		nPeaks_Acc_z	Interq_Acc_z	SMA_Acc
95	Var_Acc_x	Highest_peak_Acc_z		Highest_peak_Acc_z	Skew_Azimuth	Avr_peak_time_Acc_x
96	Vedb_Acc	Widest_Peak_Acc_z		Avr_peak_time_Acc_z	Kur_Azimuth	Kur_Pitch
97	Dfreq_Gyr_z	Dfreq_Acc_x		Widest_Peak_Azimuth	Rms_Acc_z	Cf_Gyr_y
98	Max_Gyr_y	Max_Gyr_y		Avr_peak_time_Azimuth	Mean_Azimuth	nPeaks_Acc_y
99	Cf_Acc_z	Cf_Gyr_x		nPeaks_Pitch	Interq_Pitch	Dfreq_Acc_y
100	Vedb_Gyr	Vedb_Gyr		Widest_Peak_Pitch	Avr_peak_time_Roll	Skew_Gyr_x
101	Skew_Acc_y	Highest_peak_Gyr_z		Avr_peak_time_Pitch	Kur_Acc_z	nPeaks_Pitch
102	Energy_Acc_y	Dfreq_Gyr_z		Widest_Peak_Roll	Entropy_Acc_x	Kur_Azimuth
103	Highest_peak_Acc_z	Vedb_Acc		Avr_peak_time_Roll	SMA_Angle	Max_Gyr_z
104	Widest_Peak_Acc_z	Skew_Acc_z		nPeaks_Gyr_x	Kur_Acc_x	Interq_Azimuth
105	Cf_Gyr_z	Kur_Gyr_x		Avr_peak_time_Gyr_x	nPeaks_Acc_z	Entropy_TimeD_Ang
106	Std_Gyr_x	Std_Gyr_x		nPeaks_Gyr_y	Cf_Acc_x	Cf_Pitch
107	Dfreq_Acc_x	Highest_peak_Azimuth		Highest_peak_Gyr_y	Kur_Acc_y	Highest_peak_Acc_y
108	Cf_Acc_y	Interq_Acc_y			SVM_Angle	Entropy_Pitch
109	Energy_Acc_z	Rms_Gyr_x			Vedb_Angle	nPeaks_Roll

110	Rms_Gyr_x	SVM_Gyr			Highest_peak_Gyr_x	Highest_peak_Azimuth
111	Std_Roll	SMA_Gyr			Max_Gyr_y	Highest_peak_Gyr_x
112	Min_Gyr_x	Cf_Acc_x			Kur_Gyr_x	Entropy_TimeD_Acc
113	Var_Acc_z	Energy_Gyr_y			nPeaks_Acc_x	Max_diff_Azimuth
114	Dfreq_Acc_z	Var_Gyr_y			Highest_peak_Gyr_z	DSVM_Acc
115	Interq_Acc_z	Energy_Acc_y			Skew_Gyr_x	Widest_Peak_Acc_x
116	Widest_Peak_Acc_y	Min_Acc_z			nPeaks_Acc_y	Interq_Pitch
117	SVM_Gyr	Dfreq_Acc_z			Entropy_TimeD_Acc	Widest_Peak_Pitch
118	SMA_Gyr	Var_Acc_z			mag_Ang	Highest_peak_Gyr_z
119	Dfreq_Acc_y	Dfreq_Gyr_x			Max_Azimuth	Vedb_Gyr
120	Skew_Acc_z	Var_Azimuth			Rms_Azimuth	Max_Gyr_y
121	Energy_Gyr_y	Entropy_TimeD_Gyr			Max_diff_Azimuth	Avr_peak_time_Pitch
122	Interq_Roll	Dfreq_Acc_y			Dfreq_Acc_x	SVM_Acc
123	mag_Acc	Kur_Acc_y			Avr_peak_time_Pitch	Widest_Peak_Gyr_z
124	Cf_Gyr_x	Var_Acc_x			Std_Acc_x	SMA_Angle
125	Var_Gyr_y	Highest_peak_Gyr_x			Std_Acc_y	Highest_peak_Gyr_y
126	Kur_Gyr_y	Cf_Gyr_z			Std_Acc_z	Energy_Gyr_y
127	Entropy_TimeD_Gyr	Energy_Acc_z			Std_Azimuth	Avr_peak_time_Gyr_x
128	Kur_Acc_x	Interq_Gyr_x			Std_Pitch	Max_Acc_x
129	Interq_Acc_y	Std_Roll			Std_Roll	Var_Azimuth
130	Skew_Gyr_y	Skew_Pitch			Std_Gyr_x	Avr_peak_time_Roll
131	Dfreq_Gyr_x	Kur_Acc_z			Std_Gyr_y	Widest_Peak_Gyr_x
132	Skew_Pitch	Widest_Peak_Acc_y			Std_Gyr_z	Mean_Azimuth
133	Max_diff_Roll	mag_Acc			Energy_Acc_x	Vedb_Angle
134	Cf_Azimuth	Max_diff_Roll			Energy_Acc_y	Highest_peak_Acc_z
135	Interq_Gyr_x	Skew_Gyr_z			Energy_Acc_z	Std_Gyr_z
136	Mean_Acc_z	Mean_Acc_z			Energy_Azimuth	Widest_Peak_Acc_y
137	Highest_peak_Gyr_x	Energy_Acc_x			Energy_Pitch	Mean_Gyr_x
138	Kur_Acc_y	Kur_Acc_x			Energy_Roll	Avr_peak_time_Acc_y

139	Skew_Gyr_z	Kur_Gyr_z			Energy_Gyr_x	mag_Acc
140	Widest_Peak_Azimuth	Skew_Gyr_y			Energy_Gyr_y	Highest_peak_Roll
141	Energy_Acc_x	Cf_Azimuth			Energy_Gyr_z	Rms_Acc_z
142	nPeaks_Acc_y	Kur_Gyr_y			Dfreq_Azimuth	SVM_Gyr
143	Kur_Gyr_x	Mean_Acc_x			Widest_Peak_Azimuth	mag_Gyr
144	Mean_Acc_x	Energy_Gyr_x			Interq_Acc_y	Entropy_TimeD_Gyr
145	nPeaks_Azimuth	Var_Gyr_x			Cf_Gyr_y	Cf_Acc_x
146	Var_Roll	nPeaks_Gyr_y			Mean_Gyr_x	Min_Acc_x
147	nPeaks_Acc_z	nPeaks_Acc_y			Highest_peak_Acc_x	Min_Azimuth
148	Var_Gyr_x	Skew_Gyr_x			Widest_Peak_Gyr_z	Widest_Peak_Gyr_y
149	Highest_peak_Roll	Mean_Acc_y			mag_Acc	Std_Acc_z
150	Kur_Acc_z	Widest_Peak_Acc_x			Widest_Peak_Acc_z	Energy_Acc_y
151	Energy_Gyr_x	Mean_Gyr_z			DSAM_Angle	Mean_Gyr_y
152	Mean_Gyr_z	Interq_Roll			Entropy_Acc_z	nPeaks_Acc_x
153	Avr_peak_time_Acc_y	Var_Roll			Avr_peak_time_Azimuth	Kur_Acc_x
154	Mean_Acc_y	Avr_peak_time_Acc_y			Dfreq_Gyr_y	Highest_peak_Acc_x
155	mag_Gyr	Highest_peak_Roll			Mean_Gyr_z	nPeaks_Gyr_x
156	Widest_Peak_Gyr_z	Kur_Roll			SMA_Acc	Std_Gyr_y
157	nPeaks_Gyr_y	nPeaks_Acc_x			Widest_Peak_Acc_y	Std_Pitch
158	Widest_Peak_Acc_x	nPeaks_Roll			mag_Gyr	mag_Ang
159	Avr_peak_time_Acc_z	mag_Gyr			Kur_Pitch	Max_Azimuth
160	Kur_Gyr_z	Avr_peak_time_Acc_x			Min_Azimuth	Widest_Peak_Roll
161	Mean_Gyr_y	Widest_Peak_Gyr_y			Avr_peak_time_Acc_z	Cf_Azimuth
162	Avr_peak_time_Acc_x	Skew_Azimuth			Widest_Peak_Gyr_y	Widest_Peak_Azimuth
163	Kur_Roll	Avr_peak_time_Gyr_y			Interq_Gyr_x	Energy_Pitch
164	nPeaks_Gyr_x	Kur_Pitch			SVM_Gyr	nPeaks_Gyr_z
165	Skew_Azimuth	Avr_peak_time_Acc_z			Skew_Pitch	Energy_Gyr_x
166	nPeaks_Acc_x	nPeaks_Azimuth			Entropy_Azimuth	Energy_Acc_x
167	Kur_Azimuth	Avr_peak_time_Azimuth			nPeaks_Pitch	Avr_peak_time_Gyr_y

168	Kur_Pitch	Widest_Peak_Gyr_z			Highest_peak_Roll	Energy_Roll
169	Avr_peak_time_Gyr_x	Avr_peak_time_Pitch			nPeaks_Gyr_y	Skew_Roll
170	Skew_Gyr_x	Mean_Gyr_y			Avr_peak_time_Gyr_x	nPeaks_Azimuth
171	Widest_Peak_Gyr_x	Avr_peak_time_Roll			Avr_peak_time_Gyr_y	Energy_Azimuth
172	Avr_peak_time_Gyr_y	nPeaks_Acc_z			nPeaks_Gyr_x	nPeaks_Acc_z
173	Avr_peak_time_Gyr_z	Widest_Peak_Azimuth			SMA_Gyr	Energy_Acc_z
174	nPeaks_Roll	Avr_peak_time_Gyr_x			Entropy_TimeD_Ang	Dfreq_Azimuth
175	nPeaks_Gyr_z	Avr_peak_time_Gyr_z			Vedb_Gyr	Std_Azimuth
176	Widest_Peak_Gyr_y	Kur_Azimuth			Widest_Peak_Gyr_x	Std_Roll
177	Avr_peak_time_Azimuth	nPeaks_Gyr_z			Highest_peak_Gyr_y	Std_Gyr_x
178	Widest_Peak_Roll	Widest_Peak_Gyr_x			nPeaks_Roll	nPeaks_Gyr_y
179	Avr_peak_time_Pitch	Widest_Peak_Roll			Rms_Gyr_y	Avr_peak_time_Gyr_z
180	Avr_peak_time_Roll	nPeaks_Gyr_x			Var_Azimuth	Std_Acc_y
181	nPeaks_Pitch	nPeaks_Pitch			Avr_peak_time_Acc_x	SVM_Angle
182	Widest_Peak_Pitch	Widest_Peak_Pitch			Skew_Gyr_y	Std_Acc_x
183	Mean_Gyr_x	Mean_Gyr_x			Cf_Acc_y	Energy_Gyr_z

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 2 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet1_all** over 7 $sec.window$.

	ReliefF		**GA**		**RF**	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Energy_Roll	Energy_Roll	Mean_Acc_y	Mean_Acc_y	Max_diff_Pitch	Max_diff_Pitch
2	Rms_Roll	Rms_Roll	Mean_Acc_z	Mean_Acc_z	Max_Roll	Mean_Roll
3	Dfreq_Roll	Max_Roll	Mean_Pitch	Mean_Azimuth	Mean_Roll	Max_Roll
4	Max_Roll	Dfreq_Roll	Mean_Roll	Mean_Pitch	Cf_Pitch	Max_diff_Gyr_z
5	Mean_Roll	Mean_Roll	Mean_Gyr_x	Mean_Roll	Cf_Roll	Cf_Roll
6	Min_Roll	Min_Roll	Mean_Gyr_y	Mean_Gyr_x	Rms_Roll	Rms_Pitch
7	Cf_Roll	Cf_Roll	Mean_Gyr_z	Mean_Gyr_y	Mean_Pitch	Min_Gyr_x
8	Entropy_Pitch	Entropy_Pitch	Var_Acc_x	Var_Pitch	Dfreq_Gyr_z	Min_Gyr_z
9	Max_Pitch	Max_Pitch	Var_Acc_z	Var_Roll	Kur_Gyr_x	Rms_Roll
10	Rms_Pitch	Rms_Pitch	Var_Azimuth	Std_Acc_x	Skew_Acc_y	Mean_Pitch
11	Max_diff_Pitch	Max_diff_Pitch	Var_Roll	Std_Acc_y	Var_Gyr_z	Var_Pitch
12	Energy_Pitch	Energy_Pitch	Var_Gyr_z	Std_Azimuth	Kur_Gyr_z	Min_Acc_y
13	Cf_Pitch	Cf_Pitch	Std_Azimuth	Std_Roll	Max_diff_Acc_y	Max_Gyr_x
14	Dfreq_Pitch	Entropy_Roll	Std_Pitch	Std_Gyr_z	Var_Pitch	Max_diff_Acc_z
15	Entropy_Roll	Dfreq_Pitch	Std_Gyr_x	Kur_Acc_y	Max_Acc_y	Cf_Pitch
16	MV_Gyr	MV_Gyr	Std_Gyr_y	Kur_Azimuth	Rms_Pitch	Entropy_Roll
17	Std_Pitch	Std_Pitch	Kur_Acc_y	Kur_Pitch	Max_Gyr_x	Min_Roll
18	Mean_Pitch	Min_Pitch	Kur_Roll	Kur_Roll	Var_Gyr_x	Dfreq_Roll
19	Entropy_Gyr_y	Mean_Pitch	Kur_Gyr_x	Kur_Gyr_z	Max_diff_Acc_z	Dfreq_Acc_y
20	Min_Pitch	Entropy_Gyr_y	Skew_Acc_z	Skew_Acc_y	Kur_Pitch	Min_Acc_z
21	MV_Acc	MV_Acc	Skew_Pitch	Skew_Acc_z	Min_Acc_y	Min_Pitch
22	Entropy_TimeD_Ang	Max_diff_Acc_y	Skew_Roll	Skew_Pitch	Mean_Acc_z	Highest_peak_Pitch
23	SVM_Angle	Entropy_Gyr_x	Skew_Gyr_x	Skew_Gyr_x	Dfreq_Pitch	Var_Gyr_x
24	SMA_Angle	SVM_Angle	Skew_Gyr_z	Min_Acc_y	Max_diff_Gyr_x	Interq_Acc_x
25	Entropy_Gyr_x	Entropy_TimeD_Ang	Min_Acc_z	Min_Pitch	Min_Acc_z	Skew_Pitch

26	Max_diff_Acc_y	SMA_Angle	Min_Pitch	Min_Gyr_x	Max_diff_Gyr_z	Mean_Gyr_x
27	mag_Ang	DSVM_Gyr	Min_Roll	Max_Acc_z	Min_Roll	Interq_Roll
28	DSVM_Gyr	Max_diff_Gyr_z	Min_Gyr_z	Max_Azimuth	Var_Acc_z	Var_Gyr_z
29	Max_diff_Gyr_z	Entropy_Gyr_z	Max_Acc_z	Max_Pitch	Dfreq_Roll	Kur_Gyr_x
30	Entropy_Acc_y	mag_Ang	Max_Pitch	Max_Roll	mag_Ang	Skew_Acc_y
31	Vedb_Angle	Entropy_Acc_z	Max_Gyr_z	Max_Gyr_x	Interq_Acc_x	Skew_Gyr_x
32	Entropy_Gyr_z	Min_Gyr_z	Rms_Acc_y	Max_Gyr_y	Max_Gyr_z	Entropy_Gyr_x
33	Max_Acc_y	Entropy_Acc_y	Rms_Roll	Max_Gyr_z	Entropy_Roll	Dfreq_Gyr_z
34	Entropy_Acc_z	Max_diff_Gyr_x	Rms_Gyr_x	Rms_Acc_x	Min_Gyr_x	Interq_Gyr_y
35	Min_Gyr_z	Vedb_Angle	Rms_Gyr_z	Rms_Acc_z	Vedb_Acc	Max_Gyr_z
36	Std_Gyr_z	Interq_Azimuth	Interq_Acc_y	Rms_Pitch	Skew_Acc_x	Min_Azimuth
37	Rms_Gyr_z	Std_Acc_y	Interq_Acc_z	Rms_Roll	Max_Pitch	Dfreq_Acc_x
38	Max_diff_Gyr_x	Max_diff_Acc_z	Interq_Azimuth	Rms_Gyr_x	Var_Roll	Entropy_Pitch
39	Dfreq_Gyr_y	Max_Acc_y	Interq_Roll	Interq_Acc_y	Min_Azimuth	Var_Acc_y
40	Mean_Azimuth	Highest_peak_Pitch	Interq_Gyr_y	Interq_Acc_z	Min_Pitch	Mean_Acc_z
41	Dfreq_Azimuth	Dfreq_Gyr_y	Interq_Gyr_z	Interq_Azimuth	Max_diff_Azimuth	Interq_Gyr_z
42	Highest_peak_Gyr_y	Max_Acc_x	Cf_Acc_y	Interq_Pitch	Entropy_Gyr_x	Dfreq_Gyr_x
43	Var_Pitch	Var_Pitch	Cf_Acc_z	Interq_Roll	Kur_Acc_x	Highest_peak_Azimuth
44	Std_Acc_y	Std_Gyr_z	Cf_Gyr_x	Interq_Gyr_x	Entropy_Gyr_y	Var_Azimuth
45	Interq_Azimuth	Highest_peak_Gyr_y	Cf_Gyr_y	Interq_Gyr_z	Entropy_Acc_z	Highest_peak_Roll
46	Entropy_Acc_x	Rms_Gyr_z	SMA_Angle	Cf_Acc_x	Mean_Acc_x	MV_Gyr
47	Max_diff_Acc_z	Highest_peak_Acc_x	Entropy_Acc_z	Cf_Acc_y	Entropy_Pitch	Energy_Acc_z
48	Highest_peak_Pitch	Interq_Gyr_y	Entropy_Gyr_x	Cf_Roll	DSVM_Gyr	Max_Acc_z
49	Interq_Gyr_y	Entropy_Acc_x	Entropy_Gyr_y	Cf_Gyr_x	MV_Gyr	Dfreq_Pitch
50	Rms_Acc_z	Std_Acc_x	Entropy_TimeD_Acc	Cf_Gyr_y	Highest_peak_Acc_z	Rms_Acc_y
51	AV_Ang	Dfreq_Azimuth	Entropy_TimeD_Gyr	Entropy_Acc_x	Var_Acc_y	Max_Acc_x
52	Max_Acc_x	Mean_Azimuth	Energy_Azimuth	Entropy_Azimuth	Rms_Gyr_z	Kur_Acc_z
53	Std_Acc_z	Var_Acc_y	Energy_Pitch	Entropy_Pitch	Cf_Acc_z	Var_Acc_z
54	SVM_Acc	Min_Gyr_x	Energy_Roll	Entropy_Gyr_x	Highest_peak_Gyr_y	Std_Roll

55	Min_Azimuth	Cf_Acc_x	Energy_Gyr_x	Entropy_Gyr_y	Rms_Acc_x	Cf_Acc_x
56	SMA_Acc	Skew_Acc_x	Energy_Gyr_y	Energy_Acc_x	Skew_Gyr_z	Entropy_Gyr_z
57	Rms_Acc_y	Max_Gyr_z	Energy_Gyr_z	Energy_Acc_z	Min_Acc_x	Std_Acc_z
58	Rms_Azimuth	Std_Acc_z	SVM_Angle	Energy_Azimuth	Kur_Acc_z	Energy_Roll
59	Highest_peak_Acc_x	Rms_Acc_z	DSVM_Acc	Energy_Pitch	Mean_Acc_y	Highest_peak_Gyr_y
60	Energy_Azimuth	Rms_Acc_y	Max_diff_Acc_x	Energy_Gyr_y	Max_diff_Gyr_y	Entropy_Acc_y
61	Interq_Acc_x	Max_Gyr_x	Max_diff_Acc_y	SVM_Acc	Highest_peak_Gyr_x	Rms_Gyr_x
62	Var_Acc_y	Max_Azimuth	Max_diff_Pitch	DSVM_Acc	Cf_Gyr_x	Var_Acc_x
63	Interq_Gyr_z	Min_Acc_y	Max_diff_Roll	Max_diff_Acc_y	Interq_Acc_z	Energy_Acc_y
64	Min_Acc_x	Min_Acc_x	Max_diff_Gyr_y	Max_diff_Gyr_x	Kur_Roll	Energy_Gyr_x
65	Min_Acc_y	Interq_Gyr_z	Max_diff_Gyr_z	Max_diff_Gyr_z	nPeaks_Azimuth	Dfreq_Azimuth
66	Rms_Gyr_y	Skew_Acc_y	AV_Ang	AV_Ang	Widest_Peak_Roll	Cf_Gyr_x
67	DSAM_Angle	Highest_peak_Acc_y	Vedb_Acc	MV_Gyr	nPeaks_Acc_z	Max_Acc_y
68	Entropy_TimeD_Acc	AV_Ang	Vedb_Angle	mag_Acc	MV_Acc	Highest_peak_Acc_y
69	Max_Gyr_z	Skew_Pitch	Vedb_Gyr	mag_Ang	Dfreq_Acc_x	Std_Acc_y
70	Kur_Gyr_x	Interq_Acc_x	Dfreq_Acc_y	mag_Gyr	Avr_peak_time_Acc_x	Mean_Acc_x
71	Min_Gyr_y	Rms_Azimuth	Dfreq_Azimuth	Vedb_Angle	Vedb_Angle	DSVM_Gyr
72	Std_Gyr_y	Max_diff_Gyr_y	Dfreq_Pitch	Vedb_Gyr	Mean_Gyr_x	Rms_Acc_z
73	Rms_Acc_x	Highest_peak_Gyr_z	Dfreq_Gyr_z	Dfreq_Acc_z	Max_Acc_x	Std_Azimuth
74	Max_Azimuth	Std_Gyr_y	nPeaks_Acc_x	Dfreq_Gyr_x	Cf_Acc_y	Max_Pitch
75	Highest_peak_Gyr_z	Max_Acc_z	nPeaks_Acc_y	Widest_Peak_Acc_x	Max_Gyr_y	Max_diff_Acc_x
76	Std_Acc_x	Rms_Gyr_y	Highest_peak_Acc_y	Highest_peak_Acc_x	Entropy_Gyr_z	Mean_Acc_y
77	Highest_peak_Acc_y	Min_Gyr_y	nPeaks_Acc_z	Avr_peak_time_Acc_x	Rms_Acc_y	Avr_peak_time_Pitch
78	Skew_Acc_z	Cf_Acc_z	Highest_peak_Acc_z	Highest_peak_Acc_y	Rms_Acc_z	Rms_Gyr_z
79	Cf_Acc_x	Rms_Acc_x	Avr_peak_time_Acc_z	Avr_peak_time_Acc_y	Rms_Gyr_x	Dfreq_Gyr_y
80	Max_diff_Gyr_y	Energy_Azimuth	nPeaks_Azimuth	Widest_Peak_Acc_z	Interq_Gyr_y	Std_Gyr_z
81	Entropy_Azimuth	Std_Gyr_x	Widest_Peak_Azimuth	Highest_peak_Acc_z	Interq_Gyr_z	Max_diff_Acc_y
82	Max_Gyr_x	DSAM_Angle	Highest_peak_Azimuth	Widest_Peak_Azimuth	Skew_Pitch	Interq_Gyr_x
83	Skew_Acc_x	Entropy_Azimuth	Avr_peak_time_Azimuth	Widest_Peak_Pitch	Highest_peak_Pitch	Std_Gyr_x

84	Cf_Gyr_y	Rms_Gyr_x	nPeaks_Pitch	Highest_peak_Pitch	nPeaks_Gyr_x	Energy_Azimuth
85	Cf_Acc_z	Std_Azimuth	Widest_Peak_Pitch	Avr_peak_time_Pitch	Dfreq_Acc_z	Dfreq_Acc_z
86	Min_Gyr_x	SMA_Acc	Avr_peak_time_Pitch	Widest_Peak_Roll	Min_Gyr_z	Avr_peak_time_Acc_y
87	Skew_Acc_y	Cf_Gyr_y	nPeaks_Roll	Highest_peak_Roll	DSAM_Angle	Min_Gyr_y
88	Max_Acc_z	Max_diff_Acc_x	Avr_peak_time_Roll	Avr_peak_time_Roll	SMA_Acc	Std_Pitch
89	Min_Acc_z	SVM_Acc	Widest_Peak_Gyr_x	nPeaks_Gyr_x	Cf_Gyr_y	DSAM_Angle
90	Dfreq_Acc_y	Min_Azimuth	Avr_peak_time_Gyr_x	Widest_Peak_Gyr_x	Dfreq_Acc_y	Widest_Peak_Gyr_x
91	Cf_Gyr_x	Interq_Acc_z	Avr_peak_time_Gyr_y	nPeaks_Gyr_y	Entropy_TimeD_Gyr	Mean_Gyr_y
92	Vedb_Gyr	Skew_Roll	nPeaks_Gyr_z	Widest_Peak_Gyr_z	Max_Acc_z	DSVM_Acc
93	Std_Gyr_x	Vedb_Gyr	Widest_Peak_Gyr_z	Highest_peak_Gyr_z	Highest_peak_Acc_y	Energy_Acc_x
94	Rms_Gyr_x	Skew_Acc_z	Avr_peak_time_Gyr_z	Avr_peak_time_Gyr_z	Widest_Peak_Acc_x	Std_Acc_x
95	Skew_Pitch	Highest_peak_Acc_z			Kur_Acc_y	Energy_Gyr_y
96	nPeaks_Azimuth	Cf_Gyr_x			Skew_Roll	Rms_Acc_x
97	Skew_Roll	Dfreq_Gyr_z			Dfreq_Gyr_y	MV_Acc
98	Std_Azimuth	Entropy_TimeD_Acc			Interq_Acc_y	SVM_Acc
99	Energy_Gyr_z	Dfreq_Gyr_x			DSVM_Acc	Kur_Gyr_y
100	Dfreq_Gyr_z	Interq_Pitch			Skew_Gyr_x	Skew_Acc_x
101	Energy_Acc_y	Interq_Gyr_x			Interq_Pitch	Max_diff_Roll
102	Var_Gyr_z	Cf_Gyr_z			Var_Acc_x	Max_diff_Gyr_x
103	Max_diff_Acc_x	Min_Acc_z			Vedb_Gyr	Interq_Acc_z
104	Highest_peak_Acc_z	Dfreq_Acc_z			SVM_Acc	Highest_peak_Gyr_z
105	Cf_Acc_y	Var_Acc_x			SVM_Angle	Min_Acc_x
106	Interq_Acc_y	SVM_Gyr			Kur_Azimuth	Cf_Acc_y
107	Vedb_Acc	SMA_Gyr			Entropy_TimeD_Ang	Max_diff_Gyr_y
108	Interq_Acc_z	Energy_Acc_y			Rms_Azimuth	Skew_Gyr_y
109	Dfreq_Gyr_x	Max_Gyr_y			Highest_peak_Acc_x	Rms_Gyr_y
110	Dfreq_Acc_x	Cf_Acc_y			Avr_peak_time_Pitch	Avr_peak_time_Acc_z
111	Kur_Acc_x	Highest_peak_Gyr_x			Avr_peak_time_Acc_z	Energy_Pitch
112	Dfreq_Acc_z	Kur_Pitch			SMA_Gyr	SVM_Angle

113	Highest_peak_Gyr_x	Dfreq_Acc_y			Skew_Acc_z	Avr_peak_time_Gyr_x
114	Interq_Pitch	Highest_peak_Azimuth			Mean_Azimuth	Skew_Acc_z
115	Skew_Gyr_y	Entropy_TimeD_Gyr			Widest_Peak_Gyr_x	Rms_Azimuth
116	Highest_peak_Azimuth	Vedb_Acc			nPeaks_Acc_x	Widest_Peak_Azimuth
117	Max_Gyr_y	Kur_Gyr_x			Entropy_TimeD_Acc	SMA_Angle
118	Mean_Gyr_z	Kur_Gyr_y			Mean_Gyr_z	mag_Gyr
119	DSVM_Acc	Var_Azimuth			Highest_peak_Gyr_z	Std_Gyr_y
120	SMA_Gyr	Energy_Gyr_z			Var_Gyr_y	Highest_peak_Acc_z
121	Interq_Gyr_x	Kur_Acc_x			Entropy_Azimuth	Entropy_TimeD_Ang
122	SVM_Gyr	Var_Gyr_z			Max_Azimuth	Energy_Gyr_z
123	Var_Azimuth	DSVM_Acc			Cf_Gyr_z	SMA_Gyr
124	Max_diff_Azimuth	Std_Roll			Highest_peak_Azimuth	AV_Ang
125	Var_Acc_z	Dfreq_Acc_x			Kur_Gyr_y	Kur_Roll
126	Cf_Gyr_z	Skew_Gyr_z			Avr_peak_time_Gyr_y	SVM_Gyr
127	Kur_Acc_z	Max_diff_Azimuth			Entropy_Acc_y	Interq_Azimuth
128	Entropy_TimeD_Gyr	Var_Acc_z			Cf_Acc_x	Cf_Acc_z
129	Energy_Acc_z	Mean_Acc_z			Avr_peak_time_Gyr_x	Entropy_Gyr_y
130	Var_Acc_x	Interq_Acc_y			Highest_peak_Roll	nPeaks_Pitch
131	mag_Acc	Mean_Gyr_z			Skew_Azimuth	Cf_Gyr_z
132	nPeaks_Gyr_x	Energy_Acc_z			Avr_peak_time_Acc_y	nPeaks_Acc_y
133	Kur_Gyr_y	Var_Gyr_y			SMA_Angle	nPeaks_Gyr_x
134	Energy_Gyr_y	Energy_Gyr_y			Interq_Gyr_x	Highest_peak_Acc_x
135	Std_Roll	Skew_Gyr_y			nPeaks_Gyr_z	Max_diff_Azimuth
136	Mean_Acc_z	Max_diff_Roll			SVM_Gyr	Entropy_TimeD_Gyr
137	Var_Gyr_y	Mean_Acc_y			Widest_Peak_Azimuth	Var_Roll
138	Energy_Acc_x	Energy_Gyr_x			Var_Azimuth	Mean_Azimuth
139	Mean_Acc_y	Cf_Azimuth			nPeaks_Gyr_y	Vedb_Acc
140	Kur_Gyr_z	Kur_Acc_z			Widest_Peak_Pitch	nPeaks_Roll
141	nPeaks_Acc_z	Interq_Roll			Max_diff_Acc_x	Avr_peak_time_Gyr_z

142	Kur_Pitch	Var_Gyr_x			Widest_Peak_Gyr_z	Entropy_TimeD_Acc
143	Mean_Acc_x	Energy_Acc_x			Avr_peak_time_Roll	SMA_Acc
144	Skew_Gyr_z	Widest_Peak_Acc_y			Std_Acc_x	Kur_Acc_x
145	Widest_Peak_Azimuth	Kur_Acc_y			Std_Acc_y	Interq_Acc_y
146	Avr_peak_time_Acc_z	Kur_Gyr_z			Std_Acc_z	Skew_Roll
147	Widest_Peak_Acc_z	Mean_Acc_x			Std_Azimuth	Cf_Azimuth
148	Cf_Azimuth	nPeaks_Acc_z			Std_Pitch	Cf_Gyr_y
149	Energy_Gyr_x	nPeaks_Gyr_y			Std_Roll	Vedb_Gyr
150	Widest_Peak_Gyr_x	Mean_Gyr_y			Std_Gyr_x	Mean_Gyr_z
151	Var_Gyr_x	Widest_Peak_Roll			Std_Gyr_y	nPeaks_Acc_z
152	Max_diff_Roll	Skew_Gyr_x			Std_Gyr_z	Kur_Pitch
153	Widest_Peak_Acc_x	mag_Gyr			Energy_Acc_x	Kur_Acc_y
154	Skew_Gyr_x	Skew_Azimuth			Energy_Acc_y	Var_Gyr_y
155	Interq_Roll	nPeaks_Azimuth			Energy_Acc_z	Interq_Pitch
156	Kur_Acc_y	Widest_Peak_Acc_z			Energy_Azimuth	Widest_Peak_Gyr_z
157	Skew_Azimuth	Var_Roll			Energy_Pitch	Highest_peak_Gyr_x
158	Highest_peak_Roll	Widest_Peak_Gyr_x			Energy_Roll	mag_Acc
159	Var_Roll	Avr_peak_time_Acc_z			Energy_Gyr_x	Widest_Peak_Pitch
160	Avr_peak_time_Gyr_x	Kur_Roll			Energy_Gyr_y	Entropy_Acc_z
161	Kur_Roll	Highest_peak_Roll			Energy_Gyr_z	nPeaks_Azimuth
162	Kur_Azimuth	Widest_Peak_Gyr_z			Dfreq_Azimuth	mag_Ang
163	Avr_peak_time_Acc_y	Kur_Azimuth			Dfreq_Gyr_x	Skew_Gyr_z
164	Widest_Peak_Acc_y	Widest_Peak_Azimuth			Rms_Gyr_y	Skew_Azimuth
165	Mean_Gyr_y	Avr_peak_time_Acc_x			Cf_Azimuth	Entropy_Acc_x
166	Avr_peak_time_Azimuth	nPeaks_Roll			Min_Gyr_y	Kur_Gyr_z
167	Widest_Peak_Gyr_z	Avr_peak_time_Gyr_y			Skew_Gyr_y	nPeaks_Gyr_y
168	Widest_Peak_Roll	Avr_peak_time_Azimuth			Widest_Peak_Acc_z	Widest_Peak_Acc_z
169	nPeaks_Acc_y	nPeaks_Acc_x			Max_diff_Roll	Widest_Peak_Gyr_y
170	nPeaks_Roll	Avr_peak_time_Roll			Mean_Gyr_y	Entropy_Azimuth

171	Widest_Peak_Gyr_y	Widest_Peak_Acc_x			AV_Ang	Avr_peak_time_Roll
172	Avr_peak_time_Roll	Mean_Gyr_x			Entropy_Acc_x	Widest_Peak_Acc_x
173	nPeaks_Gyr_y	mag_Acc			Interq_Azimuth	Avr_peak_time_Azimuth
174	mag_Gyr	Avr_peak_time_Acc_y			nPeaks_Acc_y	Avr_peak_time_Acc_x
175	Avr_peak_time_Acc_x	nPeaks_Gyr_x			mag_Gyr	Widest_Peak_Acc_y
176	Avr_peak_time_Gyr_y	Widest_Peak_Gyr_y			Widest_Peak_Acc_y	Kur_Azimuth
177	Widest_Peak_Pitch	Widest_Peak_Pitch			mag_Acc	Widest_Peak_Roll
178	Mean_Gyr_x	Avr_peak_time_Gyr_x			Avr_peak_time_Gyr_z	Max_Azimuth
179	nPeaks_Acc_x	nPeaks_Acc_y			nPeaks_Pitch	nPeaks_Acc_x
180	nPeaks_Gyr_z	nPeaks_Pitch			Interq_Roll	Vedb_Angle
181	Avr_peak_time_Gyr_z	Avr_peak_time_Gyr_z			Widest_Peak_Gyr_y	Max_Gyr_y
182	nPeaks_Pitch	Avr_peak_time_Pitch			Avr_peak_time_Azimuth	nPeaks_Gyr_z
183	Avr_peak_time_Pitch	nPeaks_Gyr_z			nPeaks_Roll	Avr_peak_time_Gyr_y

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 3 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet1_all** over 5 $sec.window$.

	ReliefF		**GA**		**RF**	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Energy_Roll	Energy_Roll	Mean_Acc_x	Mean_Acc_z	Mean_Roll	Mean_Roll
2	Rms_Roll	Rms_Roll	Mean_Acc_y	Mean_Azimuth	Cf_Pitch	Max_diff_Pitch
3	Dfreq_Roll	Max_Roll	Mean_Gyr_x	Mean_Pitch	Var_Gyr_z	Cf_Roll
4	Mean_Roll	Mean_Roll	Mean_Gyr_y	Mean_Roll	Kur_Gyr_x	Cf_Pitch
5	Max_Roll	Dfreq_Roll	Var_Acc_x	Var_Acc_z	Max_Gyr_x	Min_Acc_z
6	Cf_Roll	Min_Roll	Var_Acc_y	Var_Azimuth	Max_Roll	Highest_peak_Pitch
7	Min_Roll	Cf_Roll	Var_Roll	Var_Roll	Max_diff_Gyr_z	Max_diff_Gyr_z
8	Entropy_Pitch	Entropy_Pitch	Var_Gyr_y	Std_Acc_y	Max_diff_Pitch	Rms_Roll
9	Max_Pitch	Max_Pitch	Var_Gyr_z	Std_Azimuth	Entropy_Pitch	Dfreq_Roll
10	Cf_Pitch	Cf_Pitch	Std_Acc_x	Std_Roll	Mean_Acc_x	Var_Acc_z
11	Max_diff_Pitch	Max_diff_Pitch	Std_Acc_y	Std_Gyr_x	Var_Pitch	Interq_Acc_x
12	Rms_Pitch	Entropy_Roll	Std_Roll	Skew_Acc_x	Var_Gyr_x	Max_Roll
13	Entropy_Roll	Rms_Pitch	Std_Gyr_z	Skew_Acc_y	Rms_Roll	Entropy_Pitch
14	Energy_Pitch	Energy_Pitch	Kur_Acc_x	Skew_Acc_z	Min_Gyr_x	Max_Gyr_x
15	Std_Pitch	Std_Pitch	Kur_Acc_y	Skew_Pitch	Max_Pitch	Mean_Gyr_x
16	MV_Gyr	MV_Gyr	Kur_Pitch	Skew_Roll	Mean_Gyr_x	Var_Gyr_x
17	Dfreq_Pitch	Dfreq_Pitch	Kur_Gyr_z	Skew_Gyr_y	Mean_Pitch	Min_Gyr_x
18	Mean_Pitch	Mean_Pitch	Skew_Acc_x	Min_Acc_x	Var_Acc_y	Rms_Pitch
19	Entropy_Gyr_y	Min_Pitch	Skew_Acc_y	Min_Acc_z	Dfreq_Roll	Skew_Acc_y
20	Min_Pitch	Max_diff_Gyr_x	Skew_Pitch	Min_Pitch	Min_Acc_z	Var_Gyr_z
21	Max_diff_Gyr_x	Max_diff_Acc_y	Skew_Gyr_x	Min_Roll	Highest_peak_Pitch	Max_Pitch
22	Max_diff_Acc_y	Entropy_Gyr_y	Skew_Gyr_y	Min_Gyr_x	Highest_peak_Gyr_z	Max_diff_Gyr_y
23	Max_diff_Gyr_z	SVM_Angle	Min_Acc_x	Min_Gyr_y	Cf_Roll	Min_Roll
24	Entropy_Gyr_x	Entropy_TimeD_Ang	Min_Acc_y	Min_Gyr_z	Kur_Gyr_y	Var_Pitch
25	Max_Acc_y	SMA_Angle	Min_Roll	Max_Acc_x	Kur_Pitch	Mean_Pitch

26	MV_Acc	Max_Acc_y	Min_Gyr_x	Max_Acc_y	Min_Acc_y	Entropy_Roll
27	Highest_peak_Gyr_y	MV_Acc	Min_Gyr_y	Max_Acc_z	Var_Roll	Min_Gyr_z
28	Highest_peak_Pitch	Max_diff_Gyr_z	Max_Acc_y	Max_Roll	Var_Acc_x	Max_diff_Acc_y
29	Dfreq_Gyr_y	Min_Gyr_x	Max_Acc_z	Max_Gyr_x	Interq_Roll	Interq_Gyr_y
30	Min_Gyr_x	Skew_Pitch	Max_Azimuth	Max_Gyr_z	Cf_Gyr_x	Kur_Acc_z
31	Min_Gyr_z	Highest_peak_Gyr_y	Max_Gyr_x	Rms_Acc_x	Max_Acc_z	Dfreq_Acc_y
32	Rms_Acc_y	Highest_peak_Pitch	Rms_Acc_y	Rms_Acc_z	Max_Acc_x	Min_Acc_y
33	Std_Gyr_x	Max_Acc_x	Rms_Acc_z	Rms_Azimuth	Var_Acc_z	SMA_Acc
34	SVM_Angle	Interq_Gyr_z	Rms_Azimuth	Rms_Roll	Interq_Acc_z	Rms_Acc_z
35	Std_Gyr_z	Rms_Acc_x	Rms_Gyr_x	Rms_Gyr_x	Rms_Gyr_z	Interq_Gyr_z
36	Entropy_TimeD_Ang	Std_Gyr_z	Rms_Gyr_y	Rms_Gyr_y	Skew_Gyr_y	Highest_peak_Gyr_y
37	Rms_Acc_x	Rms_Gyr_z	Rms_Gyr_z	Interq_Acc_y	Max_diff_Gyr_y	MV_Gyr
38	SVM_Acc	Vedb_Angle	Interq_Acc_y	Interq_Azimuth	Entropy_Gyr_x	Max_diff_Acc_z
39	DSVM_Gyr	Min_Gyr_z	Interq_Azimuth	Interq_Pitch	Interq_Gyr_x	Min_Pitch
40	SMA_Angle	Std_Gyr_x	Interq_Roll	Interq_Gyr_x	Interq_Gyr_z	Dfreq_Gyr_x
41	Entropy_Gyr_z	Std_Acc_y	Interq_Gyr_y	Interq_Gyr_z	Dfreq_Acc_z	Max_diff_Acc_x
42	SMA_Acc	Rms_Acc_z	Interq_Gyr_z	Cf_Acc_x	Interq_Acc_x	Kur_Acc_x
43	Interq_Gyr_z	Entropy_Gyr_x	Cf_Acc_x	Cf_Acc_y	MV_Gyr	DSAM_Angle
44	Std_Acc_y	mag_Ang	Cf_Acc_z	Cf_Acc_z	Min_Roll	Var_Roll
45	Rms_Gyr_z	Max_diff_Azimuth	Cf_Pitch	Cf_Pitch	DSVM_Acc	Rms_Gyr_z
46	Rms_Acc_z	Rms_Gyr_x	Cf_Roll	Cf_Roll	Rms_Pitch	Var_Azimuth
47	AV_Ang	Var_Pitch	Cf_Gyr_y	Cf_Gyr_y	Dfreq_Acc_x	Rms_Gyr_y
48	Std_Acc_x	Interq_Gyr_y	Cf_Gyr_z	SMA_Acc	Skew_Acc_y	Max_Gyr_z
49	Rms_Gyr_x	Entropy_Acc_y	SMA_Angle	SMA_Gyr	Dfreq_Gyr_x	Dfreq_Gyr_y
50	Highest_peak_Acc_y	SVM_Acc	Entropy_Acc_x	Entropy_Acc_y	Max_diff_Gyr_x	Max_diff_Gyr_x
51	Entropy_TimeD_Acc	Rms_Acc_y	Entropy_Acc_y	Entropy_Acc_z	Var_Gyr_y	AV_Ang
52	Skew_Pitch	SMA_Acc	Entropy_Acc_z	Entropy_Pitch	Max_diff_Acc_z	MV_Acc
53	Entropy_Azimuth	Dfreq_Azimuth	Entropy_Gyr_y	Entropy_Roll	DSAM_Angle	Kur_Acc_y
54	Max_diff_Azimuth	Mean_Azimuth	Entropy_Gyr_z	Entropy_Gyr_z	Max_diff_Acc_y	Dfreq_Acc_x

55	Entropy_Acc_y	DSVM_Gyr	Energy_Roll	Energy_Azimuth	Rms_Gyr_y	Var_Acc_y
56	Interq_Azimuth	Max_diff_Acc_z	Energy_Gyr_x	Energy_Roll	Rms_Acc_x	Max_Acc_x
57	Max_diff_Acc_z	Max_Acc_z	Energy_Gyr_z	Energy_Gyr_y	SMA_Gyr	Dfreq_Acc_z
58	Skew_Acc_y	Cf_Acc_x	SVM_Acc	SVM_Acc	Rms_Gyr_x	Dfreq_Pitch
59	Cf_Acc_x	Max_Gyr_x	SVM_Gyr	Max_diff_Acc_z	Max_diff_Azimuth	Vedb_Acc
60	Max_Acc_x	DSAM_Angle	DSVM_Acc	Max_diff_Azimuth	Min_Pitch	Rms_Gyr_x
61	Max_Gyr_x	Highest_peak_Acc_y	DSAM_Angle	Max_diff_Pitch	SMA_Acc	Var_Acc_x
62	Std_Azimuth	Entropy_Acc_x	DSVM_Gyr	Max_diff_Gyr_y	Max_diff_Acc_x	Max_Acc_z
63	Cf_Gyr_y	Dfreq_Gyr_y	Max_diff_Acc_x	Max_diff_Gyr_z	Dfreq_Pitch	DSVM_Gyr
64	Interq_Gyr_y	Entropy_TimeD_Acc	Max_diff_Acc_y	MV_Acc	Dfreq_Gyr_y	Mean_Acc_y
65	Cf_Gyr_x	AV_Ang	Max_diff_Azimuth	MV_Gyr	nPeaks_Gyr_z	Highest_peak_Azimuth
66	DSAM_Angle	Skew_Acc_x	Max_diff_Pitch	Vedb_Gyr	Min_Gyr_y	Kur_Gyr_x
67	Highest_peak_Gyr_x	Std_Acc_x	Max_diff_Roll	Dfreq_Acc_y	Min_Gyr_z	Kur_Gyr_z
68	Interq_Acc_y	Entropy_Gyr_z	Max_diff_Gyr_y	Dfreq_Acc_z	mag_Ang	Entropy_Acc_z
69	Var_Pitch	Skew_Acc_y	MV_Gyr	Dfreq_Roll	Avr_peak_time_Azimuth	Rms_Azimuth
70	Interq_Gyr_x	Interq_Acc_x	mag_Ang	Dfreq_Gyr_x	Avr_peak_time_Acc_x	Max_Acc_y
71	Entropy_Acc_x	Dfreq_Acc_z	Vedb_Angle	Dfreq_Gyr_z	Interq_Acc_y	Mean_Acc_x
72	Min_Azimuth	Min_Gyr_y	Vedb_Gyr	Widest_Peak_Acc_x	Max_Acc_y	Vedb_Gyr
73	Kur_Pitch	Vedb_Gyr	Dfreq_Acc_z	Highest_peak_Acc_x	SVM_Acc	Interq_Azimuth
74	Std_Acc_z	Std_Gyr_y	Dfreq_Azimuth	Avr_peak_time_Acc_x	Min_Acc_x	Highest_peak_Gyr_z
75	Std_Gyr_y	Max_Azimuth	Dfreq_Pitch	nPeaks_Acc_y	Entropy_Acc_y	Min_Acc_x
76	Min_Acc_x	Rms_Azimuth	Dfreq_Roll	Widest_Peak_Acc_y	Dfreq_Acc_y	Interq_Roll
77	Rms_Gyr_y	Highest_peak_Gyr_z	Dfreq_Gyr_y	nPeaks_Acc_z	Highest_peak_Acc_y	SVM_Acc
78	Vedb_Angle	Max_Gyr_y	Dfreq_Gyr_z	Highest_peak_Acc_z	Rms_Acc_y	Entropy_Acc_y
79	Entropy_Acc_z	Entropy_Acc_z	nPeaks_Acc_x	nPeaks_Azimuth	Entropy_TimeD_Gyr	Var_Gyr_y
80	Vedb_Acc	Energy_Azimuth	Highest_peak_Acc_x	Widest_Peak_Azimuth	mag_Gyr	Interq_Pitch
81	Interq_Acc_x	Skew_Acc_z	nPeaks_Acc_y	Highest_peak_Azimuth	Highest_peak_Gyr_x	DSVM_Acc
82	Dfreq_Acc_z	Max_diff_Gyr_y	Widest_Peak_Acc_y	Avr_peak_time_Azimuth	MV_Acc	Skew_Gyr_y
83	mag_Ang	Cf_Gyr_y	Avr_peak_time_Acc_y	nPeaks_Pitch	Highest_peak_Gyr_y	Cf_Gyr_z

84	Highest_peak_Gyr_z	Highest_peak_Acc_x	Highest_peak_Acc_z	Widest_Peak_Pitch	Kur_Acc_x	SMA_Angle
85	Vedb_Gyr	Kur_Pitch	Avr_peak_time_Acc_z	Highest_peak_Pitch	Kur_Gyr_z	Avr_peak_time_Gyr_x
86	Var_Azimuth	Rms_Gyr_y	nPeaks_Azimuth	Widest_Peak_Roll	Mean_Acc_y	Interq_Gyr_x
87	Min_Acc_y	Min_Azimuth	nPeaks_Pitch	Avr_peak_time_Roll	Dfreq_Gyr_z	Mean_Gyr_y
88	Var_Acc_y	Dfreq_Acc_x	Avr_peak_time_Pitch	nPeaks_Gyr_x	Vedb_Gyr	Skew_Pitch
89	Skew_Acc_z	Cf_Acc_y	Highest_peak_Roll	Highest_peak_Gyr_x	Mean_Acc_z	Max_diff_Azimuth
90	Energy_Acc_y	Skew_Roll	Avr_peak_time_Roll	Highest_peak_Gyr_y	Interq_Gyr_y	Entropy_Gyr_z
91	Dfreq_Azimuth	Std_Acc_z	nPeaks_Gyr_x	nPeaks_Gyr_z	Highest_peak_Acc_z	Dfreq_Gyr_z
92	Mean_Azimuth	Var_Acc_y	Widest_Peak_Gyr_x	Widest_Peak_Gyr_z	AV_Ang	Rms_Acc_y
93	Max_Gyr_z	Entropy_Azimuth	Highest_peak_Gyr_x	Highest_peak_Gyr_z	DSVM_Gyr	Highest_peak_Acc_y
94	Highest_peak_Acc_x	Min_Acc_z	Avr_peak_time_Gyr_x		Skew_Gyr_z	mag_Acc
95	Dfreq_Acc_y	Interq_Gyr_x	nPeaks_Gyr_z		Vedb_Acc	Interq_Acc_y
96	Min_Gyr_y	Interq_Azimuth	Widest_Peak_Gyr_z		SVM_Gyr	Mean_Acc_z
97	Max_Azimuth	Cf_Acc_z	Avr_peak_time_Gyr_z		Entropy_Roll	Mean_Gyr_z
98	Cf_Acc_y	Dfreq_Gyr_z			Mean_Gyr_z	Entropy_Gyr_y
99	Skew_Acc_x	Cf_Gyr_x			Widest_Peak_Roll	Highest_peak_Acc_z
100	Max_Acc_z	Min_Acc_x			Kur_Acc_z	Rms_Acc_x
101	DSVM_Acc	Max_Gyr_z			Cf_Acc_y	Entropy_TimeD_Acc
102	Max_Gyr_y	Energy_Acc_y			Interq_Azimuth	Kur_Pitch
103	Kur_Acc_x	Dfreq_Acc_y			Highest_peak_Acc_x	Highest_peak_Acc_x
104	Min_Acc_z	SVM_Gyr			Entropy_TimeD_Ang	Cf_Gyr_y
105	Var_Gyr_x	SMA_Gyr			nPeaks_Acc_y	nPeaks_Acc_z
106	Highest_peak_Azimuth	Interq_Acc_y			Entropy_Azimuth	Min_Gyr_y
107	Dfreq_Acc_x	Dfreq_Gyr_x			Skew_Acc_x	Interq_Acc_z
108	Max_diff_Acc_x	Mean_Acc_z			Highest_peak_Azimuth	mag_Ang
109	Energy_Gyr_x	Entropy_TimeD_Gyr			Avr_peak_time_Roll	Skew_Acc_x
110	SVM_Gyr	Vedb_Acc			Interq_Pitch	Cf_Azimuth
111	Rms_Azimuth	Highest_peak_Gyr_x			Entropy_TimeD_Acc	Skew_Acc_z
112	SMA_Gyr	Var_Gyr_x			Skew_Acc_z	Skew_Gyr_x

113	Max_diff_Gyr_y	Var_Gyr_z			Cf_Gyr_y	Avr_peak_time_Pitch
114	Skew_Roll	Energy_Acc_x			Skew_Pitch	Kur_Gyr_y
115	Dfreq_Gyr_z	Max_diff_Acc_x			Mean_Azimuth	Entropy_Gyr_x
116	Energy_Gyr_z	Energy_Gyr_x			Rms_Acc_z	Cf_Gyr_x
117	Entropy_TimeD_Gyr	Energy_Gyr_z			Kur_Acc_y	nPeaks_Gyr_z
118	Cf_Acc_z	Skew_Gyr_y			Vedb_Angle	Energy_Azimuth
119	Dfreq_Gyr_x	Std_Azimuth			mag_Acc	Avr_peak_time_Gyr_z
120	Mean_Acc_y	Min_Acc_y			Mean_Gyr_y	Avr_peak_time_Acc_z
121	Energy_Azimuth	Mean_Gyr_z			nPeaks_Acc_z	Skew_Gyr_z
122	Var_Gyr_z	DSVM_Acc			Cf_Azimuth	Energy_Acc_y
123	Energy_Acc_z	Mean_Acc_y			Widest_Peak_Acc_z	Max_Azimuth
124	Cf_Gyr_z	Energy_Acc_z			Avr_peak_time_Gyr_y	Cf_Acc_x
125	Highest_peak_Acc_z	Skew_Gyr_z			Highest_peak_Roll	SVM_Gyr
126	Mean_Acc_z	Kur_Acc_x			Widest_Peak_Acc_y	SMA_Gyr
127	Var_Acc_x	Kur_Gyr_y			Avr_peak_time_Acc_y	Widest_Peak_Acc_x
128	Energy_Acc_x	Kur_Gyr_x			Avr_peak_time_Pitch	nPeaks_Pitch
129	Mean_Gyr_z	Cf_Gyr_z			Max_Gyr_z	Min_Azimuth
130	Interq_Acc_z	Highest_peak_Acc_z			Entropy_Gyr_y	Energy_Roll
131	Cf_Azimuth	Std_Roll			Avr_peak_time_Gyr_z	Widest_Peak_Azimuth
132	Skew_Gyr_y	Interq_Acc_z			nPeaks_Gyr_x	Entropy_TimeD_Gyr
133	Kur_Gyr_x	Highest_peak_Azimuth			nPeaks_Gyr_y	mag_Gyr
134	Widest_Peak_Acc_x	Var_Azimuth			Kur_Azimuth	Skew_Roll
135	Skew_Azimuth	Var_Acc_x			nPeaks_Pitch	Highest_peak_Roll
136	nPeaks_Pitch	Var_Gyr_y			Entropy_Gyr_z	Widest_Peak_Gyr_z
137	Widest_Peak_Acc_z	Mean_Acc_x			Avr_peak_time_Acc_z	Max_Gyr_y
138	Var_Acc_z	Interq_Pitch			Cf_Acc_z	Entropy_Acc_x
139	Energy_Gyr_y	Energy_Gyr_y			Std_Acc_x	Std_Gyr_z
140	Std_Roll	Interq_Roll			Std_Acc_y	nPeaks_Gyr_y
141	Var_Gyr_y	Kur_Gyr_z			Std_Acc_z	Widest_Peak_Gyr_y

142	Kur_Azimuth	Kur_Acc_z			Std_Azimuth	Widest_Peak_Acc_y
143	Kur_Acc_y	Cf_Azimuth			Std_Pitch	Std_Gyr_x
144	Skew_Gyr_z	Skew_Gyr_x			Std_Roll	Highest_peak_Gyr_x
145	Mean_Acc_x	Var_Acc_z			Std_Gyr_x	Kur_Azimuth
146	Kur_Gyr_y	mag_Acc			Std_Gyr_y	nPeaks_Acc_y
147	Widest_Peak_Azimuth	Widest_Peak_Acc_y			Std_Gyr_z	Entropy_Azimuth
148	Interq_Pitch	Max_diff_Roll			Energy_Acc_x	Energy_Acc_z
149	Kur_Gyr_z	Widest_Peak_Gyr_x			Energy_Acc_y	Energy_Gyr_y
150	Skew_Gyr_x	Widest_Peak_Azimuth			Energy_Acc_z	nPeaks_Gyr_x
151	mag_Acc	Kur_Azimuth			Energy_Azimuth	Entropy_TimeD_Ang
152	Mean_Gyr_y	Var_Roll			Energy_Pitch	Mean_Azimuth
153	nPeaks_Roll	Mean_Gyr_x			Energy_Roll	nPeaks_Azimuth
154	Highest_peak_Roll	Highest_peak_Roll			Energy_Gyr_x	Widest_Peak_Acc_z
155	Kur_Acc_z	Kur_Acc_y			Energy_Gyr_y	Widest_Peak_Roll
156	Interq_Roll	Skew_Azimuth			Energy_Gyr_z	nPeaks_Roll
157	Var_Roll	Mean_Gyr_y			Dfreq_Azimuth	Avr_peak_time_Azimuth
158	Max_diff_Roll	Kur_Roll			nPeaks_Azimuth	Skew_Azimuth
159	nPeaks_Azimuth	Widest_Peak_Gyr_z			Entropy_Acc_z	Std_Gyr_y
160	Avr_peak_time_Gyr_z	nPeaks_Pitch			Var_Azimuth	Avr_peak_time_Roll
161	nPeaks_Gyr_z	Avr_peak_time_Azimuth			Max_diff_Roll	Energy_Gyr_x
162	Widest_Peak_Acc_y	Widest_Peak_Acc_z			Max_Gyr_y	Avr_peak_time_Acc_x
163	Avr_peak_time_Azimuth	Widest_Peak_Acc_x			Rms_Azimuth	Cf_Acc_y
164	Avr_peak_time_Roll	Widest_Peak_Gyr_y			Skew_Roll	Energy_Pitch
165	mag_Gyr	nPeaks_Gyr_y			SMA_Angle	Max_diff_Roll
166	nPeaks_Acc_x	Avr_peak_time_Pitch			Widest_Peak_Acc_x	Std_Roll
167	Avr_peak_time_Acc_x	Widest_Peak_Roll			nPeaks_Acc_x	Std_Acc_x
168	Widest_Peak_Gyr_y	Avr_peak_time_Gyr_y			Cf_Gyr_z	Std_Acc_y
169	Kur_Roll	mag_Gyr			Widest_Peak_Gyr_y	Avr_peak_time_Acc_y
170	nPeaks_Gyr_y	Avr_peak_time_Acc_z			Kur_Roll	Dfreq_Azimuth

171	Avr_peak_time_Acc_y	Avr_peak_time_Acc_y			Skew_Azimuth	Std_Azimuth
172	Mean_Gyr_x	nPeaks_Gyr_z			Cf_Acc_x	Std_Acc_z
173	Widest_Peak_Gyr_z	Avr_peak_time_Gyr_x			Widest_Peak_Gyr_z	Cf_Acc_z
174	Widest_Peak_Roll	Avr_peak_time_Acc_x			Max_Azimuth	Vedb_Angle
175	nPeaks_Acc_y	nPeaks_Azimuth			nPeaks_Roll	Energy_Acc_x
176	nPeaks_Gyr_x	Widest_Peak_Pitch			Skew_Gyr_x	Avr_peak_time_Gyr_y
177	Avr_peak_time_Pitch	nPeaks_Acc_y			Avr_peak_time_Gyr_x	Energy_Gyr_z
178	Avr_peak_time_Acc_z	Avr_peak_time_Roll			Widest_Peak_Pitch	SVM_Angle
179	Avr_peak_time_Gyr_x	nPeaks_Acc_x			Widest_Peak_Gyr_x	Kur_Roll
180	Widest_Peak_Pitch	nPeaks_Gyr_x			Min_Azimuth	Widest_Peak_Gyr_x
181	nPeaks_Acc_z	nPeaks_Roll			Widest_Peak_Azimuth	Std_Pitch
182	Widest_Peak_Gyr_x	nPeaks_Acc_z			SVM_Angle	Widest_Peak_Pitch
183	Avr_peak_time_Gyr_y	Avr_peak_time_Gyr_z			Entropy_Acc_x	nPeaks_Acc_x

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 4 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet2_ac** over 10 $sec. window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Entropy_Roll	Entropy_Roll	Mean_Acc_x	Mean_Azimuth	Min_Roll	Rms_Azimuth
2	Entropy_Acc_x	Dfreq_Acc_x	Mean_Azimuth	Mean_Pitch	Mean_Roll	Min_Roll
3	Dfreq_Acc_x	Entropy_Acc_x	Mean_Roll	Var_Acc_x	Rms_Azimuth	Mean_Roll
4	Dfreq_Roll	Dfreq_Roll	Var_Acc_x	Var_Roll	Mean_Acc_x	Mean_Acc_x
5	Rms_Roll	Rms_Roll	Var_Acc_z	Std_Acc_x	Entropy_Acc_x	Interq_Azimuth
6	Mean_Roll	Mean_Roll	Var_Pitch	Std_Acc_y	Vedb_Angle	Dfreq_Roll
7	Rms_Acc_x	Mean_Acc_x	Std_Acc_y	Std_Pitch	Mean_Acc_z	Max_Pitch
8	Mean_Acc_x	Rms_Acc_x	Std_Acc_z	Kur_Acc_y	Rms_Pitch	Mean_Acc_z
9	Max_Roll	Dfreq_Acc_z	Std_Azimuth	Kur_Pitch	Max_diff_Acc_y	Entropy_Roll
10	Energy_Roll	Max_Roll	Std_Pitch	Skew_Acc_x	Entropy_Roll	Min_Acc_z
11	Min_Acc_x	Mean_Acc_z	Std_Roll	Skew_Acc_z	Rms_Roll	Rms_Pitch
12	Dfreq_Acc_z	Energy_Roll	Kur_Acc_x	Skew_Azimuth	DSAM_Angle	Rms_Roll
13	Mean_Acc_z	Min_Acc_x	Kur_Pitch	Skew_Pitch	Mean_Acc_y	Max_Acc_y
14	Energy_Acc_z	Energy_Acc_z	Skew_Acc_y	Min_Acc_x	Max_Pitch	Entropy_Acc_x
15	Rms_Acc_z	Rms_Acc_z	Skew_Acc_z	Min_Acc_z	Min_Acc_z	Mean_Acc_y
16	Energy_Acc_x	Mean_Acc_y	Skew_Azimuth	Max_Acc_x	Skew_Roll	Max_Roll
17	Mean_Acc_y	Energy_Acc_x	Skew_Roll	Max_Pitch	Rms_Acc_x	Min_Pitch
18	Cf_Acc_x	Cf_Acc_x	Min_Acc_z	Max_Roll	Cf_Roll	Dfreq_Acc_x
19	Max_Acc_y	Max_Acc_y	Min_Azimuth	Rms_Acc_x	Mean_Azimuth	Mean_Azimuth
20	Max_Pitch	Rms_Pitch	Min_Roll	Rms_Acc_y	Dfreq_Roll	Vedb_Angle
21	Cf_Roll	Max_Pitch	Max_Acc_z	Rms_Pitch	DSVM_Acc	DSAM_Angle
22	Min_Roll	Mean_Pitch	Max_Azimuth	Rms_Roll	Mean_Pitch	Rms_Acc_x
23	Rms_Pitch	Min_Roll	Max_Pitch	Interq_Azimuth	Var_Acc_x	Skew_Roll
24	Mean_Pitch	Cf_Roll	Max_Roll	Interq_Pitch	Var_Roll	Min_Acc_x
25	SVM_Angle	Energy_Pitch	Rms_Acc_x	Cf_Acc_y	Max_diff_Roll	Rms_Acc_z

26	SMA_Angle	SMA_Angle	Rms_Acc_z	Cf_Azimuth	Cf_Azimuth	Skew_Azimuth
27	Vedb_Acc	SVM_Angle	Rms_Azimuth	Cf_Roll	MV_Acc	Cf_Roll
28	Energy_Pitch	Vedb_Acc	Rms_Roll	SMA_Acc	Max_Roll	Skew_Acc_y
29	SMA_Acc	SVM_Acc	Interq_Azimuth	SMA_Angle	Rms_Acc_y	Var_Roll
30	SVM_Acc	SMA_Acc	Cf_Acc_z	Entropy_Acc_x	Rms_Acc_z	Kur_Azimuth
31	Entropy_TimeD_Ang	Entropy_TimeD_Ang	Cf_Azimuth	Entropy_Acc_y	Min_Pitch	Var_Pitch
32	Cf_Pitch	Min_Azimuth	Cf_Pitch	Entropy_Acc_z	Entropy_TimeD_Ang	Mean_Pitch
33	Max_Acc_x	Min_Acc_z	Cf_Roll	Entropy_Azimuth	Interq_Roll	SMA_Acc
34	DSAM_Angle	Min_Pitch	Entropy_Acc_x	Entropy_Pitch	Dfreq_Acc_x	Cf_Azimuth
35	Entropy_TimeD_Acc	Cf_Acc_y	Entropy_Acc_y	Entropy_Roll	Var_Pitch	Interq_Pitch
36	Cf_Acc_y	Entropy_TimeD_Acc	Entropy_Acc_z	Entropy_TimeD_Ang	Cf_Pitch	MV_Acc
37	Min_Pitch	Max_Acc_x	Entropy_TimeD_Acc	Energy_Acc_x	Max_Acc_y	Entropy_TimeD_Acc
38	Min_Azimuth	Skew_Azimuth	Entropy_TimeD_Ang	Energy_Acc_z	Kur_Acc_y	Max_diff_Roll
39	Min_Acc_z	Max_diff_Azimuth	Energy_Acc_z	SVM_Acc	Min_Acc_x	Vedb_Acc
40	Skew_Azimuth	Cf_Pitch	Energy_Azimuth	DSAM_Angle	Interq_Acc_z	Skew_Acc_x
41	Entropy_Acc_z	Entropy_Acc_z	Energy_Roll	Max_diff_Acc_z	Entropy_TimeD_Acc	DSVM_Acc
42	Max_diff_Acc_x	Rms_Acc_y	SVM_Acc	Max_diff_Pitch	Highest_peak_Acc_x	Var_Azimuth
43	Rms_Acc_y	DSAM_Angle	SVM_Angle	Max_diff_Roll	Skew_Pitch	Min_Acc_y
44	Max_diff_Azimuth	Dfreq_Pitch	DSAM_Angle	AV_Ang	Dfreq_Pitch	Cf_Acc_x
45	DSVM_Acc	Skew_Acc_y	Max_diff_Acc_y	mag_Acc	Min_Acc_y	Var_Acc_z
46	Max_Acc_z	Energy_Acc_y	Max_diff_Acc_z	Vedb_Angle	Skew_Acc_y	Max_Azimuth
47	Max_diff_Roll	DSVM_Acc	Max_diff_Azimuth	Dfreq_Acc_y	SVM_Acc	Entropy_TimeD_Ang
48	Energy_Acc_y	Max_Acc_z	AV_Ang	Dfreq_Acc_z	SMA_Angle	Skew_Pitch
49	Std_Pitch	Vedb_Angle	mag_Ang	Dfreq_Azimuth	Cf_Acc_y	Cf_Pitch
50	Vedb_Angle	Entropy_Azimuth	Vedb_Acc	Dfreq_Roll	Min_Azimuth	Rms_Acc_y
51	Entropy_Azimuth	Max_diff_Acc_x	Dfreq_Acc_x	Widest_Peak_Acc_x	Interq_Acc_x	Kur_Pitch
52	Dfreq_Pitch	Kur_Azimuth	Dfreq_Acc_y	Avr_peak_time_Acc_x	Vedb_Acc	Var_Acc_x
53	Std_Roll	Std_Roll	Dfreq_Acc_z	Avr_peak_time_Acc_y	Interq_Pitch	Interq_Roll
54	Std_Acc_x	Std_Pitch	Dfreq_Roll	nPeaks_Acc_z	SMA_Acc	Var_Acc_y

55	Dfreq_Acc_y	Min_Acc_y	Widest_Peak_Acc_x	Widest_Peak_Acc_z	Kur_Pitch	Highest_peak_Acc_x
56	Cf_Azimuth	Std_Acc_y	Highest_peak_Acc_x	Highest_peak_Acc_z	Skew_Azimuth	Max_diff_Pitch
57	Entropy_Pitch	Dfreq_Acc_y	Avr_peak_time_Acc_x	Widest_Peak_Azimuth	Dfreq_Acc_y	Max_diff_Azimuth
58	Std_Acc_y	Max_diff_Roll	nPeaks_Acc_y	Highest_peak_Azimuth	Interq_Azimuth	Kur_Acc_y
59	Skew_Pitch	Entropy_Pitch	Highest_peak_Acc_y	nPeaks_Pitch	AV_Ang	Cf_Acc_y
60	Entropy_Acc_y	Skew_Pitch	nPeaks_Acc_z	Avr_peak_time_Pitch	Cf_Acc_z	mag_Ang
61	Min_Acc_y	Entropy_Acc_y	Widest_Peak_Acc_z	Avr_peak_time_Roll	SVM_Angle	Kur_Roll
62	Skew_Acc_y	Std_Acc_x	nPeaks_Azimuth		Max_diff_Acc_x	SMA_Angle
63	Max_diff_Pitch	Rms_Azimuth	Avr_peak_time_Azimuth		Entropy_Acc_z	Dfreq_Acc_y
64	Rms_Azimuth	Max_Azimuth	Avr_peak_time_Pitch		Kur_Acc_z	Kur_Acc_x
65	Kur_Azimuth	Max_diff_Pitch	nPeaks_Roll		Kur_Azimuth	Widest_Peak_Acc_y
66	Dfreq_Azimuth	Dfreq_Azimuth	Widest_Peak_Roll		Max_Azimuth	Kur_Acc_z
67	Cf_Acc_z	Interq_Acc_y	Avr_peak_time_Roll		Entropy_Acc_y	Entropy_Pitch
68	AV_Ang	Mean_Azimuth			Cf_Acc_x	Cf_Acc_z
69	Mean_Azimuth	AV_Ang			Highest_peak_Azimuth	Max_diff_Acc_y
70	Interq_Acc_y	Cf_Azimuth			Avr_peak_time_Acc_z	SVM_Acc
71	Max_Azimuth	Cf_Acc_z			nPeaks_Pitch	Interq_Acc_x
72	Std_Acc_z	Var_Acc_y			Highest_peak_Roll	Max_Acc_x
73	Var_Acc_y	Skew_Roll			Max_diff_Pitch	Highest_peak_Acc_y
74	Max_diff_Acc_y	Std_Acc_z			Var_Acc_z	Highest_peak_Pitch
75	Highest_peak_Acc_y	Max_diff_Acc_z			Max_Acc_z	Highest_peak_Azimuth
76	Var_Pitch	Max_diff_Acc_y			Avr_peak_time_Pitch	Dfreq_Pitch
77	MV_Acc	Highest_peak_Acc_x			Max_Acc_x	AV_Ang
78	Highest_peak_Pitch	mag_Ang			Entropy_Pitch	Avr_peak_time_Pitch
79	mag_Ang	MV_Acc			mag_Ang	Entropy_Acc_y
80	Var_Roll	Std_Azimuth			mag_Acc	Interq_Acc_z
81	Interq_Pitch	Highest_peak_Acc_y			Widest_Peak_Azimuth	Max_diff_Acc_x
82	Highest_peak_Acc_x	Interq_Pitch			Entropy_Azimuth	Skew_Acc_z
83	Var_Acc_x	Var_Pitch			Avr_peak_time_Acc_x	Highest_peak_Acc_z

84	Std_Azimuth	Highest_peak_Roll			Max_diff_Azimuth	Highest_peak_Roll
85	Max_diff_Acc_z	Var_Roll			Max_diff_Acc_z	Widest_Peak_Acc_x
86	Skew_Roll	Interq_Acc_z			nPeaks_Roll	SVM_Angle
87	Highest_peak_Acc_z	Kur_Acc_z			Avr_peak_time_Azimuth	Max_diff_Acc_z
88	Skew_Acc_z	Skew_Acc_x			Interq_Acc_y	Interq_Acc_y
89	Interq_Roll	Skew_Acc_z			Highest_peak_Acc_z	Entropy_Acc_z
90	Kur_Acc_x	Var_Acc_x			Var_Acc_y	Avr_peak_time_Acc_z
91	Interq_Acc_z	Highest_peak_Acc_z			Var_Azimuth	Widest_Peak_Pitch
92	Highest_peak_Roll	Interq_Roll			Skew_Acc_x	Min_Azimuth
93	Skew_Acc_x	Highest_peak_Pitch			Highest_peak_Pitch	Max_Acc_z
94	Var_Acc_z	Interq_Azimuth			Highest_peak_Acc_y	Entropy_Azimuth
95	Interq_Acc_x	Kur_Pitch			Widest_Peak_Acc_y	nPeaks_Roll
96	Kur_Acc_z	Var_Acc_z			nPeaks_Acc_z	Avr_peak_time_Acc_x
97	nPeaks_Acc_y	Kur_Acc_x			nPeaks_Acc_x	nPeaks_Acc_x
98	Interq_Azimuth	Kur_Acc_y			Kur_Roll	nPeaks_Pitch
99	Energy_Azimuth	Interq_Acc_x			Avr_peak_time_Acc_y	Widest_Peak_Roll
100	nPeaks_Acc_z	Energy_Azimuth			nPeaks_Acc_y	mag_Acc
101	Kur_Pitch	nPeaks_Acc_x			Std_Acc_x	Avr_peak_time_Acc_y
102	Kur_Roll	nPeaks_Acc_z			Std_Acc_y	Avr_peak_time_Roll
103	Avr_peak_time_Acc_z	Kur_Roll			Std_Acc_z	Std_Acc_x
104	nPeaks_Azimuth	Var_Azimuth			Std_Azimuth	Std_Acc_y
105	Var_Azimuth	Avr_peak_time_Acc_z			Std_Pitch	Std_Acc_z
106	Highest_peak_Azimuth	Highest_peak_Azimuth			Std_Roll	Std_Azimuth
107	Kur_Acc_y	nPeaks_Acc_y			Energy_Acc_x	Std_Pitch
108	Widest_Peak_Acc_y	Avr_peak_time_Acc_x			Energy_Acc_y	Std_Roll
109	nPeaks_Acc_x	Widest_Peak_Acc_z			Energy_Acc_z	Energy_Acc_x
110	nPeaks_Pitch	Avr_peak_time_Acc_y			Energy_Azimuth	Energy_Acc_y
111	Widest_Peak_Pitch	Widest_Peak_Pitch			Energy_Pitch	Energy_Acc_z
112	Widest_Peak_Acc_x	Widest_Peak_Acc_y			Energy_Roll	Energy_Azimuth

113	Avr_peak_time_Acc_y	nPeaks_Roll			Dfreq_Acc_z	Energy_Pitch
114	Avr_peak_time_Azimuth	Avr_peak_time_Roll			Dfreq_Azimuth	Energy_Roll
115	Avr_peak_time_Pitch	nPeaks_Pitch			Widest_Peak_Acc_x	Dfreq_Acc_z
116	Avr_peak_time_Acc_x	nPeaks_Azimuth			nPeaks_Azimuth	Dfreq_Azimuth
117	Widest_Peak_Roll	mag_Acc			Avr_peak_time_Roll	nPeaks_Azimuth
118	nPeaks_Roll	Widest_Peak_Acc_x			Kur_Acc_x	nPeaks_Acc_z
119	Widest_Peak_Azimuth	Avr_peak_time_Azimuth			Widest_Peak_Acc_z	Widest_Peak_Acc_z
120	Avr_peak_time_Roll	Widest_Peak_Azimuth			Widest_Peak_Pitch	nPeaks_Acc_y
121	Widest_Peak_Acc_z	Widest_Peak_Roll			Widest_Peak_Roll	Widest_Peak_Azimuth
122	mag_Acc	Avr_peak_time_Pitch			Skew_Acc_z	Avr_peak_time_Azimuth

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 5 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet2_ac** over 7 $sec.window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Entropy_Roll	Entropy_Roll	Mean_Acc_x	Mean_Acc_y	Min_Roll	Min_Roll
2	Dfreq_Acc_x	Dfreq_Acc_x	Mean_Acc_z	Mean_Roll	Mean_Roll	Mean_Roll
3	Entropy_Acc_x	Dfreq_Roll	Mean_Azimuth	Var_Azimuth	Mean_Acc_x	Mean_Acc_z
4	Dfreq_Roll	Entropy_Acc_x	Var_Acc_x	Std_Acc_z	Mean_Acc_z	Entropy_Acc_x
5	Rms_Roll	Rms_Roll	Var_Acc_z	Std_Azimuth	Max_Pitch	Min_Acc_z
6	Mean_Roll	Mean_Roll	Var_Azimuth	Std_Pitch	Entropy_Roll	Max_Azimuth
7	Dfreq_Acc_z	Dfreq_Acc_z	Var_Pitch	Kur_Acc_y	Mean_Azimuth	Mean_Acc_x
8	Mean_Acc_z	Max_Roll	Var_Roll	Kur_Acc_z	Min_Acc_z	Rms_Roll
9	Min_Acc_x	Mean_Acc_x	Std_Pitch	Kur_Pitch	Kur_Acc_x	Max_Pitch
10	Max_Roll	Min_Acc_x	Kur_Acc_y	Kur_Roll	Rms_Roll	Entropy_Roll
11	Mean_Acc_x	Mean_Acc_z	Kur_Acc_z	Skew_Acc_x	Entropy_Acc_x	Rms_Azimuth
12	Rms_Acc_x	Rms_Acc_x	Kur_Azimuth	Skew_Azimuth	Rms_Azimuth	Mean_Azimuth
13	Energy_Roll	Energy_Roll	Skew_Acc_y	Min_Acc_x	Max_Azimuth	Dfreq_Acc_x
14	Rms_Acc_z	Min_Roll	Skew_Azimuth	Min_Acc_z	Dfreq_Acc_x	Mean_Acc_y
15	Energy_Acc_z	Rms_Acc_z	Min_Acc_x	Min_Pitch	Dfreq_Roll	Vedb_Angle
16	Min_Roll	Energy_Acc_z	Min_Pitch	Min_Roll	Mean_Pitch	Kur_Azimuth
17	Cf_Roll	Cf_Acc_x	Max_Acc_x	Max_Acc_x	Min_Acc_x	Dfreq_Roll
18	Cf_Acc_x	Cf_Roll	Max_Acc_z	Max_Acc_z	Skew_Roll	Max_diff_Acc_y
19	Energy_Acc_x	Energy_Acc_x	Max_Pitch	Max_Pitch	Var_Pitch	DSVM_Acc
20	Max_Acc_x	Max_Acc_x	Max_Roll	Max_Roll	Cf_Acc_x	Var_Pitch
21	Max_Pitch	Entropy_Azimuth	Rms_Acc_y	Rms_Azimuth	Rms_Pitch	Var_Roll
22	Entropy_Azimuth	Rms_Pitch	Rms_Acc_z	Rms_Pitch	SMA_Angle	Min_Pitch
23	Rms_Pitch	SVM_Angle	Rms_Roll	Interq_Acc_x	Max_Acc_x	Max_Acc_x
24	Mean_Pitch	Entropy_Acc_z	Interq_Acc_x	Interq_Pitch	Rms_Acc_x	Min_Acc_x
25	Entropy_Acc_z	SMA_Angle	Interq_Acc_y	Cf_Azimuth	DSVM_Acc	Min_Acc_y

26	Energy_Pitch	Max_Pitch	Interq_Acc_z	Cf_Pitch	Entropy_TimeD_Ang	Mean_Pitch
27	Min_Azimuth	Max_diff_Azimuth	Interq_Pitch	SMA_Acc	DSAM_Angle	Entropy_TimeD_Ang
28	SVM_Angle	Min_Pitch	Cf_Acc_y	SMA_Angle	Var_Acc_x	Var_Acc_x
29	Min_Pitch	Energy_Pitch	Cf_Acc_z	Entropy_Acc_x	Max_diff_Roll	DSAM_Angle
30	SMA_Angle	Entropy_TimeD_Ang	Cf_Azimuth	Entropy_Acc_z	Var_Azimuth	Interq_Roll
31	Max_diff_Azimuth	Max_diff_Acc_x	Cf_Pitch	Entropy_Azimuth	Cf_Azimuth	Skew_Roll
32	DSAM_Angle	Vedb_Acc	Cf_Roll	Entropy_Pitch	Max_diff_Acc_x	Rms_Pitch
33	Max_diff_Acc_x	Min_Azimuth	SMA_Acc	Entropy_Roll	Rms_Acc_z	SMA_Angle
34	Vedb_Acc	Min_Acc_z	Entropy_Acc_x	Entropy_TimeD_Acc	Min_Pitch	Skew_Acc_y
35	Entropy_TimeD_Ang	Mean_Pitch	Entropy_Acc_y	Entropy_TimeD_Ang	Max_Acc_y	Interq_Acc_x
36	Max_diff_Roll	Max_diff_Roll	Entropy_Acc_z	Energy_Acc_y	Skew_Acc_y	Var_Acc_z
37	Max_Acc_z	Mean_Acc_y	Entropy_TimeD_Ang	Energy_Acc_z	Dfreq_Acc_z	Rms_Acc_x
38	Min_Acc_z	Std_Roll	Energy_Acc_x	Energy_Azimuth	Mean_Acc_y	Cf_Roll
39	Dfreq_Pitch	Std_Acc_x	Energy_Acc_y	SVM_Angle	Var_Acc_y	Var_Azimuth
40	Skew_Azimuth	Entropy_Pitch	Energy_Pitch	DSVM_Acc	Cf_Acc_z	mag_Ang
41	Mean_Acc_y	DSAM_Angle	DSAM_Angle	Max_diff_Acc_x	Entropy_Pitch	Var_Acc_y
42	Std_Roll	Max_Acc_z	Max_diff_Acc_x	Max_diff_Acc_y	Interq_Azimuth	Max_diff_Azimuth
43	Entropy_Pitch	Dfreq_Pitch	Max_diff_Acc_y	MV_Acc	Kur_Azimuth	Kur_Acc_x
44	Std_Acc_x	Max_Acc_y	Max_diff_Azimuth	AV_Ang	Max_diff_Azimuth	Max_Acc_y
45	Cf_Pitch	Cf_Pitch	Max_diff_Roll	mag_Acc	Skew_Azimuth	Cf_Azimuth
46	Std_Pitch	Std_Pitch	MV_Acc	Vedb_Acc	SVM_Angle	Max_diff_Acc_x
47	Max_diff_Pitch	Skew_Azimuth	AV_Ang	Vedb_Angle	Cf_Roll	Max_diff_Roll
48	Max_Acc_y	Max_diff_Pitch	mag_Ang	Dfreq_Acc_x	Highest_peak_Azimuth	Skew_Azimuth
49	Vedb_Angle	Skew_Acc_y	Vedb_Acc	Dfreq_Azimuth	Vedb_Angle	Min_Azimuth
50	Skew_Acc_y	Dfreq_Acc_y	Dfreq_Acc_z	Dfreq_Roll	Max_Roll	Max_diff_Acc_z
51	Dfreq_Acc_y	Vedb_Angle	Dfreq_Azimuth	Highest_peak_Acc_x	MV_Acc	Highest_peak_Acc_y
52	Cf_Acc_z	Min_Acc_y	Dfreq_Pitch	nPeaks_Acc_y	Kur_Acc_y	Cf_Acc_z
53	DSVM_Acc	Kur_Azimuth	nPeaks_Acc_x	Widest_Peak_Acc_y	Min_Azimuth	Rms_Acc_z
54	Kur_Azimuth	Skew_Roll	Highest_peak_Acc_x	nPeaks_Acc_z	Interq_Roll	Interq_Acc_z

55	Min_Acc_y	DSVM_Acc	Avr_peak_time_Acc_x	Widest_Peak_Acc_z	Var_Roll	Rms_Acc_y
56	Cf_Azimuth	Entropy_Acc_y	nPeaks_Acc_y	Highest_peak_Acc_z	Kur_Pitch	Interq_Pitch
57	Rms_Acc_y	Std_Azimuth	Highest_peak_Acc_z	Avr_peak_time_Azimuth	Max_diff_Acc_z	Dfreq_Pitch
58	AV_Ang	Rms_Azimuth	Avr_peak_time_Acc_z	Widest_Peak_Pitch	Interq_Acc_x	Entropy_Acc_z
59	Std_Azimuth	mag_Ang	nPeaks_Azimuth	Avr_peak_time_Pitch	Entropy_TimeD_Acc	Cf_Pitch
60	Std_Acc_z	Cf_Azimuth	Widest_Peak_Azimuth	Widest_Peak_Roll	Vedb_Acc	Skew_Pitch
61	Rms_Azimuth	AV_Ang	Highest_peak_Azimuth	Highest_peak_Roll	Kur_Roll	Kur_Roll
62	Entropy_Acc_y	Max_Azimuth	Avr_peak_time_Azimuth		SMA_Acc	MV_Acc
63	Dfreq_Azimuth	Dfreq_Azimuth	Highest_peak_Pitch		Highest_peak_Roll	Kur_Pitch
64	Energy_Acc_y	Rms_Acc_y	Widest_Peak_Roll		Var_Acc_z	Interq_Azimuth
65	Skew_Roll	Kur_Acc_x	Highest_peak_Roll		Cf_Pitch	Max_diff_Pitch
66	Var_Pitch	Var_Pitch	Avr_peak_time_Roll		Entropy_Acc_z	Highest_peak_Acc_x
67	Highest_peak_Roll	SVM_Acc			mag_Ang	Entropy_Pitch
68	SMA_Acc	SMA_Acc			Rms_Acc_y	Interq_Acc_y
69	SVM_Acc	Std_Acc_z			Skew_Pitch	SVM_Angle
70	Mean_Azimuth	Max_diff_Acc_y			Widest_Peak_Azimuth	Cf_Acc_y
71	mag_Ang	Entropy_TimeD_Acc			Kur_Acc_z	Cf_Acc_x
72	Highest_peak_Acc_y	Var_Acc_x			Avr_peak_time_Azimuth	Highest_peak_Azimuth
73	Var_Roll	Var_Roll			Interq_Pitch	Max_Roll
74	Max_Azimuth	Max_diff_Acc_z			AV_Ang	Entropy_Azimuth
75	Var_Acc_x	Mean_Azimuth			Min_Acc_y	Kur_Acc_z
76	Highest_peak_Pitch	Cf_Acc_z			Max_diff_Pitch	Entropy_TimeD_Acc
77	Entropy_TimeD_Acc	Energy_Acc_y			Highest_peak_Acc_x	Vedb_Acc
78	Kur_Acc_x	Interq_Pitch			Widest_Peak_Acc_z	Max_Acc_z
79	Highest_peak_Acc_x	Highest_peak_Pitch			Highest_peak_Acc_z	Widest_Peak_Azimuth
80	Interq_Acc_y	Highest_peak_Acc_x			Widest_Peak_Pitch	Highest_peak_Acc_z
81	Max_diff_Acc_z	Highest_peak_Acc_y			Interq_Acc_y	SMA_Acc
82	Kur_Acc_z	Cf_Acc_y			SVM_Acc	Kur_Acc_y
83	Interq_Pitch	Interq_Roll			Interq_Acc_z	nPeaks_Acc_x

84	Cf_Acc_y	Skew_Acc_z			Entropy_Acc_y	SVM_Acc
85	Max_diff_Acc_y	Highest_peak_Roll			Dfreq_Pitch	Widest_Peak_Pitch
86	Skew_Pitch	Skew_Pitch			Max_Acc_z	Dfreq_Acc_z
87	Skew_Acc_z	Kur_Roll			nPeaks_Pitch	Skew_Acc_x
88	Std_Acc_y	Interq_Acc_y			Skew_Acc_x	Highest_peak_Pitch
89	Skew_Acc_x	Skew_Acc_x			Avr_peak_time_Acc_y	nPeaks_Acc_y
90	Highest_peak_Acc_z	Highest_peak_Acc_z			Avr_peak_time_Roll	nPeaks_Roll
91	MV_Acc	Std_Acc_y			Highest_peak_Acc_y	Avr_peak_time_Acc_x
92	Var_Acc_z	MV_Acc			Highest_peak_Pitch	Highest_peak_Roll
93	Interq_Roll	Interq_Azimuth			nPeaks_Roll	AV_Ang
94	Kur_Roll	Var_Acc_z			Dfreq_Acc_y	Widest_Peak_Acc_y
95	Var_Acc_y	Kur_Acc_z			Skew_Acc_z	Dfreq_Acc_y
96	Highest_peak_Azimuth	Highest_peak_Azimuth			Entropy_Azimuth	Entropy_Acc_y
97	Interq_Acc_z	Var_Acc_y			nPeaks_Azimuth	Widest_Peak_Acc_z
98	Interq_Azimuth	Interq_Acc_x			Avr_peak_time_Acc_x	Avr_peak_time_Roll
99	Var_Azimuth	Var_Azimuth			Widest_Peak_Roll	Widest_Peak_Roll
100	Interq_Acc_x	Interq_Acc_z			Max_diff_Acc_y	Avr_peak_time_Pitch
101	Kur_Pitch	Kur_Pitch			Avr_peak_time_Acc_z	Std_Acc_x
102	Energy_Azimuth	Energy_Azimuth			Cf_Acc_y	Energy_Azimuth
103	nPeaks_Acc_z	Widest_Peak_Acc_x			Widest_Peak_Acc_y	Avr_peak_time_Acc_z
104	Kur_Acc_y	Kur_Acc_y			nPeaks_Acc_z	nPeaks_Acc_z
105	Avr_peak_time_Acc_z	nPeaks_Acc_x			Std_Roll	Energy_Acc_y
106	nPeaks_Acc_y	nPeaks_Pitch			nPeaks_Acc_y	Skew_Acc_z
107	Widest_Peak_Pitch	Widest_Peak_Pitch			Energy_Pitch	mag_Acc
108	Widest_Peak_Acc_y	nPeaks_Acc_y			Widest_Peak_Acc_x	Std_Roll
109	mag_Acc	Avr_peak_time_Acc_y			nPeaks_Acc_x	Avr_peak_time_Azimuth
110	Avr_peak_time_Acc_x	Avr_peak_time_Acc_x			mag_Acc	Std_Acc_z
111	nPeaks_Pitch	Avr_peak_time_Pitch			Energy_Acc_z	Std_Pitch
112	Avr_peak_time_Acc_y	Widest_Peak_Acc_y			Dfreq_Azimuth	Dfreq_Azimuth

113	nPeaks_Azimuth	nPeaks_Acc_z			Energy_Acc_x	Energy_Acc_z
114	Widest_Peak_Acc_x	nPeaks_Roll			Avr_peak_time_Pitch	Widest_Peak_Acc_x
115	Avr_peak_time_Pitch	Avr_peak_time_Roll			Energy_Acc_y	nPeaks_Pitch
116	nPeaks_Acc_x	Avr_peak_time_Acc_z			Std_Azimuth	nPeaks_Azimuth
117	Avr_peak_time_Azimuth	Avr_peak_time_Azimuth			Energy_Roll	Avr_peak_time_Acc_y
118	Avr_peak_time_Roll	mag_Acc			Std_Acc_x	Energy_Pitch
119	Widest_Peak_Roll	Widest_Peak_Roll			Std_Acc_y	Std_Azimuth
120	Widest_Peak_Acc_z	Widest_Peak_Azimuth			Std_Acc_z	Energy_Acc_x
121	Widest_Peak_Azimuth	nPeaks_Azimuth			Std_Pitch	Energy_Roll
122	nPeaks_Roll	Widest_Peak_Acc_z			Energy_Azimuth	Std_Acc_y

Appendix E. 6 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet2_ac** over 5 *sec.window*.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Entropy_Roll	Entropy_Roll	Mean_Acc_x	Mean_Acc_y	Min_Roll	Min_Roll
2	Entropy_Acc_x	Entropy_Acc_x	Mean_Acc_y	Mean_Azimuth	Mean_Roll	Mean_Roll
3	Rms_Roll	Dfreq_Roll	Mean_Acc_z	Mean_Pitch	Mean_Acc_x	Mean_Acc_x
4	Dfreq_Roll	Dfreq_Acc_x	Mean_Azimuth	Mean_Roll	Mean_Azimuth	Entropy_Acc_x
5	Dfreq_Acc_x	Rms_Roll	Mean_Pitch	Var_Acc_x	Max_Pitch	Mean_Acc_z
6	Dfreq_Acc_z	Mean_Roll	Mean_Roll	Var_Acc_y	Rms_Azimuth	Mean_Azimuth
7	Mean_Roll	Mean_Acc_x	Var_Acc_x	Var_Acc_z	Mean_Acc_z	Dfreq_Roll
8	Max_Roll	Min_Acc_x	Var_Azimuth	Var_Azimuth	Max_Acc_x	Rms_Azimuth
9	Mean_Acc_x	Max_Roll	Var_Pitch	Var_Roll	Rms_Roll	Var_Acc_x
10	Mean_Acc_z	Dfreq_Acc_z	Std_Acc_x	Std_Roll	Entropy_Roll	Entropy_Roll
11	Min_Acc_x	Mean_Acc_z	Std_Acc_y	Kur_Azimuth	Dfreq_Acc_x	Rms_Roll
12	Energy_Roll	Energy_Roll	Std_Acc_z	Kur_Pitch	Entropy_Acc_x	Max_Azimuth
13	Min_Roll	Min_Roll	Std_Pitch	Skew_Acc_x	Var_Acc_x	Min_Acc_z
14	Rms_Acc_x	Rms_Acc_x	Kur_Acc_x	Skew_Acc_y	Var_Roll	Max_Acc_x
15	Cf_Acc_x	Cf_Acc_x	Kur_Acc_z	Skew_Acc_z	Dfreq_Acc_z	Min_Pitch
16	Cf_Roll	Energy_Acc_z	Kur_Roll	Skew_Pitch	Mean_Pitch	Cf_Azimuth
17	Rms_Acc_z	Rms_Acc_z	Skew_Acc_x	Min_Azimuth	Min_Acc_z	Skew_Acc_y
18	Energy_Acc_z	Max_Acc_x	Skew_Acc_y	Max_Acc_x	Max_diff_Azimuth	Rms_Acc_z
19	Max_Acc_x	Cf_Roll	Skew_Acc_z	Max_Acc_y	Cf_Acc_x	Var_Pitch
20	Energy_Acc_x	Entropy_Azimuth	Skew_Azimuth	Max_Pitch	Cf_Azimuth	Min_Azimuth
21	Entropy_Azimuth	Energy_Acc_x	Min_Acc_x	Max_Roll	Min_Azimuth	Max_Pitch
22	Std_Acc_x	Std_Roll	Min_Acc_z	Rms_Acc_x	Dfreq_Roll	Var_Roll
23	Entropy_Pitch	Entropy_Acc_z	Min_Roll	Rms_Acc_y	Rms_Acc_x	Mean_Acc_y
24	Min_Pitch	Min_Pitch	Max_Acc_x	Rms_Azimuth	Skew_Acc_y	Dfreq_Acc_x
25	Max_Pitch	Rms_Pitch	Max_Acc_y	Rms_Roll	Vedb_Angle	Min_Acc_x

26	Std_Roll	Max_Pitch	Max_Acc_z	Interq_Pitch	Var_Pitch	Rms_Acc_x
27	Entropy_Acc_z	Entropy_Pitch	Rms_Acc_y	Interq_Roll	SMA_Angle	Max_diff_Acc_x
28	Mean_Pitch	Std_Acc_x	Rms_Acc_z	Cf_Acc_y	Skew_Roll	Kur_Roll
29	SVM_Angle	SVM_Angle	Rms_Azimuth	Cf_Pitch	Max_Azimuth	Highest_peak_Acc_x
30	SMA_Angle	SMA_Angle	Rms_Roll	Cf_Roll	Min_Acc_x	Mean_Pitch
31	Max_diff_Acc_x	DSAM_Angle	Interq_Acc_x	SMA_Acc	Rms_Acc_y	Kur_Acc_x
32	Max_diff_Azimuth	Max_diff_Acc_x	Interq_Acc_y	SMA_Angle	mag_Ang	SVM_Angle
33	Rms_Pitch	Energy_Pitch	Interq_Acc_z	Entropy_Acc_y	Max_Roll	Cf_Acc_x
34	DSAM_Angle	Min_Acc_z	Interq_Azimuth	Entropy_Acc_z	Var_Azimuth	SMA_Angle
35	Entropy_TimeD_Ang	Cf_Pitch	Interq_Roll	Entropy_Azimuth	Mean_Acc_y	Skew_Roll
36	Max_diff_Roll	Max_diff_Azimuth	Cf_Acc_y	Entropy_Pitch	Skew_Azimuth	DSAM_Angle
37	Cf_Pitch	Mean_Pitch	Cf_Azimuth	Entropy_TimeD_Acc	Cf_Pitch	Interq_Pitch
38	Min_Azimuth	Max_Acc_z	SMA_Angle	SVM_Angle	Cf_Acc_y	Kur_Pitch
39	Min_Acc_z	Entropy_TimeD_Ang	Entropy_Acc_x	DSVM_Acc	Rms_Acc_z	Max_Roll
40	Energy_Pitch	Std_Pitch	Entropy_Acc_z	DSAM_Angle	Max_diff_Pitch	Max_Acc_y
41	Dfreq_Pitch	Max_diff_Roll	Entropy_Roll	Max_diff_Acc_x	Kur_Acc_y	Entropy_TimeD_Ang
42	Max_diff_Pitch	Min_Azimuth	Entropy_TimeD_Acc	Max_diff_Azimuth	Min_Pitch	Dfreq_Acc_z
43	Vedb_Acc	Dfreq_Pitch	Entropy_TimeD_Ang	Max_diff_Roll	Max_diff_Roll	Var_Acc_y
44	Std_Pitch	Max_Acc_y	Energy_Acc_x	MV_Acc	Interq_Pitch	Max_diff_Azimuth
45	Max_Acc_y	Vedb_Acc	Energy_Acc_y	mag_Acc	Cf_Roll	Entropy_Pitch
46	Max_Acc_z	Max_diff_Pitch	Energy_Azimuth	Vedb_Angle	Highest_peak_Acc_x	Highest_peak_Roll
47	Entropy_Acc_y	Min_Acc_y	DSAM_Angle	Dfreq_Acc_z	Entropy_TimeD_Ang	Var_Azimuth
48	Dfreq_Acc_y	Dfreq_Acc_y	Max_diff_Acc_x	Dfreq_Pitch	Interq_Roll	Cf_Acc_z
49	Min_Acc_y	Var_Roll	Max_diff_Pitch	Dfreq_Roll	Vedb_Acc	Skew_Acc_z
50	Var_Acc_x	Skew_Azimuth	MV_Acc	nPeaks_Acc_x	Interq_Acc_x	Rms_Pitch
51	Vedb_Angle	Entropy_Acc_y	mag_Ang	Widest_Peak_Acc_x	Highest_peak_Roll	Max_Acc_z
52	Var_Roll	Mean_Acc_y	Vedb_Acc	Avr_peak_time_Acc_x	Max_diff_Acc_y	Vedb_Angle
53	Cf_Azimuth	Std_Acc_z	Dfreq_Acc_y	nPeaks_Acc_y	Var_Acc_z	Entropy_Acc_z
54	Mean_Acc_y	Kur_Acc_x	Dfreq_Acc_z	Widest_Peak_Acc_y	Min_Acc_y	Highest_peak_Azimuth

55	Skew_Azimuth	Vedb_Angle	Widest_Peak_Acc_x	Avr_peak_time_Acc_y	Interq_Azimuth	Cf_Pitch
56	Std_Acc_z	Skew_Acc_y	nPeaks_Acc_y	nPeaks_Acc_z	Kur_Pitch	Max_diff_Pitch
57	Skew_Acc_y	AV_Ang	Widest_Peak_Acc_y	Highest_peak_Acc_z	Entropy_Pitch	Avr_peak_time_Acc_y
58	Kur_Acc_x	Interq_Roll	nPeaks_Acc_z	Avr_peak_time_Acc_z	DSAM_Angle	Interq_Acc_x
59	Std_Azimuth	Max_diff_Acc_z	Widest_Peak_Acc_z	nPeaks_Azimuth	Entropy_Azimuth	Min_Acc_y
60	Rms_Azimuth	Var_Acc_x	Highest_peak_Acc_z	nPeaks_Pitch	Highest_peak_Acc_z	Rms_Acc_y
61	Interq_Roll	Cf_Acc_z	Avr_peak_time_Acc_z	Highest_peak_Pitch	Skew_Acc_z	DSVM_Acc
62	Dfreq_Azimuth	Kur_Azimuth	nPeaks_Azimuth	Widest_Peak_Roll	SVM_Angle	Cf_Roll
63	Max_Azimuth	Var_Pitch	Highest_peak_Azimuth	Highest_peak_Roll	Kur_Acc_x	Skew_Azimuth
64	AV_Ang	Std_Azimuth	nPeaks_Pitch	Avr_peak_time_Roll	Entropy_Acc_z	Entropy_Azimuth
65	Highest_peak_Pitch	Cf_Azimuth	Widest_Peak_Roll		Max_diff_Acc_x	Cf_Acc_y
66	Cf_Acc_y	Skew_Roll	Highest_peak_Roll		Avr_peak_time_Acc_z	Dfreq_Pitch
67	Mean_Azimuth	Kur_Roll			DSVM_Acc	Max_diff_Roll
68	Rms_Acc_y	Rms_Acc_y			Rms_Pitch	mag_Ang
69	Cf_Acc_z	Interq_Pitch			Max_Acc_y	Interq_Roll
70	Max_diff_Acc_z	Rms_Azimuth			MV_Acc	Interq_Azimuth
71	Highest_peak_Acc_x	Highest_peak_Pitch			Cf_Acc_z	SVM_Acc
72	DSVM_Acc	Dfreq_Azimuth			Var_Acc_y	AV_Ang
73	Highest_peak_Roll	Highest_peak_Roll			mag_Acc	Skew_Pitch
74	Energy_Acc_y	Max_diff_Acc_y			Max_Acc_z	Entropy_Acc_y
75	Var_Pitch	DSVM_Acc			Skew_Pitch	Interq_Acc_z
76	Skew_Roll	mag_Ang			Kur_Azimuth	nPeaks_Pitch
77	Skew_Pitch	Max_Azimuth			Highest_peak_Acc_y	Max_diff_Acc_z
78	mag_Ang	Highest_peak_Acc_x			Kur_Roll	Vedb_Acc
79	Kur_Azimuth	Energy_Acc_y			Highest_peak_Azimuth	MV_Acc
80	Highest_peak_Azimuth	Highest_peak_Azimuth			AV_Ang	Var_Acc_z
81	Interq_Pitch	Mean_Azimuth			Entropy_TimeD_Acc	Highest_peak_Pitch
82	Skew_Acc_z	Cf_Acc_y			Entropy_Acc_y	Kur_Acc_z
83	Kur_Roll	Highest_peak_Acc_y			Interq_Acc_z	Widest_Peak_Acc_z

84	Max_diff_Acc_y	Skew_Acc_z			Interq_Acc_y	Max_diff_Acc_y
85	Var_Acc_z	Std_Acc_y			Dfreq_Azimuth	nPeaks_Roll
86	Interq_Acc_x	Skew_Pitch			Dfreq_Pitch	Highest_peak_Acc_y
87	SVM_Acc	Skew_Acc_x			Highest_peak_Pitch	Kur_Azimuth
88	Std_Acc_y	Var_Acc_z			Skew_Acc_x	Dfreq_Acc_y
89	SMA_Acc	SVM_Acc			Avr_peak_time_Roll	Skew_Acc_x
90	MV_Acc	Interq_Acc_x			Widest_Peak_Acc_y	Avr_peak_time_Roll
91	Entropy_TimeD_Acc	Highest_peak_Acc_z			SVM_Acc	SMA_Acc
92	Highest_peak_Acc_y	SMA_Acc			Widest_Peak_Azimuth	Widest_Peak_Azimuth
93	Interq_Azimuth	Entropy_TimeD_Acc			Dfreq_Acc_y	Avr_peak_time_Acc_x
94	Skew_Acc_x	MV_Acc			nPeaks_Acc_y	nPeaks_Acc_z
95	Kur_Pitch	Var_Acc_y			SMA_Acc	Highest_peak_Acc_z
96	Highest_peak_Acc_z	Interq_Acc_z			nPeaks_Acc_x	nPeaks_Azimuth
97	Var_Acc_y	Kur_Acc_z			Kur_Acc_z	Widest_Peak_Acc_x
98	Interq_Acc_z	Interq_Acc_y			nPeaks_Pitch	Interq_Acc_y
99	Widest_Peak_Acc_x	Interq_Azimuth			Std_Acc_x	Entropy_TimeD_Acc
100	Interq_Acc_y	Widest_Peak_Acc_x			Std_Acc_y	Std_Acc_x
101	Kur_Acc_z	Var_Azimuth			Std_Acc_z	Std_Acc_y
102	Energy_Azimuth	Kur_Pitch			Std_Azimuth	Std_Acc_z
103	Var_Azimuth	nPeaks_Roll			Std_Pitch	Std_Azimuth
104	Widest_Peak_Acc_y	Energy_Azimuth			Std_Roll	Std_Pitch
105	Kur_Acc_y	nPeaks_Azimuth			Energy_Acc_x	Std_Roll
106	nPeaks_Acc_y	Avr_peak_time_Roll			Energy_Acc_y	Energy_Acc_x
107	Widest_Peak_Roll	Kur_Acc_y			Energy_Acc_z	Energy_Acc_y
108	Avr_peak_time_Acc_y	Widest_Peak_Acc_y			Energy_Azimuth	Energy_Acc_z
109	Widest_Peak_Acc_z	Widest_Peak_Roll			Energy_Pitch	Energy_Azimuth
110	Widest_Peak_Pitch	Avr_peak_time_Pitch			Energy_Roll	Energy_Pitch
111	Avr_peak_time_Acc_z	nPeaks_Acc_z			nPeaks_Azimuth	Energy_Roll
112	nPeaks_Azimuth	Avr_peak_time_Azimuth			Widest_Peak_Acc_x	Dfreq_Azimuth

113	Avr_peak_time_Azimuth	nPeaks_Pitch			Avr_peak_time_Pitch	Widest_Peak_Acc_y
114	mag_Acc	Avr_peak_time_Acc_z			Avr_peak_time_Acc_x	Avr_peak_time_Pitch
115	Widest_Peak_Azimuth	Widest_Peak_Acc_z			nPeaks_Roll	Kur_Acc_y
116	Avr_peak_time_Pitch	Avr_peak_time_Acc_y			Widest_Peak_Roll	Avr_peak_time_Azimuth
117	nPeaks_Acc_x	nPeaks_Acc_x			Widest_Peak_Acc_z	nPeaks_Acc_y
118	Avr_peak_time_Roll	nPeaks_Acc_y			Max_diff_Acc_z	mag_Acc
119	Avr_peak_time_Acc_x	Widest_Peak_Azimuth			Widest_Peak_Pitch	Widest_Peak_Roll
120	nPeaks_Pitch	Widest_Peak_Pitch			Avr_peak_time_Azimuth	Widest_Peak_Pitch
121	nPeaks_Roll	mag_Acc			nPeaks_Acc_z	nPeaks_Acc_x
122	nPeaks_Acc_z	Avr_peak_time_Acc_x			Avr_peak_time_Acc_y	Avr_peak_time_Acc_z

Appendix E. 7 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet2_b** over 10 $sec. window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Acc_x	Mean_Roll	Mean_Acc_y	Mean_Acc_x	Mean_Acc_x	Min_Roll
2	Mean_Roll	Mean_Acc_x	Mean_Azimuth	Mean_Gyr_z	Mean_Roll	Mean_Roll
3	Min_Roll	Entropy_Roll	Mean_Pitch	Var_Acc_y	Min_Roll	Mean_Acc_x
4	Dfreq_Acc_z	Dfreq_Roll	Mean_Roll	Var_Acc_z	Mean_Acc_z	Mean_Acc_z
5	Entropy_Roll	Entropy_Acc_x	Mean_Gyr_x	Var_Azimuth	Mean_Pitch	Rms_Roll
6	Cf_Roll	Dfreq_Acc_x	Var_Acc_y	Var_Gyr_x	Vedb_Acc	Rms_Pitch
7	Mean_Acc_z	Cf_Roll	Var_Acc_z	Std_Acc_z	Rms_Pitch	Vedb_Acc
8	Dfreq_Pitch	Dfreq_Acc_z	Std_Acc_x	Std_Gyr_x	Rms_Roll	Rms_Acc_y
9	Vedb_Acc	Mean_Acc_z	Std_Acc_y	Std_Gyr_y	Min_Pitch	Mean_Pitch
10	Dfreq_Roll	Min_Roll	Std_Acc_z	Kur_Acc_x	Kur_Acc_y	Dfreq_Roll
11	Rms_Pitch	Energy_Pitch	Std_Azimuth	Kur_Azimuth	Dfreq_Pitch	Cf_Gyr_x
12	Dfreq_Acc_x	Vedb_Acc	Std_Roll	Kur_Pitch	Mean_Acc_y	Mean_Acc_y
13	Energy_Acc_y	Rms_Pitch	Std_Gyr_x	Kur_Gyr_x	Max_Roll	Kur_Gyr_x
14	Max_Roll	Rms_Roll	Std_Gyr_y	Skew_Acc_x	Entropy_Acc_x	Min_Gyr_y
15	Energy_Pitch	Max_Roll	Kur_Acc_x	Skew_Gyr_y	Cf_Pitch	Entropy_Roll
16	Dfreq_Acc_y	Dfreq_Pitch	Kur_Acc_y	Skew_Gyr_z	Dfreq_Roll	Max_diff_Acc_y
17	Rms_Acc_y	Energy_Acc_y	Kur_Acc_z	Min_Acc_z	Rms_Acc_x	Kur_Acc_y
18	Entropy_Acc_x	Rms_Acc_y	Kur_Pitch	Min_Azimuth	Max_Gyr_y	Interq_Roll
19	Rms_Roll	Dfreq_Acc_y	Kur_Roll	Min_Pitch	Min_Gyr_x	Max_Roll
20	Cf_Acc_x	Energy_Roll	Kur_Gyr_x	Min_Gyr_x	Dfreq_Acc_x	Cf_Acc_y
21	Min_Acc_x	Min_Acc_x	Skew_Roll	Min_Gyr_y	Rms_Acc_y	Max_Acc_y
22	Mean_Pitch	Mean_Pitch	Skew_Gyr_y	Min_Gyr_z	Skew_Gyr_x	Var_Acc_y
23	Interq_Gyr_x	Entropy_Acc_z	Skew_Gyr_z	Max_Acc_y	Kur_Acc_z	Var_Acc_z
24	Entropy_Acc_z	Cf_Acc_x	Min_Acc_y	Max_Acc_z	Max_Acc_x	Max_Gyr_z
25	Energy_Acc_z	Min_Acc_y	Min_Gyr_x	Max_Azimuth	Min_Acc_z	Dfreq_Acc_x

26	Energy_Roll	Rms_Acc_z	Min_Gyr_z	Max_Gyr_z	Entropy_Acc_z	Min_Acc_z
27	Rms_Acc_z	Max_Acc_x	Max_Acc_y	Rms_Acc_y	Interq_Gyr_z	Skew_Gyr_x
28	Entropy_Acc_y	Energy_Acc_z	Max_Acc_z	Rms_Azimuth	SVM_Acc	Var_Acc_x
29	Rms_Gyr_x	Cf_Acc_y	Max_Pitch	Rms_Pitch	MV_Acc	Min_Pitch
30	Std_Gyr_x	Rms_Gyr_x	Max_Roll	Rms_Roll	Max_Acc_z	Var_Gyr_y
31	Mean_Acc_y	Std_Gyr_x	Max_Gyr_x	Rms_Gyr_x	Skew_Acc_z	Skew_Acc_y
32	Min_Acc_y	Mean_Acc_y	Max_Gyr_z	Rms_Gyr_y	Max_Pitch	Max_Acc_x
33	Var_Gyr_x	Rms_Acc_x	Rms_Acc_y	Interq_Acc_z	Max_diff_Pitch	Max_diff_Gyr_y
34	Energy_Gyr_x	Max_Acc_y	Rms_Acc_z	Interq_Gyr_x	Min_Acc_y	SVM_Acc
35	Max_Acc_x	Energy_Gyr_x	Rms_Roll	Cf_Acc_y	nPeaks_Pitch	Var_Roll
36	Interq_Gyr_z	Var_Gyr_x	Rms_Gyr_x	Cf_Pitch	Var_Acc_x	Max_diff_Gyr_z
37	Min_Pitch	Interq_Gyr_x	Rms_Gyr_z	Cf_Gyr_y	Interq_Gyr_x	Max_diff_Roll
38	Entropy_Pitch	Interq_Gyr_z	Interq_Acc_x	Cf_Gyr_z	Kur_Acc_x	Interq_Gyr_y
39	Interq_Acc_y	Min_Acc_z	Interq_Acc_z	Entropy_Acc_x	Entropy_TimeD_Acc	Mean_Gyr_z
40	Rms_Acc_x	Entropy_Acc_y	Interq_Roll	Entropy_Pitch	Mean_Gyr_z	Min_Gyr_x
41	Dfreq_Gyr_x	Min_Pitch	Interq_Gyr_z	Entropy_Roll	Rms_Acc_z	SMA_Acc
42	Highest_peak_Gyr_x	Interq_Acc_y	Cf_Acc_y	Entropy_TimeD_Acc	Entropy_Roll	Widest_Peak_Acc_z
43	Max_Acc_y	Energy_Acc_x	Cf_Azimuth	Entropy_TimeD_Gyr	Widest_Peak_Acc_z	Entropy_Gyr_y
44	Interq_Gyr_y	Dfreq_Gyr_x	Cf_Pitch	Energy_Acc_y	Cf_Acc_y	Rms_Acc_z
45	Min_Acc_z	Highest_peak_Gyr_x	Cf_Gyr_x	Energy_Azimuth	Interq_Acc_y	Entropy_Azimuth
46	Cf_Acc_y	Cf_Gyr_x	Cf_Gyr_z	Energy_Pitch	Max_Gyr_z	Interq_Gyr_z
47	Max_diff_Acc_y	Skew_Acc_y	SMA_Acc	Energy_Roll	Cf_Acc_z	Max_Gyr_x
48	MV_Acc	Interq_Gyr_y	SMA_Angle	Energy_Gyr_z	nPeaks_Roll	Max_diff_Acc_x
49	Std_Acc_z	DSVM_Acc	Entropy_Acc_x	SVM_Acc	Rms_Gyr_x	Kur_Pitch
50	Std_Roll	Interq_Acc_z	Entropy_Acc_z	DSVM_Acc	Var_Roll	Max_diff_Azimuth
51	Interq_Roll	Min_Gyr_y	Entropy_Azimuth	DSAM_Angle	SMA_Acc	Max_Pitch
52	Std_Pitch	MV_Gyr	Entropy_Gyr_x	DSVM_Gyr	Var_Gyr_z	DSVM_Gyr
53	DSVM_Acc	Entropy_Gyr_x	Entropy_Gyr_z	Max_diff_Acc_x	Interq_Roll	MV_Gyr
54	Highest_peak_Roll	Max_diff_Gyr_x	Entropy_TimeD_Ang	Max_diff_Acc_y	Max_Acc_y	Min_Acc_x

55	Energy_Acc_x	Max_diff_Acc_y	Entropy_TimeD_Gyr	Max_diff_Azimuth	Max_diff_Gyr_x	Cf_Roll
56	MV_Gyr	MV_Acc	Energy_Acc_y	Max_diff_Roll	Kur_Pitch	Entropy_Acc_y
57	Highest_peak_Acc_z	SVM_Acc	Energy_Acc_z	Max_diff_Gyr_x	nPeaks_Acc_y	Kur_Gyr_z
58	Min_Gyr_y	Entropy_Gyr_y	Energy_Azimuth	Max_diff_Gyr_z	Widest_Peak_Acc_x	Min_Gyr_z
59	Max_Pitch	Max_diff_Acc_z	Energy_Pitch	MV_Acc	Min_Gyr_y	DSVM_Acc
60	Std_Acc_y	Interq_Acc_x	Energy_Gyr_z	AV_Ang	Var_Acc_y	Interq_Acc_y
61	Interq_Acc_z	SMA_Acc	SVM_Angle	MV_Gyr	Cf_Azimuth	Widest_Peak_Acc_y
62	Max_diff_Acc_z	Std_Gyr_z	SVM_Gyr	mag_Acc	mag_Gyr	Rms_Acc_x
63	Var_Acc_y	Highest_peak_Acc_z	DSAM_Angle	mag_Ang	Max_diff_Acc_y	Interq_Gyr_x
64	Cf_Pitch	Cf_Acc_z	DSVM_Gyr	Vedb_Gyr	Max_Azimuth	Entropy_Acc_x
65	Max_diff_Gyr_x	Entropy_TimeD_Acc	Max_diff_Pitch	Dfreq_Acc_z	Kur_Gyr_y	Max_Gyr_y
66	SMA_Acc	Skew_Roll	Max_diff_Roll	Dfreq_Azimuth	Entropy_Gyr_x	Min_Acc_y
67	Mean_Gyr_z	Entropy_Gyr_z	Max_diff_Gyr_z	Dfreq_Pitch	Vedb_Gyr	Cf_Pitch
68	SVM_Acc	Highest_peak_Acc_x	MV_Acc	Dfreq_Roll	Entropy_Pitch	DSAM_Angle
69	Skew_Acc_y	Std_Roll	AV_Ang	Dfreq_Gyr_x	Dfreq_Gyr_x	Skew_Gyr_z
70	Skew_Roll	Highest_peak_Gyr_z	mag_Acc	Dfreq_Gyr_y	Var_Gyr_y	Dfreq_Pitch
71	Var_Acc_z	SVM_Gyr	Dfreq_Acc_y	Widest_Peak_Acc_x	Vedb_Angle	Dfreq_Gyr_z
72	Entropy_Gyr_x	SMA_Gyr	Dfreq_Acc_z	Highest_peak_Acc_y	Max_diff_Roll	Widest_Peak_Gyr_y
73	SVM_Gyr	Entropy_Pitch	Dfreq_Azimuth	Avr_peak_time_Acc_y	Min_Gyr_z	Rms_Gyr_y
74	Min_Gyr_x	nPeaks_Gyr_y	Dfreq_Gyr_x	nPeaks_Acc_z	Avr_peak_time_Acc_z	Kur_Azimuth
75	SMA_Gyr	Rms_Gyr_z	nPeaks_Acc_x	Widest_Peak_Acc_z	Cf_Gyr_z	Skew_Azimuth
76	Kur_Acc_y	Std_Acc_z	Widest_Peak_Acc_x	Highest_peak_Acc_z	Cf_Gyr_x	Kur_Acc_z
77	Highest_peak_Acc_x	Std_Gyr_y	Widest_Peak_Acc_y	Avr_peak_time_Acc_z	Max_diff_Gyr_y	Skew_Acc_x
78	Entropy_TimeD_Acc	Skew_Gyr_x	Avr_peak_time_Acc_y	Highest_peak_Azimuth	SVM_Gyr	Highest_peak_Azimuth
79	Widest_Peak_Acc_x	Rms_Gyr_y	Widest_Peak_Acc_z	Widest_Peak_Pitch	Interq_Acc_z	Avr_peak_time_Gyr_z
80	Rms_Gyr_y	Widest_Peak_Gyr_x	Highest_peak_Acc_z	Highest_peak_Pitch	Highest_peak_Acc_x	Entropy_Acc_z
81	Std_Gyr_y	Max_diff_Azimuth	nPeaks_Azimuth	Avr_peak_time_Pitch	AV_Ang	Skew_Acc_z
82	Entropy_TimeD_Gyr	Kur_Acc_y	nPeaks_Pitch	nPeaks_Roll	Kur_Gyr_x	Max_Azimuth
83	Interq_Pitch	Max_Pitch	Avr_peak_time_Pitch	Widest_Peak_Roll	Highest_peak_Gyr_x	Rms_Gyr_x

84	Entropy_Gyr_z	Max_Gyr_x	Avr_peak_time_Roll	Avr_peak_time_Roll	Max_Gyr_x	Var_Gyr_x
85	Std_Gyr_z	DSVM_Gyr	Widest_Peak_Gyr_y	nPeaks_Gyr_x	Dfreq_Gyr_z	Avr_peak_time_Acc_z
86	Cf_Gyr_x	Highest_peak_Roll	nPeaks_Gyr_z	nPeaks_Gyr_y	Entropy_TimeD_Gyr	Var_Gyr_z
87	Max_diff_Gyr_y	Entropy_TimeD_Gyr	Highest_peak_Gyr_z	Avr_peak_time_Gyr_y	Highest_peak_Pitch	Vedb_Gyr
88	Var_Pitch	Std_Acc_y	Avr_peak_time_Gyr_z	Widest_Peak_Gyr_z	Cf_Roll	Highest_peak_Roll
89	Entropy_Gyr_y	Min_Gyr_x			Skew_Azimuth	Min_Azimuth
90	Var_Roll	Vedb_Gyr			DSVM_Acc	Entropy_Pitch
91	Skew_Acc_x	Std_Acc_x			Cf_Acc_x	Widest_Peak_Gyr_z
92	DSVM_Gyr	Cf_Pitch			Skew_Gyr_z	Max_diff_Pitch
93	Skew_Gyr_x	Interq_Roll			Min_Acc_x	Cf_Acc_x
94	Rms_Gyr_z	Var_Acc_y			nPeaks_Acc_x	Entropy_TimeD_Acc
95	Min_Gyr_z	Max_Acc_z			Max_diff_Gyr_z	Entropy_Gyr_z
96	Kur_Gyr_x	Var_Gyr_z			Var_Acc_z	Max_diff_Acc_z
97	Highest_peak_Gyr_z	Max_diff_Gyr_y			DSVM_Gyr	nPeaks_Acc_x
98	Max_Gyr_y	Mean_Gyr_z			Dfreq_Gyr_y	Widest_Peak_Gyr_x
99	Widest_Peak_Acc_z	Min_Gyr_z			Highest_peak_Acc_y	Max_Acc_z
100	Skew_Pitch	Highest_peak_Gyr_y			SMA_Gyr	Kur_Gyr_y
101	Mean_Gyr_y	Max_Gyr_y			Cf_Gyr_y	Cf_Gyr_y
102	Dfreq_Gyr_y	Var_Roll			Widest_Peak_Azimuth	Interq_Azimuth
103	Interq_Acc_x	Std_Pitch			Avr_peak_time_Gyr_y	Kur_Roll
104	Energy_Gyr_y	Avr_peak_time_Gyr_y			mag_Acc	Var_Azimuth
105	Var_Gyr_y	Var_Acc_z			Interq_Azimuth	nPeaks_Acc_z
106	Highest_peak_Gyr_y	Var_Acc_x			Dfreq_Azimuth	Vedb_Angle
107	Vedb_Gyr	Max_diff_Roll			Highest_peak_Acc_z	Highest_peak_Gyr_z
108	Highest_peak_Azimuth	Energy_Gyr_z			Skew_Acc_x	Skew_Pitch
109	Widest_Peak_Gyr_y	Dfreq_Gyr_y			Entropy_Gyr_y	Rms_Gyr_z
110	Highest_peak_Pitch	Skew_Acc_x			Widest_Peak_Gyr_x	SVM_Gyr
111	Cf_Acc_z	Var_Gyr_y			Rms_Gyr_z	Widest_Peak_Roll
112	Dfreq_Gyr_z	Skew_Acc_z			Interq_Pitch	Kur_Acc_x

113	nPeaks_Acc_z	Widest_Peak_Roll			Min_Azimuth	Avr_peak_time_Azimuth
114	Var_Gyr_z	Energy_Gyr_y			Energy_Acc_y	Avr_peak_time_Acc_y
115	Widest_Peak_Azimuth	nPeaks_Acc_y			nPeaks_Azimuth	Highest_peak_Acc_y
116	nPeaks_Acc_y	Widest_Peak_Gyr_y			Entropy_TimeD_Ang	SMA_Gyr
117	Highest_peak_Acc_y	DSAM_Angle			Highest_peak_Roll	Widest_Peak_Pitch
118	Std_Acc_x	Highest_peak_Acc_y			Energy_Gyr_z	Highest_peak_Acc_x
119	Var_Acc_x	AV_Ang			Widest_Peak_Gyr_z	mag_Gyr
120	nPeaks_Pitch	Std_Azimuth			Var_Gyr_x	Entropy_Gyr_x
121	Widest_Peak_Roll	Max_diff_Gyr_z			Interq_Acc_x	Entropy_TimeD_Gyr
122	Kur_Roll	Kur_Gyr_z			Highest_peak_Gyr_y	Interq_Acc_x
123	Max_diff_Pitch	Max_Gyr_z			Skew_Roll	Highest_peak_Acc_z
124	Kur_Pitch	Var_Pitch			Max_diff_Acc_x	Highest_peak_Pitch
125	Skew_Acc_z	Max_diff_Pitch			DSAM_Angle	Mean_Gyr_y
126	Energy_Gyr_z	Skew_Pitch			Entropy_Gyr_z	Skew_Roll
127	Max_Gyr_x	nPeaks_Acc_z			Var_Pitch	Widest_Peak_Acc_x
128	Avr_peak_time_Acc_z	Kur_Gyr_x			Var_Azimuth	nPeaks_Pitch
129	DSAM_Angle	Max_diff_Acc_x			Widest_Peak_Acc_y	Dfreq_Gyr_y
130	Max_diff_Roll	Interq_Pitch			Entropy_Azimuth	Skew_Gyr_y
131	Max_Acc_z	Skew_Gyr_y			Avr_peak_time_Azimuth	Std_Acc_x
132	AV_Ang	Dfreq_Gyr_z			Std_Gyr_y	Std_Acc_y
133	Kur_Gyr_z	Widest_Peak_Gyr_z			Entropy_Acc_y	Std_Acc_z
134	Max_diff_Gyr_z	Skew_Gyr_z			Max_diff_Azimuth	Std_Azimuth
135	Avr_peak_time_Pitch	Avr_peak_time_Acc_z			SMA_Angle	Std_Pitch
136	nPeaks_Gyr_y	Cf_Gyr_z			Avr_peak_time_Acc_y	Std_Roll
137	nPeaks_Roll	Avr_peak_time_Gyr_x			Kur_Gyr_z	Std_Gyr_x
138	Avr_peak_time_Roll	nPeaks_Gyr_z			Rms_Azimuth	Std_Gyr_y
139	Avr_peak_time_Acc_y	mag_Ang			Mean_Azimuth	Std_Gyr_z
140	mag_Ang	Interq_Azimuth			Avr_peak_time_Roll	Energy_Acc_x
141	Max_Gyr_z	Widest_Peak_Acc_z			Mean_Gyr_y	Energy_Acc_y

142	Cf_Azimuth	Widest_Peak_Azimuth			Highest_peak_Gyr_z	Energy_Acc_z
143	Kur_Acc_x	Kur_Acc_x			Energy_Gyr_x	Energy_Azimuth
144	Max_Azimuth	Cf_Gyr_y			Avr_peak_time_Pitch	Energy_Pitch
145	Kur_Gyr_y	Mean_Gyr_y			Interq_Gyr_y	Energy_Roll
146	Cf_Gyr_z	Kur_Acc_z			Skew_Acc_y	Energy_Gyr_x
147	Skew_Azimuth	Avr_peak_time_Gyr_z			Skew_Pitch	Energy_Gyr_y
148	Kur_Acc_z	Avr_peak_time_Acc_y			Rms_Gyr_y	Energy_Gyr_z
149	Skew_Gyr_z	nPeaks_Gyr_x			Energy_Acc_z	SVM_Angle
150	Energy_Azimuth	Mean_Gyr_x			Mean_Gyr_x	Dfreq_Acc_y
151	Kur_Azimuth	Kur_Roll			nPeaks_Acc_z	Dfreq_Acc_z
152	Widest_Peak_Pitch	Kur_Gyr_y			Kur_Azimuth	Dfreq_Azimuth
153	Avr_peak_time_Gyr_y	Widest_Peak_Acc_y			MV_Gyr	Highest_peak_Gyr_x
154	nPeaks_Azimuth	Var_Azimuth			Std_Acc_z	nPeaks_Acc_y
155	Rms_Azimuth	Highest_peak_Pitch			Std_Roll	mag_Acc
156	Avr_peak_time_Gyr_x	nPeaks_Pitch			Skew_Gyr_y	Avr_peak_time_Gyr_y
157	Cf_Gyr_y	Min_Azimuth			nPeaks_Gyr_z	MV_Acc
158	Vedb_Angle	Avr_peak_time_Roll			nPeaks_Gyr_y	SMA_Angle
159	mag_Gyr	Highest_peak_Azimuth			Std_Azimuth	mag_Ang
160	mag_Acc	Kur_Pitch			Std_Acc_x	Mean_Azimuth
161	Avr_peak_time_Azimuth	Avr_peak_time_Acc_x			SVM_Angle	Widest_Peak_Azimuth
162	Widest_Peak_Acc_y	nPeaks_Acc_x			Avr_peak_time_Gyr_x	Cf_Gyr_z
163	Std_Azimuth	Widest_Peak_Acc_x			Widest_Peak_Roll	Rms_Azimuth
164	Mean_Gyr_x	Avr_peak_time_Pitch			Kur_Roll	nPeaks_Gyr_y
165	Skew_Gyr_y	Widest_Peak_Pitch			Energy_Acc_x	Interq_Acc_z
166	Max_diff_Acc_x	Cf_Azimuth			Avr_peak_time_Acc_x	Var_Pitch
167	nPeaks_Acc_x	Avr_peak_time_Azimuth			Energy_Pitch	Avr_peak_time_Acc_x
168	Avr_peak_time_Acc_x	mag_Acc			Avr_peak_time_Gyr_z	Cf_Azimuth
169	Mean_Azimuth	nPeaks_Azimuth			nPeaks_Gyr_x	Max_diff_Gyr_x
170	Dfreq_Azimuth	Kur_Azimuth			Widest_Peak_Gyr_y	Avr_peak_time_Roll

171	Widest_Peak_Gyr_z	nPeaks_Roll			Widest_Peak_Pitch	Highest_peak_Gyr_y
172	nPeaks_Gyr_x	Entropy_Azimuth			mag_Ang	nPeaks_Gyr_z
173	Entropy_TimeD_Ang	Skew_Azimuth			Highest_peak_Azimuth	nPeaks_Gyr_x
174	Avr_peak_time_Gyr_z	Max_Azimuth			Max_diff_Acc_z	nPeaks_Azimuth
175	SVM_Angle	mag_Gyr			Std_Gyr_x	Entropy_TimeD_Ang
176	SMA_Angle	Entropy_TimeD_Ang			Dfreq_Acc_y	Avr_peak_time_Pitch
177	Min_Azimuth	SVM_Angle			Std_Acc_y	Dfreq_Gyr_x
178	Interq_Azimuth	SMA_Angle			Std_Pitch	Mean_Gyr_x
179	Var_Azimuth	Vedb_Angle			Dfreq_Acc_z	Cf_Acc_z
180	Widest_Peak_Gyr_x	Rms_Azimuth			Std_Gyr_z	Interq_Pitch
181	nPeaks_Gyr_z	Energy_Azimuth			Energy_Azimuth	AV_Ang
182	Entropy_Azimuth	Mean_Azimuth			Energy_Roll	Avr_peak_time_Gyr_x
183	Max_diff_Azimuth	Dfreq_Azimuth			Energy_Gyr_y	nPeaks_Roll

Appendix E. 8 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet2_b** over 7 $sec.window$.

	ReliefF		**GA**		**RF**	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Acc_x	Mean_Acc_x	Mean_Acc_z	Mean_Acc_x	Min_Roll	Min_Roll
2	Mean_Roll	Mean_Roll	Mean_Azimuth	Mean_Gyr_x	Mean_Acc_x	Mean_Roll
3	Cf_Roll	Entropy_Roll	Mean_Pitch	Var_Roll	Mean_Roll	Rms_Roll
4	Entropy_Roll	Cf_Roll	Mean_Roll	Var_Gyr_y	Rms_Pitch	Mean_Acc_x
5	Dfreq_Acc_x	Energy_Pitch	Mean_Gyr_y	Std_Acc_z	Rms_Roll	Mean_Pitch
6	Dfreq_Roll	Rms_Pitch	Mean_Gyr_z	Std_Azimuth	Cf_Acc_x	Min_Acc_y
7	Dfreq_Pitch	Dfreq_Pitch	Var_Pitch	Std_Gyr_y	Mean_Acc_y	Mean_Acc_z
8	Min_Roll	Dfreq_Acc_z	Var_Gyr_z	Kur_Acc_y	Mean_Acc_z	Var_Roll
9	Mean_Acc_z	Mean_Acc_z	Std_Acc_x	Kur_Roll	Kur_Acc_y	Rms_Acc_y
10	Dfreq_Acc_z	Dfreq_Roll	Std_Acc_y	Skew_Acc_z	Skew_Gyr_x	Dfreq_Pitch
11	Max_Roll	Min_Roll	Std_Acc_z	Skew_Gyr_x	SVM_Acc	Rms_Pitch
12	Rms_Pitch	Dfreq_Acc_x	Std_Azimuth	Skew_Gyr_y	Mean_Pitch	Mean_Acc_y
13	Entropy_Acc_x	Max_Roll	Std_Pitch	Skew_Gyr_z	Vedb_Acc	Max_Pitch
14	Dfreq_Acc_y	Dfreq_Acc_y	Std_Gyr_y	Min_Acc_x	Dfreq_Pitch	Max_diff_Acc_y
15	Energy_Pitch	Rms_Roll	Kur_Acc_y	Min_Azimuth	Var_Roll	Dfreq_Roll
16	Cf_Acc_x	Entropy_Acc_x	Kur_Acc_z	Min_Roll	Min_Acc_y	Rms_Acc_z
17	Vedb_Acc	Vedb_Acc	Kur_Pitch	Min_Gyr_y	Entropy_TimeD_Acc	Vedb_Acc
18	Rms_Roll	Cf_Acc_x	Kur_Roll	Max_Acc_y	Dfreq_Acc_x	Entropy_Gyr_z
19	Rms_Acc_y	Mean_Pitch	Kur_Gyr_x	Max_Acc_z	Max_Acc_y	Skew_Gyr_x
20	Energy_Acc_y	Rms_Acc_y	Kur_Gyr_y	Max_Azimuth	Rms_Acc_y	Dfreq_Acc_x
21	Min_Acc_x	Max_Acc_x	Skew_Acc_y	Max_Pitch	Min_Pitch	Min_Acc_z
22	Mean_Pitch	Energy_Roll	Skew_Acc_z	Max_Roll	Kur_Pitch	Interq_Roll
23	Entropy_Acc_y	Energy_Acc_y	Skew_Azimuth	Max_Gyr_z	Energy_Pitch	Min_Gyr_y
24	Max_Acc_x	Min_Acc_x	Skew_Pitch	Rms_Acc_x	Cf_Roll	Entropy_Acc_x
25	Rms_Acc_z	Entropy_Acc_y	Skew_Gyr_z	Rms_Acc_y	Dfreq_Roll	Var_Acc_z

26	Entropy_Acc_z	Mean_Acc_y	Min_Acc_x	Rms_Azimuth	Mean_Gyr_z	Max_Roll
27	Energy_Acc_z	Rms_Acc_z	Min_Acc_y	Rms_Pitch	Interq_Gyr_y	Cf_Acc_x
28	Energy_Roll	Entropy_Acc_z	Min_Pitch	Rms_Gyr_x	Var_Pitch	Highest_peak_Gyr_z
29	Interq_Gyr_x	Min_Acc_y	Min_Gyr_x	Rms_Gyr_z	Max_diff_Gyr_y	SVM_Acc
30	Mean_Acc_y	Energy_Acc_z	Max_Acc_x	Interq_Acc_y	Max_diff_Pitch	Max_diff_Gyr_y
31	Entropy_Pitch	Interq_Gyr_x	Max_Acc_z	Interq_Pitch	Var_Azimuth	Entropy_TimeD_Acc
32	Min_Acc_y	Rms_Gyr_x	Max_Azimuth	Interq_Gyr_x	Cf_Pitch	Min_Gyr_x
33	Rms_Acc_x	Std_Gyr_x	Max_Pitch	Interq_Gyr_z	Max_Roll	Cf_Roll
34	Rms_Gyr_x	Entropy_Pitch	Max_Roll	Cf_Azimuth	Skew_Acc_y	Var_Gyr_z
35	Max_Acc_y	Rms_Acc_x	Max_Gyr_x	Cf_Roll	Interq_Roll	Max_Acc_y
36	Std_Gyr_x	Max_Acc_y	Rms_Acc_x	Cf_Gyr_z	Var_Acc_x	Var_Acc_y
37	Entropy_Gyr_y	Min_Gyr_y	Rms_Acc_y	SMA_Angle	Min_Acc_x	Highest_peak_Azimuth
38	Max_diff_Acc_y	Min_Pitch	Rms_Roll	SMA_Gyr	Interq_Acc_x	Var_Gyr_y
39	Interq_Acc_y	Energy_Gyr_x	Rms_Gyr_x	Entropy_Acc_y	Max_Gyr_y	Highest_peak_Roll
40	Min_Gyr_y	Var_Gyr_x	Rms_Gyr_z	Entropy_Azimuth	Max_diff_Roll	Max_Gyr_z
41	Max_diff_Azimuth	Cf_Acc_y	Interq_Acc_x	Entropy_Pitch	Rms_Acc_x	Interq_Acc_x
42	Min_Pitch	Max_diff_Acc_y	Interq_Acc_z	Entropy_Roll	Widest_Peak_Acc_z	Interq_Gyr_z
43	Cf_Gyr_x	Skew_Acc_y	Interq_Gyr_x	Entropy_Gyr_x	Dfreq_Gyr_y	Var_Gyr_x
44	SVM_Acc	Interq_Gyr_z	Interq_Gyr_y	Entropy_Gyr_y	Min_Gyr_x	Min_Gyr_z
45	Interq_Acc_x	Energy_Acc_x	Cf_Acc_x	Entropy_Gyr_z	Vedb_Angle	Highest_peak_Gyr_y
46	Interq_Gyr_y	Highest_peak_Gyr_x	Cf_Acc_z	Entropy_TimeD_Ang	Kur_Acc_z	Min_Acc_x
47	Cf_Acc_y	Std_Roll	Cf_Roll	Energy_Acc_y	SMA_Acc	Rms_Azimuth
48	SMA_Acc	Dfreq_Gyr_x	Cf_Gyr_y	Energy_Azimuth	Widest_Peak_Gyr_y	Max_diff_Gyr_z
49	Highest_peak_Gyr_x	Highest_peak_Roll	SMA_Acc	Energy_Pitch	Entropy_Gyr_x	Skew_Gyr_z
50	Dfreq_Gyr_x	Std_Gyr_z	SMA_Angle	Energy_Roll	Energy_Azimuth	Mean_Gyr_y
51	Var_Gyr_x	Highest_peak_Gyr_z	Entropy_Acc_x	Energy_Gyr_x	Entropy_Acc_y	Var_Azimuth
52	Energy_Gyr_x	Std_Gyr_y	Entropy_Roll	Energy_Gyr_z	Entropy_Roll	Skew_Gyr_y
53	Entropy_TimeD_Acc	Rms_Gyr_y	Entropy_Gyr_y	SVM_Angle	Kur_Acc_x	Max_diff_Pitch
54	Rms_Gyr_y	Min_Acc_z	Entropy_TimeD_Acc	SVM_Gyr	Energy_Acc_y	Rms_Acc_x

55	Entropy_Gyr_x	Max_diff_Azimuth	Entropy_TimeD_Ang	DSVM_Gyr	Cf_Gyr_x	Dfreq_Gyr_y
56	Std_Gyr_y	Max_Pitch	Energy_Acc_x	Max_diff_Acc_y	Kur_Gyr_z	nPeaks_Acc_z
57	Interq_Gyr_z	Cf_Gyr_x	Energy_Azimuth	Max_diff_Pitch	Skew_Pitch	Cf_Acc_z
58	Highest_peak_Gyr_z	Min_Gyr_x	Energy_Pitch	Max_diff_Roll	Var_Acc_y	Kur_Acc_y
59	Std_Roll	Interq_Acc_y	Energy_Roll	Max_diff_Gyr_x	Rms_Acc_z	DSVM_Acc
60	Std_Gyr_z	SVM_Acc	Energy_Gyr_x	Max_diff_Gyr_y	Dfreq_Gyr_x	Entropy_Roll
61	Interq_Roll	Max_diff_Gyr_y	Energy_Gyr_y	MV_Gyr	Cf_Gyr_z	Widest_Peak_Acc_x
62	Interq_Acc_z	Entropy_Gyr_x	DSVM_Gyr	mag_Ang	Std_Gyr_y	Max_diff_Acc_z
63	Energy_Acc_x	Std_Acc_x	Max_diff_Acc_y	Vedb_Gyr	Max_Pitch	Kur_Acc_x
64	DSAM_Angle	Entropy_Gyr_y	Max_diff_Acc_z	Dfreq_Acc_x	Rms_Gyr_z	Kur_Pitch
65	MV_Gyr	SMA_Acc	Max_diff_Azimuth	Dfreq_Pitch	Highest_peak_Acc_z	Highest_peak_Acc_x
66	Min_Acc_z	Interq_Roll	Max_diff_Pitch	Dfreq_Roll	Min_Gyr_y	Interq_Pitch
67	Mean_Gyr_z	Max_Gyr_y	Max_diff_Roll	Dfreq_Gyr_y	Max_diff_Gyr_z	Min_Pitch
68	MV_Acc	Cf_Pitch	Max_diff_Gyr_x	Dfreq_Gyr_z	Entropy_Pitch	Skew_Acc_y
69	Std_Acc_y	Entropy_TimeD_Acc	MV_Acc	Widest_Peak_Acc_x	Avr_peak_time_Acc_x	Entropy_TimeD_Gyr
70	Max_diff_Gyr_x	Highest_peak_Acc_z	AV_Ang	Highest_peak_Acc_x	MV_Gyr	Dfreq_Gyr_z
71	AV_Ang	Rms_Gyr_z	mag_Ang	Avr_peak_time_Acc_x	Avr_peak_time_Gyr_x	Skew_Acc_z
72	Highest_peak_Acc_x	Interq_Acc_x	mag_Gyr	Widest_Peak_Acc_z	Widest_Peak_Azimuth	Entropy_Acc_y
73	Var_Acc_y	Skew_Gyr_x	Dfreq_Acc_x	Highest_peak_Acc_z	Highest_peak_Gyr_x	nPeaks_Gyr_x
74	Highest_peak_Gyr_y	Std_Pitch	Dfreq_Acc_z	nPeaks_Azimuth	Rms_Gyr_x	Max_diff_Gyr_x
75	Max_Pitch	MV_Gyr	Dfreq_Pitch	Highest_peak_Azimuth	Mean_Gyr_y	Entropy_Acc_z
76	Rms_Gyr_z	MV_Acc	nPeaks_Acc_x	Avr_peak_time_Azimuth	Energy_Acc_z	Avr_peak_time_Gyr_x
77	Highest_peak_Roll	Interq_Gyr_y	Widest_Peak_Acc_x	nPeaks_Pitch	Max_diff_Gyr_x	Avr_peak_time_Acc_y
78	Dfreq_Gyr_y	Var_Roll	Avr_peak_time_Acc_x	Widest_Peak_Gyr_x	DSAM_Angle	Max_Acc_x
79	Max_diff_Gyr_y	Mean_Gyr_z	Widest_Peak_Acc_y	Avr_peak_time_Gyr_x	DSVM_Gyr	Skew_Roll
80	Skew_Acc_y	Std_Acc_y	Avr_peak_time_Acc_y	Avr_peak_time_Gyr_z	Max_diff_Acc_z	mag_Ang
81	DSVM_Acc	Var_Acc_x	Widest_Peak_Acc_z		Std_Gyr_z	Mean_Gyr_z
82	nPeaks_Azimuth	Interq_Acc_z	nPeaks_Azimuth		Highest_peak_Acc_x	Dfreq_Acc_y
83	Skew_Gyr_z	Cf_Acc_z	Widest_Peak_Azimuth		Avr_peak_time_Azimuth	Highest_peak_Acc_z

84	Std_Pitch	SVM_Gyr	Avr_peak_time_Azimuth		SVM_Angle	Cf_Pitch
85	Cf_Pitch	Var_Acc_y	Avr_peak_time_Pitch		Dfreq_Acc_y	SMA_Gyr
86	Skew_Gyr_x	SMA_Gyr	nPeaks_Roll		Highest_peak_Azimuth	Interq_Acc_z
87	Std_Acc_z	Vedb_Gyr	Widest_Peak_Roll		Highest_peak_Gyr_z	Skew_Pitch
88	Std_Acc_x	Max_Gyr_x	Avr_peak_time_Roll		Entropy_Azimuth	Avr_peak_time_Acc_x
89	nPeaks_Gyr_y	DSVM_Acc	Avr_peak_time_Gyr_x		Std_Acc_y	Cf_Gyr_x
90	Cf_Gyr_z	Max_diff_Gyr_x	nPeaks_Gyr_z		Entropy_Gyr_y	nPeaks_Gyr_z
91	Min_Gyr_z	Var_Gyr_z	Widest_Peak_Gyr_z		nPeaks_Pitch	Avr_peak_time_Pitch
92	Max_Gyr_y	Highest_peak_Acc_x	Avr_peak_time_Gyr_z		Std_Gyr_x	Widest_Peak_Gyr_z
93	Max_Gyr_x	Std_Acc_z			Kur_Gyr_x	mag_Gyr
94	Min_Gyr_x	DSVM_Gyr			Entropy_Acc_x	AV_Ang
95	Cf_Acc_z	Entropy_TimeD_Gyr			Highest_peak_Gyr_y	Max_diff_Acc_x
96	DSVM_Gyr	Interq_Pitch			Interq_Gyr_x	Var_Acc_x
97	Highest_peak_Acc_z	Kur_Gyr_x			Interq_Acc_y	SMA_Acc
98	SVM_Gyr	DSAM_Angle			Std_Acc_x	Max_Gyr_x
99	SMA_Gyr	Max_diff_Acc_z			Interq_Acc_z	mag_Acc
100	Energy_Gyr_y	Entropy_Gyr_z			Max_diff_Acc_y	MV_Acc
101	Vedb_Gyr	Highest_peak_Acc_y			Max_Acc_x	Entropy_TimeD_Ang
102	Var_Gyr_y	Dfreq_Gyr_y			Widest_Peak_Roll	Widest_Peak_Gyr_y
103	Widest_Peak_Gyr_x	Skew_Acc_x			Interq_Pitch	Max_Azimuth
104	Var_Gyr_z	Min_Gyr_z			Entropy_TimeD_Gyr	Kur_Gyr_x
105	Dfreq_Gyr_z	Highest_peak_Gyr_y			Skew_Acc_z	Cf_Acc_y
106	Var_Roll	AV_Ang			Kur_Gyr_y	Cf_Gyr_z
107	Widest_Peak_Acc_y	Dfreq_Gyr_z			Cf_Acc_y	Interq_Acc_y
108	Entropy_TimeD_Gyr	Skew_Pitch			Min_Acc_z	Vedb_Gyr
109	Var_Acc_x	Energy_Gyr_y			Widest_Peak_Pitch	Entropy_Pitch
110	Widest_Peak_Acc_z	Var_Gyr_y			Energy_Roll	Var_Pitch
111	Widest_Peak_Azimuth	Skew_Roll			Mean_Gyr_x	DSVM_Gyr
112	Energy_Gyr_z	Kur_Acc_y			DSVM_Acc	nPeaks_Acc_y

113	Max_diff_Roll	Max_Acc_z			Std_Azimuth	Interq_Gyr_y
114	Kur_Gyr_x	Energy_Gyr_z			Interq_Azimuth	Max_diff_Roll
115	Var_Acc_z	nPeaks_Gyr_z			Energy_Acc_x	Mean_Azimuth
116	Max_diff_Acc_z	Kur_Acc_x			SMA_Angle	Highest_peak_Pitch
117	Entropy_Gyr_z	Std_Azimuth			Max_Gyr_z	Skew_Azimuth
118	Kur_Gyr_z	Skew_Acc_z			Entropy_TimeD_Ang	Avr_peak_time_Acc_z
119	Interq_Pitch	Var_Pitch			Mean_Azimuth	Kur_Acc_z
120	Kur_Acc_y	Max_diff_Roll			Std_Pitch	Rms_Gyr_y
121	Avr_peak_time_Azimuth	Max_diff_Pitch			Dfreq_Gyr_z	Interq_Azimuth
122	Skew_Acc_z	Var_Acc_z			Cf_Azimuth	Max_Gyr_y
123	nPeaks_Gyr_z	nPeaks_Gyr_x			Entropy_Acc_z	Mean_Gyr_x
124	Widest_Peak_Gyr_y	Max_Gyr_z			Std_Acc_z	Widest_Peak_Acc_z
125	Max_Gyr_z	nPeaks_Pitch			Min_Gyr_z	Avr_peak_time_Gyr_y
126	Avr_peak_time_Gyr_y	Avr_peak_time_Gyr_z			Vedb_Gyr	Rms_Gyr_z
127	Var_Pitch	Mean_Gyr_y			Var_Gyr_x	Kur_Azimuth
128	mag_Ang	Skew_Gyr_z			Cf_Gyr_y	SVM_Gyr
129	Kur_Acc_z	nPeaks_Acc_z			Max_Acc_z	Vedb_Angle
130	Skew_Pitch	Max_diff_Gyr_z			Skew_Gyr_z	Dfreq_Gyr_x
131	Max_diff_Pitch	Widest_Peak_Acc_y			Dfreq_Acc_z	Interq_Gyr_x
132	Highest_peak_Acc_y	Widest_Peak_Acc_z			Cf_Acc_z	nPeaks_Pitch
133	mag_Acc	Highest_peak_Azimuth			nPeaks_Gyr_z	DSAM_Angle
134	Max_diff_Gyr_z	Mean_Gyr_x			Std_Roll	Std_Acc_x
135	Highest_peak_Azimuth	Skew_Gyr_y			Var_Gyr_z	Std_Acc_y
136	mag_Gyr	Kur_Roll			Var_Gyr_y	Std_Acc_z
137	Max_Acc_z	Cf_Gyr_z			Skew_Gyr_y	Std_Azimuth
138	Widest_Peak_Gyr_z	Max_diff_Acc_x			nPeaks_Acc_z	Std_Pitch
139	Std_Azimuth	Cf_Gyr_y			MV_Acc	Std_Roll
140	nPeaks_Acc_y	Kur_Acc_z			SVM_Gyr	Std_Gyr_x
141	Kur_Roll	mag_Gyr			Energy_Gyr_x	Std_Gyr_y

142	Skew_Acc_x	Avr_peak_time_Gyr_x			Avr_peak_time_Gyr_z	Std_Gyr_z
143	Interq_Azimuth	Avr_peak_time_Pitch			Skew_Azimuth	Energy_Acc_x
144	Mean_Gyr_y	Kur_Pitch			Kur_Roll	Energy_Acc_y
145	Skew_Roll	Min_Azimuth			Widest_Peak_Acc_x	Energy_Acc_z
146	Cf_Gyr_y	Kur_Gyr_z			SMA_Gyr	Energy_Azimuth
147	Kur_Acc_x	Widest_Peak_Gyr_z			Avr_peak_time_Gyr_y	Energy_Pitch
148	nPeaks_Gyr_x	Avr_peak_time_Acc_z			Var_Acc_z	Energy_Roll
149	Widest_Peak_Acc_x	Kur_Gyr_y			Avr_peak_time_Pitch	Energy_Gyr_x
150	Highest_peak_Pitch	Avr_peak_time_Azimuth			Interq_Gyr_z	Energy_Gyr_y
151	Avr_peak_time_Acc_z	Avr_peak_time_Roll			Highest_peak_Pitch	Energy_Gyr_z
152	Cf_Azimuth	Widest_Peak_Azimuth			Avr_peak_time_Acc_y	SVM_Angle
153	Kur_Pitch	Highest_peak_Pitch			Entropy_Gyr_z	Dfreq_Acc_z
154	Min_Azimuth	Widest_Peak_Pitch			Widest_Peak_Gyr_x	Dfreq_Azimuth
155	Max_diff_Acc_x	Widest_Peak_Acc_x			Energy_Gyr_z	Entropy_Gyr_x
156	Avr_peak_time_Gyr_x	mag_Acc			AV_Ang	Rms_Gyr_x
157	Kur_Azimuth	nPeaks_Azimuth			Dfreq_Azimuth	Widest_Peak_Acc_y
158	Var_Azimuth	nPeaks_Roll			nPeaks_Acc_y	Entropy_Gyr_y
159	nPeaks_Acc_z	nPeaks_Gyr_y			Skew_Acc_x	SMA_Angle
160	Avr_peak_time_Roll	Var_Azimuth			Energy_Gyr_y	Kur_Gyr_y
161	Avr_peak_time_Gyr_z	Entropy_Azimuth			Widest_Peak_Gyr_z	Kur_Roll
162	Avr_peak_time_Pitch	Widest_Peak_Gyr_y			Rms_Gyr_y	Entropy_Azimuth
163	nPeaks_Roll	Kur_Azimuth			nPeaks_Azimuth	Cf_Gyr_y
164	Avr_peak_time_Acc_y	Cf_Azimuth			Kur_Azimuth	Kur_Gyr_z
165	Entropy_Azimuth	Avr_peak_time_Gyr_y			Max_Azimuth	Max_diff_Azimuth
166	Skew_Azimuth	Skew_Azimuth			Highest_peak_Roll	Avr_peak_time_Azimuth
167	Skew_Gyr_y	Interq_Azimuth			Min_Azimuth	nPeaks_Roll
168	Kur_Gyr_y	Avr_peak_time_Acc_y			Widest_Peak_Acc_y	Cf_Azimuth
169	nPeaks_Pitch	Widest_Peak_Gyr_x			nPeaks_Acc_x	Highest_peak_Acc_y
170	Avr_peak_time_Acc_x	nPeaks_Acc_y			mag_Acc	Min_Azimuth

171	nPeaks_Acc_x	mag_Ang			nPeaks_Roll	Avr_peak_time_Roll
172	Widest_Peak_Roll	Widest_Peak_Roll			Avr_peak_time_Acc_z	MV_Gyr
173	Max_Azimuth	Max_Azimuth			Max_diff_Acc_x	Widest_Peak_Pitch
174	Mean_Gyr_x	Avr_peak_time_Acc_x			Max_diff_Azimuth	Widest_Peak_Azimuth
175	Widest_Peak_Pitch	Energy_Azimuth			Skew_Roll	nPeaks_Azimuth
176	Energy_Azimuth	nPeaks_Acc_x			nPeaks_Gyr_y	Widest_Peak_Gyr_x
177	SVM_Angle	Rms_Azimuth			Max_Gyr_x	nPeaks_Acc_x
178	SMA_Angle	Vedb_Angle			mag_Ang	nPeaks_Gyr_y
179	Entropy_TimeD_Ang	Dfreq_Azimuth			Rms_Azimuth	Highest_peak_Gyr_x
180	Rms_Azimuth	Mean_Azimuth			Highest_peak_Acc_y	Max_Acc_z
181	Mean_Azimuth	SVM_Angle			mag_Gyr	Avr_peak_time_Gyr_z
182	Dfreq_Azimuth	SMA_Angle			nPeaks_Gyr_x	Widest_Peak_Roll
183	Vedb_Angle	Entropy_TimeD_Ang			Avr_peak_time_Roll	Skew_Acc_x

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 9 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet2_b** over 5 *sec. window*.

	ReliefF		**GA**		**RF**	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Acc_x	Mean_Acc_x	Mean_Acc_x	Mean_Acc_x	Min_Roll	Min_Roll
2	Mean_Roll	Mean_Roll	Mean_Acc_y	Mean_Pitch	Mean_Acc_x	Mean_Roll
3	Cf_Roll	Cf_Roll	Mean_Acc_z	Mean_Roll	Mean_Roll	Mean_Acc_x
4	Min_Roll	Dfreq_Acc_z	Mean_Azimuth	Mean_Gyr_y	Rms_Pitch	Mean_Acc_z
5	Dfreq_Roll	Dfreq_Roll	Mean_Roll	Mean_Gyr_z	Mean_Pitch	Mean_Acc_y
6	Entropy_Roll	Min_Roll	Mean_Gyr_y	Var_Acc_x	Mean_Acc_z	Rms_Pitch
7	Dfreq_Pitch	Max_Roll	Var_Roll	Var_Acc_y	Dfreq_Pitch	Max_diff_Acc_y
8	Rms_Pitch	Entropy_Roll	Var_Gyr_y	Var_Azimuth	Rms_Roll	Min_Acc_y
9	Max_Roll	Dfreq_Acc_y	Std_Acc_y	Var_Gyr_x	Dfreq_Acc_x	Mean_Pitch
10	Dfreq_Acc_z	Dfreq_Pitch	Std_Acc_z	Var_Gyr_y	Mean_Acc_y	Rms_Roll
11	Dfreq_Acc_y	Rms_Roll	Std_Gyr_x	Var_Gyr_z	Vedb_Acc	Var_Roll
12	Energy_Pitch	Mean_Acc_z	Kur_Acc_x	Std_Azimuth	Kur_Acc_y	Min_Pitch
13	Mean_Acc_z	Rms_Pitch	Kur_Acc_y	Std_Pitch	Max_diff_Gyr_z	Skew_Gyr_x
14	Dfreq_Acc_x	Dfreq_Acc_x	Kur_Acc_z	Std_Gyr_x	Cf_Pitch	Dfreq_Pitch
15	Rms_Roll	Rms_Acc_y	Kur_Azimuth	Std_Gyr_y	Cf_Roll	Entropy_Roll
16	Cf_Acc_x	Cf_Acc_x	Kur_Gyr_y	Kur_Acc_x	Interq_Roll	Kur_Acc_y
17	Min_Acc_x	Energy_Pitch	Skew_Acc_x	Kur_Acc_z	Rms_Acc_z	Max_Pitch
18	Vedb_Acc	Energy_Acc_y	Skew_Pitch	Kur_Azimuth	Interq_Gyr_z	Rms_Acc_z
19	Rms_Acc_y	Min_Acc_x	Skew_Roll	Kur_Gyr_z	Min_Pitch	Min_Gyr_x
20	Entropy_Acc_y	Max_Acc_x	Skew_Gyr_z	Skew_Acc_x	Var_Acc_x	Max_Roll
21	Max_Acc_x	Entropy_Acc_y	Min_Acc_x	Skew_Acc_y	Var_Roll	Rms_Acc_y
22	Energy_Acc_y	Entropy_Acc_x	Min_Azimuth	Skew_Azimuth	Var_Gyr_x	Min_Acc_z
23	Mean_Pitch	Vedb_Acc	Min_Pitch	Skew_Pitch	Max_Acc_x	Var_Gyr_z
24	Entropy_Acc_x	Rms_Acc_z	Min_Gyr_z	Skew_Roll	Dfreq_Roll	Max_Gyr_z
25	Rms_Acc_z	Energy_Roll	Max_Acc_x	Skew_Gyr_x	Entropy_Acc_y	Entropy_Acc_x

26	Energy_Roll	Mean_Pitch	Max_Azimuth	Skew_Gyr_y	Var_Gyr_y	Mean_Azimuth
27	Entropy_Acc_z	Energy_Acc_z	Max_Roll	Min_Acc_z	Highest_peak_Gyr_z	Interq_Roll
28	Mean_Acc_y	Max_Acc_y	Max_Gyr_x	Min_Azimuth	Max_diff_Gyr_y	Dfreq_Roll
29	Energy_Acc_z	Mean_Acc_y	Max_Gyr_y	Min_Roll	Max_Roll	Skew_Azimuth
30	Interq_Gyr_x	Entropy_Acc_z	Max_Gyr_z	Min_Gyr_x	Max_Acc_z	Cf_Acc_y
31	Min_Pitch	Min_Acc_y	Rms_Acc_x	Min_Gyr_z	Interq_Acc_x	Rms_Acc_x
32	Min_Acc_y	Min_Pitch	Rms_Acc_z	Max_Acc_x	Cf_Acc_x	Max_diff_Acc_x
33	Rms_Acc_x	Interq_Gyr_x	Rms_Pitch	Max_Acc_y	Rms_Acc_y	Max_diff_Gyr_y
34	Max_Acc_y	Rms_Acc_x	Rms_Roll	Max_Roll	Rms_Acc_x	Vedb_Acc
35	Entropy_Gyr_z	Rms_Gyr_x	Rms_Gyr_x	Max_Gyr_x	Interq_Azimuth	Interq_Gyr_z
36	Std_Gyr_x	Std_Gyr_x	Rms_Gyr_y	Rms_Acc_x	Rms_Gyr_y	Skew_Acc_y
37	Rms_Gyr_x	Min_Acc_z	Rms_Gyr_z	Rms_Acc_z	Cf_Acc_y	Mean_Gyr_y
38	Entropy_Pitch	Entropy_Pitch	Interq_Acc_x	Rms_Azimuth	Min_Acc_z	Widest_Peak_Pitch
39	Min_Acc_z	Cf_Acc_y	Interq_Acc_z	Rms_Pitch	Widest_Peak_Gyr_x	Cf_Roll
40	Std_Roll	Std_Gyr_z	Interq_Pitch	Rms_Roll	Var_Pitch	Widest_Peak_Acc_y
41	Entropy_Gyr_y	Interq_Gyr_z	Interq_Roll	Rms_Gyr_x	Interq_Gyr_x	Interq_Acc_x
42	Min_Gyr_y	Dfreq_Gyr_z	Interq_Gyr_z	Rms_Gyr_z	Dfreq_Gyr_z	Dfreq_Gyr_z
43	Std_Gyr_z	Var_Gyr_x	Cf_Acc_x	Interq_Acc_x	Entropy_Acc_x	Cf_Acc_z
44	Interq_Gyr_z	Entropy_Gyr_y	Cf_Acc_z	Interq_Acc_y	Interq_Acc_y	Entropy_Acc_y
45	Highest_peak_Gyr_z	Energy_Gyr_x	Cf_Gyr_x	Interq_Acc_z	Entropy_Pitch	DSAM_Angle
46	Dfreq_Gyr_x	Max_Acc_z	Cf_Gyr_y	Interq_Azimuth	Max_diff_Acc_z	Highest_peak_Pitch
47	Highest_peak_Gyr_x	Mean_Gyr_z	SMA_Gyr	Interq_Roll	Highest_peak_Acc_z	Max_Acc_x
48	Energy_Acc_x	Rms_Gyr_z	Entropy_Acc_y	Interq_Gyr_y	Kur_Gyr_x	Entropy_Pitch
49	Interq_Acc_x	Min_Gyr_y	Entropy_Roll	Cf_Acc_x	SMA_Acc	Var_Gyr_y
50	Highest_peak_Acc_z	Max_diff_Acc_y	Entropy_Gyr_y	Cf_Acc_y	Max_Gyr_z	Max_Gyr_x
51	Skew_Acc_y	Dfreq_Gyr_x	Entropy_TimeD_Acc	Cf_Acc_z	Mean_Gyr_x	Var_Acc_x
52	Max_diff_Acc_y	Max_Pitch	Entropy_TimeD_Ang	Cf_Azimuth	Mean_Gyr_z	AV_Ang
53	Interq_Gyr_y	Highest_peak_Gyr_z	Energy_Acc_x	Cf_Pitch	Entropy_TimeD_Gyr	Cf_Acc_x
54	Cf_Acc_y	Cf_Gyr_z	Energy_Acc_z	Cf_Roll	Entropy_Acc_z	Skew_Roll

55	Highest_peak_Roll	Interq_Gyr_y	Energy_Azimuth	Cf_Gyr_z	Entropy_Roll	Max_diff_Gyr_x
56	Interq_Acc_y	Highest_peak_Gyr_x	Energy_Roll	SMA_Gyr	Min_Acc_y	Max_Acc_y
57	Std_Acc_x	Cf_Gyr_x	Energy_Gyr_y	Entropy_Acc_z	DSAM_Angle	Max_diff_Gyr_z
58	Max_Pitch	Std_Roll	Energy_Gyr_z	Entropy_Azimuth	nPeaks_Gyr_x	Rms_Gyr_z
59	Entropy_Gyr_x	Skew_Acc_y	SVM_Angle	Entropy_Pitch	Highest_peak_Gyr_y	Kur_Acc_x
60	Var_Gyr_x	Min_Gyr_x	DSAM_Angle	Entropy_Gyr_x	Interq_Pitch	Dfreq_Acc_x
61	SVM_Acc	Interq_Acc_y	DSVM_Gyr	Entropy_Gyr_y	Cf_Gyr_x	Max_Acc_z
62	Cf_Gyr_x	Entropy_Gyr_z	Max_diff_Gyr_z	Entropy_TimeD_Acc	Entropy_Gyr_y	mag_Ang
63	Rms_Gyr_y	Skew_Gyr_z	AV_Ang	Entropy_TimeD_Ang	Interq_Gyr_y	Rms_Azimuth
64	SMA_Acc	Energy_Acc_x	mag_Acc	Energy_Acc_x	Mean_Gyr_y	Min_Gyr_y
65	Energy_Gyr_x	Highest_peak_Acc_x	Vedb_Acc	Energy_Acc_y	Skew_Gyr_x	Highest_peak_Roll
66	nPeaks_Acc_x	Highest_peak_Roll	Vedb_Gyr	Energy_Acc_z	Entropy_TimeD_Acc	Highest_peak_Acc_z
67	DSVM_Acc	Interq_Acc_x	Dfreq_Acc_x	Energy_Azimuth	Vedb_Angle	Kur_Gyr_z
68	Std_Gyr_y	Highest_peak_Acc_z	Dfreq_Azimuth	Energy_Pitch	nPeaks_Gyr_z	Max_diff_Azimuth
69	Rms_Gyr_z	Std_Acc_x	Dfreq_Pitch	Energy_Roll	Var_Acc_y	Kur_Acc_z
70	MV_Gyr	Rms_Gyr_y	Dfreq_Roll	Energy_Gyr_y	Min_Gyr_z	Entropy_TimeD_Ang
71	Std_Acc_y	Std_Gyr_y	Dfreq_Gyr_y	Energy_Gyr_z	Widest_Peak_Pitch	Kur_Gyr_y
72	Skew_Gyr_x	Std_Acc_y	Dfreq_Gyr_z	SVM_Acc	Max_Pitch	Var_Acc_y
73	Kur_Acc_y	Interq_Roll	Widest_Peak_Acc_x	SVM_Angle	Min_Gyr_x	Min_Gyr_z
74	Cf_Acc_z	Cf_Acc_z	Highest_peak_Acc_x	DSAM_Angle	Skew_Gyr_z	Avr_peak_time_Acc_x
75	Entropy_TimeD_Acc	Max_Gyr_y	Avr_peak_time_Acc_x	Max_diff_Acc_x	SMA_Gyr	Mean_Gyr_x
76	Highest_peak_Acc_x	Max_Gyr_x	Widest_Peak_Acc_y	Max_diff_Acc_z	Min_Gyr_y	Cf_Azimuth
77	Interq_Acc_z	Max_diff_Gyr_y	Highest_peak_Acc_y	Max_diff_Gyr_x	Max_diff_Acc_y	Interq_Gyr_y
78	Skew_Gyr_z	Skew_Gyr_x	Avr_peak_time_Acc_y	Max_diff_Gyr_y	Max_Azimuth	Dfreq_Acc_y
79	Var_Roll	Max_diff_Gyr_x	nPeaks_Acc_z	Max_diff_Gyr_z	Avr_peak_time_Acc_z	Var_Gyr_x
80	MV_Acc	Cf_Pitch	Widest_Peak_Acc_z	MV_Gyr	Max_diff_Acc_x	Avr_peak_time_Gyr_y
81	Cf_Gyr_z	Skew_Acc_x	Avr_peak_time_Acc_z	mag_Acc	Rms_Gyr_x	Cf_Gyr_x
82	Var_Acc_x	SVM_Acc	nPeaks_Azimuth	mag_Ang	Highest_peak_Roll	Cf_Pitch
83	Widest_Peak_Roll	Max_Gyr_z	Widest_Peak_Azimuth	mag_Gyr	Cf_Acc_z	Highest_peak_Azimuth

84	nPeaks_Acc_z	Max_diff_Acc_z	Highest_peak_Azimuth	Vedb_Angle	nPeaks_Acc_x	SMA_Acc
85	Interq_Roll	Highest_peak_Gyr_y	Avr_peak_time_Azimuth	Dfreq_Acc_z	Entropy_Gyr_z	Skew_Pitch
86	Var_Acc_y	SMA_Acc	nPeaks_Pitch	Dfreq_Azimuth	AV_Ang	Interq_Acc_y
87	Max_diff_Gyr_y	Var_Gyr_z	nPeaks_Roll	Dfreq_Roll	Max_Acc_y	SVM_Angle
88	Std_Pitch	SVM_Gyr	nPeaks_Gyr_x	Dfreq_Gyr_y	Max_diff_Pitch	Max_Azimuth
89	Skew_Acc_x	Var_Acc_y	Widest_Peak_Gyr_x	Widest_Peak_Acc_x	SMA_Angle	Entropy_TimeD_Gyr
90	Dfreq_Gyr_z	MV_Gyr	Highest_peak_Gyr_x	Highest_peak_Acc_x	Min_Acc_x	nPeaks_Acc_x
91	Widest_Peak_Acc_z	Kur_Acc_y	Avr_peak_time_Gyr_x	Avr_peak_time_Acc_x	Skew_Acc_y	Skew_Acc_z
92	Max_diff_Acc_z	Min_Gyr_z	nPeaks_Gyr_y	Avr_peak_time_Acc_y	Highest_peak_Gyr_x	Cf_Gyr_z
93	Cf_Gyr_y	SMA_Gyr	Widest_Peak_Gyr_y	Highest_peak_Acc_z	Highest_peak_Acc_y	Vedb_Angle
94	Max_Gyr_y	Max_diff_Acc_x	Widest_Peak_Gyr_z	nPeaks_Azimuth	Entropy_Gyr_x	Dfreq_Gyr_x
95	SVM_Gyr	Dfreq_Gyr_y	Avr_peak_time_Gyr_z	Highest_peak_Azimuth	Cf_Gyr_z	Entropy_Gyr_y
96	Var_Gyr_z	Entropy_TimeD_Acc		Widest_Peak_Pitch	MV_Acc	Highest_peak_Gyr_z
97	Mean_Gyr_z	DSAM_Angle		Avr_peak_time_Pitch	mag_Ang	Entropy_Gyr_z
98	SMA_Gyr	Skew_Roll		Widest_Peak_Roll	DSVM_Gyr	Highest_peak_Gyr_y
99	Cf_Pitch	Std_Pitch		Highest_peak_Roll	Widest_Peak_Roll	Var_Pitch
100	Std_Acc_z	Entropy_Azimuth		nPeaks_Gyr_x	Max_diff_Azimuth	Cf_Gyr_y
101	Dfreq_Gyr_y	Highest_peak_Acc_y		Widest_Peak_Gyr_x	Vedb_Gyr	Var_Acc_z
102	Min_Gyr_x	Var_Roll		Widest_Peak_Gyr_y	MV_Gyr	Mean_Gyr_z
103	Skew_Acc_z	AV_Ang		nPeaks_Gyr_z	Widest_Peak_Azimuth	Var_Azimuth
104	Mean_Gyr_y	Entropy_TimeD_Gyr		Widest_Peak_Gyr_z	Kur_Azimuth	Entropy_Acc_z
105	Highest_peak_Gyr_y	Energy_Gyr_z		Avr_peak_time_Gyr_z	Highest_peak_Acc_x	nPeaks_Azimuth
106	Max_Gyr_x	Mean_Gyr_y			Var_Azimuth	Highest_peak_Acc_x
107	Max_Gyr_z	Interq_Acc_z			Var_Gyr_z	MV_Acc
108	Min_Gyr_z	Std_Acc_z			Widest_Peak_Gyr_z	Max_diff_Pitch
109	DSAM_Angle	Entropy_Gyr_x			nPeaks_Acc_z	Entropy_TimeD_Acc
110	Entropy_TimeD_Gyr	DSVM_Acc			Avr_peak_time_Gyr_z	Min_Acc_x
111	Max_diff_Gyr_z	Interq_Pitch			Cf_Azimuth	Interq_Pitch
112	Highest_peak_Acc_y	Skew_Acc_z			Rms_Azimuth	Avr_peak_time_Gyr_x

113	DSVM_Gyr	Vedb_Gyr			Avr_peak_time_Acc_y	Max_Gyr_y
114	AV_Ang	Var_Acc_x			Skew_Gyr_y	Widest_Peak_Acc_x
115	Widest_Peak_Azimuth	Std_Azimuth			Var_Acc_z	Interq_Acc_z
116	Avr_peak_time_Acc_x	DSVM_Gyr			Interq_Acc_z	Max_diff_Roll
117	Max_diff_Pitch	MV_Acc			Dfreq_Gyr_y	Rms_Gyr_x
118	Widest_Peak_Gyr_x	Cf_Gyr_y			SVM_Acc	Interq_Gyr_x
119	Interq_Pitch	Kur_Acc_z			Mean_Azimuth	nPeaks_Gyr_z
120	Max_diff_Azimuth	Widest_Peak_Gyr_z			Widest_Peak_Gyr_y	Rms_Gyr_y
121	Energy_Gyr_y	Max_diff_Roll			Kur_Roll	Interq_Azimuth
122	Vedb_Gyr	Max_diff_Gyr_z			Dfreq_Gyr_x	SVM_Acc
123	Var_Gyr_y	Energy_Gyr_y			Max_diff_Gyr_x	Avr_peak_time_Roll
124	Max_Acc_z	Var_Gyr_y			nPeaks_Acc_y	MV_Gyr
125	Kur_Acc_z	Skew_Pitch			Widest_Peak_Acc_z	nPeaks_Gyr_y
126	Skew_Gyr_y	Max_diff_Pitch			Min_Azimuth	Entropy_Azimuth
127	Max_diff_Acc_x	Skew_Gyr_y			Avr_peak_time_Acc_x	Widest_Peak_Gyr_x
128	Energy_Gyr_z	Kur_Gyr_x			Avr_peak_time_Pitch	Avr_peak_time_Gyr_z
129	Max_diff_Gyr_x	Widest_Peak_Gyr_x			DSVM_Acc	Dfreq_Gyr_y
130	Kur_Gyr_y	Widest_Peak_Acc_z			Kur_Gyr_z	Kur_Gyr_x
131	Max_diff_Roll	Kur_Acc_x			Max_Gyr_y	Highest_peak_Acc_y
132	Kur_Gyr_x	nPeaks_Acc_x			Std_Acc_x	Widest_Peak_Gyr_z
133	Std_Azimuth	Highest_peak_Pitch			Std_Acc_y	SMA_Angle
134	Kur_Acc_x	Interq_Azimuth			Std_Acc_z	DSVM_Gyr
135	Var_Pitch	Var_Acc_z			Std_Azimuth	Min_Azimuth
136	Mean_Gyr_x	Kur_Roll			Std_Pitch	Avr_peak_time_Pitch
137	Widest_Peak_Pitch	Var_Azimuth			Std_Roll	Vedb_Gyr
138	nPeaks_Azimuth	Cf_Azimuth			Std_Gyr_x	Avr_peak_time_Acc_y
139	Kur_Roll	Avr_peak_time_Gyr_y			Std_Gyr_y	Kur_Pitch
140	Var_Acc_z	Avr_peak_time_Gyr_x			Std_Gyr_z	mag_Acc
141	Widest_Peak_Acc_x	Mean_Gyr_x			Energy_Acc_x	Skew_Gyr_z

142	Skew_Roll	Var_Pitch			Energy_Acc_y	Widest_Peak_Roll
143	Skew_Pitch	nPeaks_Gyr_x			Energy_Acc_z	SMA_Gyr
144	Avr_peak_time_Acc_z	nPeaks_Acc_z			Energy_Azimuth	Std_Acc_x
145	Entropy_TimeD_Ang	Kur_Gyr_z			Energy_Pitch	Std_Acc_y
146	SVM_Angle	Skew_Azimuth			Energy_Roll	Std_Acc_z
147	Kur_Pitch	Highest_peak_Azimuth			Energy_Gyr_x	Std_Azimuth
148	SMA_Angle	Widest_Peak_Azimuth			Energy_Gyr_y	Std_Pitch
149	Kur_Gyr_z	SMA_Angle			Energy_Gyr_z	Std_Roll
150	Highest_peak_Pitch	Entropy_TimeD_Ang			Dfreq_Acc_z	Std_Gyr_x
151	Var_Azimuth	SVM_Angle			Dfreq_Azimuth	Std_Gyr_y
152	Energy_Azimuth	nPeaks_Gyr_y			Entropy_Azimuth	Std_Gyr_z
153	Interq_Azimuth	Max_diff_Azimuth			Kur_Pitch	Energy_Acc_x
154	Min_Azimuth	Widest_Peak_Roll			Max_diff_Roll	Energy_Acc_y
155	nPeaks_Acc_y	Min_Azimuth			Dfreq_Acc_y	Energy_Acc_z
156	Dfreq_Azimuth	Avr_peak_time_Roll			Kur_Acc_x	Energy_Azimuth
157	Mean_Azimuth	Kur_Pitch			Kur_Acc_z	Energy_Pitch
158	Vedb_Angle	Dfreq_Azimuth			Skew_Roll	Energy_Roll
159	mag_Gyr	Mean_Azimuth			Entropy_TimeD_Ang	Energy_Gyr_x
160	nPeaks_Gyr_y	Avr_peak_time_Acc_x			mag_Acc	Energy_Gyr_y
161	Avr_peak_time_Azimuth	Energy_Azimuth			Max_Gyr_x	Energy_Gyr_z
162	Avr_peak_time_Roll	Widest_Peak_Pitch			Widest_Peak_Acc_x	Dfreq_Acc_z
163	Widest_Peak_Gyr_y	mag_Gyr			Rms_Gyr_z	Dfreq_Azimuth
164	Rms_Azimuth	nPeaks_Gyr_z			Widest_Peak_Acc_y	Highest_peak_Gyr_x
165	nPeaks_Roll	nPeaks_Acc_y			nPeaks_Azimuth	nPeaks_Acc_y
166	Skew_Azimuth	Avr_peak_time_Acc_y			Highest_peak_Azimuth	Skew_Gyr_y
167	Widest_Peak_Acc_y	Vedb_Angle			Skew_Azimuth	Avr_peak_time_Acc_z
168	Highest_peak_Azimuth	Avr_peak_time_Acc_z			Kur_Gyr_y	Avr_peak_time_Azimuth
169	Cf_Azimuth	Avr_peak_time_Azimuth			Highest_peak_Pitch	DSVM_Acc
170	Avr_peak_time_Acc_y	mag_Ang			Avr_peak_time_Roll	Max_diff_Acc_z

171	Avr_peak_time_Pitch	Rms_Azimuth			Avr_peak_time_Azimuth	Entropy_Gyr_x
172	Avr_peak_time_Gyr_y	Kur_Gyr_y			nPeaks_Roll	nPeaks_Gyr_x
173	Avr_peak_time_Gyr_z	Widest_Peak_Gyr_y			SVM_Angle	Kur_Roll
174	Avr_peak_time_Gyr_x	Avr_peak_time_Gyr_z			Skew_Acc_z	Widest_Peak_Gyr_y
175	Kur_Azimuth	Kur_Azimuth			Cf_Gyr_y	SVM_Gyr
176	Entropy_Azimuth	Widest_Peak_Acc_y			mag_Gyr	Widest_Peak_Azimuth
177	nPeaks_Gyr_z	Avr_peak_time_Pitch			nPeaks_Pitch	Kur_Azimuth
178	mag_Acc	Max_Azimuth			Avr_peak_time_Gyr_y	nPeaks_Acc_z
179	mag_Ang	nPeaks_Roll			Skew_Pitch	nPeaks_Pitch
180	nPeaks_Gyr_x	mag_Acc			SVM_Gyr	mag_Gyr
181	Widest_Peak_Gyr_z	nPeaks_Pitch			nPeaks_Gyr_y	Widest_Peak_Acc_z
182	Max_Azimuth	nPeaks_Azimuth			Skew_Acc_x	Skew_Acc_x
183	nPeaks_Pitch	Widest_Peak_Acc_x			Avr_peak_time_Gyr_x	nPeaks_Roll

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 10 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet3_all** over 10 $sec.window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	nPeaks_Gyr_z	Mean_Roll	Mean_Acc_x	Mean_Acc_x	Mean_Pitch	Cf_Pitch
2	Var_Pitch	Cf_Pitch	Mean_Pitch	Mean_Azimuth	Skew_Acc_z	Mean_Acc_x
3	Mean_Roll	Mean_Acc_x	Mean_Roll	Mean_Roll	Entropy_TimeD_Gyr	Var_Acc_y
4	Mean_Pitch	Max_Pitch	Mean_Gyr_x	Mean_Gyr_x	Highest_peak_Azimuth	Min_Acc_y
5	Std_Pitch	Std_Pitch	Var_Azimuth	Mean_Gyr_y	Std_Gyr_x	Mean_Roll
6	Entropy_Roll	Var_Pitch	Var_Gyr_x	Mean_Gyr_z	Std_Acc_z	Var_Pitch
7	Mean_Acc_y	Cf_Roll	Var_Gyr_z	Var_Acc_x	Energy_Gyr_z	Var_Gyr_x
8	Dfreq_Roll	Min_Roll	Std_Azimuth	Var_Pitch	Std_Acc_x	Cf_Azimuth
9	Rms_Roll	Min_Acc_y	Std_Pitch	Var_Gyr_z	nPeaks_Acc_z	Min_Acc_z
10	Widest_Peak_Gyr_x	Entropy_Pitch	Std_Roll	Std_Acc_x	Energy_Roll	Mean_Gyr_y
11	Cf_Pitch	Max_diff_Azimuth	Std_Gyr_x	Std_Acc_z	Std_Gyr_z	Mean_Gyr_z
12	Mean_Acc_x	Max_Roll	Std_Gyr_y	Std_Azimuth	Max_Gyr_y	Max_Pitch
13	Max_diff_Azimuth	Interq_Pitch	Std_Gyr_z	Std_Pitch	Entropy_Roll	Interq_Acc_z
14	Max_Pitch	Mean_Acc_y	Kur_Acc_x	Std_Roll	Max_diff_Azimuth	Dfreq_Roll
15	Avr_peak_time_Gyr_z	Mean_Pitch	Kur_Acc_y	Std_Gyr_x	Min_Pitch	Min_Acc_x
16	Cf_Roll	Highest_peak_Gyr_x	Kur_Azimuth	Std_Gyr_y	Skew_Azimuth	Max_Acc_y
17	Min_Roll	Dfreq_Gyr_x	Kur_Pitch	Std_Gyr_z	Std_Azimuth	Min_Gyr_y
18	Skew_Acc_y	Interq_Acc_y	Kur_Roll	Kur_Acc_x	AV_Ang	Highest_peak_Acc_z
19	Energy_Roll	Dfreq_Roll	Kur_Gyr_y	Kur_Acc_z	Kur_Gyr_x	Min_Roll
20	Dfreq_Acc_x	Var_Acc_y	Skew_Acc_y	Kur_Roll	Vedb_Angle	Vedb_Acc
21	Min_Acc_x	Entropy_Gyr_x	Skew_Pitch	Kur_Gyr_x	Max_Pitch	Max_Acc_x
22	Interq_Pitch	Rms_Roll	Skew_Roll	Kur_Gyr_z	Rms_Acc_x	Dfreq_Pitch
23	Highest_peak_Pitch	Std_Acc_y	Skew_Gyr_z	Skew_Acc_x	Energy_Acc_x	Kur_Acc_x
24	Skew_Acc_z	Mean_Gyr_y	Min_Acc_x	Skew_Acc_y	Rms_Roll	Mean_Gyr_x
25	Var_Acc_y	Mean_Gyr_x	Min_Azimuth	Skew_Acc_z	Cf_Acc_z	Max_diff_Acc_y

26	Min_Pitch	Rms_Gyr_y	Min_Gyr_y	Skew_Azimuth	SVM_Gyr	Dfreq_Acc_x
27	Highest_peak_Acc_y	Max_diff_Roll	Min_Gyr_z	Skew_Pitch	Min_Gyr_x	Skew_Acc_z
28	Kur_Gyr_x	Max_diff_Gyr_y	Max_Acc_x	Skew_Roll	Rms_Gyr_z	mag_Acc
29	Min_Acc_z	Var_Gyr_x	Max_Azimuth	Skew_Gyr_x	Dfreq_Roll	SMA_Acc
30	Energy_Acc_y	Min_Acc_x	Max_Gyr_x	Skew_Gyr_y	DSVM_Gyr	Entropy_Pitch
31	Skew_Acc_x	Energy_Gyr_x	Max_Gyr_y	Skew_Gyr_z	Skew_Acc_x	Rms_Acc_x
32	Interq_Acc_x	Entropy_Roll	Rms_Azimuth	Min_Acc_x	Kur_Gyr_y	MV_Acc
33	Rms_Acc_x	Energy_Roll	Rms_Pitch	Min_Acc_y	Std_Pitch	Entropy_Acc_y
34	Max_diff_Gyr_x	Skew_Acc_z	Rms_Roll	Min_Azimuth	SVM_Angle	Min_Azimuth
35	Rms_Acc_y	Dfreq_Pitch	Rms_Gyr_x	Min_Pitch	Widest_Peak_Gyr_z	Dfreq_Gyr_x
36	Std_Acc_y	Min_Pitch	Rms_Gyr_z	Min_Roll	Kur_Acc_y	Widest_Peak_Gyr_z
37	Max_diff_Gyr_y	Std_Gyr_y	Interq_Acc_x	Min_Gyr_y	Cf_Roll	Max_Gyr_z
38	Cf_Acc_x	Mean_Gyr_z	Interq_Acc_y	Max_Acc_x	Entropy_Gyr_z	Interq_Gyr_y
39	nPeaks_Gyr_y	Std_Gyr_x	Interq_Acc_z	Max_Acc_z	Avr_peak_time_Azimuth	Kur_Pitch
40	Min_Acc_y	Rms_Gyr_x	Interq_Gyr_y	Max_Pitch	Std_Roll	Avr_peak_time_Gyr_x
41	nPeaks_Pitch	Max_Acc_z	Interq_Gyr_z	Max_Roll	DSAM_Angle	Kur_Acc_y
42	Interq_Gyr_y	Energy_Gyr_y	Cf_Gyr_x	Max_Gyr_z	nPeaks_Roll	Skew_Pitch
43	Highest_peak_Roll	Max_Acc_x	Cf_Gyr_y	Rms_Acc_y	Interq_Azimuth	Max_Acc_z
44	Widest_Peak_Gyr_z	Cf_Acc_x	SMA_Acc	Rms_Acc_z	Mean_Gyr_z	SVM_Acc
45	Entropy_Gyr_z	Min_Acc_z	SMA_Angle	Rms_Azimuth	Std_Acc_y	Highest_peak_Acc_y
46	Interq_Acc_y	Max_diff_Gyr_x	Entropy_Acc_x	Rms_Pitch	Var_Gyr_z	Interq_Azimuth
47	Var_Gyr_z	Highest_peak_Pitch	Entropy_Acc_y	Rms_Roll	Min_Acc_x	Entropy_Gyr_x
48	Kur_Acc_x	Entropy_Acc_y	Entropy_Acc_z	Rms_Gyr_x	Energy_Acc_z	Max_diff_Acc_x
49	Std_Gyr_y	Skew_Acc_x	Entropy_Azimuth	Rms_Gyr_y	Vedb_Acc	Avr_peak_time_Pitch
50	Cf_Gyr_x	Var_Gyr_y	Entropy_Roll	Interq_Acc_y	Dfreq_Acc_z	Cf_Acc_y
51	Kur_Roll	Energy_Pitch	Entropy_Gyr_y	Interq_Acc_z	Dfreq_Acc_x	Interq_Gyr_z
52	nPeaks_Roll	Rms_Pitch	Entropy_Gyr_z	Interq_Gyr_y	Max_diff_Gyr_x	Max_Gyr_y
53	Max_Roll	Interq_Azimuth	Entropy_TimeD_Ang	Interq_Gyr_z	Energy_Azimuth	Rms_Acc_z
54	Std_Gyr_z	Energy_Acc_z	Entropy_TimeD_Gyr	Cf_Acc_z	Skew_Gyr_y	MV_Gyr

55	Rms_Gyr_y	Dfreq_Acc_y	Energy_Acc_y	Cf_Azimuth	nPeaks_Acc_y	Min_Gyr_x
56	Dfreq_Acc_z	Max_Gyr_y	Energy_Acc_z	Cf_Pitch	Mean_Azimuth	Skew_Acc_x
57	Mean_Acc_z	Rms_Acc_z	Energy_Pitch	Cf_Roll	Skew_Acc_y	Max_diff_Gyr_z
58	Std_Roll	Rms_Acc_x	Energy_Roll	Cf_Gyr_y	Max_Roll	Max_diff_Gyr_x
59	Var_Gyr_y	Max_Gyr_z	Energy_Gyr_x	SMA_Angle	nPeaks_Gyr_y	Var_Roll
60	Entropy_Pitch	Max_diff_Pitch	SVM_Acc	SMA_Gyr	SMA_Gyr	Interq_Acc_y
61	Entropy_Gyr_x	Var_Acc_z	SVM_Angle	Entropy_Pitch	SMA_Acc	Max_Roll
62	Highest_peak_Acc_x	Interq_Gyr_y	SVM_Gyr	Entropy_Gyr_x	Widest_Peak_Pitch	Interq_Pitch
63	Energy_Gyr_z	Min_Azimuth	Max_diff_Acc_y	Entropy_Gyr_y	Kur_Gyr_z	Cf_Roll
64	Energy_Gyr_y	Entropy_TimeD_Acc	Max_diff_Azimuth	Entropy_TimeD_Acc	Min_Gyr_z	Dfreq_Gyr_z
65	Max_diff_Roll	SVM_Acc	Max_diff_Gyr_x	Entropy_TimeD_Gyr	Rms_Azimuth	Highest_peak_Gyr_x
66	Rms_Gyr_z	Kur_Acc_x	Max_diff_Gyr_y	Energy_Acc_x	Widest_Peak_Acc_x	Highest_peak_Gyr_y
67	Widest_Peak_Roll	Highest_peak_Gyr_y	MV_Gyr	Energy_Azimuth	Vedb_Gyr	Rms_Pitch
68	Max_Gyr_z	Dfreq_Acc_x	mag_Ang	Energy_Pitch	SMA_Angle	Mean_Acc_y
69	Highest_peak_Gyr_x	Min_Gyr_x	mag_Gyr	Energy_Gyr_x	Avr_peak_time_Acc_y	Mean_Pitch
70	Widest_Peak_Gyr_y	Entropy_Acc_z	Vedb_Gyr	Energy_Gyr_y	Widest_Peak_Azimuth	Skew_Roll
71	Var_Roll	Dfreq_Gyr_y	Dfreq_Acc_y	SVM_Acc	Entropy_Acc_z	Var_Gyr_y
72	Energy_Pitch	Interq_Acc_x	Dfreq_Acc_z	DSAM_Angle	Skew_Roll	Entropy_Acc_x
73	Max_diff_Pitch	Highest_peak_Acc_x	Dfreq_Gyr_x	DSVM_Gyr	Dfreq_Gyr_x	Widest_Peak_Gyr_y
74	Rms_Pitch	Std_Acc_z	Dfreq_Gyr_z	Max_diff_Acc_y	Dfreq_Gyr_z	Std_Acc_x
75	Skew_Gyr_x	Dfreq_Acc_z	Widest_Peak_Acc_x	Max_diff_Roll	Dfreq_Pitch	Std_Acc_y
76	Dfreq_Gyr_x	Mean_Acc_z	Highest_peak_Acc_x	Max_diff_Gyr_y	Kur_Acc_z	Std_Acc_z
77	Mean_Gyr_z	Energy_Acc_y	nPeaks_Acc_y	MV_Acc	Max_Acc_y	Std_Azimuth
78	Min_Azimuth	Dfreq_Gyr_z	Widest_Peak_Acc_y	AV_Ang	Cf_Gyr_x	Std_Pitch
79	Dfreq_Gyr_z	Var_Azimuth	Highest_peak_Acc_y	Vedb_Acc	mag_Gyr	Std_Roll
80	Dfreq_Pitch	SMA_Acc	nPeaks_Acc_z	Vedb_Gyr	Interq_Gyr_x	Std_Gyr_x
81	Avr_peak_time_Pitch	Rms_Acc_y	Highest_peak_Acc_z	Dfreq_Acc_x	Highest_peak_Acc_y	Std_Gyr_y
82	Entropy_Acc_z	Highest_peak_Roll	Avr_peak_time_Acc_z	Dfreq_Acc_y	nPeaks_Gyr_x	Std_Gyr_z
83	SMA_Acc	Highest_peak_Acc_z	Widest_Peak_Azimuth	Dfreq_Acc_z	Interq_Roll	Kur_Azimuth

84	Highest_peak_Gyr_z	Avr_peak_time_Azimuth	Widest_Peak_Pitch	Dfreq_Pitch	Avr_peak_time_Gyr_x	Rms_Gyr_x
85	Avr_peak_time_Gyr_y	Energy_Azimuth	Widest_Peak_Roll	Dfreq_Roll	Max_diff_Roll	Interq_Gyr_x
86	Skew_Azimuth	Cf_Acc_z	Avr_peak_time_Roll	Dfreq_Gyr_x	Highest_peak_Roll	Cf_Acc_x
87	Energy_Acc_x	MV_Acc	nPeaks_Gyr_y	Dfreq_Gyr_y	Max_diff_Acc_x	Cf_Gyr_x
88	Rms_Acc_z	Max_diff_Acc_z	Highest_peak_Gyr_y	Dfreq_Gyr_z	Cf_Pitch	Cf_Gyr_y
89	Kur_Acc_z	Min_Gyr_y	Widest_Peak_Gyr_z	nPeaks_Acc_x	Entropy_TimeD_Acc	Cf_Gyr_z
90	AV_Ang	Highest_peak_Acc_y		Widest_Peak_Acc_x	Mean_Acc_x	SMA_Angle
91	Entropy_Acc_x	Max_Azimuth		Highest_peak_Acc_x	Energy_Gyr_x	SMA_Gyr
92	Max_Acc_x	Max_diff_Gyr_z		Avr_peak_time_Acc_x	Widest_Peak_Acc_y	Entropy_Azimuth
93	Highest_peak_Gyr_y	Cf_Gyr_z		nPeaks_Acc_y	Min_Acc_z	Entropy_Roll
94	Highest_peak_Azimuth	Min_Gyr_z		Widest_Peak_Acc_y	Rms_Gyr_x	Entropy_TimeD_Ang
95	Mean_Gyr_y	Rms_Gyr_z		Widest_Peak_Acc_z	Rms_Acc_y	Entropy_TimeD_Gyr
96	DSAM_Angle	nPeaks_Pitch		Avr_peak_time_Acc_z	Dfreq_Acc_y	Energy_Acc_x
97	Energy_Acc_z	Energy_Acc_x		nPeaks_Azimuth	Mean_Acc_z	Energy_Acc_y
98	Avr_peak_time_Roll	DSVM_Acc		Widest_Peak_Azimuth	Interq_Acc_y	Energy_Acc_z
99	Max_Gyr_y	Entropy_Gyr_z		Avr_peak_time_Azimuth	Dfreq_Gyr_y	Energy_Azimuth
100	Cf_Gyr_y	Std_Gyr_z		nPeaks_Pitch	Energy_Gyr_y	Energy_Pitch
101	Avr_peak_time_Gyr_x	Avr_peak_time_Pitch		Highest_peak_Pitch	Rms_Acc_z	Energy_Roll
102	Widest_Peak_Acc_x	Avr_peak_time_Acc_y		Highest_peak_Roll	Min_Acc_y	Energy_Gyr_x
103	Max_diff_Acc_x	Skew_Acc_y		Avr_peak_time_Roll	Max_diff_Gyr_y	Energy_Gyr_y
104	Dfreq_Gyr_y	Skew_Gyr_y		nPeaks_Gyr_x	Interq_Acc_x	Energy_Gyr_z
105	SVM_Acc	Std_Azimuth		Widest_Peak_Gyr_x	Energy_Pitch	SVM_Angle
106	Entropy_TimeD_Acc	Cf_Acc_y		Highest_peak_Gyr_x	Var_Azimuth	SVM_Gyr
107	Std_Acc_x	Vedb_Angle		Highest_peak_Gyr_y	Entropy_Gyr_x	DSVM_Acc
108	Kur_Gyr_y	nPeaks_Acc_y		nPeaks_Gyr_z	Var_Acc_y	DSVM_Gyr
109	DSVM_Gyr	DSVM_Gyr		Avr_peak_time_Gyr_z	Kur_Acc_x	Max_diff_Acc_z
110	Mean_Gyr_x	Rms_Azimuth			Entropy_Acc_y	Max_diff_Gyr_y
111	Interq_Roll	Kur_Gyr_z			Max_Gyr_z	AV_Ang
112	Max_diff_Gyr_z	Std_Roll			Highest_peak_Pitch	mag_Ang

113	Avr_peak_time_Acc_y	Energy_Gyr_z			Widest_Peak_Gyr_y	Vedb_Angle
114	Skew_Roll	mag_Ang			nPeaks_Gyr_z	Vedb_Gyr
115	Entropy_Azimuth	Skew_Roll			Std_Gyr_y	Dfreq_Acc_z
116	Var_Acc_z	Interq_Gyr_x			Entropy_TimeD_Ang	Dfreq_Azimuth
117	mag_Gyr	Cf_Gyr_y			Energy_Acc_y	Dfreq_Gyr_y
118	Cf_Gyr_z	Avr_peak_time_Acc_z			Mean_Roll	nPeaks_Acc_x
119	Max_Acc_y	nPeaks_Gyr_x			Dfreq_Azimuth	Avr_peak_time_Acc_y
120	Std_Acc_z	Var_Gyr_z			Var_Pitch	nPeaks_Acc_z
121	Var_Gyr_x	MV_Gyr			nPeaks_Azimuth	Widest_Peak_Acc_z
122	Max_Gyr_x	SVM_Gyr			Cf_Azimuth	Avr_peak_time_Acc_z
123	Max_Azimuth	mag_Acc			Interq_Gyr_z	nPeaks_Azimuth
124	Cf_Acc_z	Max_Acc_y			Max_diff_Gyr_z	Widest_Peak_Azimuth
125	Energy_Gyr_x	SMA_Gyr			SVM_Acc	Avr_peak_time_Azimuth
126	Max_diff_Acc_z	Vedb_Acc			Interq_Gyr_y	Widest_Peak_Pitch
127	nPeaks_Gyr_x	Widest_Peak_Azimuth			Mean_Acc_y	nPeaks_Roll
128	Kur_Gyr_z	Skew_Gyr_z			Avr_peak_time_Acc_x	Avr_peak_time_Roll
129	Skew_Gyr_y	Var_Roll			Var_Gyr_y	nPeaks_Gyr_x
130	Entropy_Gyr_y	Skew_Gyr_x			Kur_Pitch	nPeaks_Gyr_y
131	Energy_Azimuth	Kur_Roll			Mean_Gyr_x	Avr_peak_time_Gyr_y
132	Interq_Acc_z	Avr_peak_time_Roll			Kur_Roll	nPeaks_Gyr_z
133	Var_Acc_x	nPeaks_Azimuth			Max_Acc_x	Highest_peak_Gyr_z
134	Cf_Acc_y	Std_Acc_x			Highest_peak_Acc_x	Avr_peak_time_Gyr_z
135	Max_diff_Acc_y	Entropy_TimeD_Gyr			Mean_Gyr_y	Cf_Acc_z
136	Avr_peak_time_Acc_x	Highest_peak_Azimuth			Min_Azimuth	Var_Gyr_z
137	Std_Gyr_x	Cf_Azimuth			Max_Acc_z	Min_Pitch
138	Cf_Azimuth	Skew_Pitch			Rms_Gyr_y	Highest_peak_Pitch
139	Interq_Gyr_x	nPeaks_Acc_x			Entropy_Azimuth	Skew_Azimuth
140	Rms_Gyr_x	Vedb_Gyr			Min_Gyr_y	nPeaks_Acc_y
141	Max_Acc_z	Var_Acc_x			Max_Azimuth	Max_diff_Azimuth

142	Dfreq_Acc_y	Entropy_Azimuth			Interq_Acc_z	Kur_Gyr_z
143	Min_Gyr_y	Skew_Azimuth			Var_Gyr_x	Interq_Roll
144	nPeaks_Acc_x	Interq_Gyr_z			Max_Gyr_x	Var_Azimuth
145	nPeaks_Acc_y	AV_Ang			Widest_Peak_Gyr_x	Max_Azimuth
146	Avr_peak_time_Azimuth	Highest_peak_Gyr_z			Skew_Gyr_x	Interq_Acc_x
147	Skew_Gyr_z	DSAM_Angle			Cf_Acc_x	Mean_Acc_z
148	Interq_Gyr_z	Max_diff_Acc_y			Kur_Azimuth	Min_Gyr_z
149	nPeaks_Acc_z	Entropy_Acc_x			Max_diff_Acc_z	Kur_Acc_z
150	Kur_Pitch	Avr_peak_time_Gyr_y			Highest_peak_Gyr_z	Max_diff_Pitch
151	Var_Azimuth	Interq_Acc_z			Highest_peak_Gyr_x	Mean_Azimuth
152	Skew_Pitch	Entropy_TimeD_Ang			Cf_Gyr_y	Widest_Peak_Acc_y
153	Min_Gyr_x	Cf_Gyr_x			Skew_Pitch	Max_diff_Roll
154	MV_Gyr	Max_diff_Acc_x			Avr_peak_time_Gyr_y	Skew_Acc_y
155	Min_Gyr_z	SMA_Angle			Highest_peak_Acc_z	Avr_peak_time_Acc_x
156	SVM_Gyr	Mean_Azimuth			Avr_peak_time_Gyr_z	Widest_Peak_Gyr_x
157	Avr_peak_time_Acc_z	Dfreq_Azimuth			Widest_Peak_Roll	Skew_Gyr_x
158	mag_Acc	SVM_Angle			Cf_Acc_y	Max_Gyr_x
159	Kur_Azimuth	Interq_Roll			Min_Roll	Rms_Acc_y
160	Entropy_TimeD_Gyr	Widest_Peak_Pitch			Highest_peak_Gyr_y	nPeaks_Pitch
161	Widest_Peak_Acc_z	Kur_Acc_z			Entropy_Gyr_y	Entropy_Gyr_z
162	Vedb_Gyr	Kur_Gyr_y			DSVM_Acc	Entropy_Acc_z
163	SMA_Gyr	Max_Gyr_x			Entropy_Acc_x	Rms_Gyr_z
164	Interq_Azimuth	Kur_Gyr_x			Cf_Gyr_z	Dfreq_Acc_y
165	Highest_peak_Acc_z	Avr_peak_time_Acc_x			Avr_peak_time_Acc_z	Widest_Peak_Roll
166	Std_Azimuth	Kur_Pitch			nPeaks_Acc_x	Highest_peak_Azimuth
167	Vedb_Angle	Widest_Peak_Acc_y			mag_Acc	DSAM_Angle
168	nPeaks_Azimuth	Avr_peak_time_Gyr_z			Avr_peak_time_Roll	Kur_Roll
169	Rms_Azimuth	Avr_peak_time_Gyr_x			Widest_Peak_Acc_z	Rms_Gyr_y
170	Dfreq_Azimuth	Kur_Azimuth			MV_Acc	Highest_peak_Roll

171	Mean_Azimuth	Widest_Peak_Acc_x			Entropy_Pitch	Rms_Roll
172	Entropy_TimeD_Ang	nPeaks_Gyr_y			MV_Gyr	Highest_peak_Acc_x
173	SMA_Angle	nPeaks_Acc_z			Var_Acc_z	Entropy_Gyr_y
174	SVM_Angle	mag_Gyr			Avr_peak_time_Pitch	Entropy_TimeD_Acc
175	Vedb_Acc	nPeaks_Gyr_z			Max_diff_Pitch	Var_Acc_z
176	Entropy_Acc_y	Widest_Peak_Gyr_z			mag_Ang	Skew_Gyr_z
177	Kur_Acc_y	Kur_Acc_y			Interq_Pitch	mag_Gyr
178	Widest_Peak_Acc_y	Widest_Peak_Acc_z			Var_Acc_x	Skew_Gyr_y
179	Widest_Peak_Pitch	Widest_Peak_Gyr_x			nPeaks_Pitch	Kur_Gyr_y
180	mag_Ang	Entropy_Gyr_y			Rms_Pitch	Rms_Azimuth
181	DSVM_Acc	Widest_Peak_Roll			Max_diff_Acc_y	Var_Acc_x
182	Widest_Peak_Azimuth	Widest_Peak_Gyr_y			Var_Roll	Widest_Peak_Acc_x
183	MV_Acc	nPeaks_Roll			Skew_Gyr_z	Kur_Gyr_x

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 11 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet3_all** over 7 $sec. window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Roll	Mean_Roll	Mean_Acc_x	Mean_Acc_x	Dfreq_Roll	Rms_Roll
2	Rms_Roll	Mean_Acc_x	Mean_Pitch	Mean_Acc_y	Mean_Pitch	Dfreq_Roll
3	Mean_Acc_x	Cf_Roll	Var_Acc_x	Mean_Acc_z	Mean_Acc_x	Max_Pitch
4	Cf_Roll	Rms_Roll	Var_Acc_y	Mean_Pitch	Cf_Pitch	Mean_Roll
5	Dfreq_Roll	Std_Gyr_y	Var_Acc_z	Mean_Roll	Min_Gyr_y	Max_diff_Pitch
6	Energy_Roll	Max_Roll	Var_Gyr_x	Mean_Gyr_x	Min_Acc_x	Min_Gyr_y
7	Mean_Gyr_z	Rms_Gyr_y	Std_Acc_x	Var_Azimuth	Min_Azimuth	Mean_Gyr_z
8	Std_Gyr_y	Mean_Gyr_z	Std_Acc_z	Var_Pitch	Interq_Gyr_y	Highest_peak_Gyr_y
9	Entropy_Roll	Dfreq_Roll	Std_Pitch	Var_Gyr_y	Mean_Roll	Min_Acc_x
10	Rms_Gyr_y	Std_Pitch	Std_Roll	Std_Acc_x	Rms_Roll	Min_Pitch
11	Max_Roll	Max_diff_Gyr_y	Std_Gyr_y	Std_Acc_y	Mean_Gyr_y	Interq_Gyr_y
12	Max_diff_Gyr_y	Min_Roll	Kur_Acc_x	Std_Acc_z	Var_Acc_y	Min_Roll
13	Var_Gyr_y	Var_Gyr_y	Kur_Azimuth	Std_Azimuth	Min_Gyr_z	Mean_Acc_x
14	Min_Roll	Energy_Roll	Kur_Pitch	Std_Pitch	Entropy_Roll	Interq_Gyr_x
15	Energy_Gyr_y	Var_Pitch	Kur_Roll	Std_Roll	Max_diff_Pitch	Max_diff_Gyr_x
16	Min_Acc_x	Entropy_Roll	Kur_Gyr_y	Kur_Acc_x	mag_Ang	Mean_Gyr_y
17	Interq_Gyr_y	Energy_Gyr_y	Skew_Acc_x	Kur_Azimuth	Var_Acc_z	Mean_Acc_z
18	Interq_Gyr_x	Rms_Acc_x	Skew_Azimuth	Kur_Gyr_y	Max_diff_Gyr_y	Var_Pitch
19	Min_Pitch	Min_Acc_y	Skew_Gyr_x	Kur_Gyr_z	Interq_Pitch	Var_Gyr_y
20	Highest_peak_Gyr_y	Dfreq_Gyr_y	Skew_Gyr_y	Skew_Acc_x	Interq_Gyr_x	Max_Acc_y
21	mag_Ang	Interq_Gyr_y	Min_Gyr_x	Skew_Pitch	Max_Acc_z	Entropy_Gyr_x
22	Max_Acc_y	Highest_peak_Gyr_y	Min_Gyr_y	Skew_Gyr_x	Min_Pitch	Cf_Gyr_x
23	Skew_Acc_y	nPeaks_Azimuth	Min_Gyr_z	Min_Roll	Min_Acc_z	Kur_Gyr_x
24	Entropy_Pitch	Max_Pitch	Max_Acc_x	Min_Gyr_x	Cf_Gyr_z	Max_Gyr_y
25	Rms_Acc_x	Cf_Pitch	Max_Azimuth	Min_Gyr_y	Rms_Gyr_y	Dfreq_Gyr_y

26	Cf_Pitch	Std_Gyr_z	Max_Pitch	Min_Gyr_z	Skew_Acc_z	Rms_Gyr_y
27	Entropy_Gyr_z	Min_Gyr_z	Rms_Acc_y	Max_Acc_x	Mean_Gyr_z	DSVM_Acc
28	Highest_peak_Gyr_x	Std_Roll	Rms_Acc_z	Max_Acc_z	Skew_Acc_x	Cf_Pitch
29	Var_Gyr_x	Cf_Acc_x	Rms_Azimuth	Max_Azimuth	Dfreq_Gyr_x	Dfreq_Acc_x
30	Energy_Gyr_x	Dfreq_Acc_x	Rms_Gyr_y	Max_Pitch	Widest_Peak_Gyr_x	Var_Roll
31	Dfreq_Gyr_x	Rms_Gyr_z	Rms_Gyr_z	Max_Roll	Mean_Acc_z	Entropy_Acc_y
32	Std_Gyr_x	Min_Pitch	Interq_Azimuth	Max_Gyr_z	Entropy_TimeD_Gyr	Var_Gyr_z
33	Dfreq_Gyr_y	Min_Acc_x	Interq_Pitch	Rms_Azimuth	Highest_peak_Roll	Avr_peak_time_Pitch
34	Rms_Gyr_x	Entropy_Pitch	Interq_Roll	Rms_Gyr_x	Kur_Acc_y	Max_diff_Gyr_y
35	Min_Acc_z	Interq_Pitch	Cf_Acc_x	Interq_Acc_x	Highest_peak_Gyr_y	Rms_Acc_x
36	Max_Pitch	Max_Acc_y	Cf_Azimuth	Interq_Acc_y	Mean_Gyr_x	Entropy_Pitch
37	Max_Acc_x	Max_diff_Pitch	Cf_Pitch	Interq_Acc_z	Min_Roll	Cf_Roll
38	Min_Acc_y	Energy_Acc_x	Cf_Roll	Interq_Azimuth	Var_Gyr_y	Var_Gyr_x
39	Std_Roll	Std_Acc_x	Cf_Gyr_x	Interq_Gyr_y	Max_diff_Azimuth	nPeaks_Gyr_x
40	Max_diff_Gyr_z	Interq_Gyr_z	Cf_Gyr_z	Interq_Gyr_z	Avr_peak_time_Acc_y	Highest_peak_Roll
41	Max_Gyr_y	Highest_peak_Roll	SMA_Acc	Cf_Acc_z	Max_diff_Gyr_z	AV_Ang
42	Max_diff_Gyr_x	Var_Gyr_z	Entropy_Acc_y	Cf_Pitch	Interq_Azimuth	Max_diff_Roll
43	Cf_Acc_x	Highest_peak_Acc_x	Entropy_Acc_z	Cf_Gyr_x	Entropy_TimeD_Ang	Entropy_Roll
44	Min_Gyr_y	Max_diff_Acc_y	Entropy_Azimuth	Cf_Gyr_z	Highest_peak_Pitch	Min_Gyr_x
45	Mean_Acc_y	Max_diff_Gyr_z	Entropy_Roll	Entropy_Acc_z	Max_Gyr_y	DSAM_Angle
46	Skew_Acc_z	Energy_Gyr_z	Entropy_Gyr_z	Entropy_Azimuth	Vedb_Acc	Mean_Acc_y
47	Std_Acc_z	Entropy_Gyr_z	Entropy_TimeD_Acc	Entropy_Roll	Cf_Acc_y	Var_Acc_x
48	Std_Gyr_z	Max_Gyr_y	Energy_Acc_y	Entropy_Gyr_z	MV_Acc	Kur_Roll
49	Mean_Pitch	Dfreq_Acc_z	Energy_Azimuth	Entropy_TimeD_Acc	nPeaks_Roll	Var_Acc_z
50	Interq_Gyr_z	Mean_Acc_z	Energy_Pitch	Energy_Acc_z	Avr_peak_time_Gyr_x	Highest_peak_Gyr_z
51	Energy_Acc_x	Interq_Roll	Energy_Roll	Energy_Roll	Widest_Peak_Acc_y	Rms_Azimuth
52	Cf_Gyr_y	DSVM_Acc	Energy_Gyr_y	Energy_Gyr_y	Skew_Azimuth	Interq_Gyr_z
53	Std_Acc_x	Entropy_Gyr_x	Energy_Gyr_z	Energy_Gyr_z	Kur_Pitch	Highest_peak_Gyr_x
54	Kur_Roll	Max_diff_Gyr_x	SVM_Angle	SVM_Acc	Skew_Roll	mag_Ang

55	Dfreq_Acc_x	Min_Gyr_y	DSVM_Acc	SVM_Angle	Cf_Roll	Widest_Peak_Gyr_y
56	nPeaks_Azimuth	Max_Acc_x	DSAM_Angle	SVM_Gyr	Highest_peak_Acc_x	Widest_Peak_Gyr_z
57	Cf_Gyr_z	Min_Acc_z	Max_diff_Acc_z	Max_diff_Acc_y	Max_Acc_x	Avr_peak_time_Gyr_x
58	Var_Acc_z	Max_diff_Roll	Max_diff_Azimuth	Max_diff_Pitch	Interq_Gyr_z	mag_Acc
59	Avr_peak_time_Azimuth	Skew_Acc_y	Max_diff_Roll	Max_diff_Roll	Widest_Peak_Gyr_z	Widest_Peak_Azimuth
60	DSVM_Acc	Var_Gyr_x	Max_diff_Gyr_y	Max_diff_Gyr_x	Entropy_Pitch	Avr_peak_time_Acc_z
61	Entropy_TimeD_Gyr	Var_Acc_x	MV_Acc	Max_diff_Gyr_y	Entropy_Acc_x	Rms_Gyr_x
62	Rms_Gyr_z	Energy_Gyr_x	mag_Ang	MV_Gyr	Max_diff_Roll	Skew_Gyr_x
63	Highest_peak_Acc_x	mag_Ang	mag_Gyr	Vedb_Acc	Skew_Acc_y	nPeaks_Roll
64	Entropy_Acc_z	Cf_Gyr_y	Vedb_Angle	Vedb_Gyr	Min_Acc_y	Avr_peak_time_Acc_y
65	Interq_Roll	Widest_Peak_Pitch	Vedb_Gyr	Dfreq_Azimuth	Widest_Peak_Pitch	Dfreq_Pitch
66	Vedb_Gyr	Dfreq_Gyr_z	Dfreq_Acc_x	nPeaks_Acc_x	Kur_Acc_z	SMA_Gyr
67	Rms_Acc_y	Var_Roll	Dfreq_Gyr_x	Widest_Peak_Acc_x	Entropy_Gyr_z	Skew_Acc_y
68	MV_Gyr	Std_Gyr_x	Dfreq_Gyr_y	nPeaks_Acc_y	Max_diff_Gyr_x	Interq_Azimuth
69	Max_diff_Pitch	Rms_Gyr_x	Dfreq_Gyr_z	Widest_Peak_Acc_y	Var_Roll	Avr_peak_time_Roll
70	DSVM_Gyr	Entropy_Acc_z	nPeaks_Acc_x	Highest_peak_Acc_y	Interq_Acc_x	Interq_Roll
71	Highest_peak_Roll	Highest_peak_Pitch	Highest_peak_Acc_x	Avr_peak_time_Acc_y	Cf_Gyr_y	mag_Gyr
72	SMA_Gyr	Highest_peak_Gyr_z	Avr_peak_time_Acc_x	Avr_peak_time_Acc_z	Interq_Roll	Highest_peak_Acc_y
73	SVM_Gyr	Mean_Gyr_y	nPeaks_Acc_y	nPeaks_Azimuth	Cf_Acc_z	Max_diff_Azimuth
74	Std_Pitch	Highest_peak_Gyr_x	Highest_peak_Acc_y	Highest_peak_Azimuth	Max_diff_Acc_x	Kur_Gyr_y
75	Energy_Acc_y	Skew_Acc_z	Avr_peak_time_Acc_y	Widest_Peak_Pitch	Mean_Azimuth	Widest_Peak_Pitch
76	Max_diff_Roll	Rms_Acc_y	nPeaks_Acc_z	Avr_peak_time_Pitch	Cf_Acc_x	Entropy_Acc_x
77	Entropy_Gyr_x	Interq_Gyr_x	Widest_Peak_Acc_z	nPeaks_Gyr_x	AV_Ang	Widest_Peak_Gyr_x
78	Var_Acc_x	Energy_Acc_y	Avr_peak_time_Acc_z	Widest_Peak_Gyr_x	Widest_Peak_Gyr_y	Min_Gyr_z
79	Interq_Acc_z	Dfreq_Gyr_x	Highest_peak_Azimuth	Highest_peak_Gyr_x	Interq_Acc_y	Skew_Acc_z
80	Var_Gyr_z	Skew_Gyr_y	Avr_peak_time_Azimuth	Avr_peak_time_Gyr_x	SVM_Acc	nPeaks_Azimuth
81	Min_Gyr_x	MV_Acc	nPeaks_Pitch	nPeaks_Gyr_y	Std_Acc_x	Cf_Acc_x
82	Var_Roll	DSVM_Gyr	Highest_peak_Pitch	Highest_peak_Gyr_y	Std_Acc_y	Mean_Azimuth
83	Energy_Gyr_z	Dfreq_Pitch	Highest_peak_Roll	Widest_Peak_Gyr_z	Std_Acc_z	Kur_Pitch

84	Widest_Peak_Pitch	Dfreq_Acc_y	Avr_peak_time_Roll	Avr_peak_time_Gyr_z	Std_Azimuth	Entropy_Gyr_y
85	Kur_Gyr_z	Vedb_Acc	Widest_Peak_Gyr_x		Std_Pitch	Highest_peak_Pitch
86	Interq_Acc_x	Entropy_Acc_x	Highest_peak_Gyr_x		Std_Roll	Skew_Roll
87	Min_Gyr_z	Energy_Acc_z	Avr_peak_time_Gyr_x		Std_Gyr_x	MV_Acc
88	Avr_peak_time_Pitch	Rms_Acc_z	Widest_Peak_Gyr_z		Std_Gyr_y	Vedb_Gyr
89	Kur_Gyr_x	Interq_Acc_z	Highest_peak_Gyr_z		Std_Gyr_z	Avr_peak_time_Azimuth
90	Widest_Peak_Gyr_z	Vedb_Gyr			Max_Gyr_z	MV_Gyr
91	Vedb_Acc	Avr_peak_time_Azimuth			Rms_Gyr_x	SVM_Gyr
92	Max_Gyr_z	Min_Gyr_x			Cf_Azimuth	nPeaks_Pitch
93	Max_Azimuth	Energy_Pitch			SMA_Gyr	Highest_peak_Acc_x
94	MV_Acc	Mean_Acc_y			Entropy_Acc_y	Kur_Acc_x
95	Entropy_Gyr_y	Mean_Pitch			Entropy_Gyr_y	Cf_Gyr_y
96	Dfreq_Pitch	Skew_Gyr_x			Energy_Acc_x	Rms_Pitch
97	Dfreq_Acc_y	Max_diff_Acc_x			Energy_Acc_y	Entropy_Acc_z
98	Max_Acc_z	Max_Gyr_z			Energy_Acc_z	Min_Acc_y
99	Skew_Gyr_y	mag_Acc			Energy_Azimuth	Max_Roll
100	mag_Gyr	Interq_Acc_x			Energy_Pitch	Widest_Peak_Roll
101	Highest_peak_Pitch	Rms_Pitch			Energy_Roll	Avr_peak_time_Gyr_z
102	Var_Pitch	Entropy_Acc_y			Energy_Gyr_x	Max_Azimuth
103	Rms_Pitch	nPeaks_Gyr_z			Energy_Gyr_y	Mean_Gyr_x
104	Dfreq_Gyr_z	Entropy_TimeD_Gyr			Energy_Gyr_z	Min_Azimuth
105	Energy_Azimuth	Max_Acc_z			SVM_Gyr	Skew_Azimuth
106	Vedb_Angle	Cf_Gyr_x			DSAM_Angle	Dfreq_Gyr_x
107	Rms_Azimuth	Var_Acc_y			DSVM_Gyr	nPeaks_Gyr_z
108	Interq_Azimuth	Cf_Gyr_z			Vedb_Angle	Rms_Acc_y
109	Kur_Acc_y	Std_Acc_z			Vedb_Gyr	Interq_Acc_y
110	Widest_Peak_Gyr_y	Std_Acc_y			Dfreq_Acc_y	Cf_Azimuth
111	Cf_Acc_z	SVM_Gyr			Dfreq_Acc_z	Max_diff_Acc_y
112	Max_diff_Acc_y	Mean_Gyr_x			Dfreq_Azimuth	Max_Gyr_z

113	Energy_Pitch	Avr_peak_time_Roll			Dfreq_Pitch	Var_Acc_y
114	Entropy_TimeD_Ang	mag_Gyr			nPeaks_Acc_y	Max_Gyr_x
115	SVM_Angle	Widest_Peak_Acc_y			nPeaks_Acc_z	Highest_peak_Azimuth
116	Max_diff_Acc_x	Kur_Gyr_x			nPeaks_Pitch	Avr_peak_time_Acc_x
117	Mean_Azimuth	Kur_Roll			nPeaks_Gyr_y	Std_Acc_z
118	Dfreq_Azimuth	SMA_Gyr			nPeaks_Gyr_z	Skew_Pitch
119	Max_Gyr_x	Highest_peak_Acc_y			Avr_peak_time_Gyr_z	Highest_peak_Acc_z
120	Mean_Gyr_y	Var_Acc_z			DSVM_Acc	Min_Acc_z
121	Entropy_Acc_x	Avr_peak_time_Gyr_z			Rms_Pitch	Mean_Pitch
122	Skew_Acc_x	Kur_Acc_z			Var_Pitch	Rms_Gyr_z
123	SMA_Angle	Max_Gyr_x			Dfreq_Acc_x	Entropy_TimeD_Acc
124	nPeaks_Acc_y	Min_Azimuth			Dfreq_Gyr_z	Dfreq_Acc_y
125	nPeaks_Gyr_x	Cf_Acc_y			Max_Pitch	nPeaks_Acc_y
126	Cf_Gyr_x	Kur_Pitch			Var_Acc_x	Max_diff_Gyr_z
127	Widest_Peak_Acc_y	Entropy_TimeD_Acc			Min_Gyr_x	Max_diff_Acc_x
128	Interq_Pitch	Kur_Gyr_y			Avr_peak_time_Azimuth	Skew_Acc_x
129	Skew_Gyr_z	Cf_Acc_z			Max_Acc_y	Energy_Pitch
130	Highest_peak_Gyr_z	Vedb_Angle			Cf_Gyr_x	Cf_Acc_y
131	Kur_Acc_z	SMA_Acc			Interq_Acc_z	Std_Gyr_z
132	Widest_Peak_Gyr_x	Energy_Azimuth			Dfreq_Gyr_y	Widest_Peak_Acc_y
133	Mean_Gyr_x	Rms_Azimuth			mag_Gyr	Interq_Acc_x
134	Var_Acc_y	SVM_Acc			Entropy_Gyr_x	Max_Acc_z
135	Widest_Peak_Acc_z	Entropy_Gyr_y			Kur_Gyr_z	Energy_Acc_y
136	Kur_Gyr_y	Avr_peak_time_Pitch			Widest_Peak_Azimuth	Energy_Gyr_z
137	Highest_peak_Acc_z	Skew_Roll			SMA_Acc	Max_diff_Acc_z
138	Energy_Acc_z	Avr_peak_time_Acc_z			Var_Gyr_x	Energy_Azimuth
139	Mean_Acc_z	Kur_Acc_y			Var_Azimuth	Var_Azimuth
140	Dfreq_Acc_z	Entropy_TimeD_Ang			Highest_peak_Acc_z	Kur_Acc_z
141	Cf_Acc_y	Highest_peak_Acc_z			Entropy_Acc_z	Interq_Acc_z

142	Avr_peak_time_Gyr_x	Cf_Azimuth			Kur_Azimuth	Energy_Gyr_x
143	Avr_peak_time_Acc_y	Avr_peak_time_Gyr_y			Skew_Gyr_y	Cf_Acc_z
144	nPeaks_Gyr_y	SVM_Angle			Rms_Acc_x	Energy_Acc_x
145	Rms_Acc_z	SMA_Angle			Rms_Gyr_z	Widest_Peak_Acc_x
146	Avr_peak_time_Acc_z	Kur_Gyr_z			Widest_Peak_Acc_z	Vedb_Angle
147	Std_Acc_y	Mean_Azimuth			Widest_Peak_Roll	Widest_Peak_Acc_z
148	Widest_Peak_Roll	Dfreq_Azimuth			Rms_Acc_z	Cf_Gyr_z
149	Interq_Acc_y	AV_Ang			Avr_peak_time_Pitch	Avr_peak_time_Gyr_y
150	Max_diff_Acc_z	DSAM_Angle			MV_Gyr	Entropy_Azimuth
151	Highest_peak_Acc_y	Max_diff_Acc_z			Avr_peak_time_Acc_x	Vedb_Acc
152	Min_Azimuth	Widest_Peak_Gyr_z			Avr_peak_time_Gyr_y	Rms_Acc_z
153	nPeaks_Pitch	Skew_Acc_x			Kur_Gyr_x	Skew_Gyr_z
154	AV_Ang	MV_Gyr			Highest_peak_Acc_y	Max_Acc_x
155	Widest_Peak_Acc_x	nPeaks_Gyr_y			nPeaks_Gyr_x	SMA_Angle
156	Widest_Peak_Azimuth	Interq_Acc_y			SVM_Angle	Entropy_TimeD_Ang
157	mag_Acc	Skew_Pitch			nPeaks_Azimuth	SVM_Angle
158	nPeaks_Acc_z	Skew_Gyr_z			Avr_peak_time_Acc_z	nPeaks_Acc_z
159	DSAM_Angle	nPeaks_Roll			Avr_peak_time_Roll	Std_Acc_y
160	Avr_peak_time_Gyr_y	Skew_Azimuth			Widest_Peak_Acc_x	Energy_Gyr_y
161	Entropy_Acc_y	Max_Azimuth			Max_diff_Acc_z	Interq_Pitch
162	Skew_Roll	Kur_Acc_x			Max_Roll	nPeaks_Gyr_y
163	nPeaks_Roll	Kur_Azimuth			Skew_Gyr_z	nPeaks_Acc_x
164	SMA_Acc	Widest_Peak_Acc_z			mag_Acc	SMA_Acc
165	Highest_peak_Azimuth	Widest_Peak_Roll			Highest_peak_Azimuth	Kur_Acc_y
166	Entropy_TimeD_Acc	nPeaks_Acc_y			Rms_Azimuth	Energy_Acc_z
167	Kur_Acc_x	Highest_peak_Azimuth			Kur_Gyr_y	Dfreq_Acc_z
168	SVM_Acc	Entropy_Azimuth			Entropy_TimeD_Acc	Skew_Gyr_y
169	Var_Azimuth	Widest_Peak_Acc_x			nPeaks_Acc_x	Dfreq_Gyr_z
170	Skew_Pitch	Interq_Azimuth			Rms_Acc_y	Entropy_Gyr_z

171	Cf_Azimuth	nPeaks_Pitch			Mean_Acc_y	Std_Azimuth
172	Skew_Gyr_x	Avr_peak_time_Acc_y			Var_Gyr_z	Dfreq_Azimuth
173	nPeaks_Gyr_z	Widest_Peak_Azimuth			SMA_Angle	Std_Roll
174	Avr_peak_time_Acc_x	Avr_peak_time_Acc_x			Max_diff_Acc_y	Std_Acc_x
175	nPeaks_Acc_x	Widest_Peak_Gyr_y			Max_Gyr_x	SVM_Acc
176	Avr_peak_time_Gyr_z	Widest_Peak_Gyr_x			Highest_peak_Gyr_z	Entropy_TimeD_Gyr
177	Std_Azimuth	Avr_peak_time_Gyr_x			Kur_Acc_x	Std_Gyr_y
178	Kur_Azimuth	nPeaks_Gyr_x			Skew_Gyr_x	Std_Gyr_x
179	Avr_peak_time_Roll	nPeaks_Acc_z			Highest_peak_Gyr_x	Energy_Roll
180	Skew_Azimuth	Std_Azimuth			Entropy_Azimuth	DSVM_Gyr
181	Entropy_Azimuth	nPeaks_Acc_x			Max_Azimuth	Kur_Gyr_z
182	Max_diff_Azimuth	Var_Azimuth			Skew_Pitch	Kur_Azimuth
183	Kur_Pitch	Max_diff_Azimuth			Kur_Roll	Std_Pitch

Appendix E. Ranked Features Tables for Sheep DataSets

Appendix E. 12 Ranked features from (ReliefF, GA, and RF) FS methods for **DataSet3_all** over 5 $sec.window$.

	ReliefF		GA		RF	
#	FNSW	FOSW	FNSW	FOSW	FNSW	FOSW
1	Mean_Roll	Mean_Roll	Mean_Acc_x	Mean_Acc_y	Dfreq_Roll	Rms_Roll
2	Mean_Acc_x	Cf_Roll	Mean_Acc_y	Mean_Acc_z	Mean_Gyr_z	Mean_Roll
3	Cf_Roll	Mean_Acc_x	Mean_Pitch	Mean_Azimuth	Mean_Roll	Mean_Gyr_z
4	Dfreq_Roll	Max_Roll	Var_Acc_y	Mean_Roll	Mean_Acc_x	Dfreq_Roll
5	Mean_Gyr_z	Entropy_Roll	Var_Acc_z	Mean_Gyr_x	Rms_Roll	Skew_Acc_z
6	Rms_Roll	Rms_Roll	Var_Azimuth	Mean_Gyr_y	Min_Roll	Max_diff_Gyr_y
7	Entropy_Roll	Dfreq_Roll	Var_Pitch	Mean_Gyr_z	Skew_Acc_z	Mean_Acc_x
8	Cf_Acc_x	Min_Roll	Var_Roll	Var_Acc_x	Interq_Gyr_x	Max_Roll
9	Min_Roll	Max_diff_Gyr_y	Var_Gyr_x	Var_Acc_y	Min_Pitch	Max_diff_Gyr_x
10	Dfreq_Acc_x	Cf_Acc_x	Var_Gyr_y	Var_Acc_z	Var_Gyr_y	Var_Acc_y
11	Max_Roll	Mean_Gyr_z	Var_Gyr_z	Var_Pitch	Highest_peak_Gyr_x	Var_Gyr_y
12	Std_Gyr_y	Rms_Gyr_y	Std_Acc_x	Var_Roll	mag_Ang	Min_Roll
13	Rms_Acc_x	Std_Gyr_y	Std_Acc_y	Std_Acc_y	Kur_Gyr_x	Max_diff_Pitch
14	Std_Pitch	Max_Acc_x	Std_Acc_z	Std_Acc_z	Rms_Acc_x	Dfreq_Gyr_x
15	Max_diff_Gyr_y	Max_diff_Pitch	Std_Azimuth	Std_Azimuth	Max_Acc_x	Min_Acc_x
16	Energy_Roll	Energy_Roll	Std_Pitch	Std_Gyr_x	Skew_Acc_y	Max_Azimuth
17	Rms_Gyr_y	Max_Acc_y	Std_Gyr_x	Std_Gyr_y	Entropy_Acc_y	Interq_Roll
18	Min_Acc_z	Rms_Acc_x	Std_Gyr_y	Std_Gyr_z	Min_Gyr_y	Skew_Roll
19	Highest_peak_Acc_x	Highest_peak_Acc_x	Std_Gyr_z	Kur_Acc_x	Interq_Pitch	mag_Acc
20	Max_Acc_x	Energy_Gyr_y	Kur_Acc_z	Kur_Acc_y	Skew_Gyr_x	Entropy_Gyr_x
21	Min_Acc_x	Std_Roll	Kur_Azimuth	Kur_Acc_z	Var_Pitch	Rms_Gyr_y
22	Std_Roll	Var_Gyr_y	Kur_Gyr_x	Kur_Azimuth	Max_Gyr_z	SMA_Angle
23	Entropy_Pitch	Entropy_Pitch	Kur_Gyr_z	Kur_Gyr_z	Widest_Peak_Acc_z	Entropy_Pitch
24	Highest_peak_Gyr_y	Interq_Gyr_y	Skew_Acc_z	Skew_Acc_x	Var_Gyr_x	Entropy_Roll
25	Interq_Gyr_z	Std_Gyr_z	Skew_Azimuth	Skew_Gyr_x	Max_Acc_z	Max_diff_Gyr_z

26	Min_Pitch	Entropy_Gyr_x	Skew_Pitch	Min_Acc_x	Interq_Gyr_y	Min_Acc_z
27	Interq_Acc_x	Dfreq_Gyr_y	Min_Acc_x	Min_Azimuth	Rms_Gyr_z	Min_Azimuth
28	Var_Pitch	Min_Acc_x	Min_Acc_y	Min_Pitch	Rms_Acc_y	Max_Acc_y
29	Var_Gyr_y	Rms_Gyr_z	Min_Azimuth	Min_Gyr_y	Cf_Gyr_x	Highest_peak_Acc_x
30	Interq_Pitch	Max_diff_Gyr_z	Min_Pitch	Min_Gyr_z	Entropy_Roll	Max_Gyr_y
31	Std_Acc_x	Interq_Gyr_z	Min_Roll	Max_Acc_x	Interq_Roll	Max_Pitch
32	Skew_Acc_z	Skew_Acc_y	Min_Gyr_x	Max_Acc_z	Dfreq_Gyr_y	Skew_Acc_y
33	Energy_Gyr_y	Dfreq_Acc_x	Max_Acc_x	Max_Roll	Max_diff_Roll	Avr_peak_time_Acc_z
34	Highest_peak_Gyr_x	Min_Pitch	Max_Acc_y	Max_Gyr_y	Min_Acc_x	Var_Acc_x
35	Cf_Pitch	DSVM_Gyr	Max_Acc_z	Rms_Acc_x	Max_diff_Pitch	Min_Gyr_x
36	Energy_Acc_x	Highest_peak_Gyr_x	Max_Roll	Rms_Acc_y	Max_diff_Gyr_x	Kur_Acc_z
37	mag_Ang	Dfreq_Gyr_z	Max_Gyr_x	Rms_Acc_z	Entropy_Azimuth	Rms_Acc_z
38	Dfreq_Gyr_x	Dfreq_Gyr_x	Max_Gyr_z	Rms_Azimuth	Interq_Acc_x	Dfreq_Gyr_y
39	Entropy_Acc_z	Min_Acc_z	Rms_Acc_y	Rms_Gyr_y	Min_Azimuth	DSVM_Gyr
40	Max_Gyr_y	Highest_peak_Gyr_y	Rms_Azimuth	Rms_Gyr_z	Dfreq_Gyr_z	Highest_peak_Gyr_y
41	Rms_Pitch	Min_Gyr_y	Rms_Pitch	Interq_Acc_y	Avr_peak_time_Roll	Max_diff_Acc_z
42	Min_Gyr_y	Skew_Acc_z	Rms_Gyr_x	Interq_Azimuth	Min_Gyr_z	Interq_Gyr_x
43	Max_diff_Pitch	Max_Gyr_y	Rms_Gyr_y	Interq_Gyr_x	Entropy_Acc_x	Avr_peak_time_Acc_x
44	Energy_Pitch	Highest_peak_Gyr_z	Rms_Gyr_z	Interq_Gyr_y	Rms_Gyr_y	Mean_Azimuth
45	Interq_Gyr_y	Std_Pitch	Interq_Acc_x	Cf_Acc_z	Skew_Azimuth	Mean_Pitch
46	Dfreq_Pitch	Cf_Gyr_z	Interq_Pitch	Cf_Pitch	Max_diff_Gyr_y	Highest_peak_Acc_z
47	Max_Pitch	Rms_Acc_y	Interq_Roll	Cf_Gyr_y	Cf_Acc_y	Interq_Gyr_z
48	Avr_peak_time_Gyr_z	Interq_Pitch	Interq_Gyr_x	SMA_Acc	SMA_Acc	Vedb_Angle
49	Cf_Gyr_z	Max_diff_Gyr_x	Interq_Gyr_y	SMA_Angle	Cf_Roll	Cf_Roll
50	Std_Gyr_z	Entropy_Acc_z	Cf_Acc_x	Entropy_Acc_z	Interq_Acc_z	Kur_Gyr_z
51	Interq_Acc_z	Std_Acc_x	Cf_Acc_y	Entropy_Azimuth	Interq_Gyr_z	Interq_Acc_y
52	DSVM_Gyr	nPeaks_Roll	Cf_Azimuth	Entropy_Pitch	Highest_peak_Pitch	Entropy_Gyr_z
53	Min_Acc_y	Cf_Pitch	Cf_Roll	Entropy_Gyr_x	Max_Acc_y	Highest_peak_Gyr_z
54	Cf_Gyr_x	Energy_Acc_x	Cf_Gyr_y	Entropy_Gyr_z	Interq_Azimuth	Max_Gyr_z

55	Interq_Gyr_x	Energy_Pitch	SMA_Acc	Entropy_TimeD_Ang	Widest_Peak_Acc_x	Cf_Azimuth
56	Dfreq_Gyr_y	Rms_Pitch	SMA_Angle	Entropy_TimeD_Gyr	Cf_Pitch	AV_Ang
57	Rms_Gyr_z	Energy_Acc_y	Entropy_Acc_x	Energy_Acc_y	Max_diff_Gyr_z	Var_Gyr_x
58	Cf_Gyr_y	Max_Pitch	Entropy_Acc_y	Energy_Acc_z	Max_Pitch	Rms_Acc_x
59	Std_Gyr_x	Interq_Roll	Entropy_Azimuth	Energy_Gyr_x	Mean_Pitch	SVM_Angle
60	Std_Acc_z	Max_Gyr_z	Entropy_Pitch	SVM_Acc	Dfreq_Gyr_x	Avr_peak_time_Acc_y
61	Var_Gyr_x	Min_Acc_y	Entropy_Roll	DSVM_Acc	Skew_Roll	nPeaks_Gyr_x
62	nPeaks_Gyr_z	Entropy_TimeD_Gyr	Entropy_Gyr_x	DSVM_Gyr	Mean_Gyr_y	Rms_Pitch
63	Var_Acc_x	Dfreq_Pitch	Entropy_Gyr_y	Max_diff_Acc_y	Max_Roll	Highest_peak_Acc_y
64	Rms_Gyr_x	Vedb_Gyr	Entropy_Gyr_z	Max_diff_Azimuth	nPeaks_Acc_y	Min_Pitch
65	Max_diff_Acc_y	Dfreq_Acc_z	Entropy_TimeD_Ang	Max_diff_Gyr_y	SVM_Acc	Avr_peak_time_Roll
66	Highest_peak_Pitch	Mean_Pitch	Entropy_TimeD_Gyr	Max_diff_Gyr_z	Avr_peak_time_Gyr_y	Rms_Azimuth
67	Interq_Roll	Entropy_Gyr_z	Energy_Acc_x	AV_Ang	Vedb_Acc	nPeaks_Roll
68	Entropy_Azimuth	Rms_Gyr_x	Energy_Azimuth	MV_Gyr	Avr_peak_time_Acc_x	Min_Acc_y
69	Kur_Gyr_x	Energy_Gyr_z	Energy_Gyr_y	mag_Ang	Dfreq_Pitch	Skew_Gyr_x
70	Energy_Gyr_x	SVM_Gyr	Energy_Gyr_z	mag_Gyr	Avr_peak_time_Pitch	Widest_Peak_Gyr_z
71	Skew_Acc_y	Std_Gyr_x	DSAM_Angle	Vedb_Angle	Entropy_Gyr_x	Max_diff_Azimuth
72	Vedb_Acc	MV_Gyr	DSVM_Gyr	Vedb_Gyr	Entropy_Gyr_z	Widest_Peak_Roll
73	Max_diff_Gyr_z	SMA_Gyr	Max_diff_Pitch	Dfreq_Acc_x	Cf_Gyr_y	Interq_Acc_z
74	Mean_Pitch	Var_Roll	Max_diff_Roll	Dfreq_Azimuth	Entropy_Acc_z	Var_Gyr_z
75	Min_Gyr_z	Mean_Acc_z	Max_diff_Gyr_z	Dfreq_Pitch	Min_Acc_z	Min_Gyr_y
76	Dfreq_Acc_z	Var_Gyr_z	MV_Acc	nPeaks_Acc_x	Widest_Peak_Gyr_x	Mean_Acc_y
77	Mean_Acc_z	Max_Acc_z	MV_Gyr	Widest_Peak_Acc_x	Cf_Gyr_z	Mean_Acc_z
78	Min_Azimuth	Avr_peak_time_Roll	mag_Ang	Widest_Peak_Acc_y	Max_Gyr_y	Dfreq_Pitch
79	Max_Gyr_z	Dfreq_Acc_y	Vedb_Angle	Highest_peak_Acc_y	Kur_Pitch	MV_Gyr
80	Dfreq_Gyr_z	Energy_Gyr_x	Vedb_Gyr	Avr_peak_time_Acc_y	AV_Ang	Skew_Pitch
81	Highest_peak_Gyr_z	Var_Gyr_x	Dfreq_Acc_y	nPeaks_Acc_z	Highest_peak_Acc_z	Kur_Gyr_y
82	Var_Roll	mag_Ang	Dfreq_Azimuth	Widest_Peak_Acc_z	Highest_peak_Acc_y	Skew_Azimuth
83	Avr_peak_time_Gyr_x	Interq_Acc_x	Dfreq_Gyr_y	Highest_peak_Acc_z	Interq_Acc_y	Interq_Gyr_y

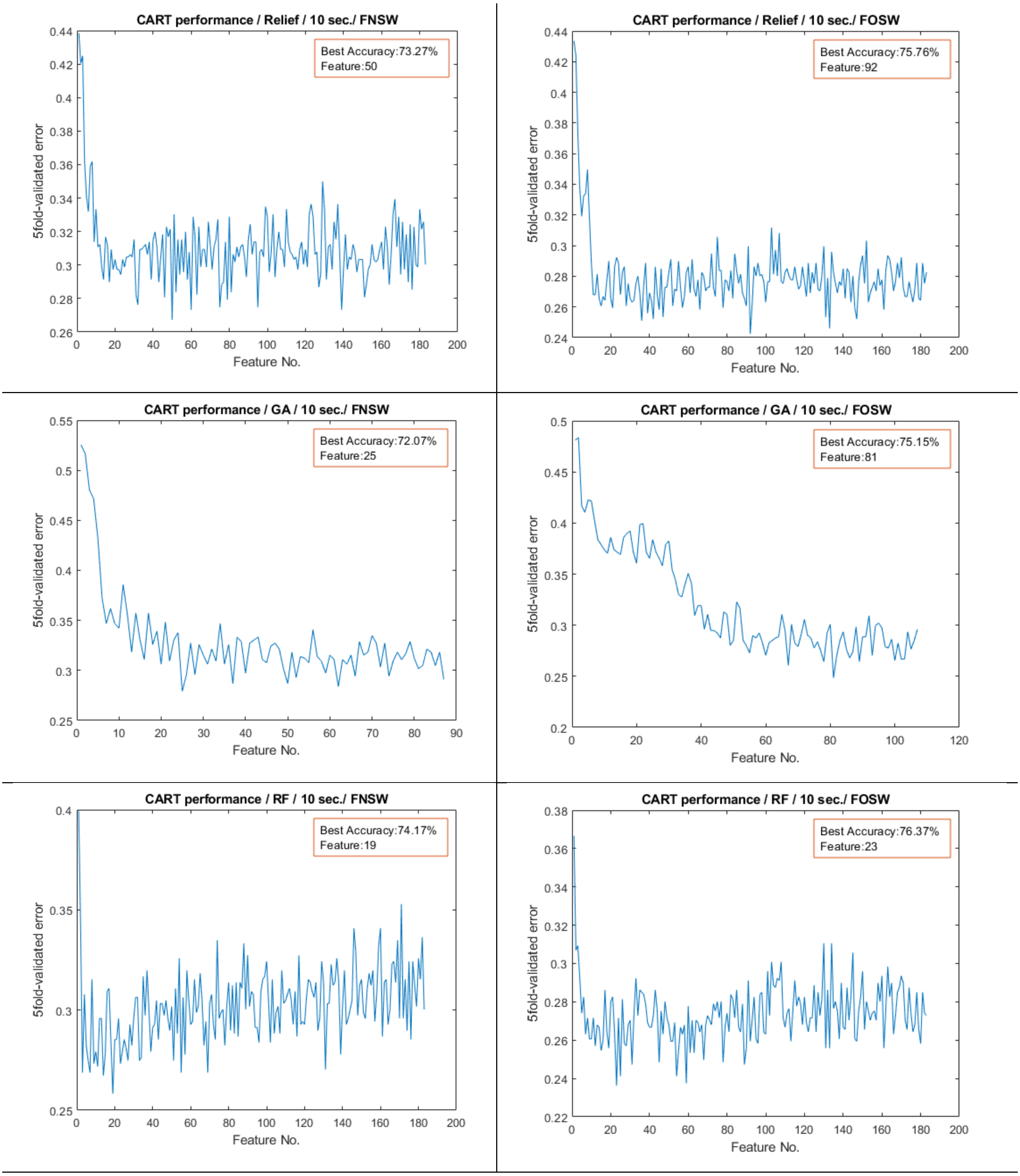
84	Entropy_Gyr_y	Max_diff_Acc_y	Dfreq_Gyr_z	Avr_peak_time_Acc_z	Highest_peak_Acc_x	Dfreq_Acc_y
85	Var_Acc_z	Var_Pitch	Widest_Peak_Acc_x	nPeaks_Azimuth	Widest_Peak_Pitch	Highest_peak_Azimuth
86	Max_diff_Gyr_x	Mean_Acc_y	Avr_peak_time_Acc_x	nPeaks_Pitch	Var_Acc_x	Avr_peak_time_Gyr_z
87	Rms_Acc_y	Kur_Pitch	Widest_Peak_Acc_y	Widest_Peak_Pitch	Highest_peak_Azimuth	Cf_Pitch
88	Var_Gyr_z	Highest_peak_Acc_y	Highest_peak_Acc_y	nPeaks_Roll	Mean_Acc_y	SMA_Gyr
89	Max_Gyr_x	Mean_Gyr_y	nPeaks_Acc_z	Widest_Peak_Roll	Max_Gyr_x	Cf_Acc_y
90	Max_Acc_z	Std_Acc_z	Widest_Peak_Acc_z	Highest_peak_Roll	Kur_Acc_z	Interq_Pitch
91	Energy_Gyr_z	Max_diff_Roll	Highest_peak_Azimuth	Avr_peak_time_Roll	Widest_Peak_Gyr_z	Entropy_Gyr_y
92	Entropy_Acc_x	Var_Acc_x	nPeaks_Pitch	nPeaks_Gyr_x	DSVM_Gyr	Kur_Azimuth
93	Energy_Acc_y	Interq_Gyr_x	Widest_Peak_Pitch	Highest_peak_Gyr_x	Mean_Azimuth	Var_Acc_z
94	Avr_peak_time_Acc_z	Energy_Acc_z	Avr_peak_time_Pitch	Avr_peak_time_Gyr_x	SMA_Angle	Widest_Peak_Acc_x
95	Max_diff_Acc_z	Avr_peak_time_Gyr_x	nPeaks_Roll	Widest_Peak_Gyr_y	Avr_peak_time_Azimuth	Var_Pitch
96	Highest_peak_Roll	Skew_Pitch	Highest_peak_Roll	nPeaks_Gyr_z	Max_diff_Acc_y	Max_diff_Roll
97	Mean_Gyr_y	Highest_peak_Roll	Avr_peak_time_Roll	Avr_peak_time_Gyr_z	nPeaks_Azimuth	Dfreq_Gyr_z
98	Highest_peak_Acc_y	Rms_Acc_z	nPeaks_Gyr_x		Highest_peak_Gyr_y	DSAM_Angle
99	Dfreq_Acc_y	Skew_Acc_x	Widest_Peak_Gyr_x		Dfreq_Acc_x	Entropy_Acc_y
100	Max_diff_Azimuth	Entropy_Acc_y	nPeaks_Gyr_y		Highest_peak_Gyr_z	Skew_Gyr_z
101	Max_diff_Roll	Var_Acc_z	Avr_peak_time_Gyr_y		Rms_Azimuth	Entropy_Azimuth
102	Skew_Gyr_y	Highest_peak_Pitch	nPeaks_Gyr_z		Min_Acc_y	Cf_Gyr_z
103	Cf_Azimuth	DSVM_Acc	Widest_Peak_Gyr_z		mag_Acc	Max_diff_Acc_x
104	Interq_Azimuth	nPeaks_Gyr_x	Highest_peak_Gyr_z		Rms_Gyr_x	Rms_Gyr_z
105	Entropy_Gyr_x	Rms_Azimuth			Skew_Acc_x	Interq_Acc_x
106	Mean_Acc_y	Kur_Acc_y			Cf_Acc_z	Avr_peak_time_Azimuth
107	Kur_Acc_y	Vedb_Angle			Var_Roll	Cf_Acc_x
108	Skew_Pitch	Energy_Azimuth			Avr_peak_time_Acc_y	Std_Acc_x
109	Vedb_Gyr	Min_Gyr_z			Var_Azimuth	Std_Acc_y
110	Widest_Peak_Roll	Dfreq_Azimuth			Avr_peak_time_Gyr_x	Std_Acc_z
111	Max_Acc_y	Mean_Azimuth			Std_Acc_x	Std_Azimuth
112	nPeaks_Gyr_x	Kur_Roll			Std_Acc_y	Std_Roll

113	Min_Gyr_x	AV_Ang			Std_Acc_z	Std_Gyr_x
114	Cf_Acc_z	Widest_Peak_Acc_z			Std_Azimuth	Std_Gyr_y
115	MV_Gyr	nPeaks_Azimuth			Std_Pitch	Std_Gyr_z
116	Skew_Roll	Interq_Acc_z			Std_Roll	Entropy_TimeD_Acc
117	Interq_Acc_y	Min_Azimuth			Std_Gyr_x	Energy_Acc_x
118	SVM_Gyr	Max_Gyr_x			Std_Gyr_y	Energy_Acc_y
119	DSVM_Acc	Cf_Acc_z			Std_Gyr_z	Energy_Acc_z
120	Entropy_TimeD_Gyr	nPeaks_Acc_x			Cf_Azimuth	Energy_Azimuth
121	Energy_Acc_z	DSAM_Angle			Entropy_TimeD_Ang	Energy_Pitch
122	Kur_Acc_x	Avr_peak_time_Acc_z			Entropy_TimeD_Gyr	Energy_Roll
123	Dfreq_Azimuth	Entropy_Acc_x			Energy_Acc_x	Energy_Gyr_x
124	Mean_Azimuth	Kur_Acc_z			Energy_Acc_y	Energy_Gyr_y
125	SMA_Gyr	Cf_Gyr_y			Energy_Acc_z	Energy_Gyr_z
126	Kur_Acc_z	Vedb_Acc			Energy_Azimuth	Dfreq_Acc_z
127	Max_diff_Acc_x	Kur_Gyr_x			Energy_Pitch	Dfreq_Azimuth
128	Entropy_Gyr_z	Cf_Gyr_x			Energy_Roll	nPeaks_Acc_y
129	Rms_Acc_z	Skew_Gyr_x			Energy_Gyr_x	nPeaks_Acc_z
130	Avr_peak_time_Acc_x	Max_diff_Acc_z			Energy_Gyr_y	nPeaks_Gyr_z
131	Widest_Peak_Gyr_x	Entropy_TimeD_Ang			Energy_Gyr_z	Avr_peak_time_Gyr_x
132	Entropy_TimeD_Acc	Skew_Gyr_y			SVM_Angle	Var_Roll
133	SMA_Angle	Min_Gyr_x			MV_Acc	nPeaks_Acc_x
134	Var_Azimuth	SMA_Angle			MV_Gyr	Max_Acc_x
135	Energy_Azimuth	SVM_Angle			Vedb_Gyr	Kur_Roll
136	SVM_Angle	Std_Acc_y			Dfreq_Acc_z	Var_Azimuth
137	Kur_Gyr_z	Max_Azimuth			Dfreq_Azimuth	Max_Acc_z
138	SVM_Acc	Var_Acc_y			nPeaks_Acc_x	Kur_Pitch
139	SMA_Acc	Widest_Peak_Azimuth			nPeaks_Pitch	nPeaks_Azimuth
140	MV_Acc	Widest_Peak_Gyr_x			nPeaks_Gyr_y	Max_Gyr_x
141	Entropy_TimeD_Ang	MV_Acc			Kur_Gyr_z	Cf_Acc_z

142	Avr_peak_time_Azimuth	Kur_Gyr_y			Var_Acc_y	Vedb_Acc
143	Kur_Roll	Avr_peak_time_Acc_x			Entropy_TimeD_Acc	Highest_peak_Pitch
144	Kur_Pitch	Skew_Roll			Avr_peak_time_Gyr_z	Cf_Gyr_x
145	Rms_Azimuth	Highest_peak_Acc_z			Vedb_Angle	Mean_Gyr_y
146	Vedb_Angle	Cf_Acc_y			Cf_Acc_x	Min_Gyr_z
147	Std_Azimuth	Skew_Gyr_z			Var_Gyr_z	Widest_Peak_Acc_y
148	Skew_Acc_x	mag_Acc			Highest_peak_Roll	Entropy_TimeD_Ang
149	Skew_Azimuth	Interq_Acc_y			SMA_Gyr	Avr_peak_time_Pitch
150	Widest_Peak_Acc_z	Max_diff_Acc_x			Rms_Pitch	SMA_Acc
151	Avr_peak_time_Gyr_y	nPeaks_Pitch			Kur_Acc_x	Skew_Gyr_y
152	Skew_Gyr_x	Cf_Azimuth			Kur_Roll	Rms_Acc_y
153	Entropy_Acc_y	Entropy_Gyr_y			Skew_Gyr_y	MV_Acc
154	AV_Ang	Widest_Peak_Gyr_z			Mean_Acc_z	Widest_Peak_Pitch
155	nPeaks_Azimuth	nPeaks_Gyr_z			Entropy_Pitch	Highest_peak_Roll
156	Widest_Peak_Acc_x	Mean_Gyr_x			mag_Gyr	Kur_Acc_y
157	DSAM_Angle	SMA_Acc			Rms_Acc_z	Cf_Gyr_y
158	Highest_peak_Acc_z	Kur_Gyr_z			Avr_peak_time_Acc_z	Kur_Gyr_x
159	nPeaks_Acc_x	Entropy_TimeD_Acc			nPeaks_Roll	Entropy_TimeD_Gyr
160	Widest_Peak_Pitch	SVM_Acc			nPeaks_Gyr_x	Std_Pitch
161	Kur_Gyr_y	Widest_Peak_Pitch			nPeaks_Gyr_z	Rms_Gyr_x
162	Std_Acc_y	Widest_Peak_Acc_y			Widest_Peak_Roll	SVM_Acc
163	Var_Acc_y	Widest_Peak_Acc_x			SVM_Gyr	Avr_peak_time_Gyr_y
164	Highest_peak_Azimuth	Highest_peak_Azimuth			Dfreq_Acc_y	SVM_Gyr
165	Widest_Peak_Gyr_y	mag_Gyr			nPeaks_Acc_z	Widest_Peak_Gyr_x
166	Avr_peak_time_Pitch	Avr_peak_time_Gyr_z			Max_diff_Acc_x	Entropy_Acc_x
167	Avr_peak_time_Roll	Max_diff_Azimuth			Entropy_Gyr_y	Max_diff_Acc_y
168	mag_Gyr	Avr_peak_time_Azimuth			Var_Acc_z	Dfreq_Acc_x
169	Max_Azimuth	Entropy_Azimuth			Widest_Peak_Azimuth	Widest_Peak_Azimuth
170	Widest_Peak_Gyr_z	Kur_Azimuth			DSAM_Angle	Interq_Azimuth

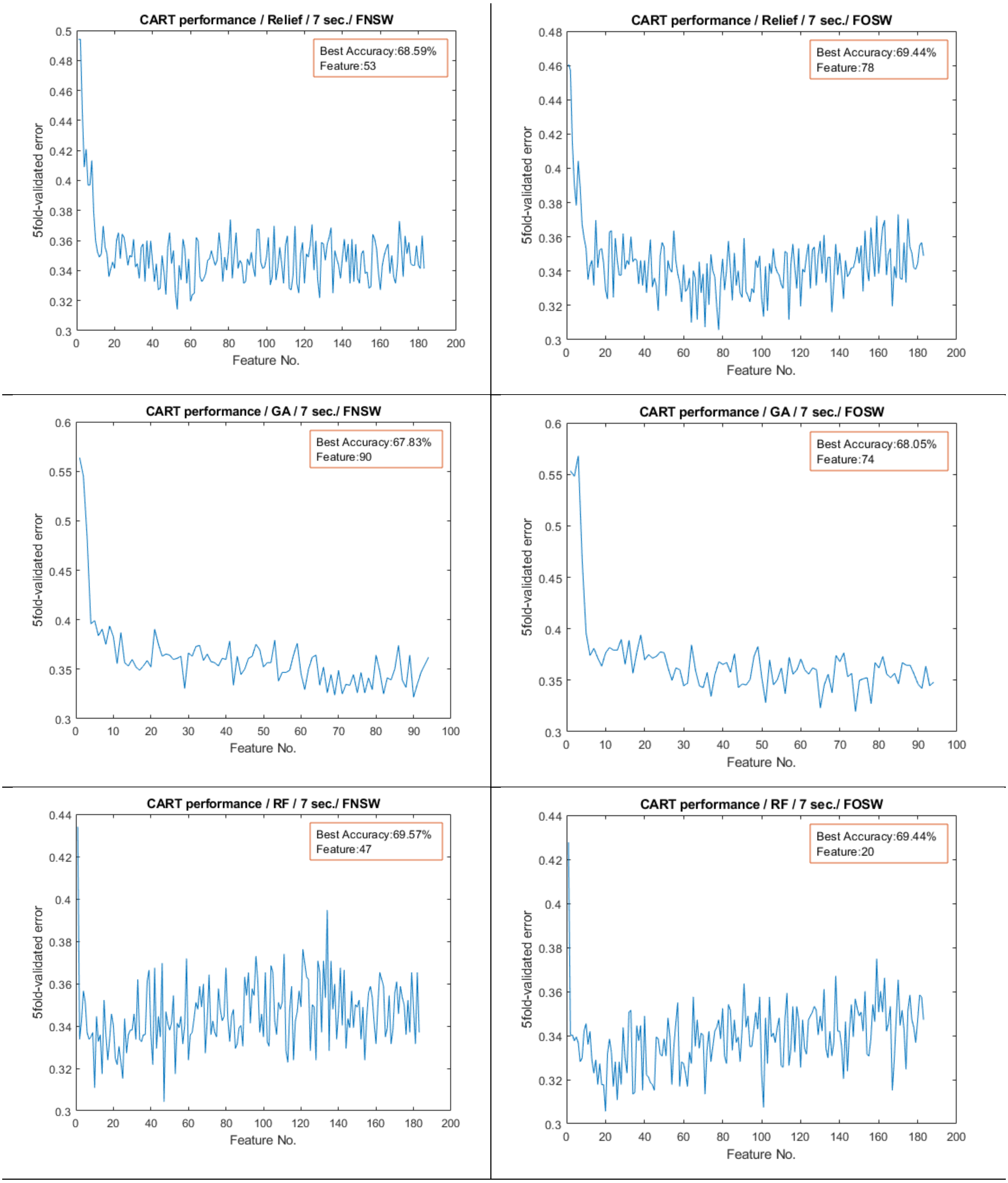
171	Skew_Gyr_z	Avr_peak_time_Acc_y			Kur_Gyr_y	Kur_Acc_x
172	nPeaks_Acc_z	Avr_peak_time_Pitch			Min_Gyr_x	Vedb_Gyr
173	Mean_Gyr_x	nPeaks_Acc_y			Max_diff_Acc_z	DSVM_Acc
174	Widest_Peak_Azimuth	Kur_Acc_x			Kur_Azimuth	Skew_Acc_x
175	Avr_peak_time_Acc_y	Skew_Azimuth			Skew_Pitch	mag_Gyr
176	nPeaks_Roll	Widest_Peak_Gyr_y			Max_Azimuth	nPeaks_Pitch
177	Cf_Acc_y	Avr_peak_time_Gyr_y			DSVM_Acc	mag_Ang
178	nPeaks_Gyr_y	Var_Azimuth			Max_diff_Azimuth	Highest_peak_Gyr_x
179	mag_Acc	Widest_Peak_Roll			Kur_Acc_y	nPeaks_Gyr_y
180	Kur_Azimuth	nPeaks_Acc_z			Mean_Gyr_x	Entropy_Acc_z
181	nPeaks_Acc_y	Std_Azimuth			Widest_Peak_Gyr_y	Mean_Gyr_x
182	Widest_Peak_Acc_y	Interq_Azimuth			Skew_Gyr_z	Widest_Peak_Gyr_y
183	nPeaks_Pitch	nPeaks_Gyr_y			Widest_Peak_Acc_y	Widest_Peak_Acc_z

Appendix F. CART Performance Results to Test for the Best Number of Features



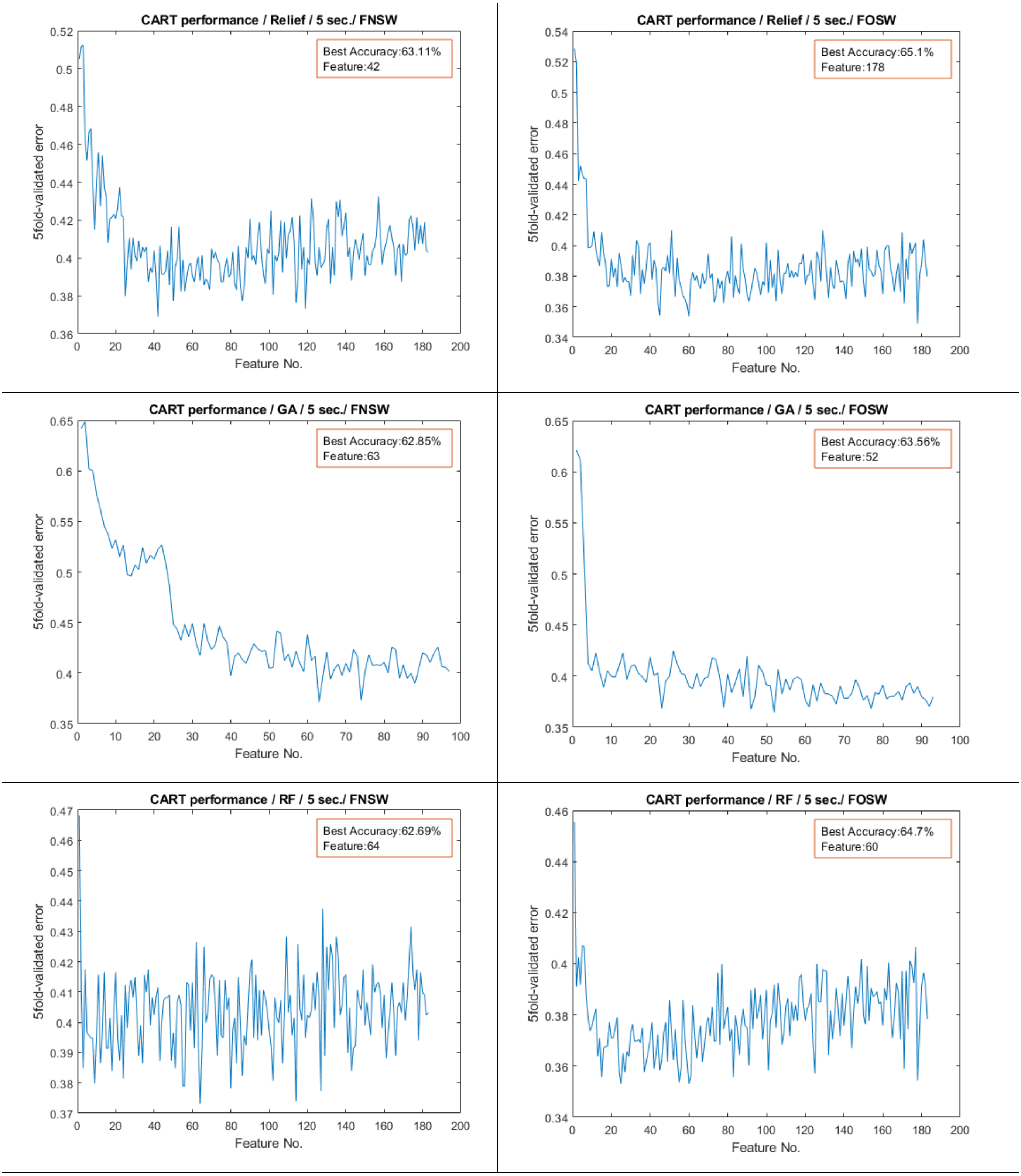
Appendix F. 1 Best no. of features ranked by 3 feature selection methods for **DataSet1_all** over **10** $sec. window$.

Appendix F. CART Performance Results to Test for the Best Number of Features



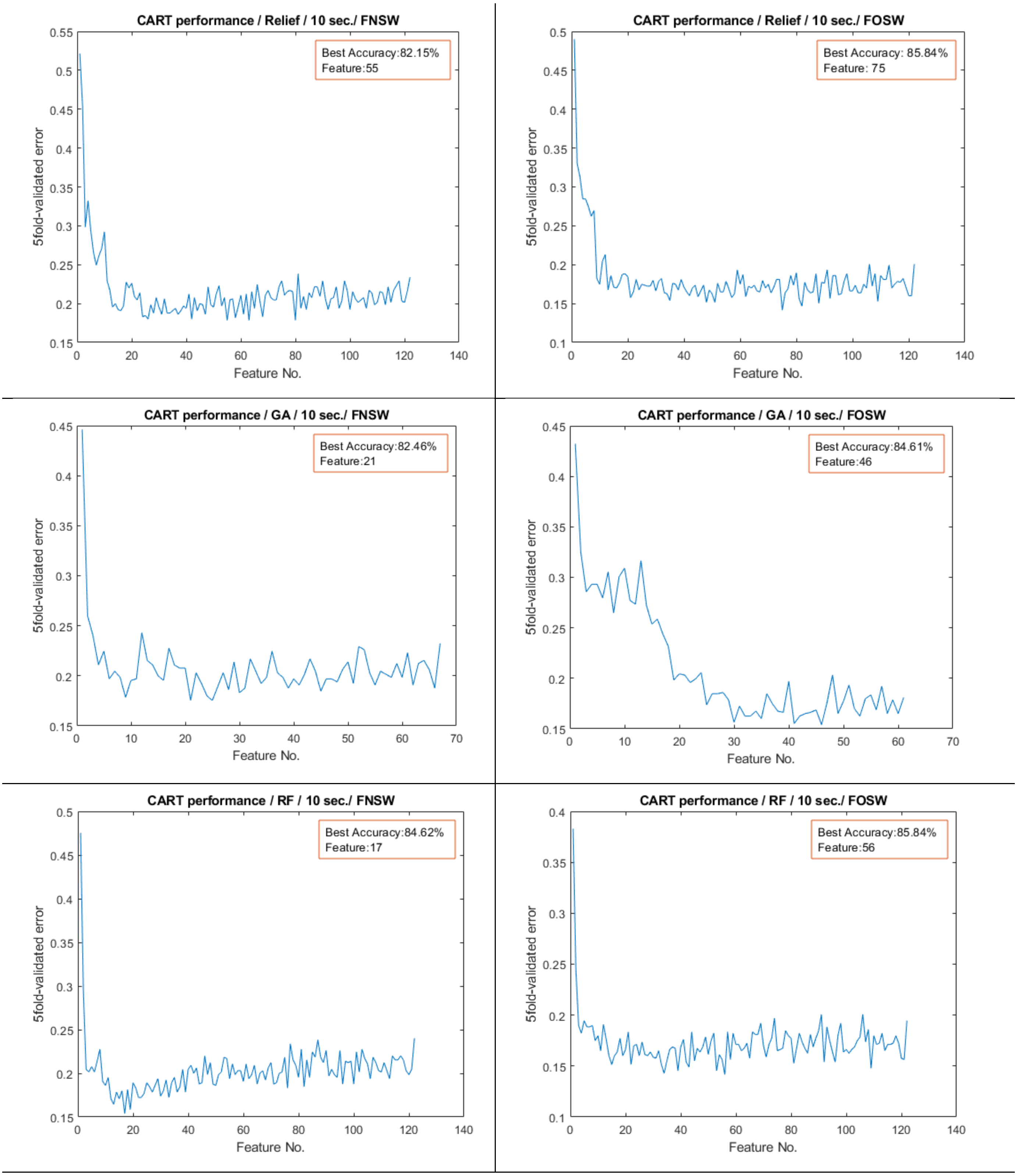
Appendix F. 2 Best no. of features ranked by 3 feature selection methods for **DataSet1_all** over **7** $sec. window$.

Appendix F. CART Performance Results to Test for the Best Number of Features



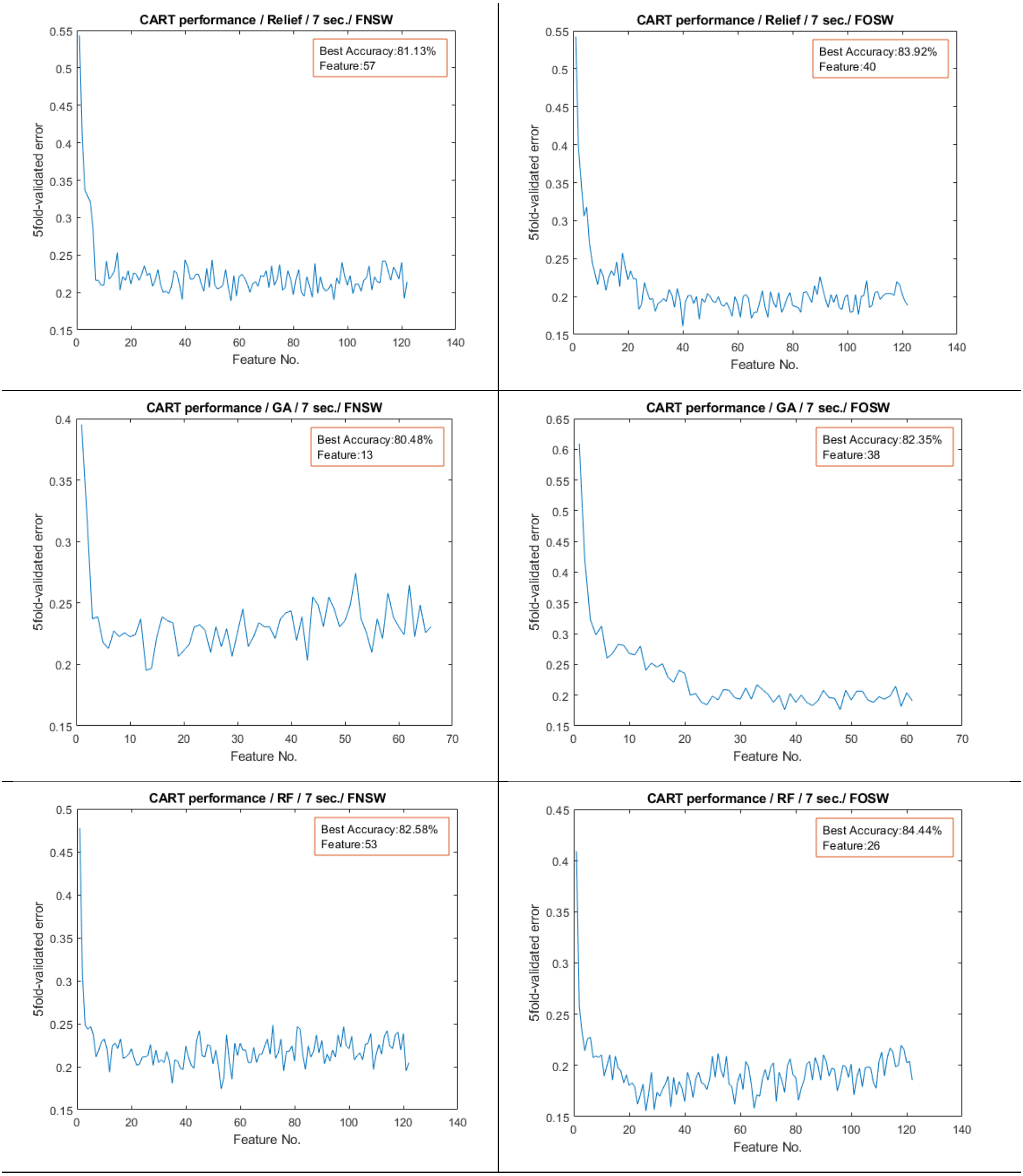
Appendix F. 3 Best no. of features ranked by 3 feature selection methods for **DataSet1_all** over **5** $sec. window$.

Appendix F. CART Performance Results to Test for the Best Number of Features



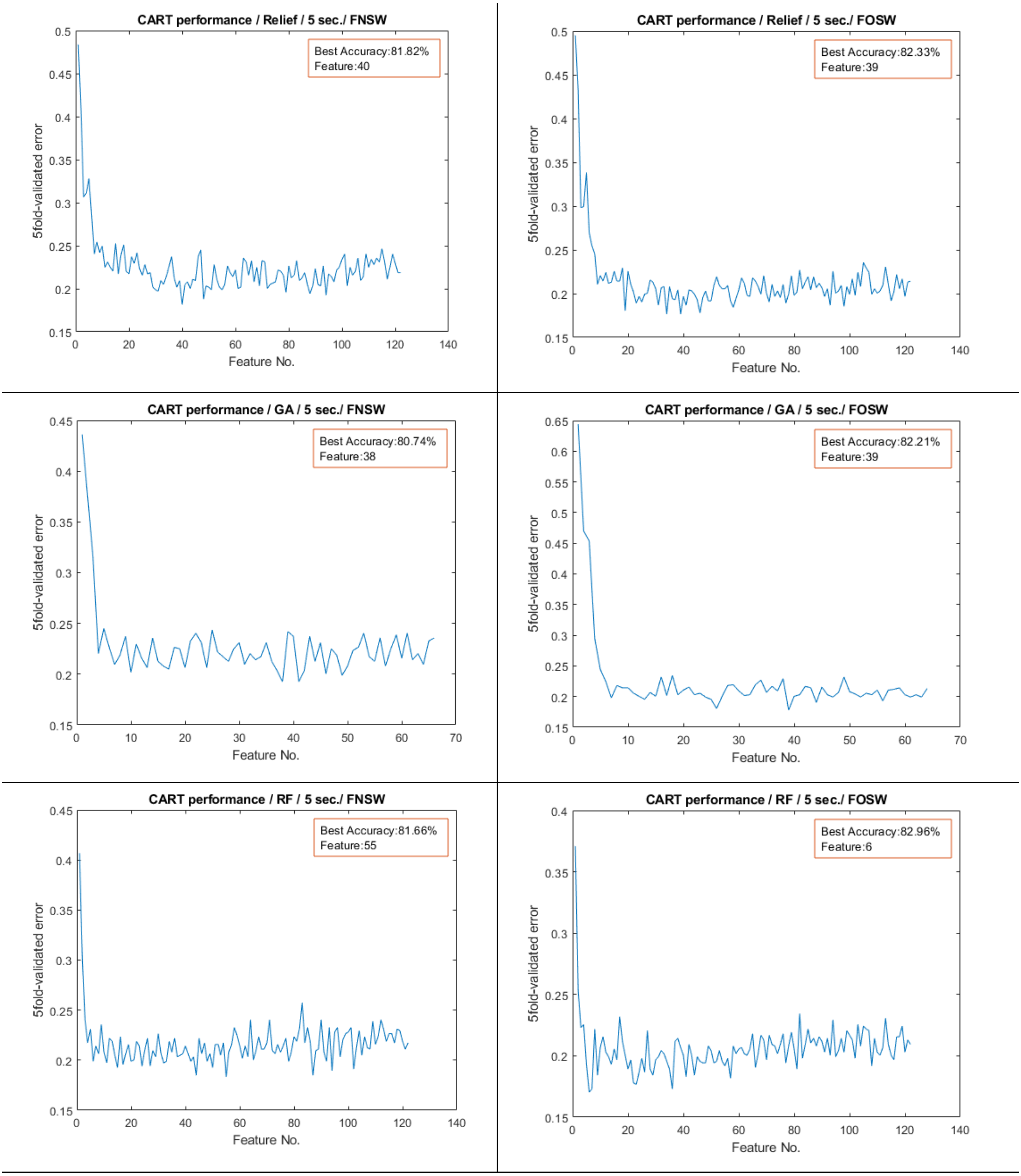
Appendix F. 4 Best no. of features ranked by 3 feature selection methods for **DataSet2_ac** over $\mathbf{10}\ sec.window$.

Appendix F. CART Performance Results to Test for the Best Number of Features

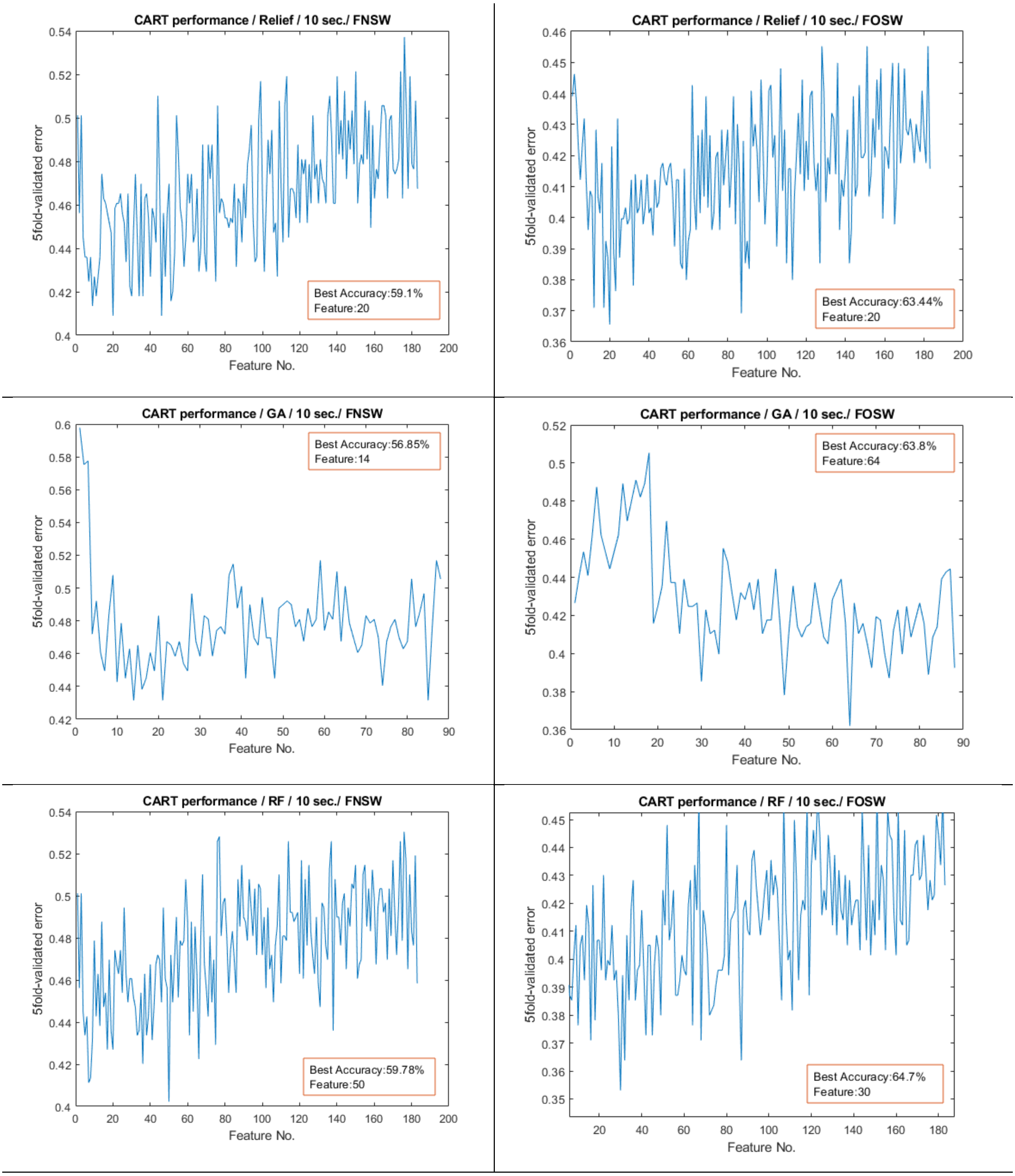


Appendix F. 5 Best no. of features ranked by 3 feature selection methods for **DataSet2_ac** over **7** *sec. window*.

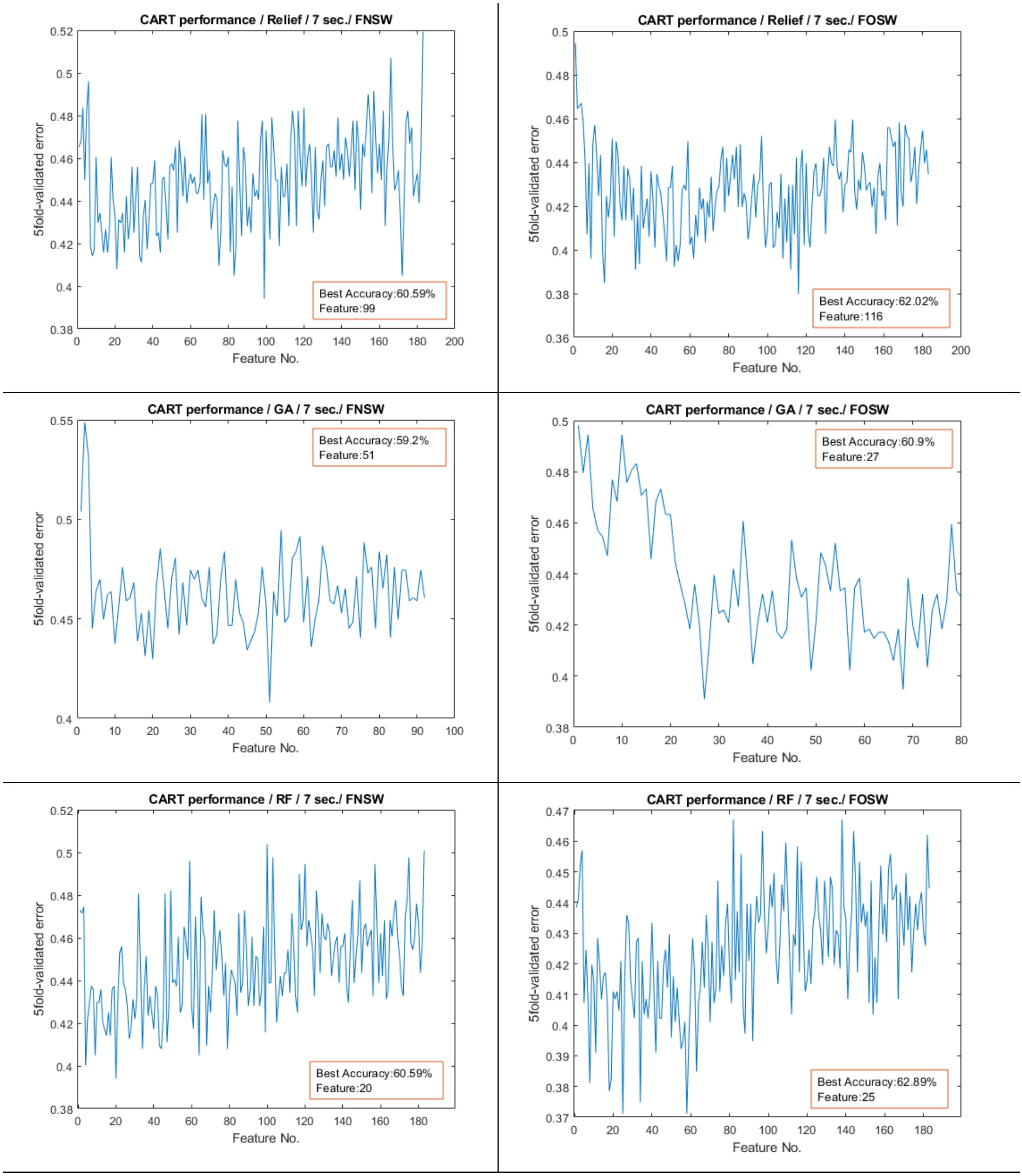
Appendix F. CART Performance Results to Test for the Best Number of Features



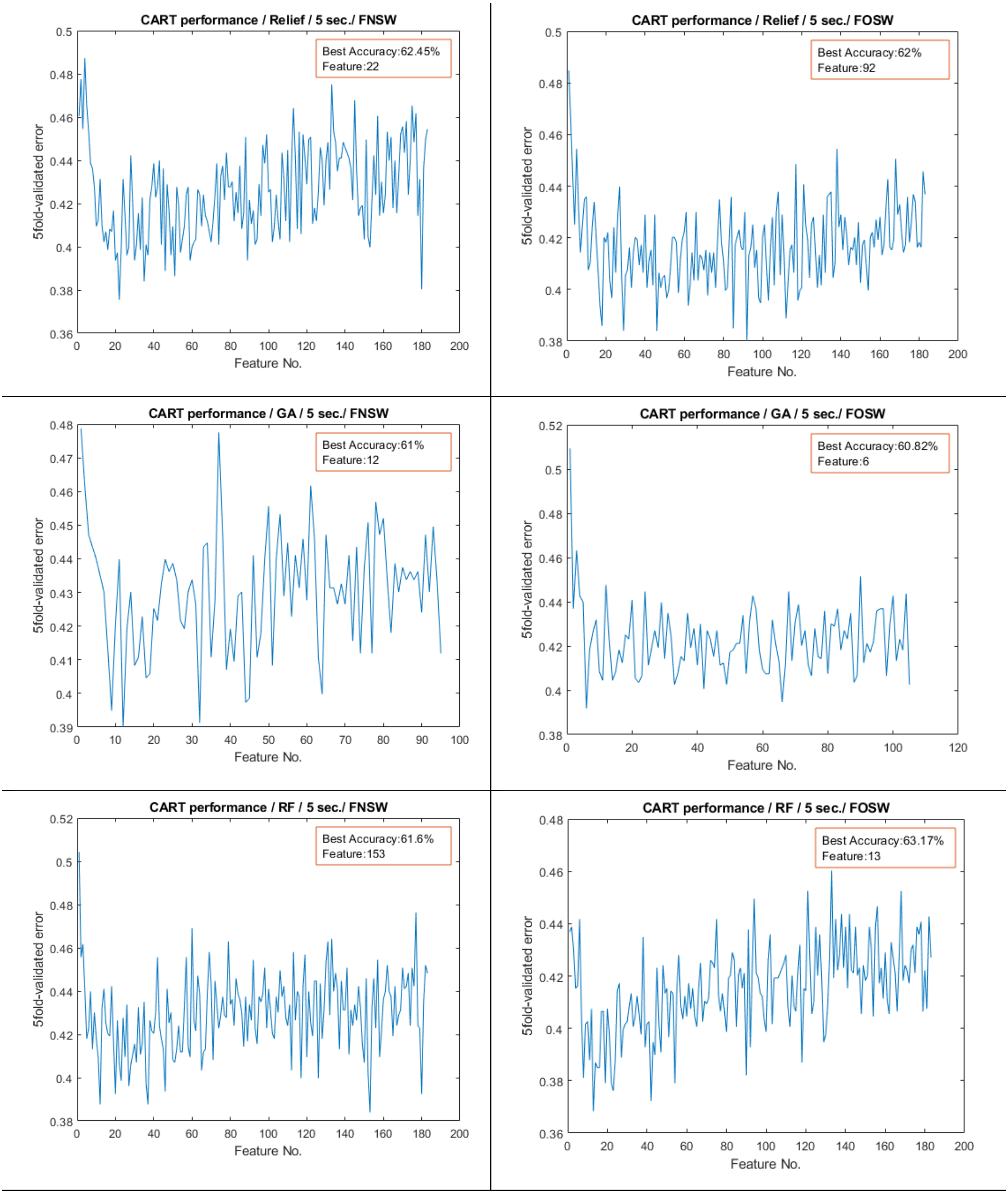
Appendix F. 6 Best no. of features ranked by 3 feature selection methods for **DataSet2_ac** over **5** $sec.window$.



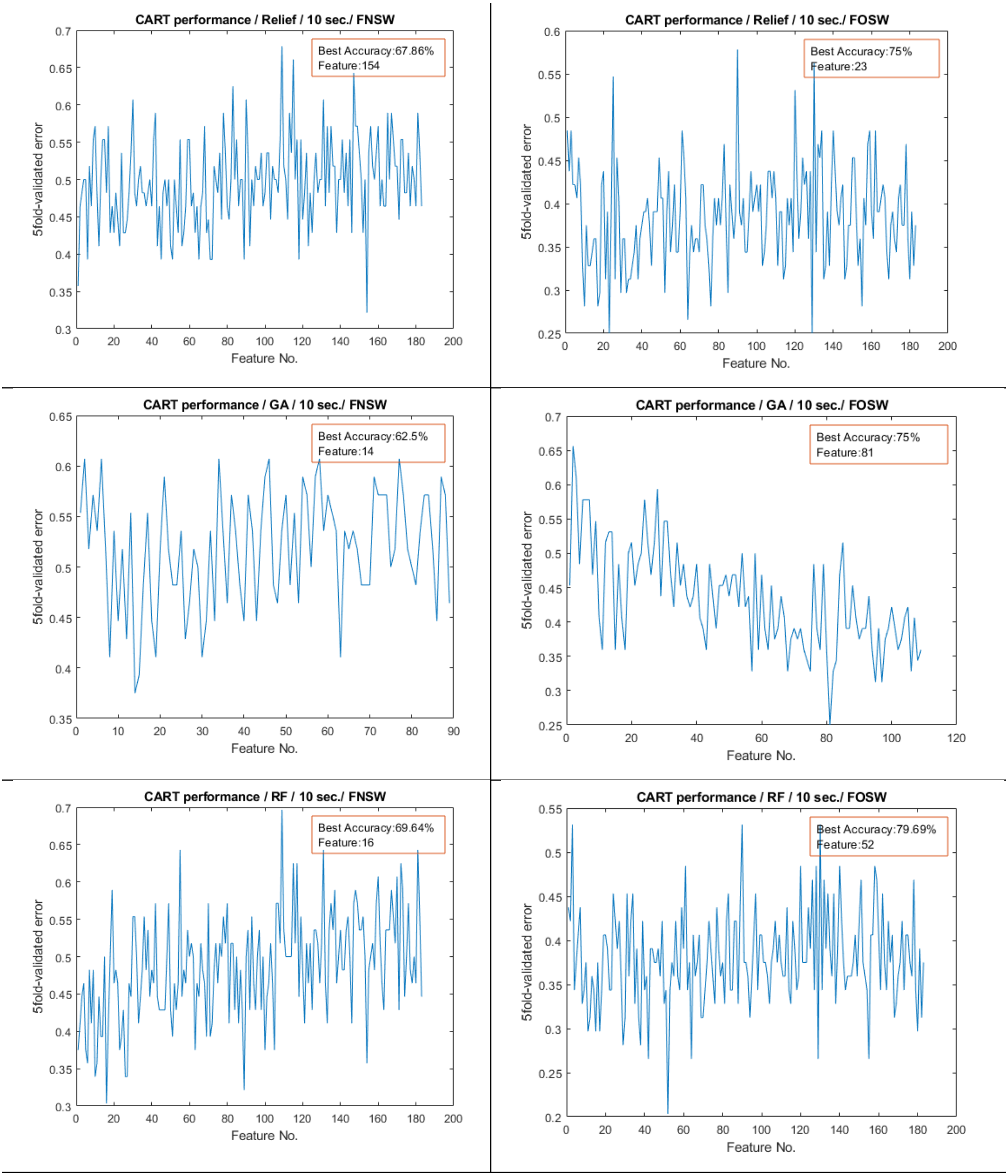
Appendix F. 7 Best no. of features ranked by 3 feature selection methods for **DataSet2_b** over **10** *sec. window*.



Appendix F. 8 Best no. of features ranked by 3 feature selection methods for **DataSet2_b** over **7** *sec. window*.

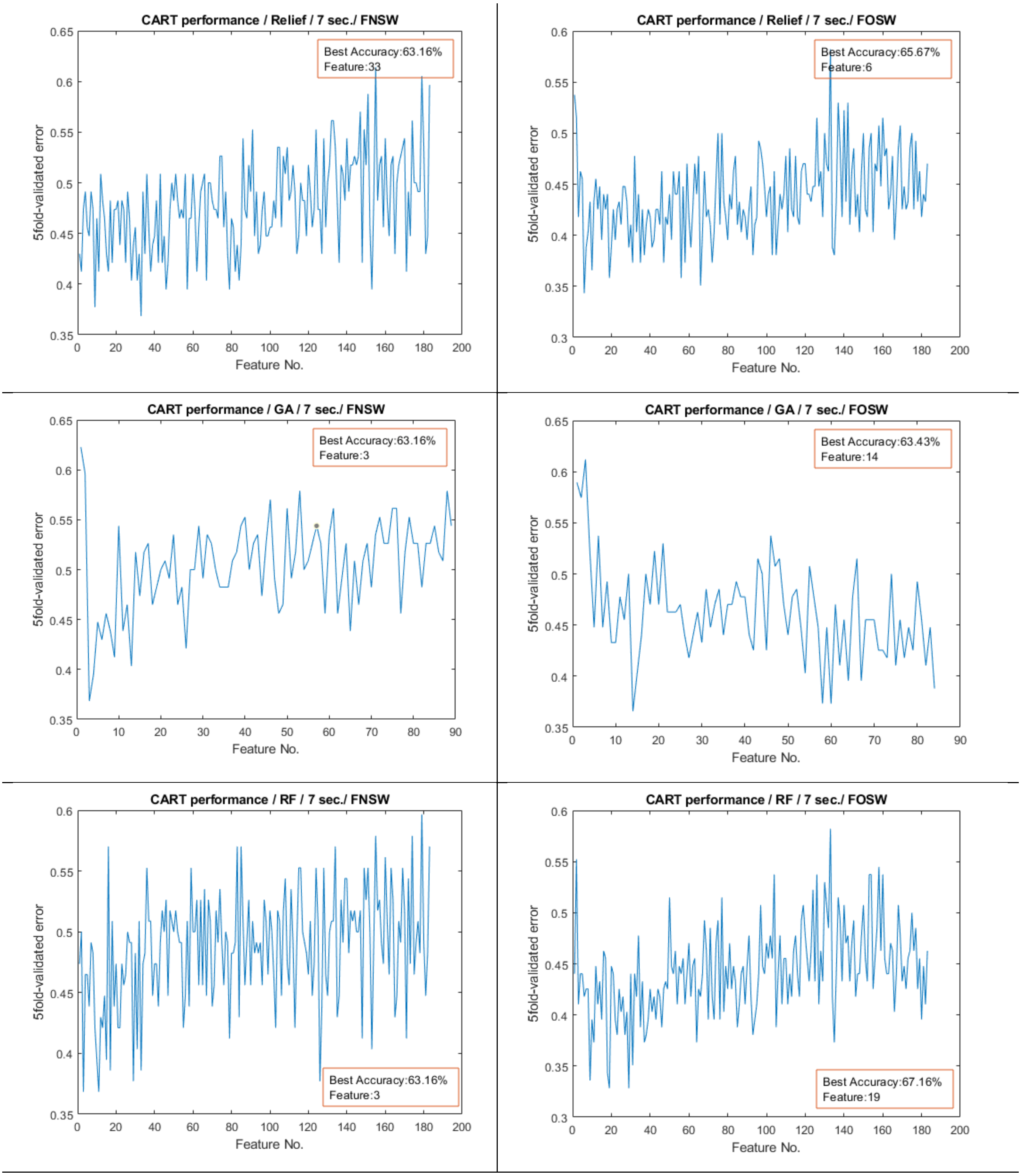


Appendix F. 9 Best no. of features ranked by 3 feature selection methods for **DataSet2_b** over **5** *sec. window*.



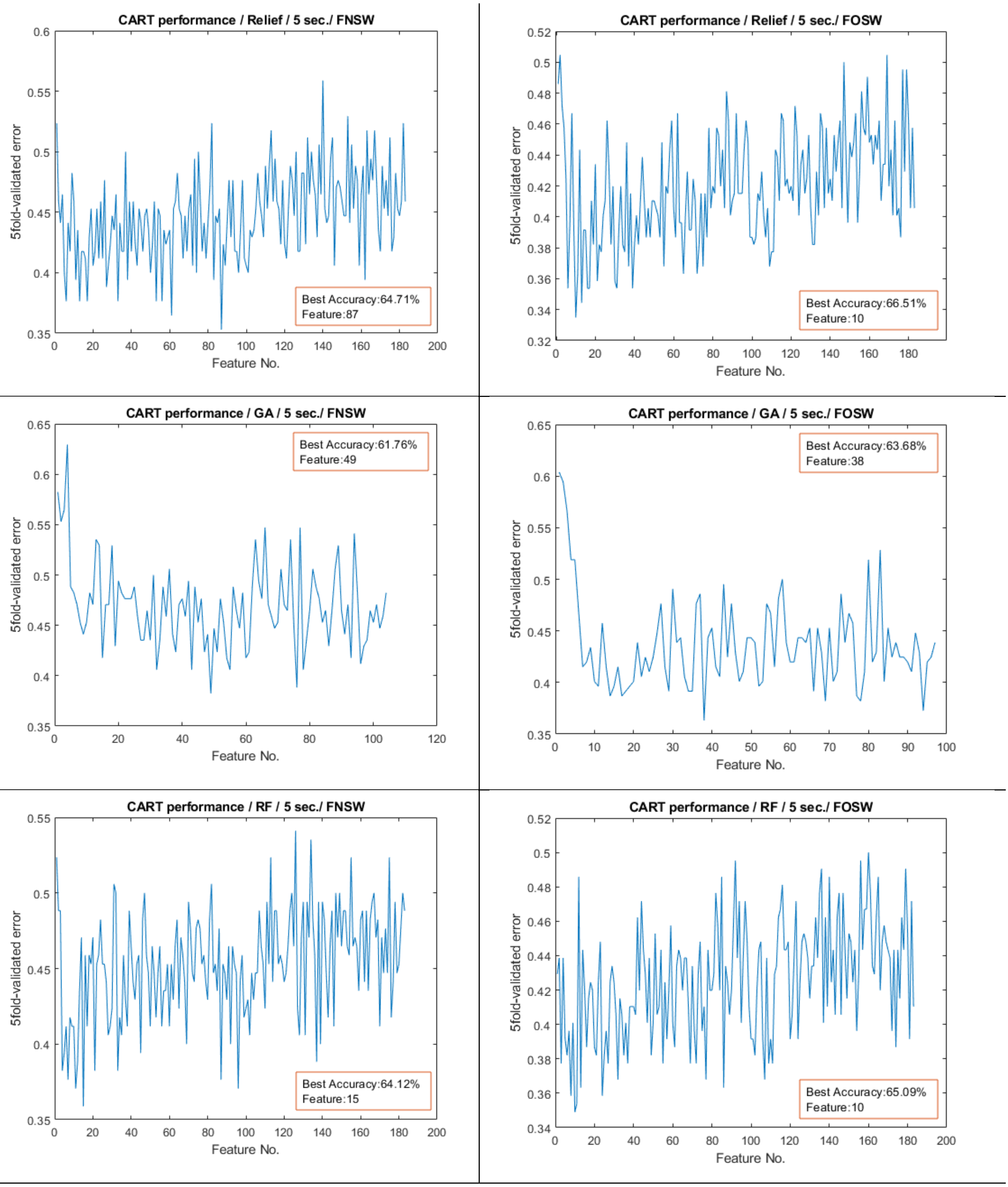
Appendix F. 10 Best no. of features ranked by 3 feature selection methods for **DataSet3_all** over **10** *sec. window*.

Appendix F. CART Performance Results to Test for the Best Number of Features



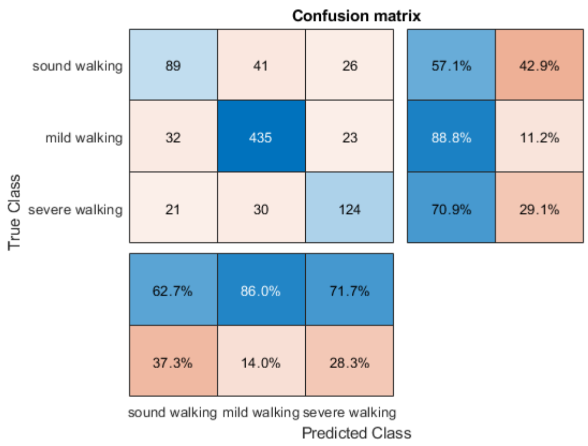
Appendix F. 11 Best no. of features ranked by 3 feature selection methods for **DataSet3_all** over **7** *sec. window*.

Appendix F. CART Performance Results to Test for the Best Number of Features

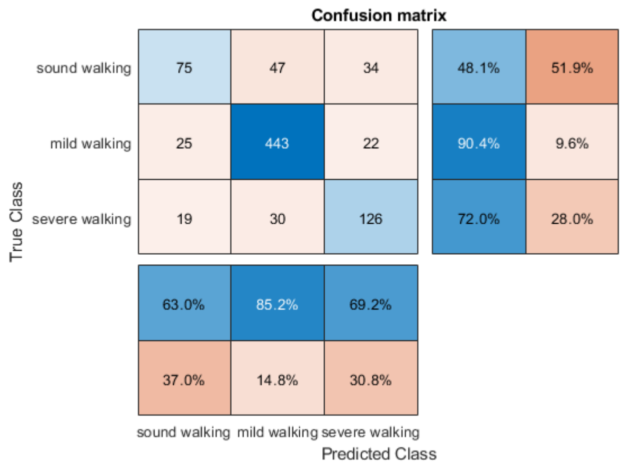


Appendix F. 12 Best no. of features ranked by 3 feature selection methods for **DataSet3_all** over **5** $sec.window$.

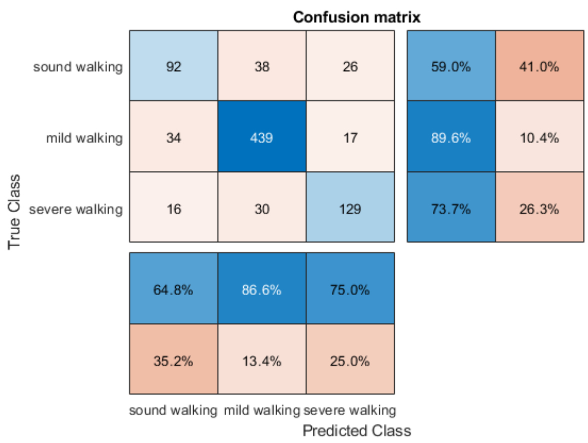
Appendix G. Confusion Matrices of Ensemble Classifiers for Sheep DataSets



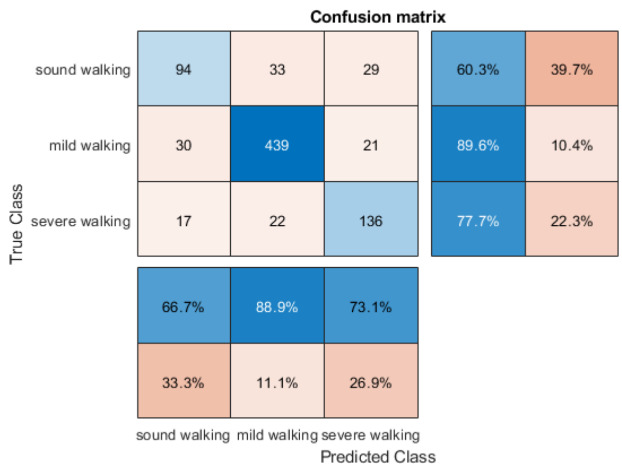
(a) Bagging ensemble /20- features ReliefF



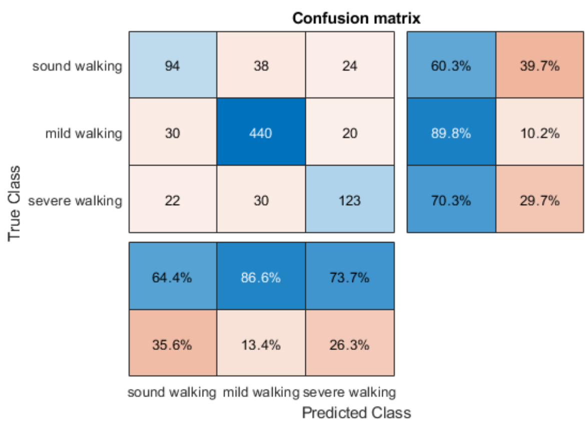
(b) Boosting ensemble /20- features ReliefF



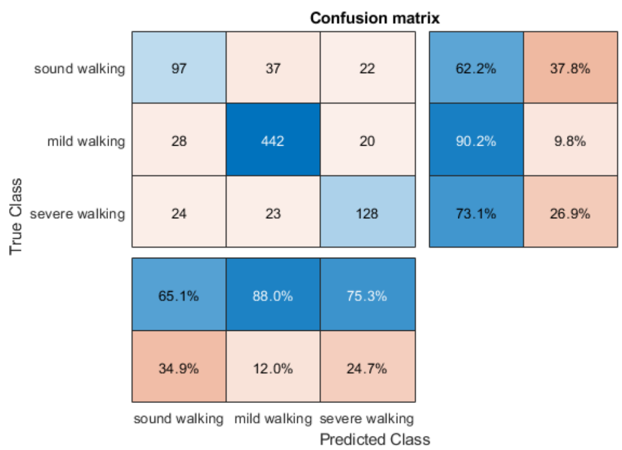
(c) Bagging ensemble /81- features GA



(d) Boosting ensemble /81- features GA

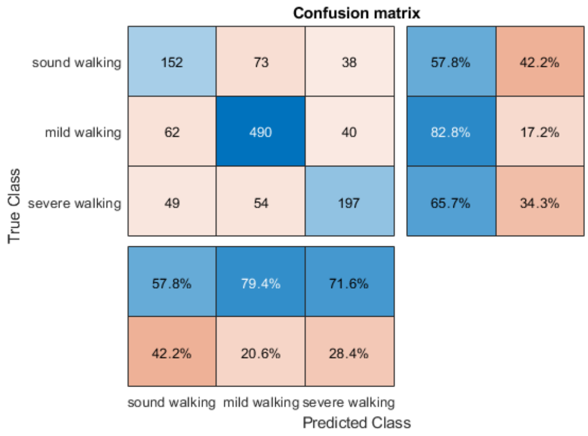


(e) Bagging ensemble /20- features RF

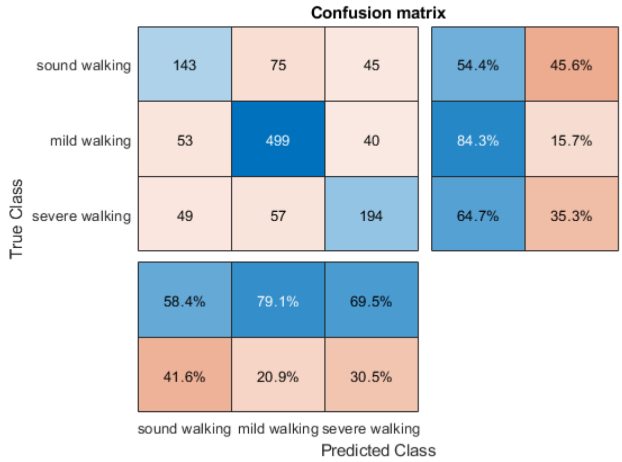


(f) Boosting ensemble /20- features RF

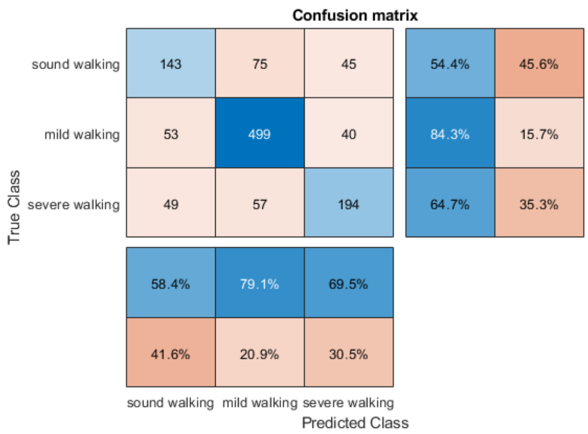
Appendix G. 1 Confusion Matrices for **5-fold** validation method of **DataSet1_all**, FOSW segmentation method over **10** $sec. window$.



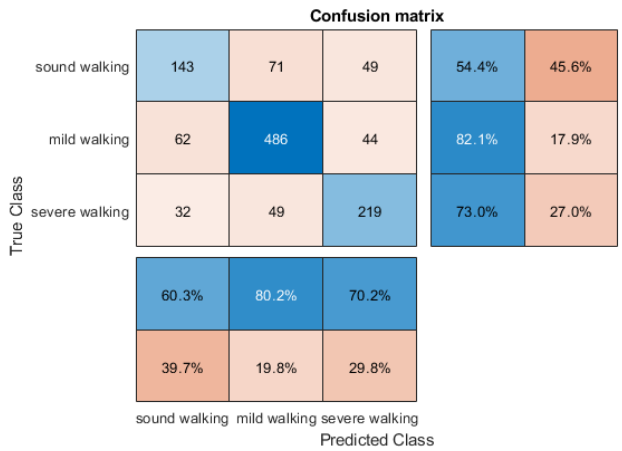
(a) Bagging ensemble /20- features ReliefF



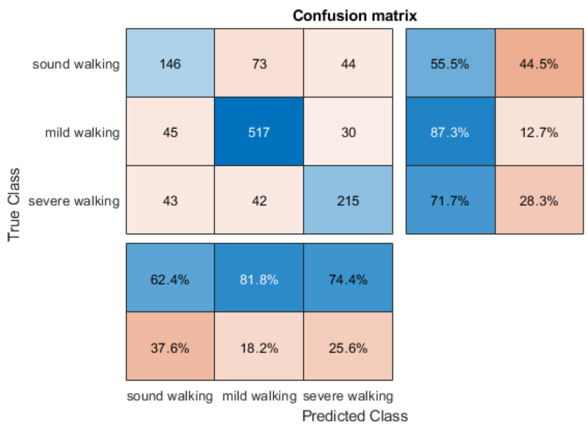
(b) Boosting ensemble /20- features ReliefF



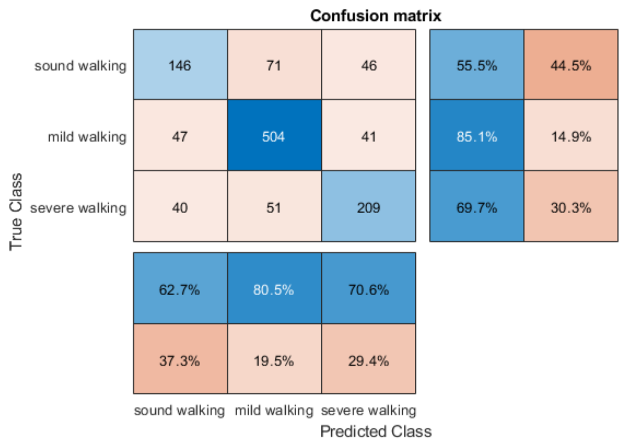
(c) Bagging ensemble /74- features GA



(d) Boosting ensemble /74- features GA

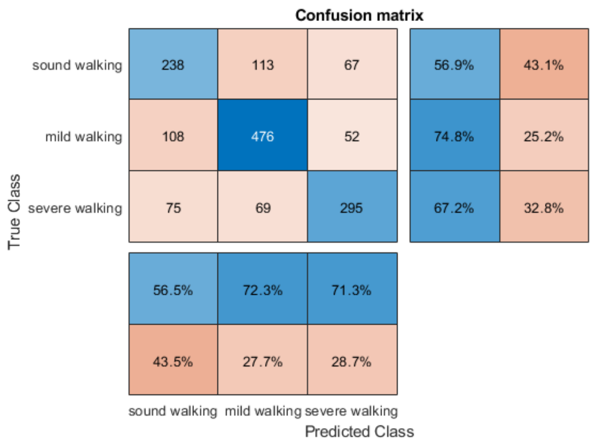


(e) Bagging ensemble /20- features RF

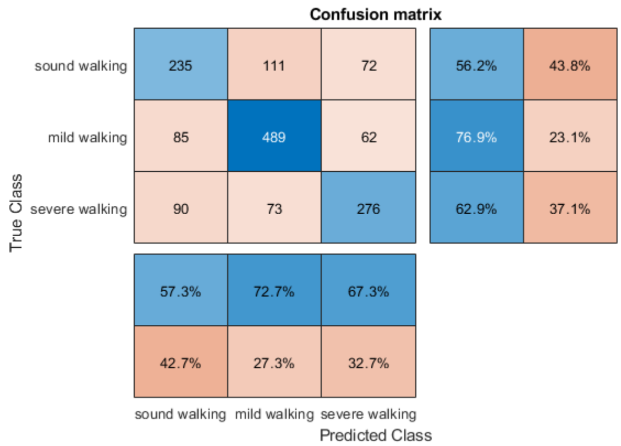


(f) Boosting ensemble /20- features RF

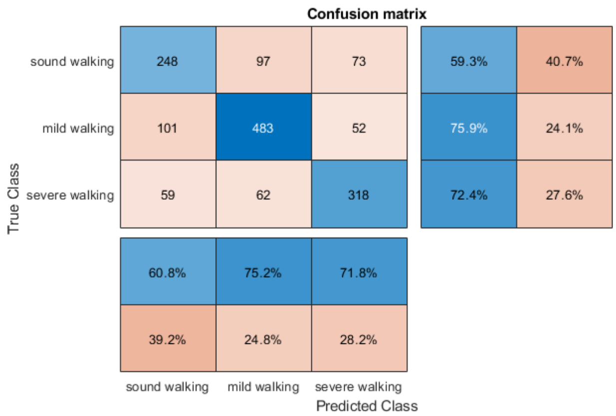
Appendix G. 2 Confusion Matrices for **5-fold** validation method of **DataSet1_all**, FOSW segmentation method over **7** $sec.window$.



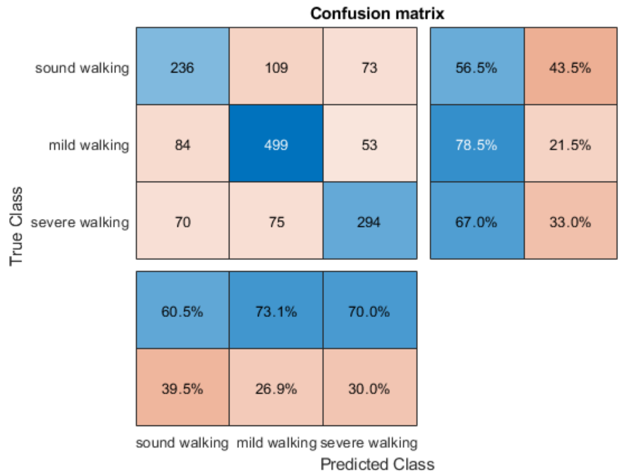
(a) Bagging ensemble /20- features ReliefF



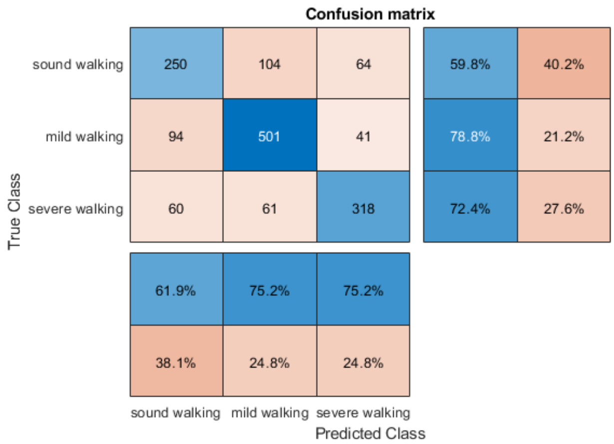
(b) Boosting ensemble /20- features ReliefF



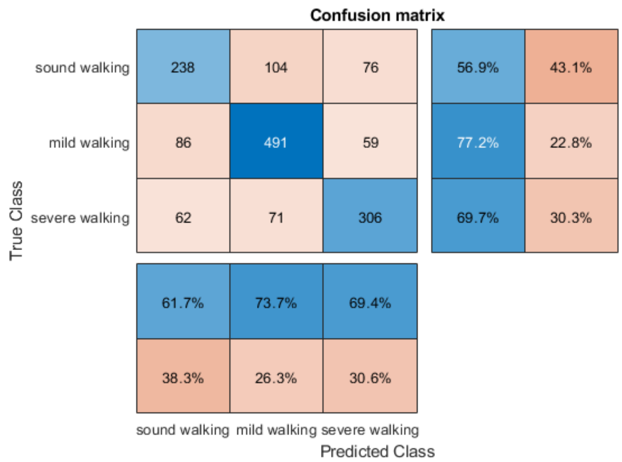
(c) Bagging ensemble /52- features GA



(d) Boosting ensemble /52- features GA

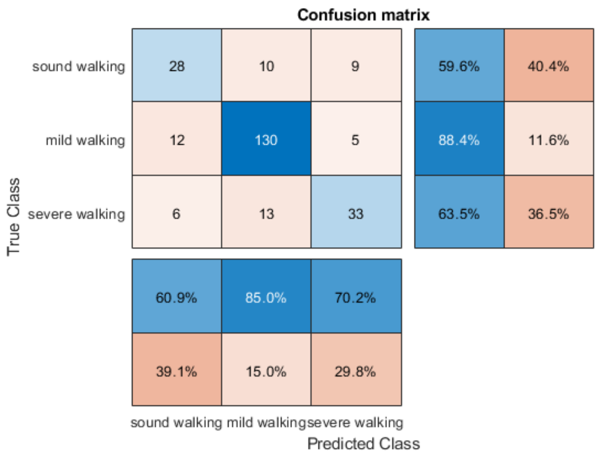


(e) Bagging ensemble /20- features RF

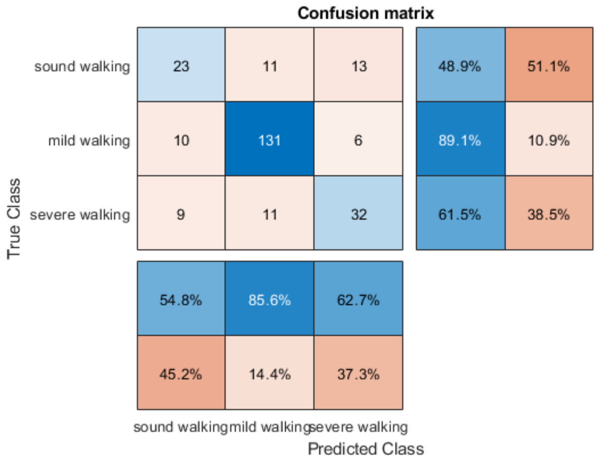


(f) Boosting ensemble /20- features RF

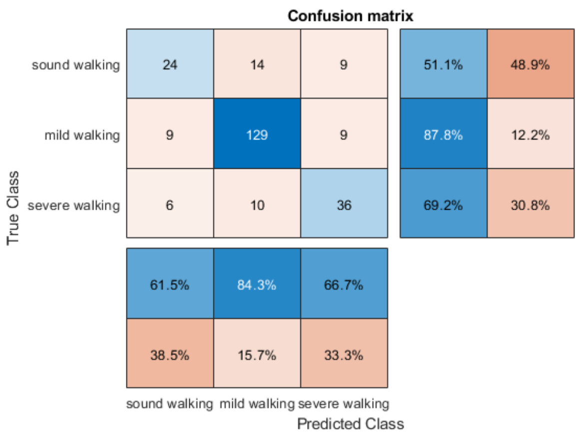
Appendix G. 3 Confusion Matrices for **5-fold** validation method of **DataSet1_all**, FOSW segmentation method over **5** $sec. window$.



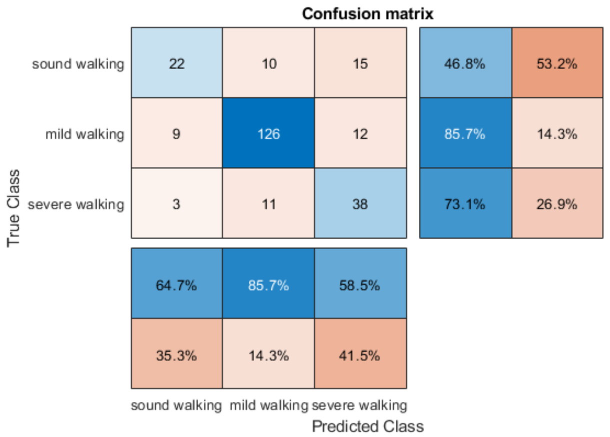
(a) Bagging ensemble /20- features ReliefF



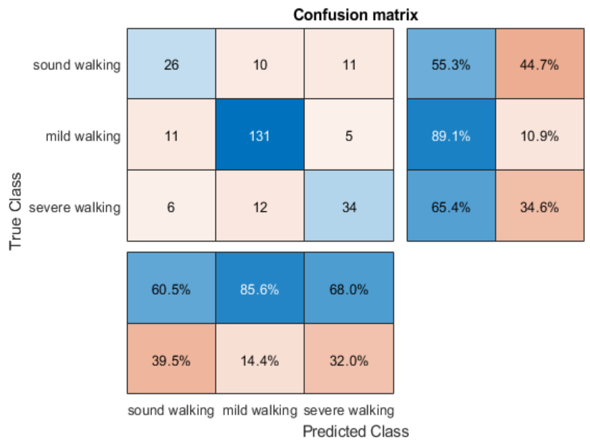
(b) Boosting ensemble /20- features ReliefF



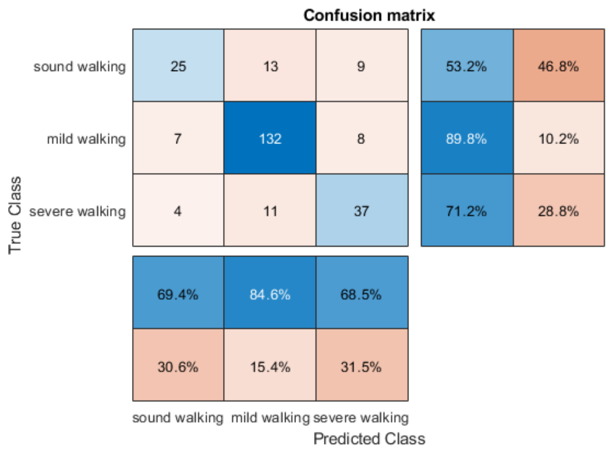
(c) Bagging ensemble /81- features GA



(d) Boosting ensemble /81- features GA

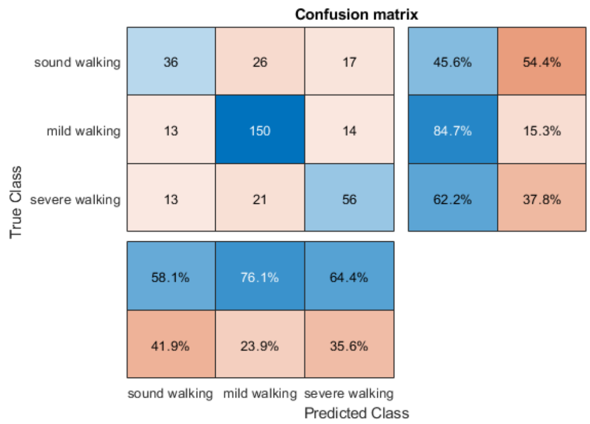


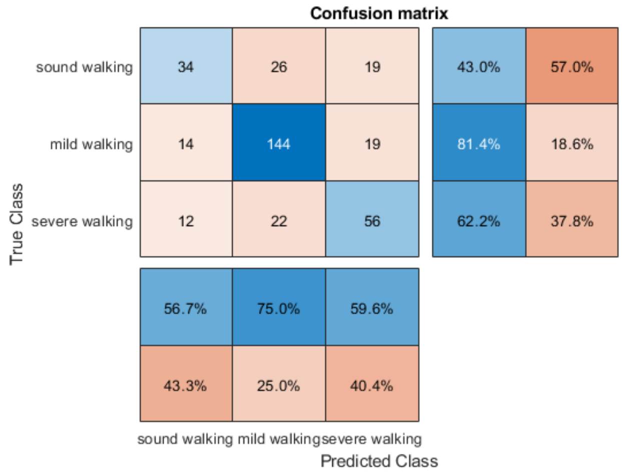
(e) Bagging ensemble /20- features RF



(f) Boosting ensemble /20- features RF

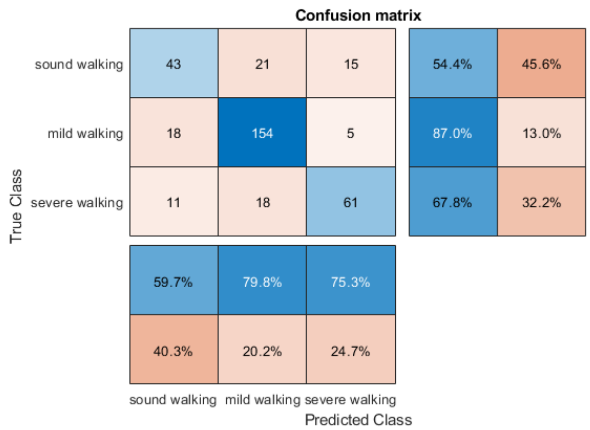
Appendix G. 4 Confusion Matrices for **0.3 hold-out** validation method of **DataSet1_all**, FOSW segmentation method over **10** $sec. window$.

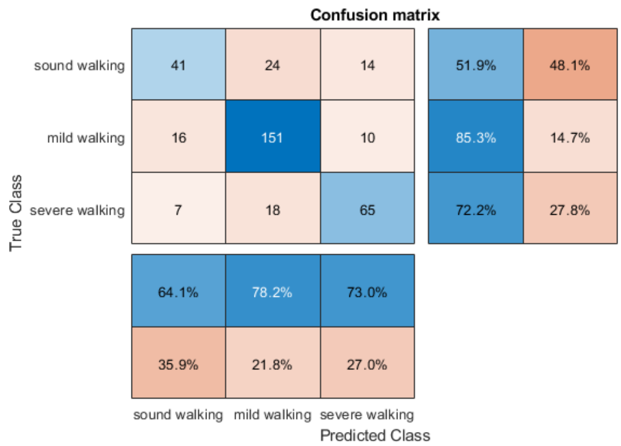




(a) Bagging ensemble /20- features ReliefF

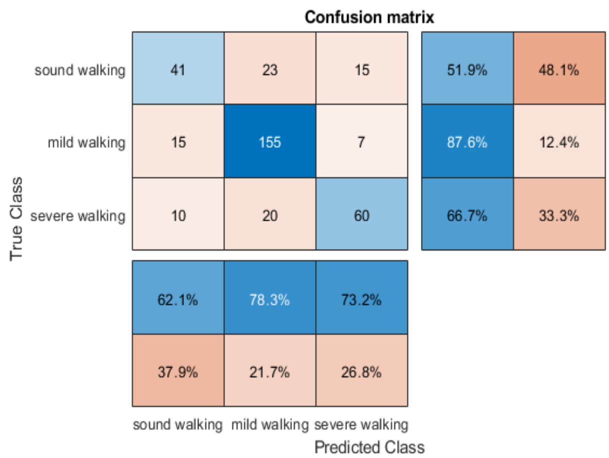
(b) Boosting ensemble /20- features ReliefF

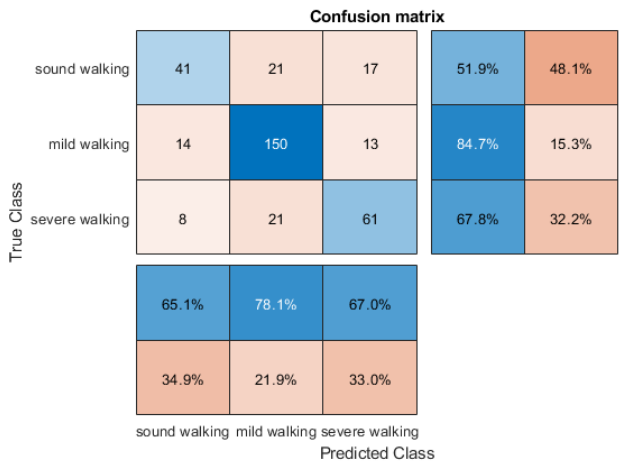




(c) Bagging ensemble /74- features GA

(d) Boosting ensemble /74- features GA

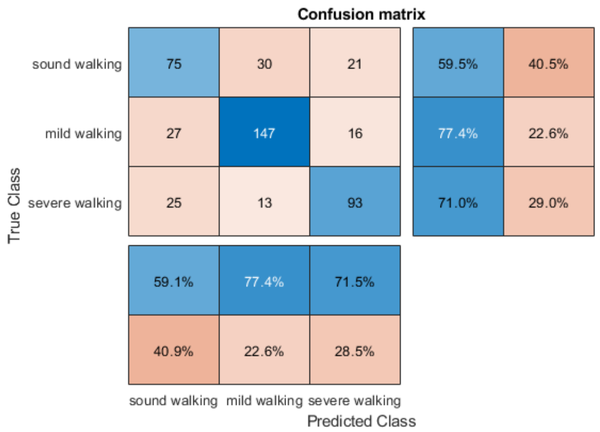




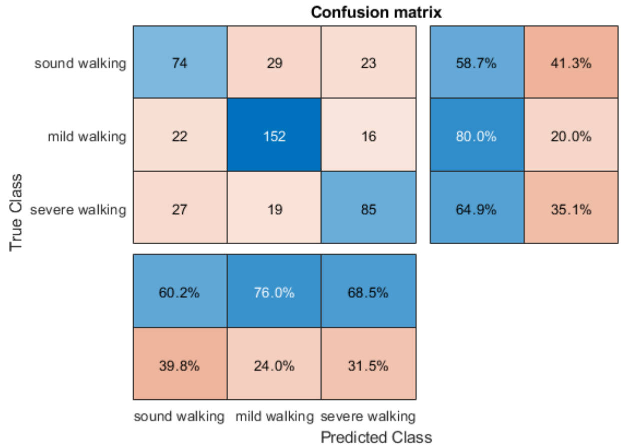
(e) Bagging ensemble /20- features RF

(f) Boosting ensemble /20- features RF

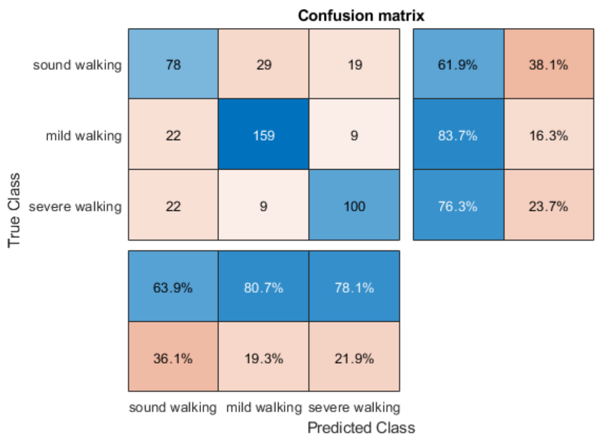
Appendix G. 5 Confusion Matrices for **0.3 hold-out** validation method of **DataSet1_all**, FOSW segmentation method over **7** $sec.window$.



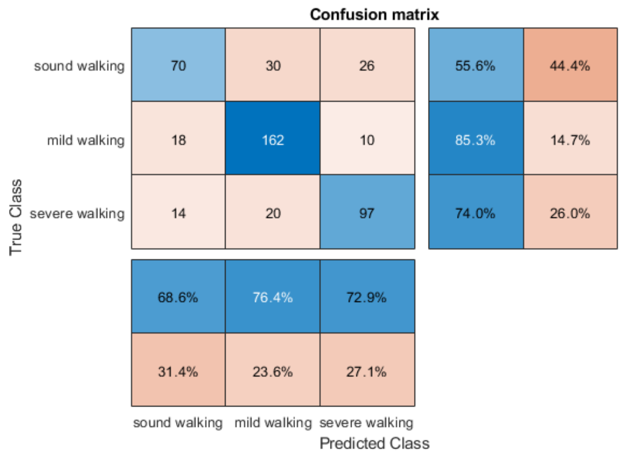
(a) Bagging ensemble /20- features ReliefF



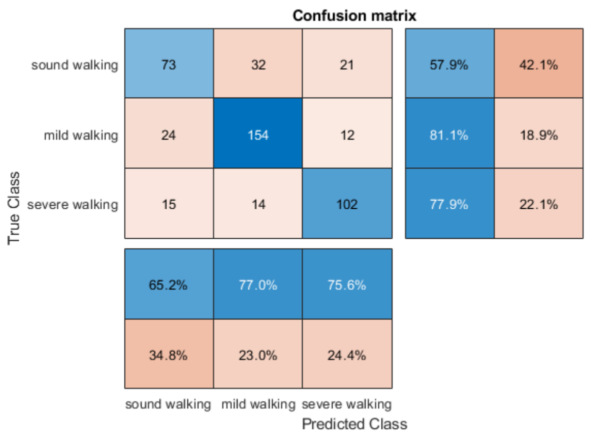
(b) Boosting ensemble /20- features ReliefF



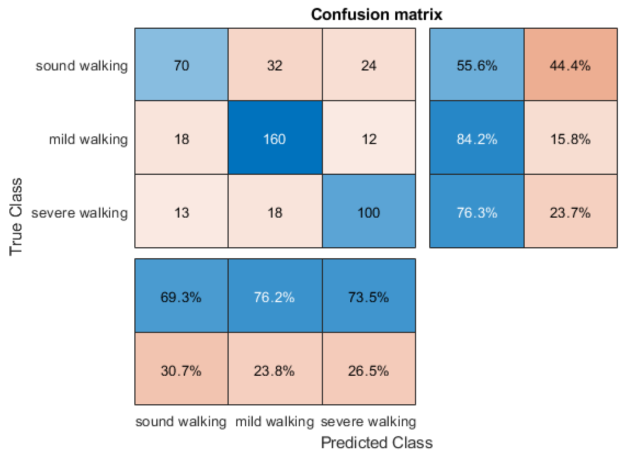
(c) Bagging ensemble /52- features GA



(d) Boosting ensemble /52- features GA

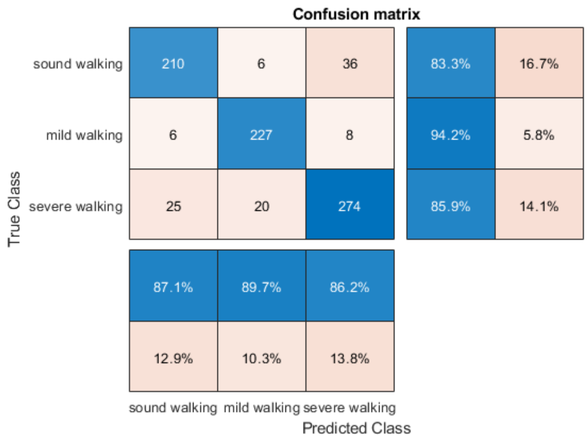


(e) Bagging ensemble /20- features RF

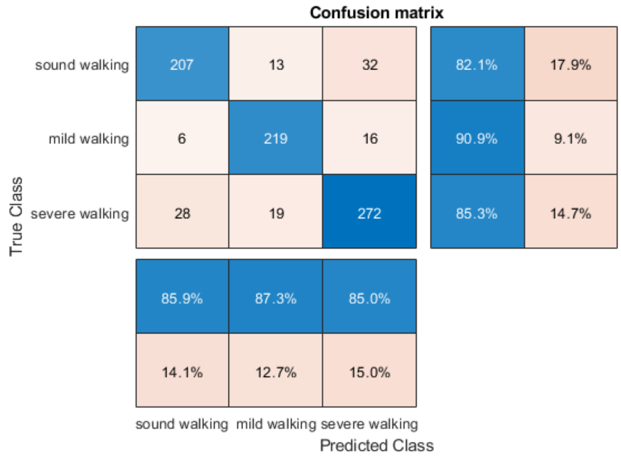


(f) Boosting ensemble /20- features RF

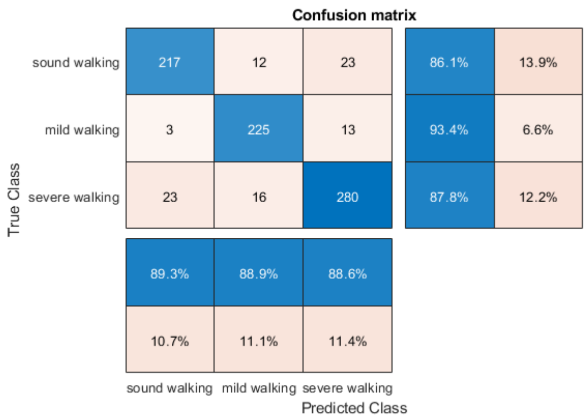
Appendix G. 6 Confusion Matrices for **0.3 hold-out** validation method of **DataSet1_all**, FOSW segmentation method over **5** $sec.window$.



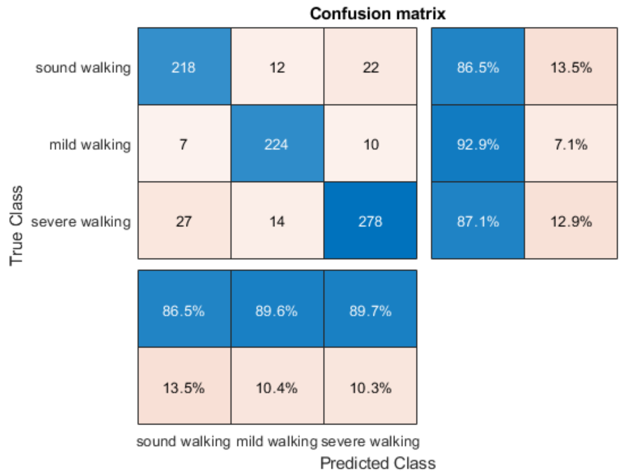
(a) Bagging ensemble /20 - features ReliefF



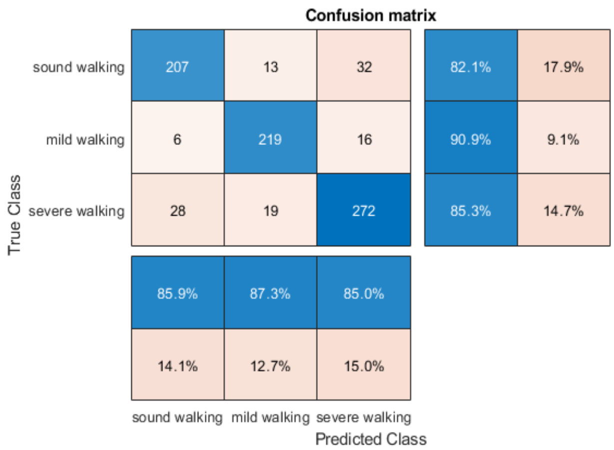
(b) Boosting ensemble /20 - features ReliefF



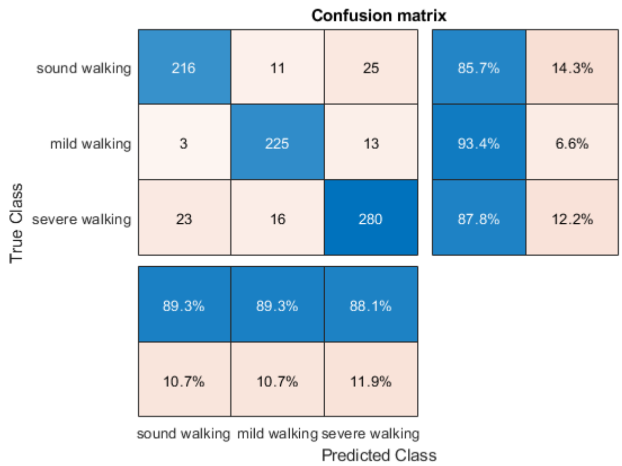
(c) Bagging ensemble /46 - features GA



(d) Boosting ensemble /46 - features GA

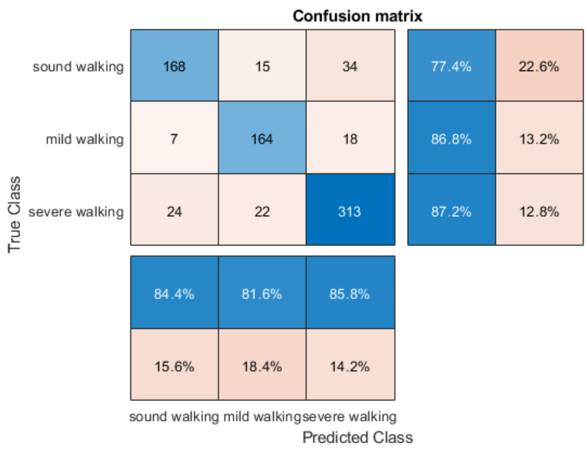


(e) Bagging ensemble /20 - features RF

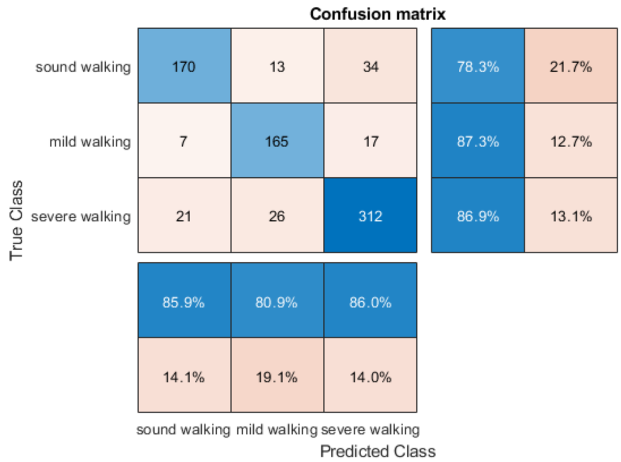


(f) Boosting ensemble /20 - features RF

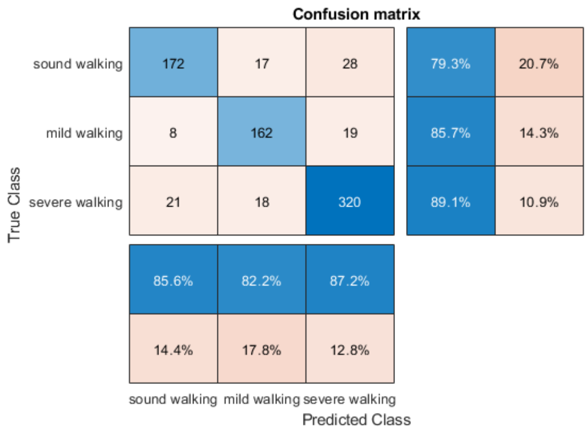
Appendix G. 7 Confusion Matrices for **5-fold** validation method of **DataSet2_ac**, FOSW segmentation method over **10** $sec.window$.



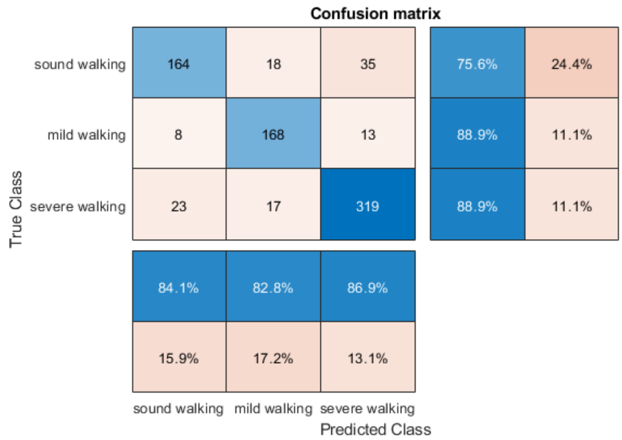
(a) Bagging ensemble /20 - features ReliefF



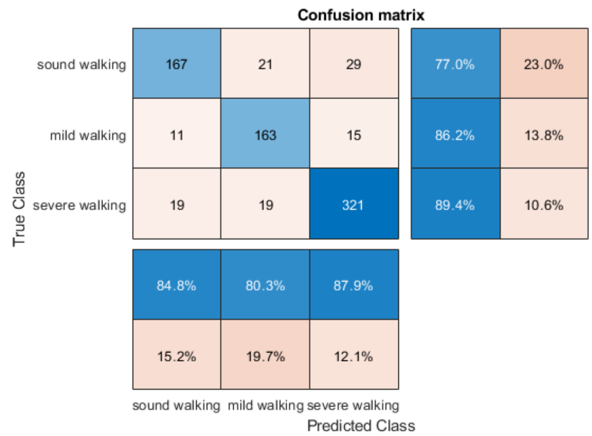
(b) Boosting ensemble /20 - features ReliefF



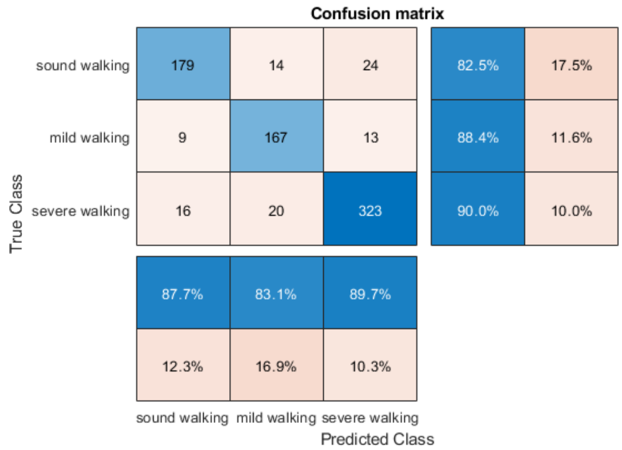
(c) Bagging ensemble /38 - features GA



(d) Boosting ensemble /38 - features GA

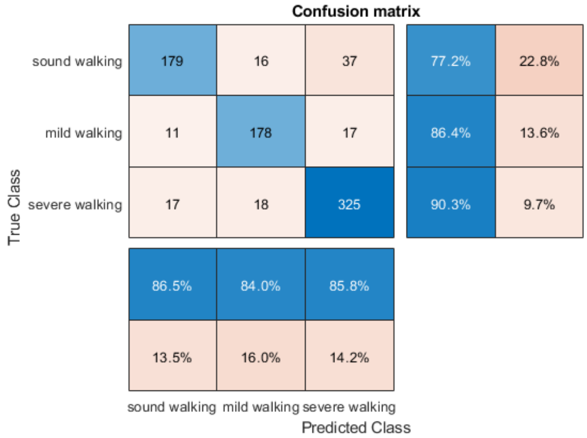


(e) Bagging ensemble /20 - features RF

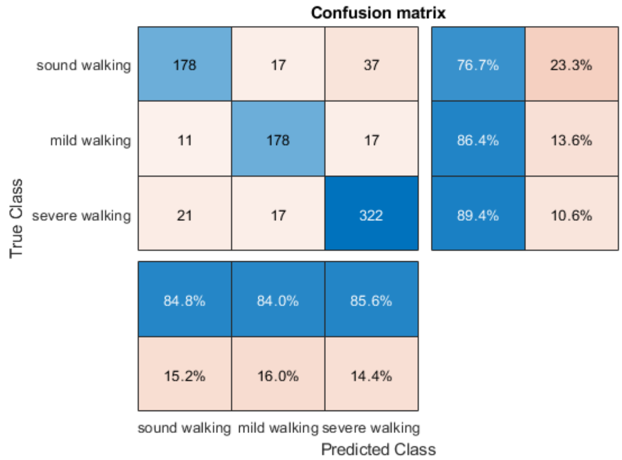


(f) Boosting ensemble /20 - features RF

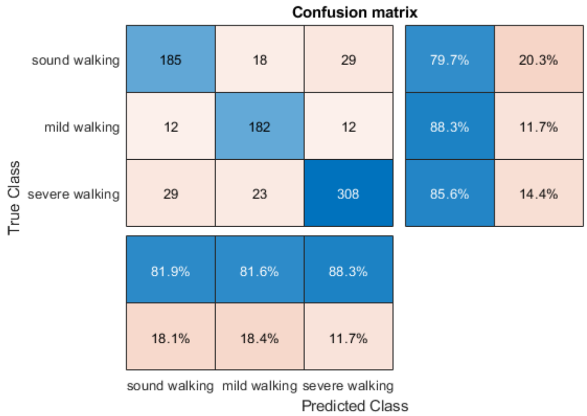
Appendix G. 8 Confusion Matrices for **5-fold** validation method of **DataSet2_ac**, FOSW segmentation method over **7** $sec. window$.



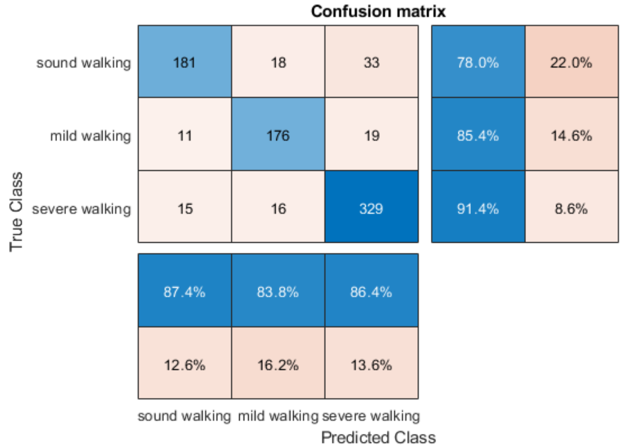
(a) Bagging ensemble /20 - features ReliefF



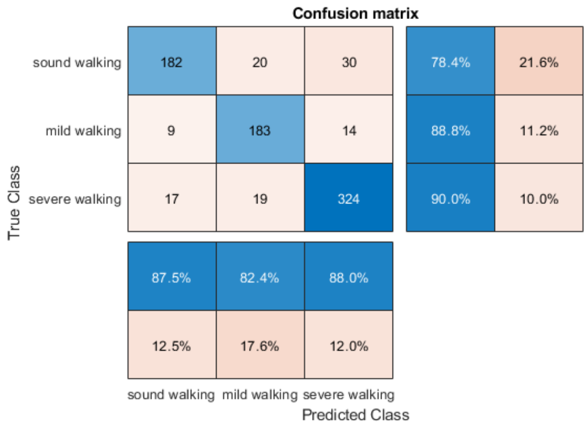
(b) Boosting ensemble /20 - features ReliefF



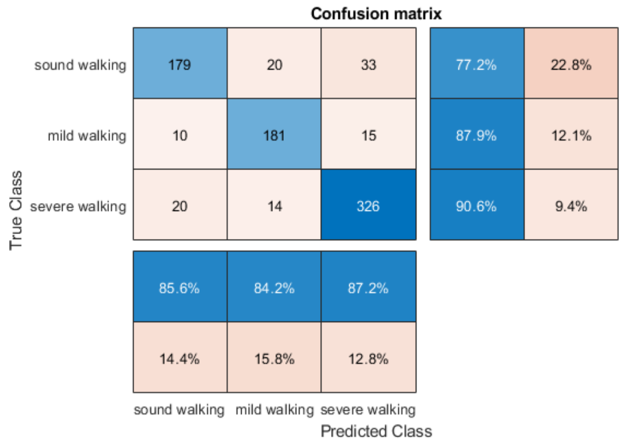
(c) Bagging ensemble /39 - features GA



(d) Boosting ensemble /39 - features GA

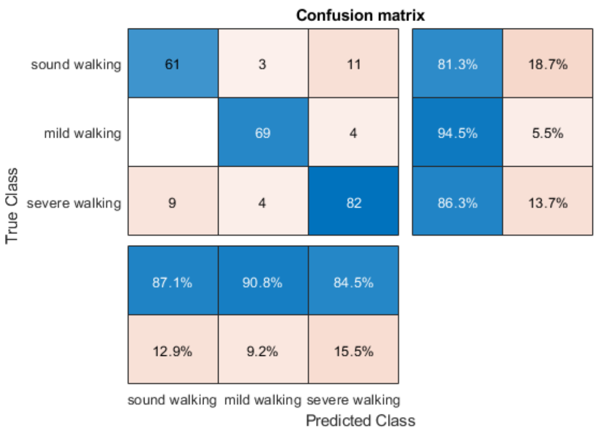


(e) Bagging ensemble /20 - features RF

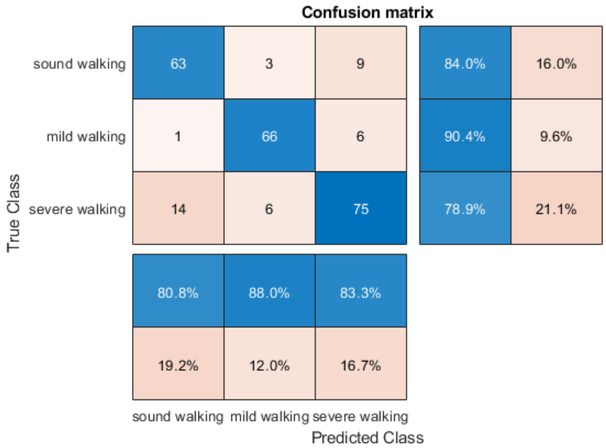


(f) Boosting ensemble /20 - features RF

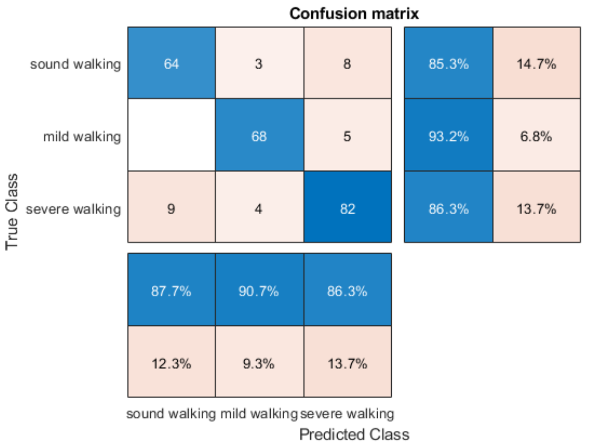
Appendix G. 9 Confusion Matrices for **5-fold** validation method of **DataSet2_ac**, FOSW segmentation method over **5** $sec. window$.



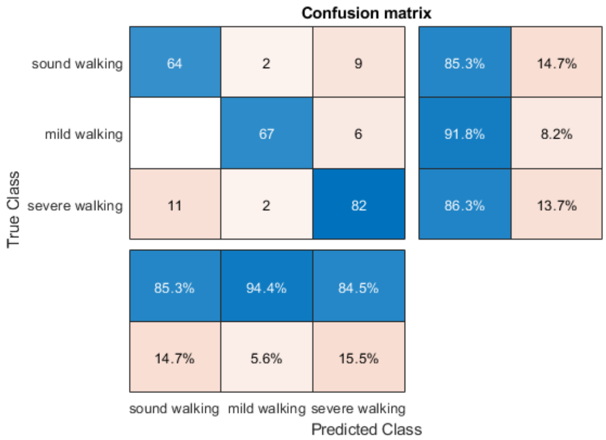
(a) Bagging ensemble /20 - features ReliefF



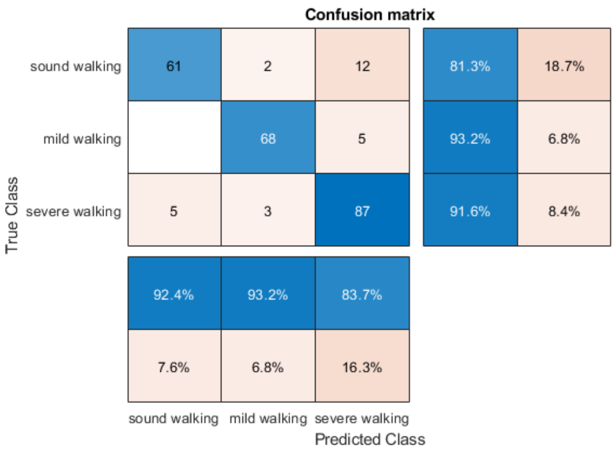
(b) Boosting ensemble /20 - features ReliefF



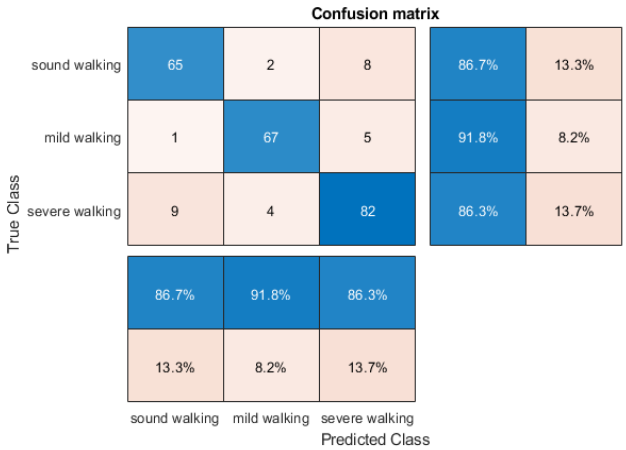
(c) Bagging ensemble /46 - features GA



(d) Boosting ensemble /46 - features GA

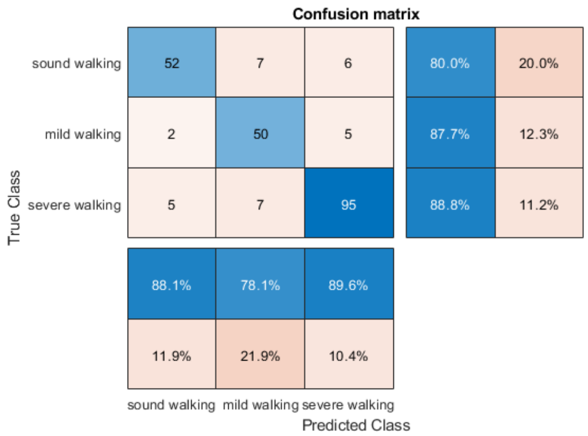


(e) Bagging ensemble /20 - features RF

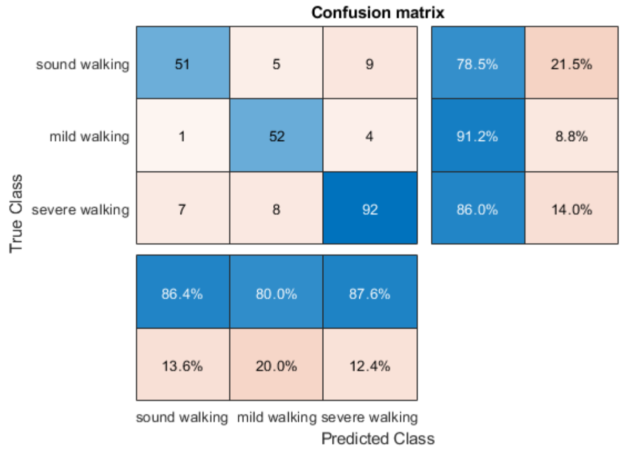


(f) Boosting ensemble /20 - features RF

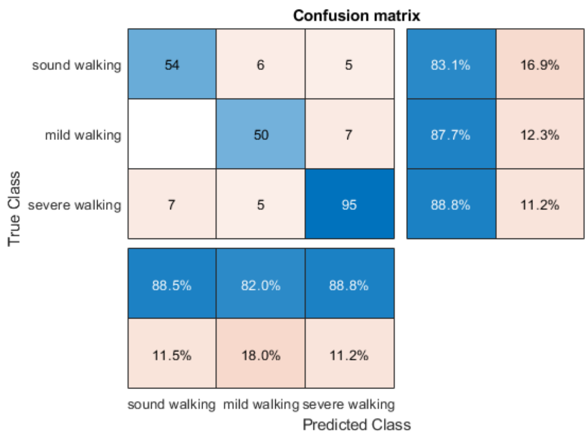
Appendix G. 10 Confusion Matrices for **0.3 hold-out** validation method of **DataSet2_ac**, FOSW segmentation method over **10** $sec. window$.



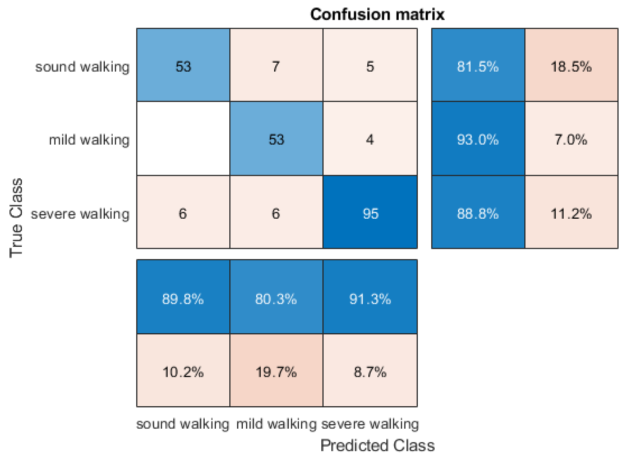
(a) Bagging ensemble /20 - features ReliefF



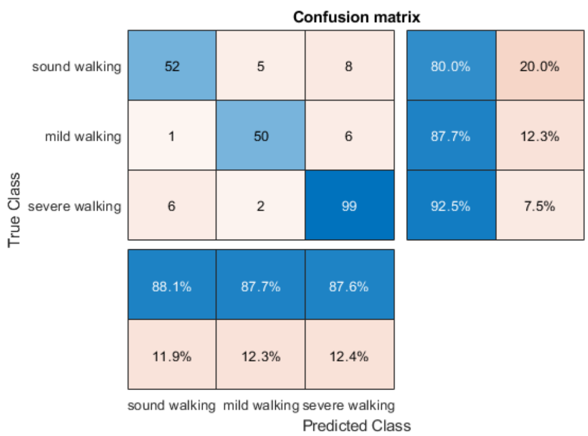
(b) Boosting ensemble /20 - features ReliefF



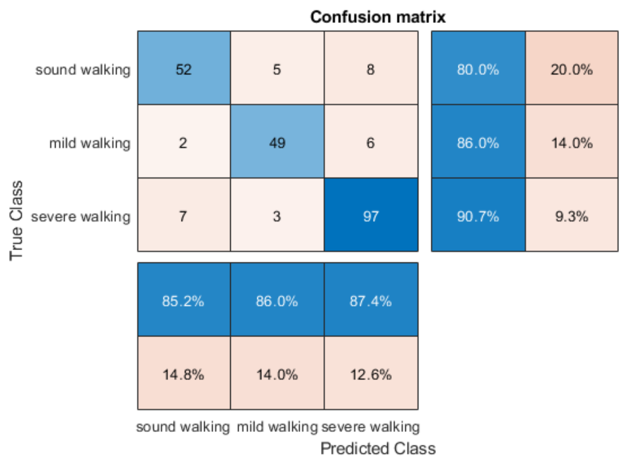
(c) Bagging ensemble /38 - features GA



(d) Boosting ensemble /38 - features GA

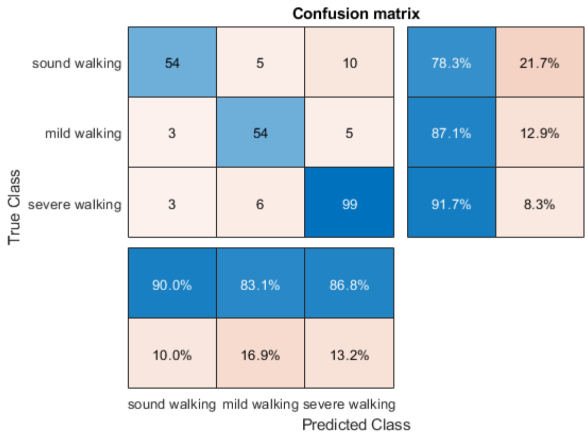


(e) Bagging ensemble /20 - features RF

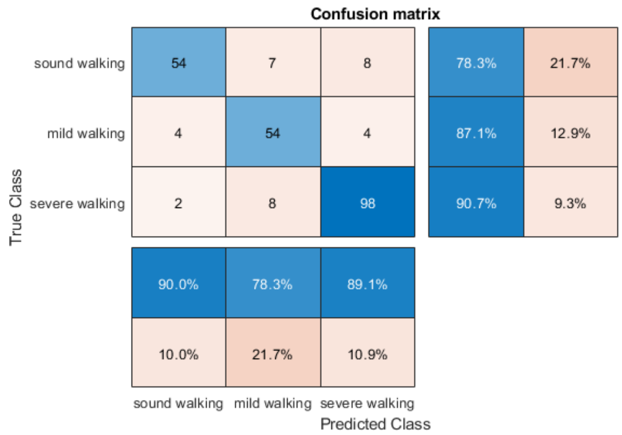


(f) Boosting ensemble /20 - features RF

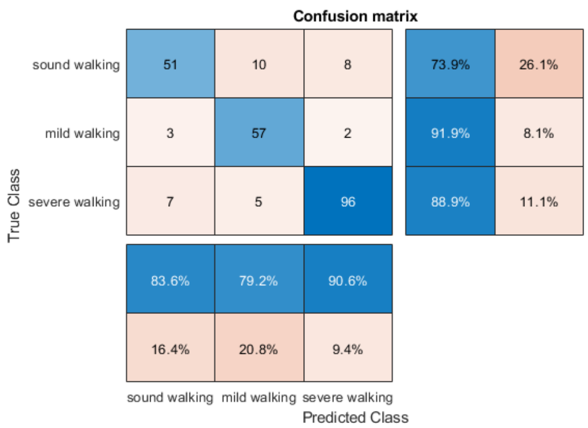
Appendix G. 11 Confusion Matrices for **0.3 hold-out** validation method of **DataSet2_ac**, FOSW segmentation method over **7** $sec.window$.



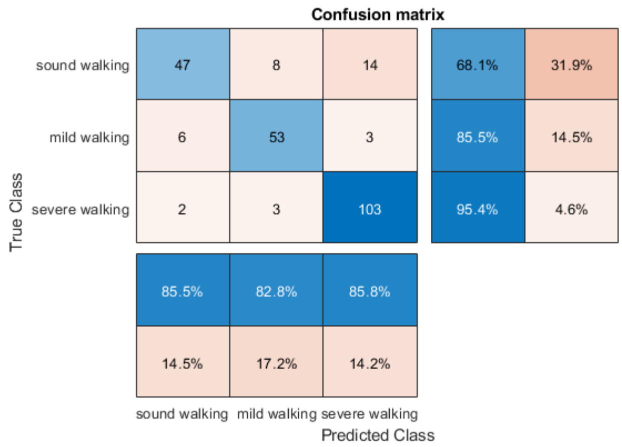
(g) Bagging ensemble /20- features ReliefF



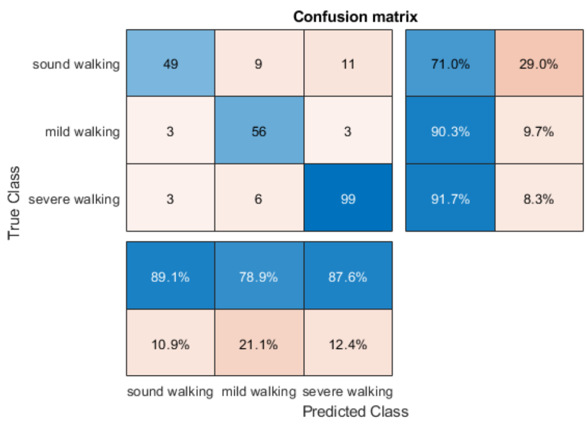
(h) Boosting ensemble /20- features ReliefF



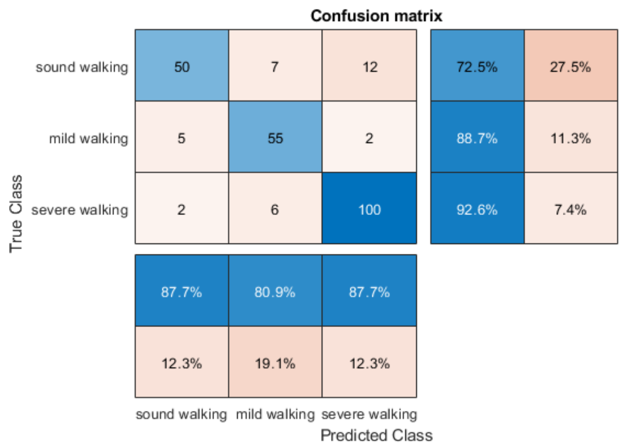
(i) Bagging ensemble /39- features GA



(j) Boosting ensemble /39- features GA

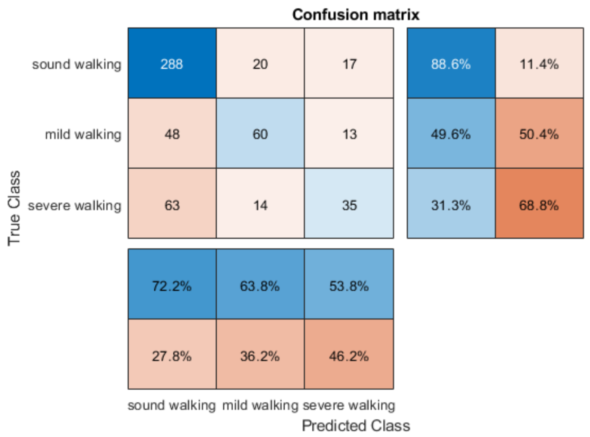


(k) Bagging ensemble /20- features RF

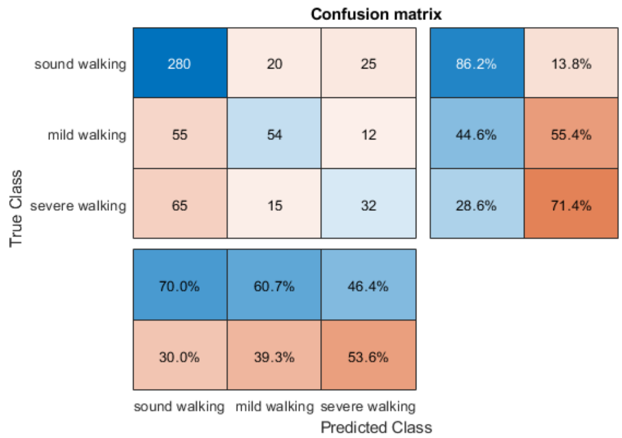


(l) Boosting ensemble /20- features RF

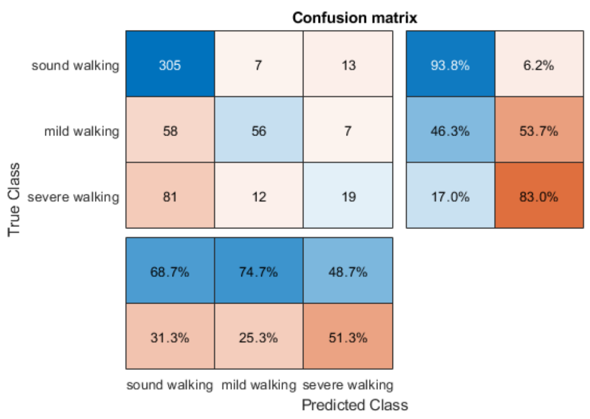
Appendix G. 12 Confusion Matrices for **0.3 hold-out** validation method of **DataSet2_ac**, FOSW segmentation method over **5** $sec.window$.



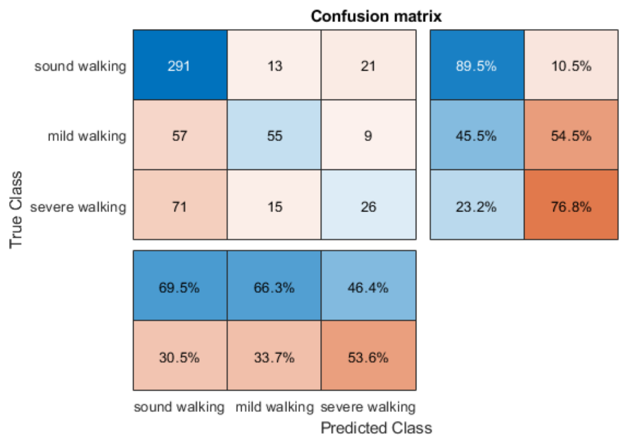
(a) Bagging ensemble /20 - features ReliefF



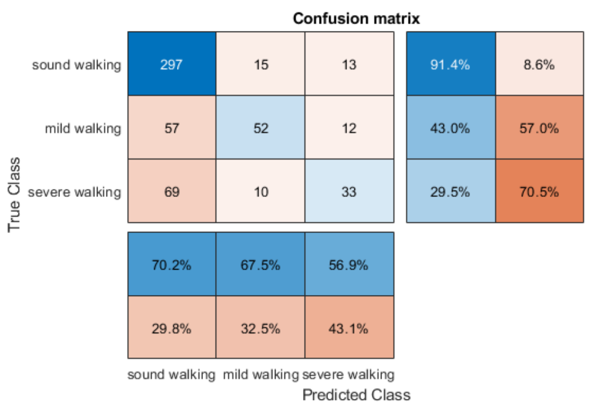
(b) Boosting ensemble /20 - features ReliefF



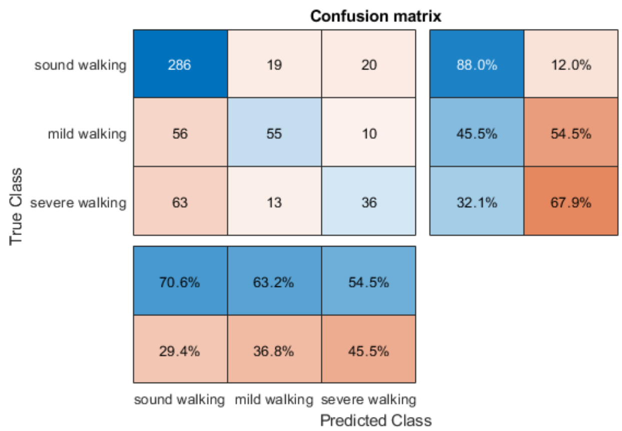
(c) Bagging ensemble /64 - features GA



(d) Boosting ensemble /64 - features GA

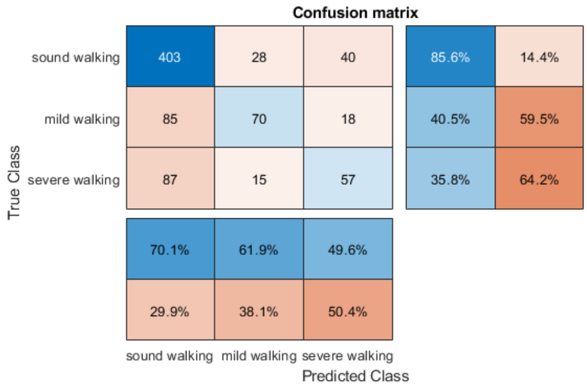


(e) Bagging ensemble /20 - features RF

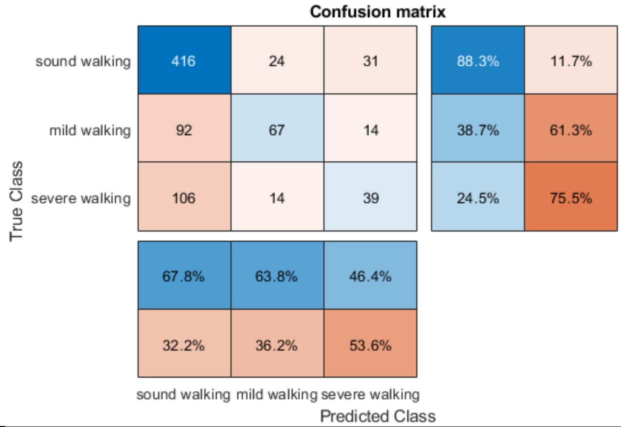


(f) Boosting ensemble /20 - features RF

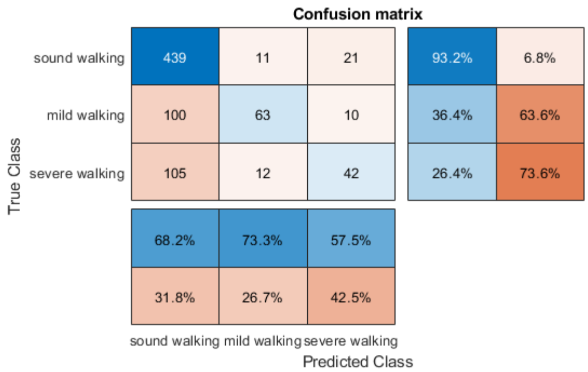
Appendix G. 13 Confusion Matrices for **5-fold** validation method of **DataSet2_b**, FOSW segmentation method over **10** $sec. window$.



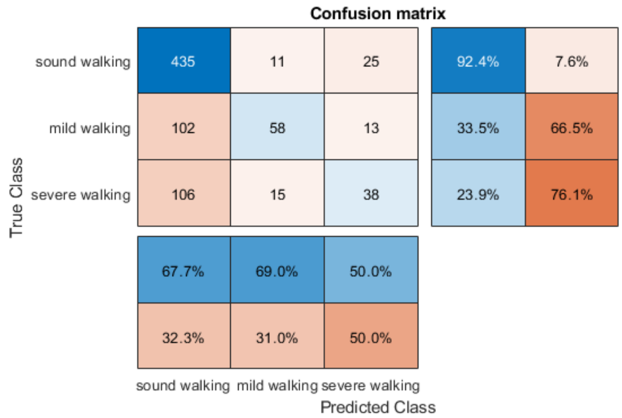
(a) Bagging ensemble /20 - features ReliefF



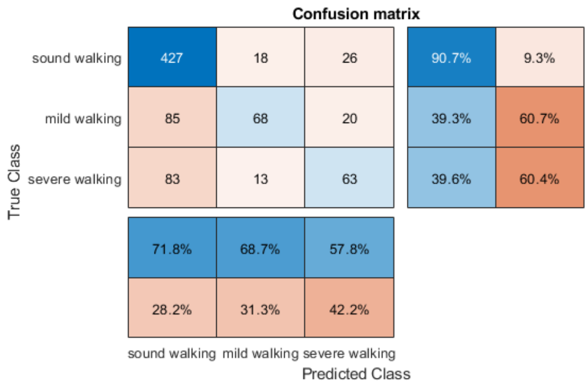
(b) Boosting ensemble /20 - features ReliefF



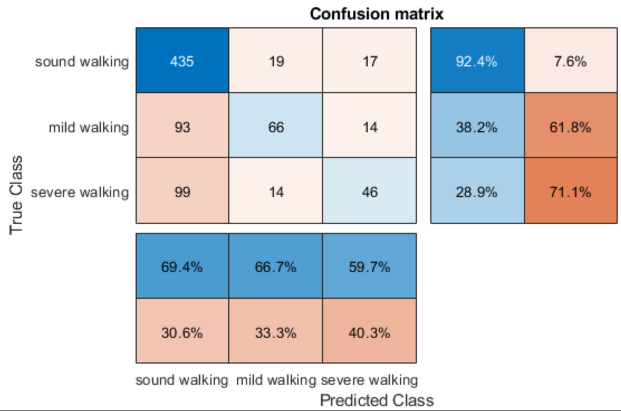
(c) Bagging ensemble /27 - features GA



(d) Boosting ensemble /27 - features GA

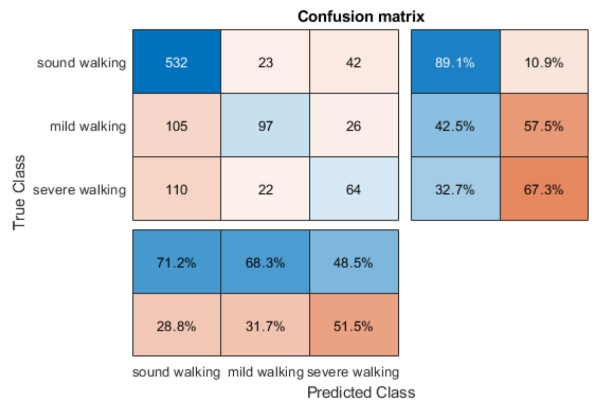


(e) Bagging ensemble /20 - features RF

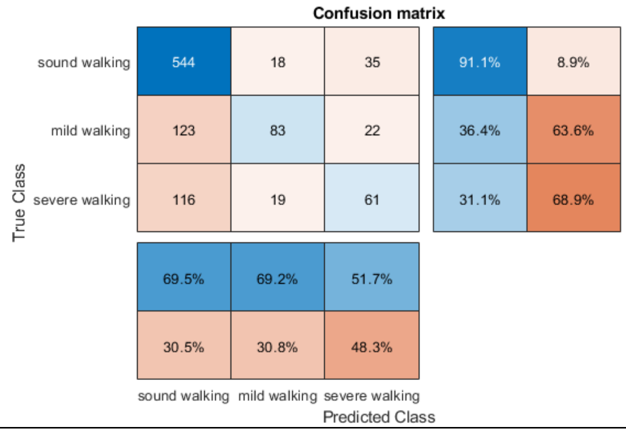


(f) Boosting ensemble /20 - features RF

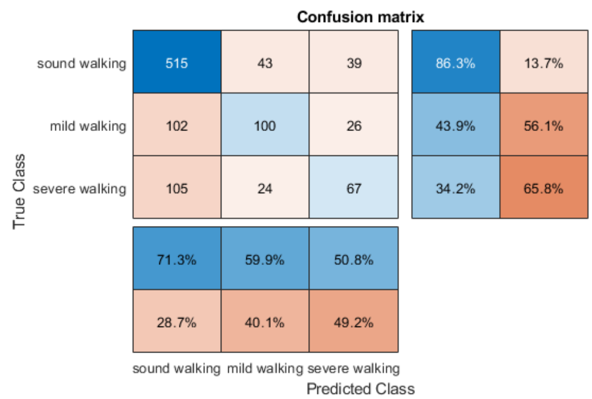
Appendix G. 14 Confusion Matrices for **5-fold** validation method of **DataSet2_b**, FOSW segmentation method over **7** $sec. window$.



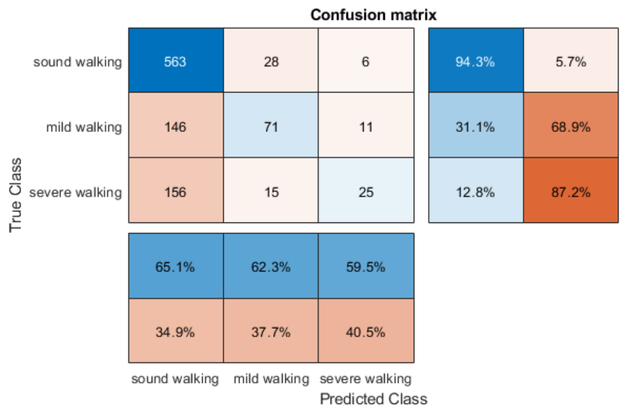
(a) Bagging ensemble /20 - features ReliefF



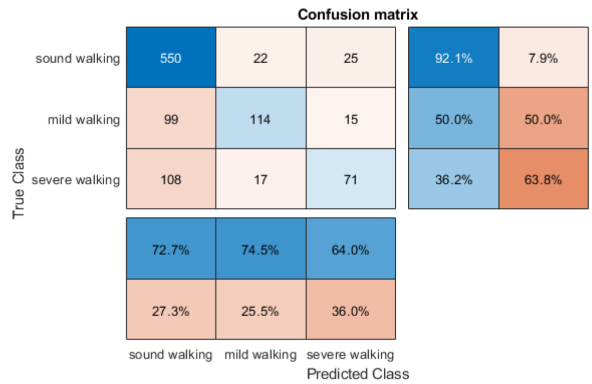
(b) Boosting ensemble /20 - features ReliefF



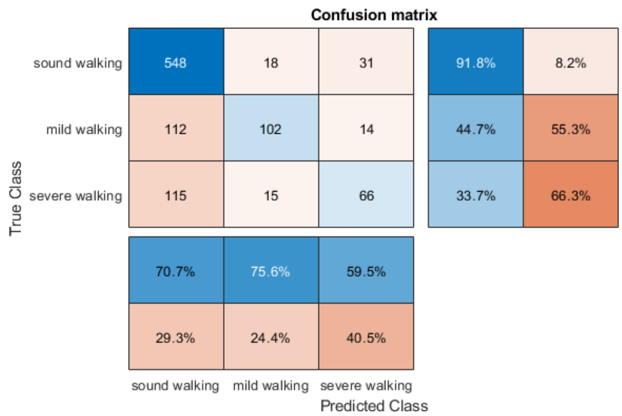
(c) Bagging ensemble /6 - features GA



(d) Boosting ensemble /6 - features GA

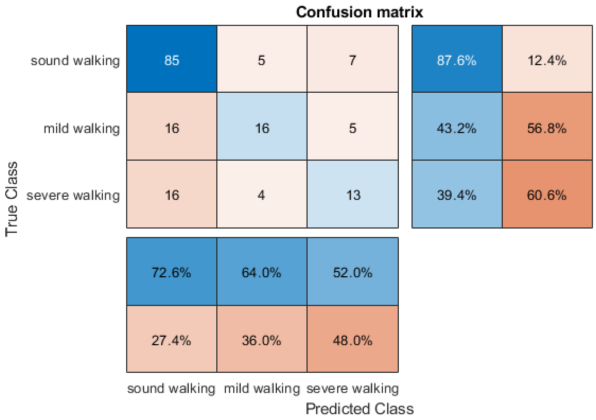


(e) Bagging ensemble /20 - features RF

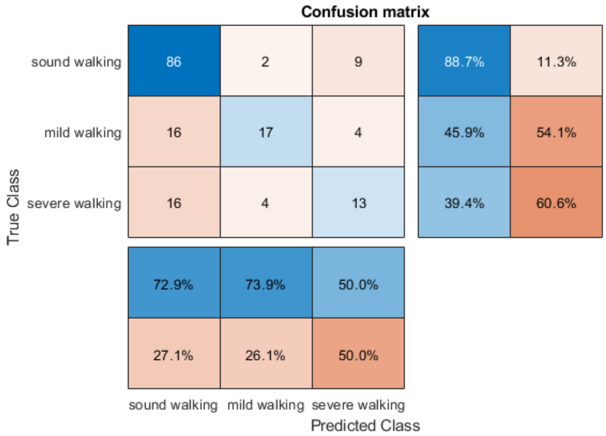


(f) Boosting ensemble /20 - features RF

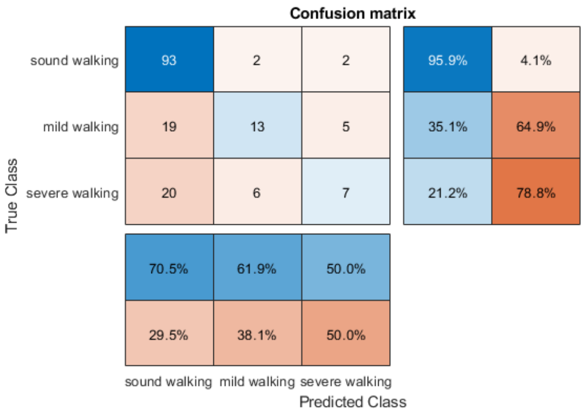
Appendix G. 15 Confusion Matrices for **5-fold** validation method of **DataSet2_b**, FOSW segmentation method over **5** $sec.window$.



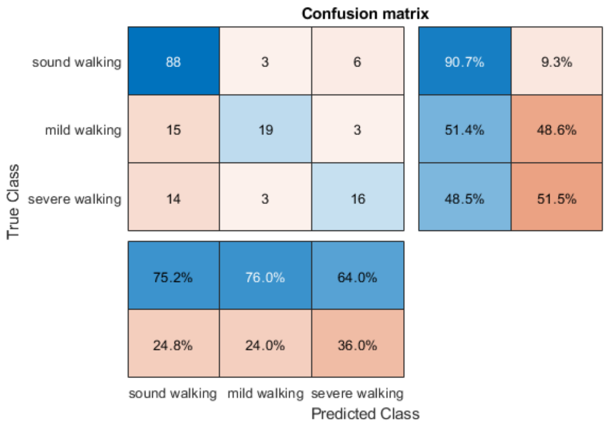
(a) Bagging ensemble /20 - features ReliefF



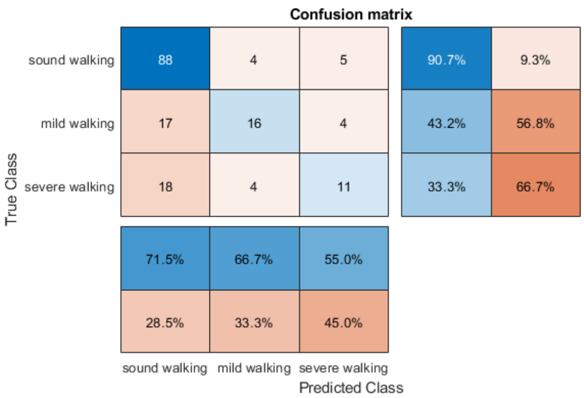
(b) Boosting ensemble /20 - features ReliefF



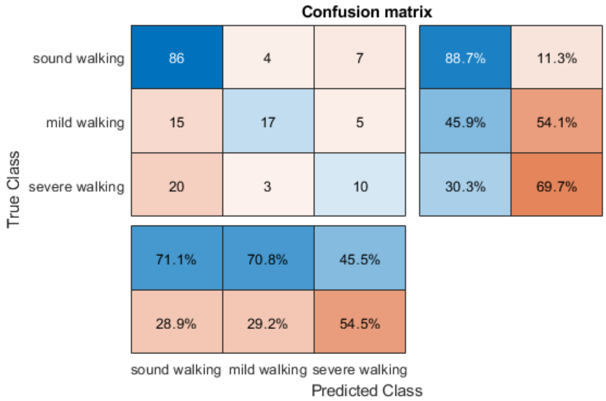
(c) Bagging ensemble /64 - features GA



(d) Boosting ensemble /64 - features GA

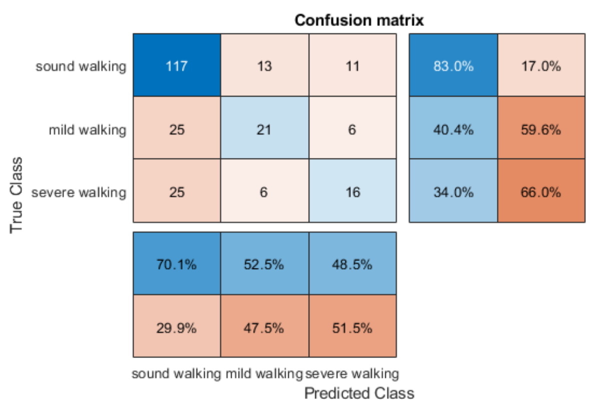


(e) Bagging ensemble /20 - features RF

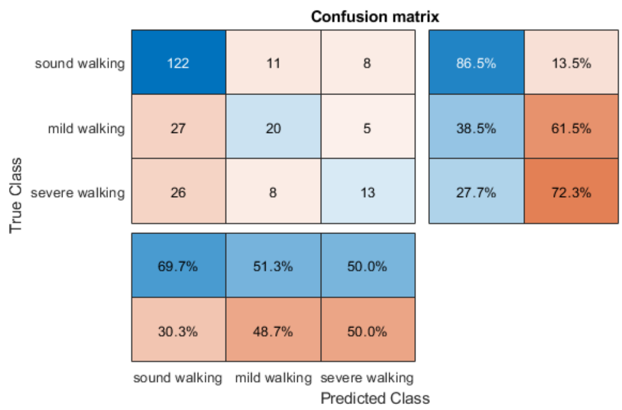


(f) Boosting ensemble /20 - features RF

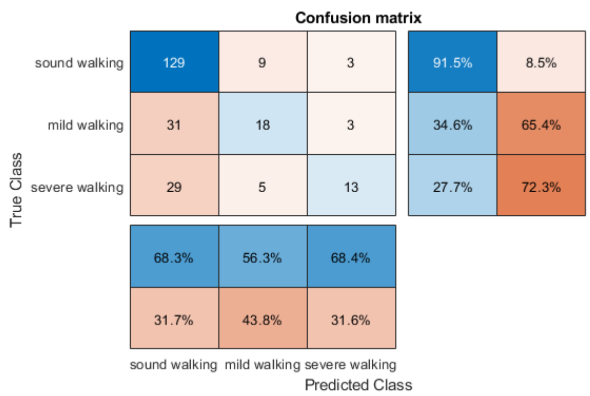
Appendix G. 16 Confusion Matrices for **0.3 hold-out** validation method of **DataSet2_b**, FOSW segmentation method over **10** $sec. window$.



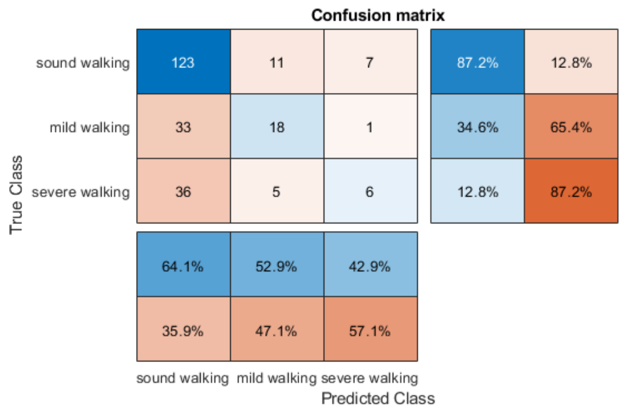
(a) Bagging ensemble /20 - features ReliefF



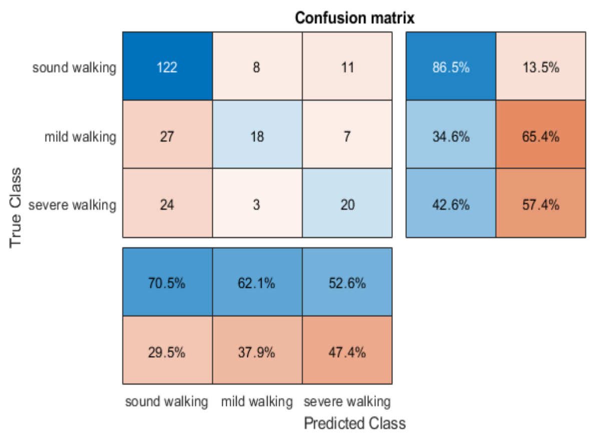
(b) Boosting ensemble /20 - features ReliefF



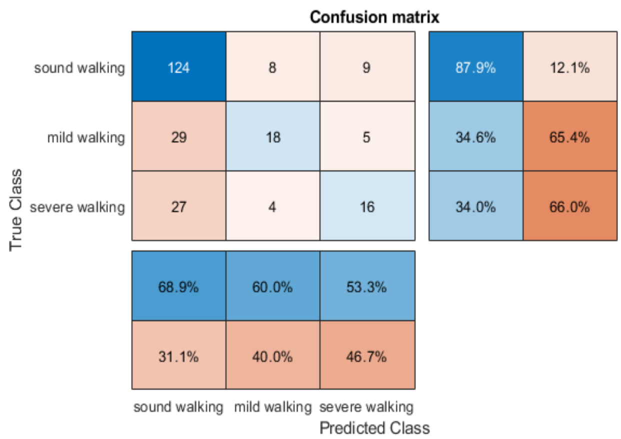
(c) Bagging ensemble /27 - features GA



(d) Boosting ensemble /27 - features GA

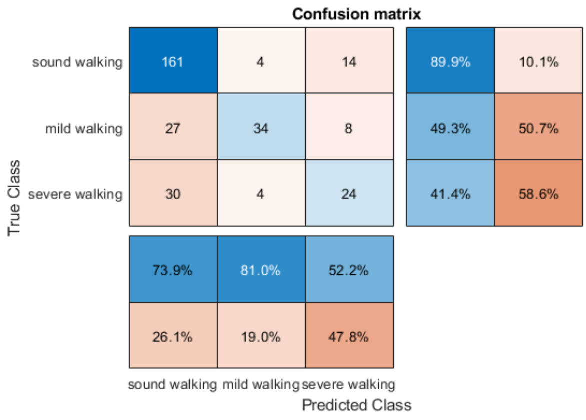


(e) Bagging ensemble /20 - features RF

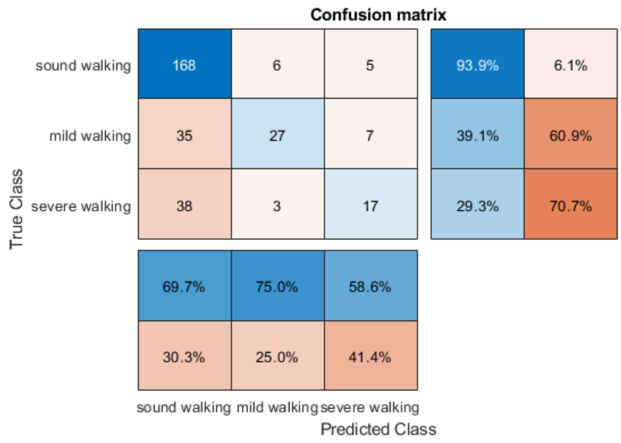


(f) Boosting ensemble /20 - features RF

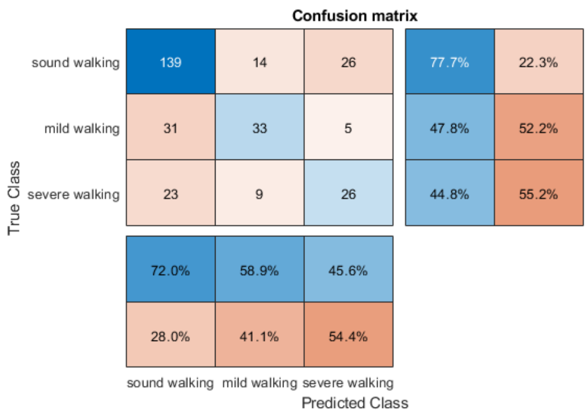
Appendix G. 17 Confusion Matrices for **0.3 hold-out** validation method of **DataSet2_b**, FOSW segmentation method over **7** $sec.window$.



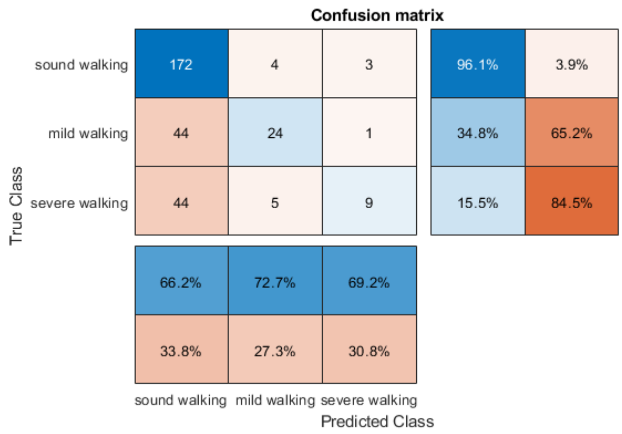
(a) Bagging ensemble /20- features ReliefF



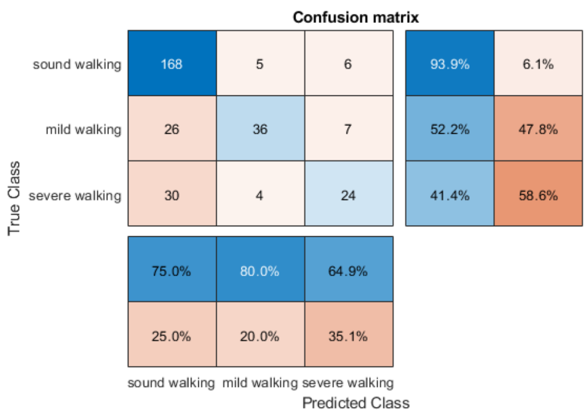
(b) Boosting ensemble /20- features ReliefF



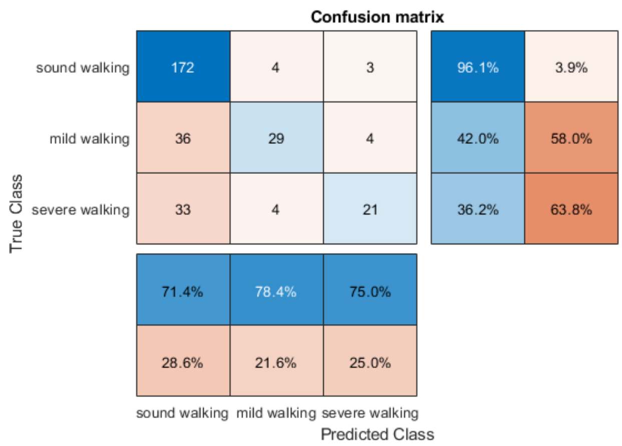
(c) Bagging ensemble /6 - features GA



(d) Boosting ensemble /6 - features GA

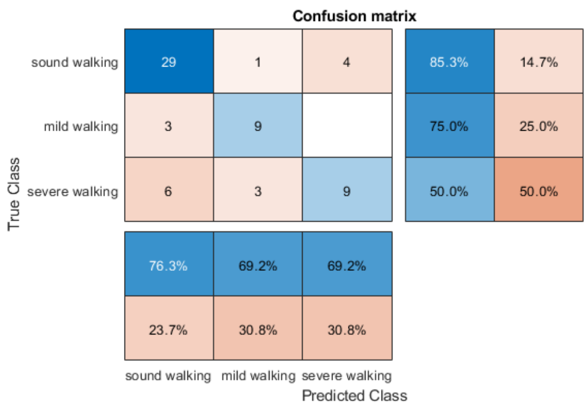


(e) Bagging ensemble /20- features RF

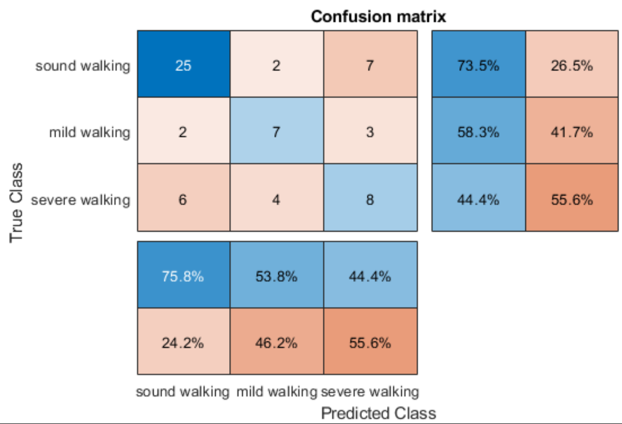


(f) Boosting ensemble /20- features RF

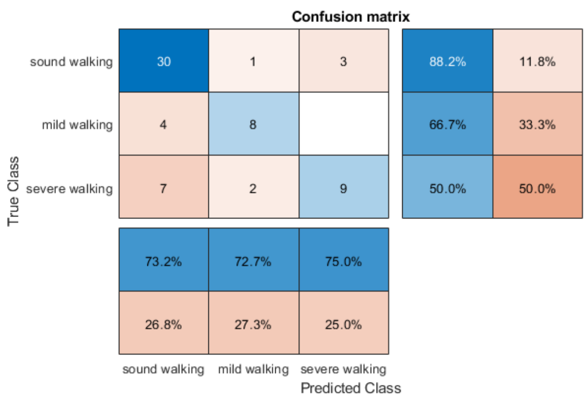
Appendix G. 18 Confusion Matrices for **0.3 hold-out** validation method of **DataSet2_b**, FOSW segmentation method over **5** $sec.window$.



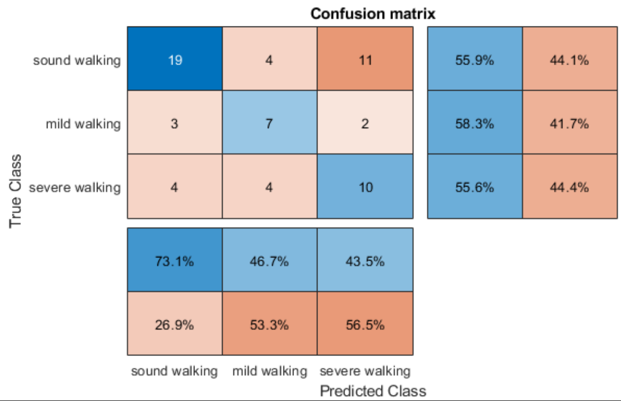
(a) Bagging ensemble /20 - features ReliefF



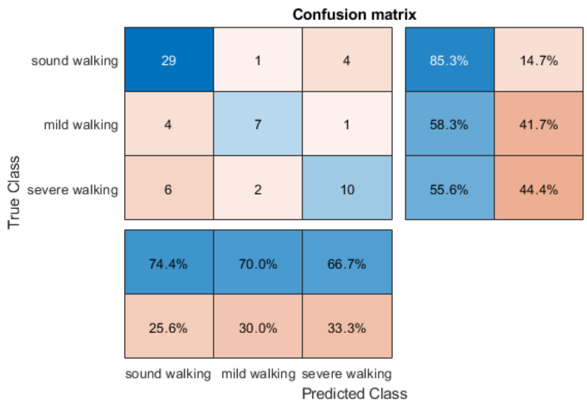
(b) RusBoosting ensemble /20 - features ReliefF



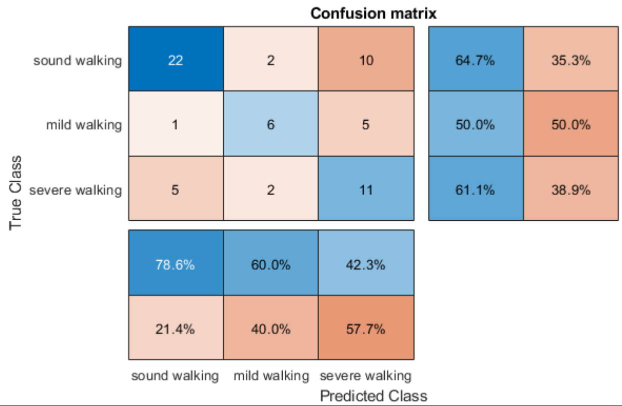
(c) Bagging ensemble /81 - features GA



(d) RusBoosting ensemble /81 - features GA

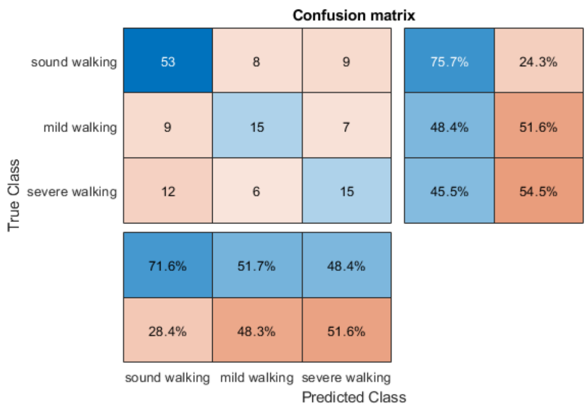


(e) Bagging ensemble /20 - features RF

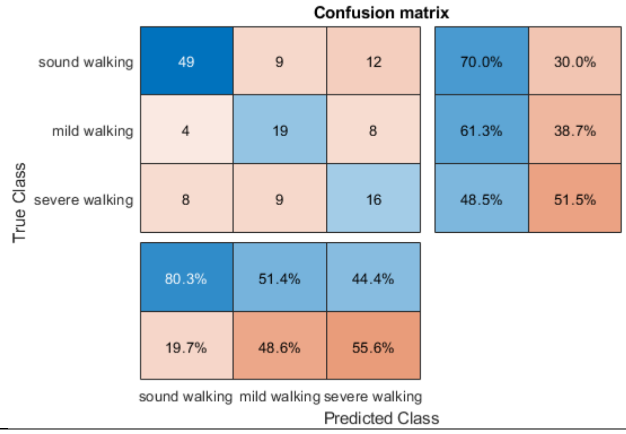


(f) RusBoosting ensemble /20 - features RF

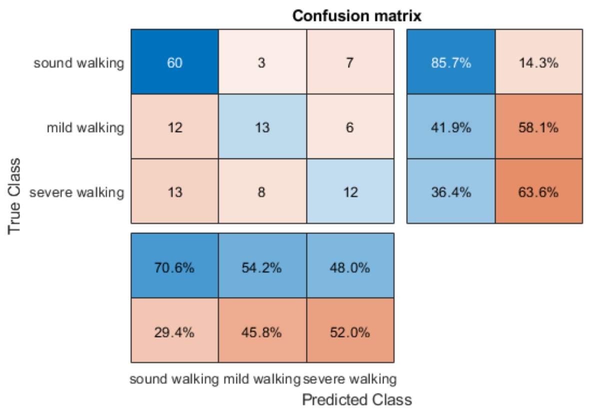
Appendix G. 19 Confusion Matrices for **5-fold** validation method of **DataSet3_all**, FOSW segmentation method over **10** $sec.window$.



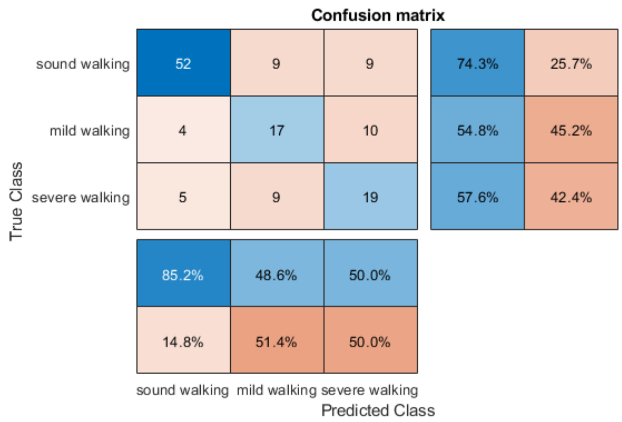
(a) Bagging ensemble /20 - features ReliefF



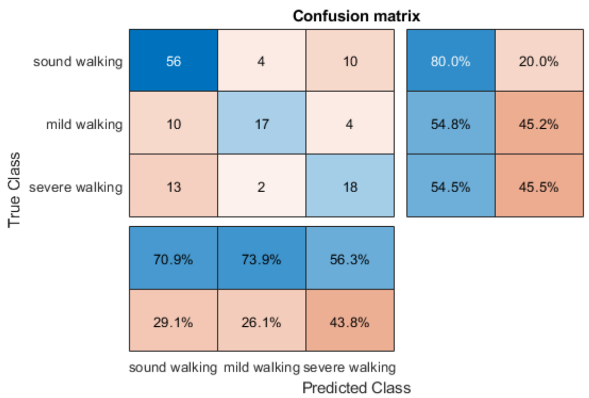
(b) RusBoosting ensemble /20 - features ReliefF



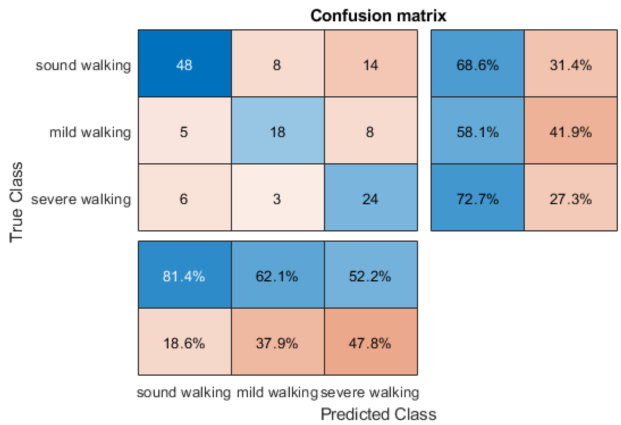
(c) Bagging ensemble /14 - features GA



(d) RusBoosting ensemble /14 - features GA

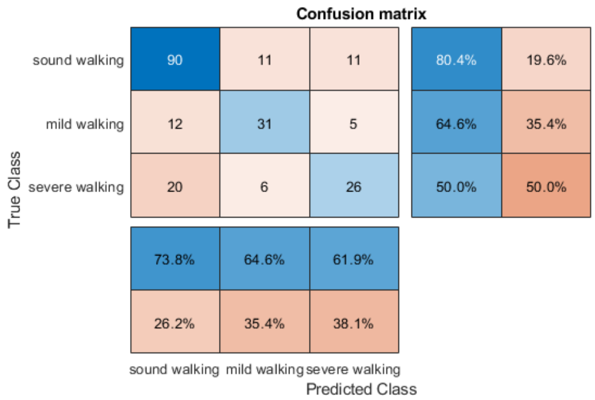


(e) Bagging ensemble /20 - features RF

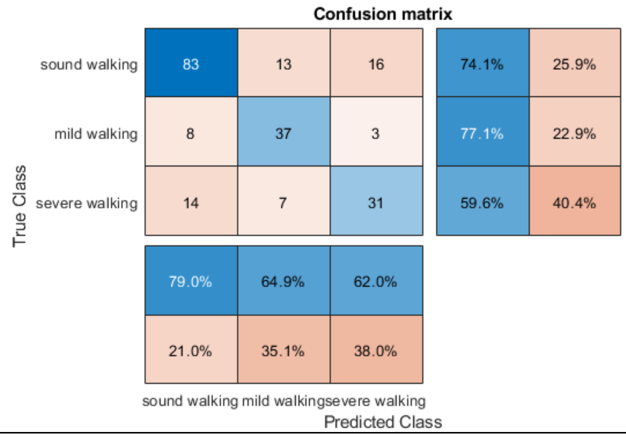


(f) RusBoosting ensemble /20 - features RF

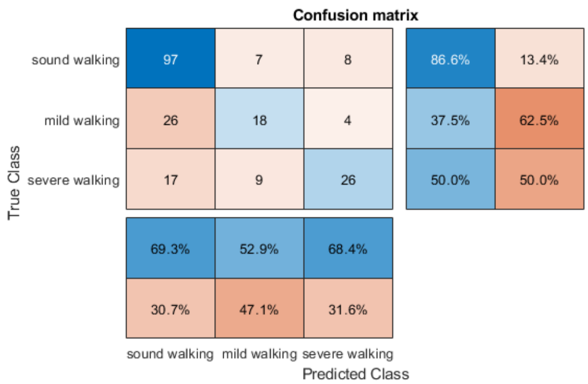
Appendix G. 20 Confusion Matrices for **5-fold** validation method of **DataSet3_all**, FOSW segmentation method over **7** $sec.window$.



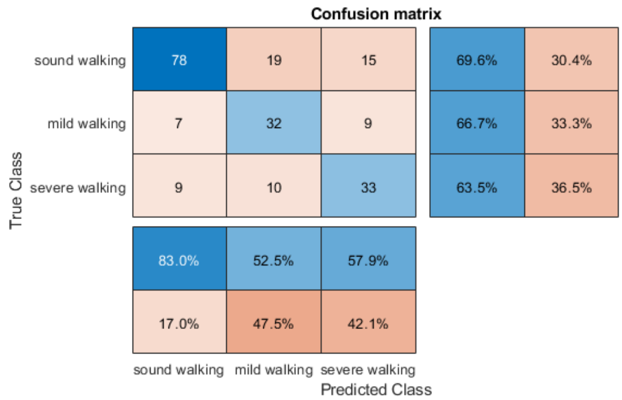
(a) Bagging ensemble /20 - features ReliefF



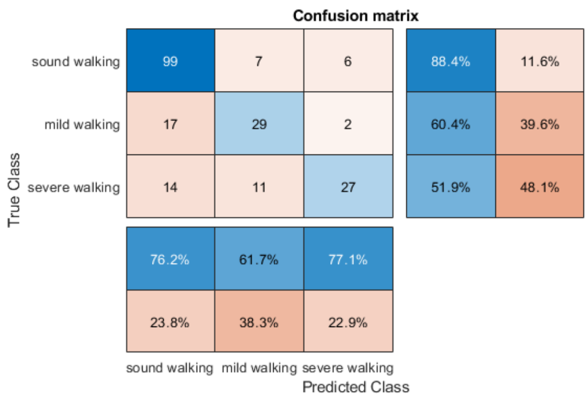
(b) RusBoosting ensemble /20 - features ReliefF



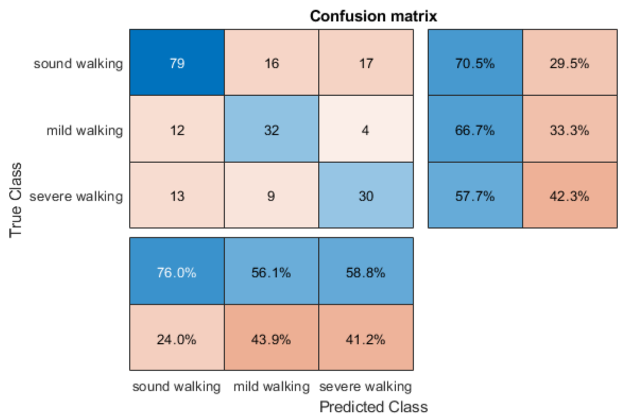
(c) Bagging ensemble /38 - features GA



(d) RusBoosting ensemble /38 - features GA

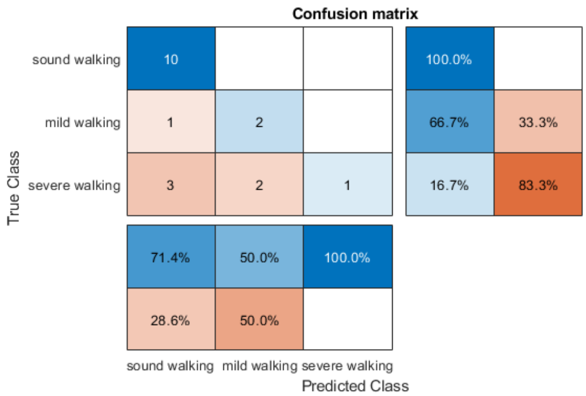


(e) Bagging ensemble /20 - features RF

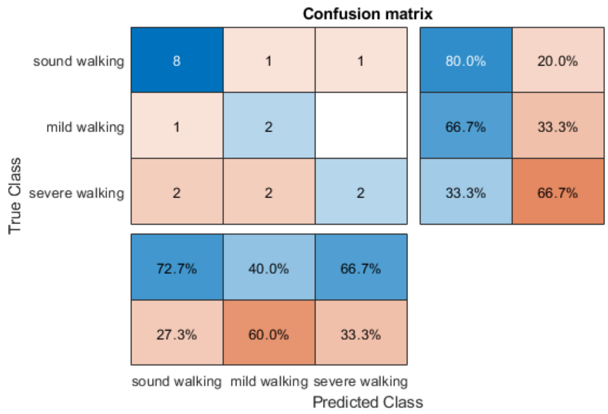


(f) RusBoosting ensemble /20 - features RF

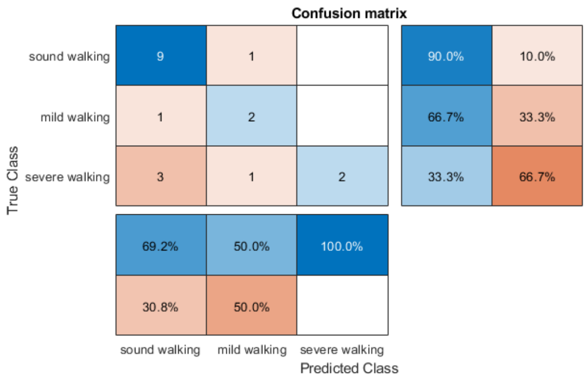
Appendix G. 21 Confusion Matrices for **5-fold** validation method of **DataSet3_all**, FOSW segmentation method over **5** $sec.window$.



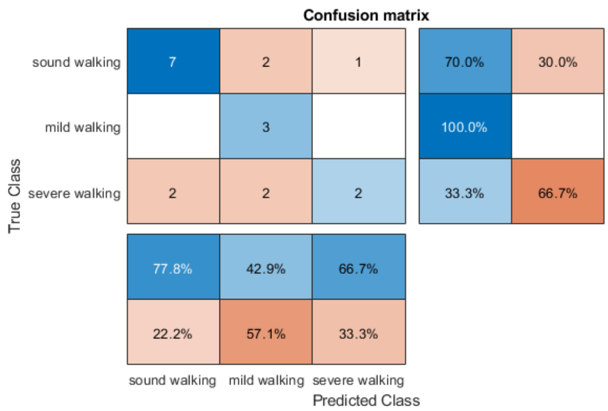
(a) Bagging ensemble /20 - features ReliefF



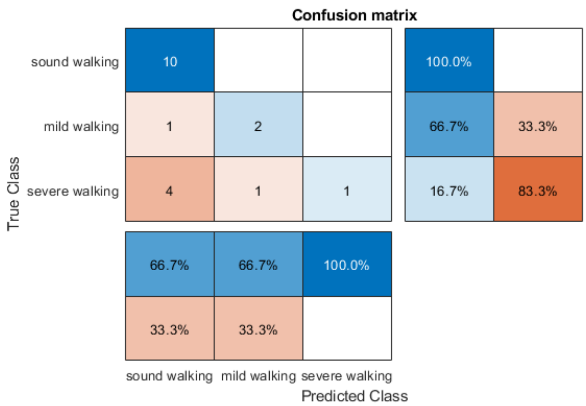
(b) RusBoosting ensemble /20 - features ReliefF



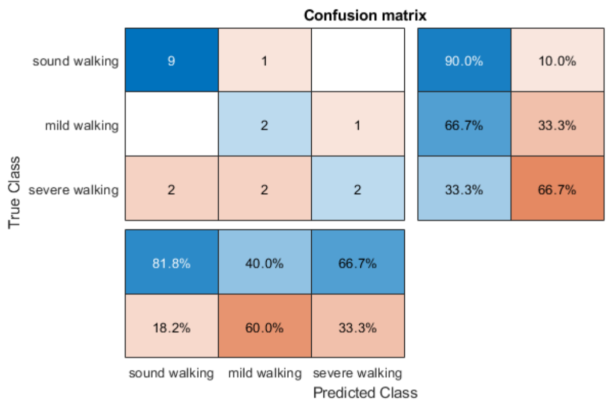
(c) Bagging ensemble /81 - features GA



(d) RusBoosting ensemble /81 - features GA

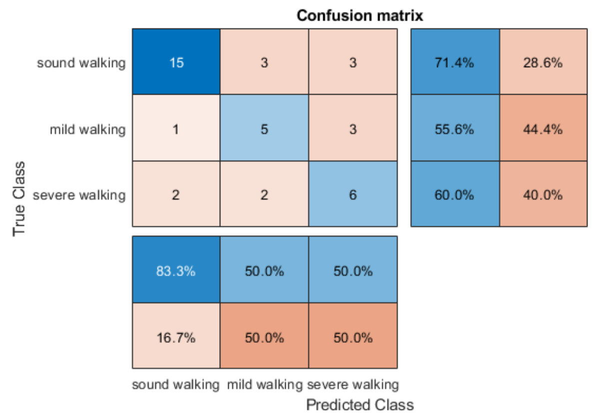


(e) Bagging ensemble /20 - features RF

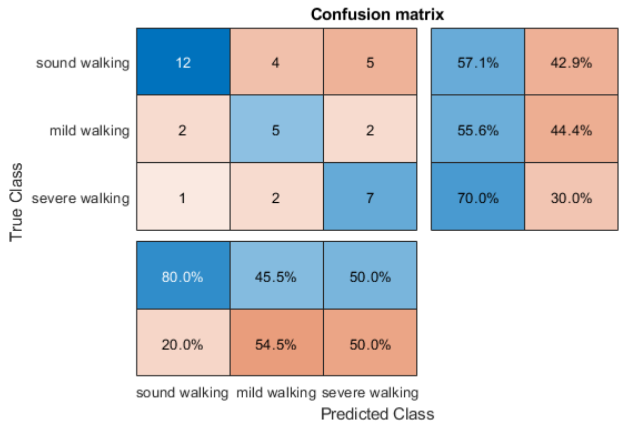


(f) RusBoosting ensemble /20 - features RF

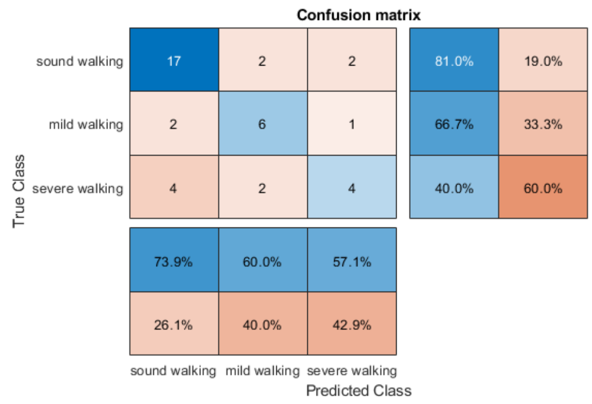
Appendix G. 22 Confusion Matrices for **0.3 hold-out** validation method of **DataSet3_all**, FOSW segmentation method over **10** $sec. window$.



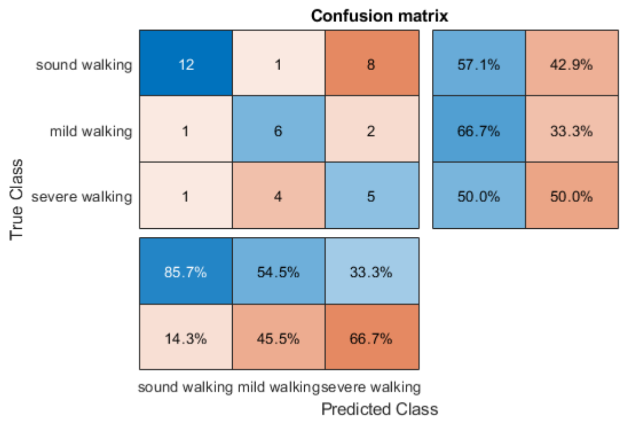
(a) Bagging ensemble /20 - features ReliefF



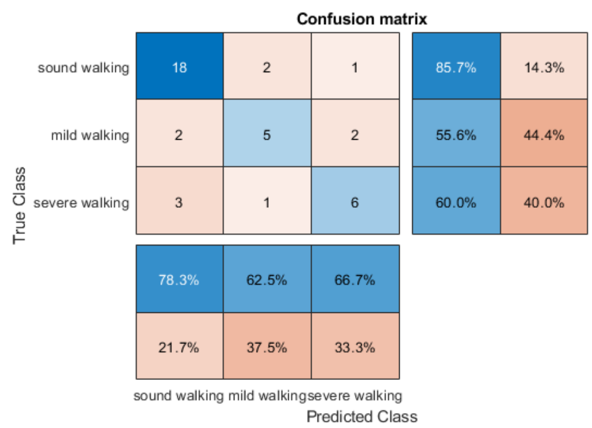
(b) RusBoosting ensemble /20 - features ReliefF



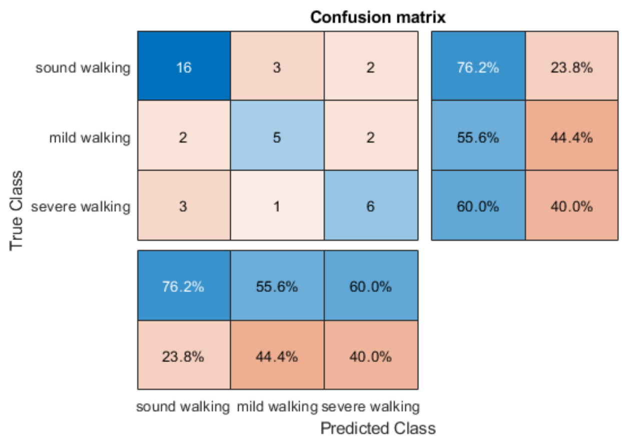
(c) Bagging ensemble /14 - features GA



(d) RusBoosting ensemble /14 - features GA

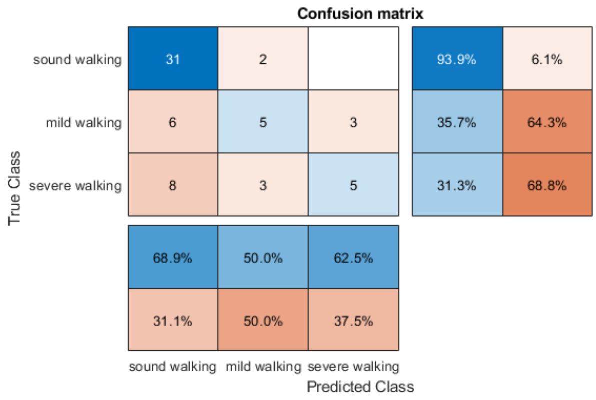


(e) Bagging ensemble /20 - features RF

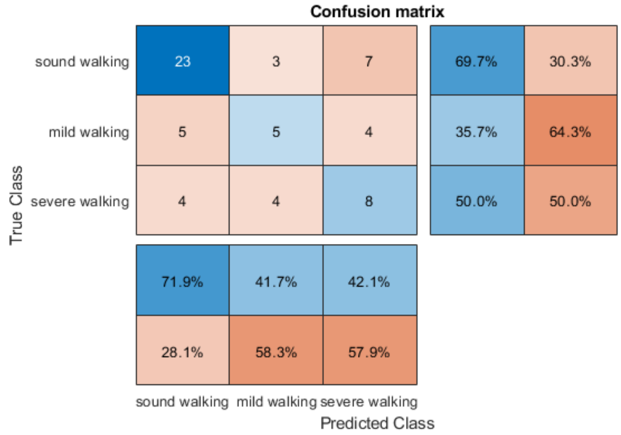


(f) RusBoosting ensemble /20 - features RF

Appendix G. 23 Confusion Matrices for **0.3 hold-out** validation method of **DataSet3_all**, FOSW segmentation method over **7** $sec.window$.

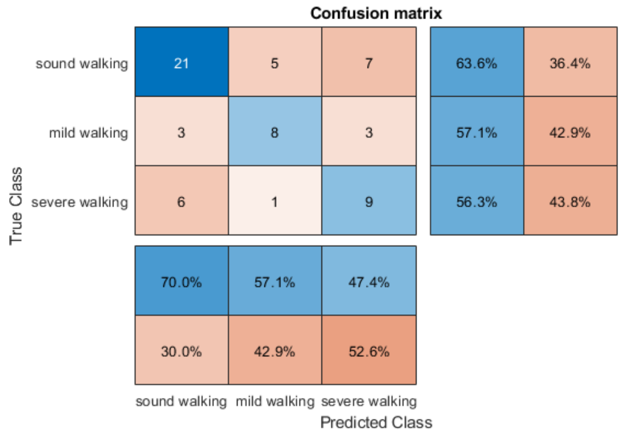


(a) Bagging ensemble /20- features ReliefF

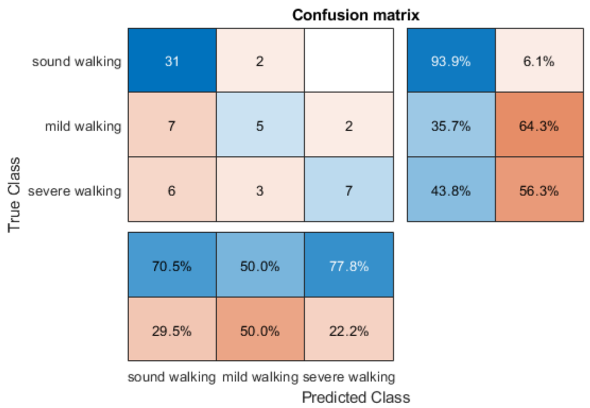


(b) RusBoosting ensemble /20- features ReliefF

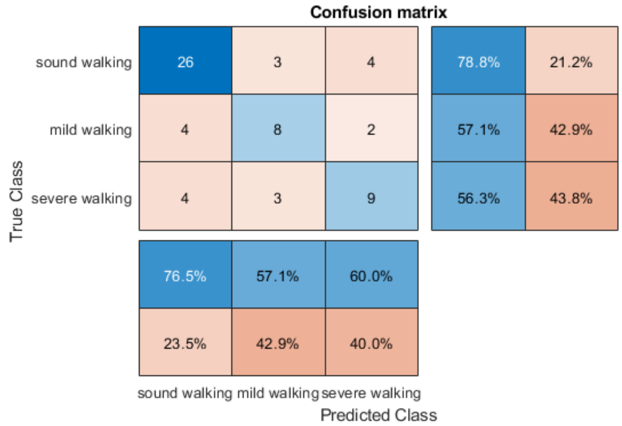
(c) Bagging ensemble /38 - features GA



(d) RusBoosting ensemble /38 - features GA



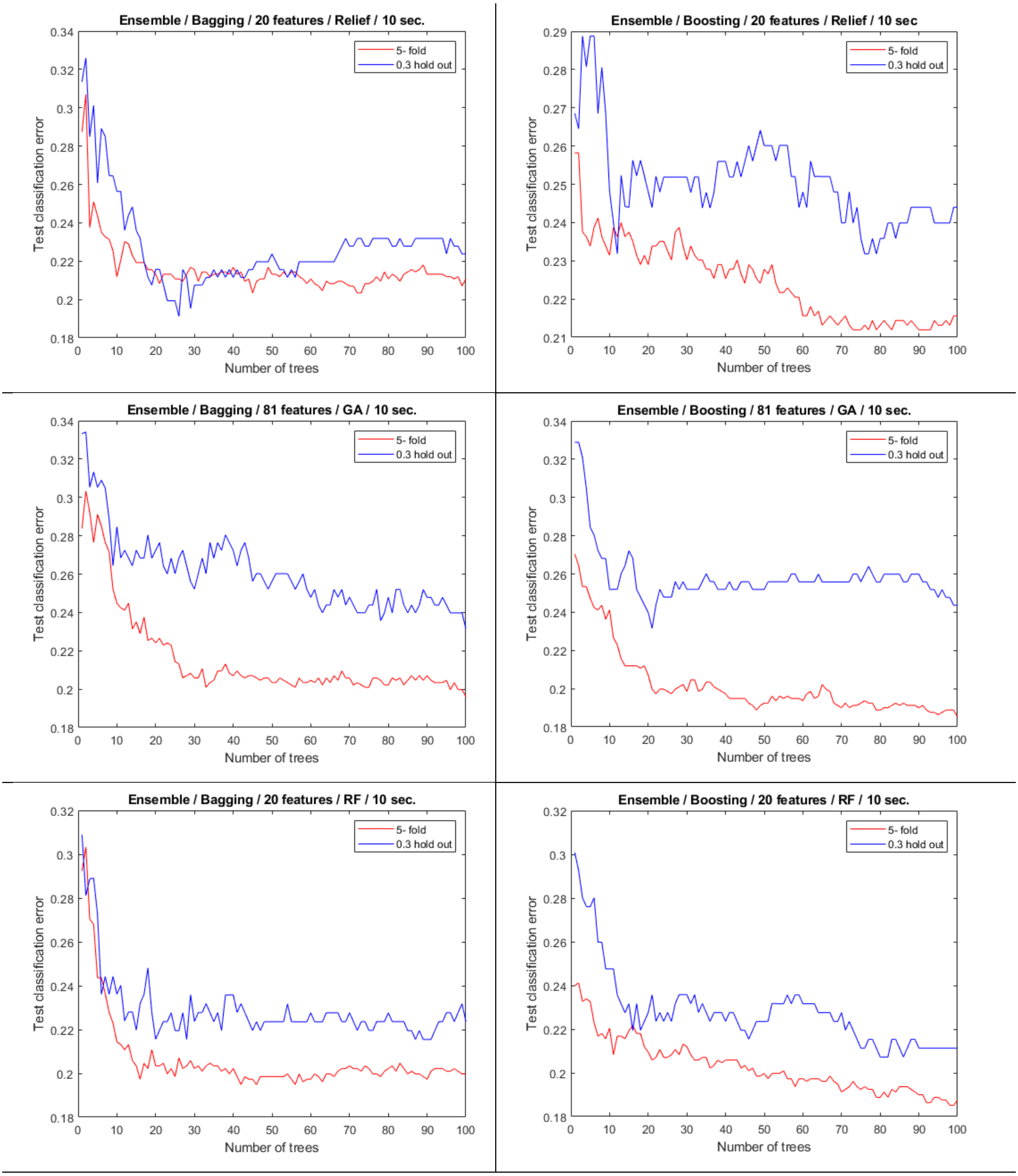
(e) Bagging ensemble /20- features RF



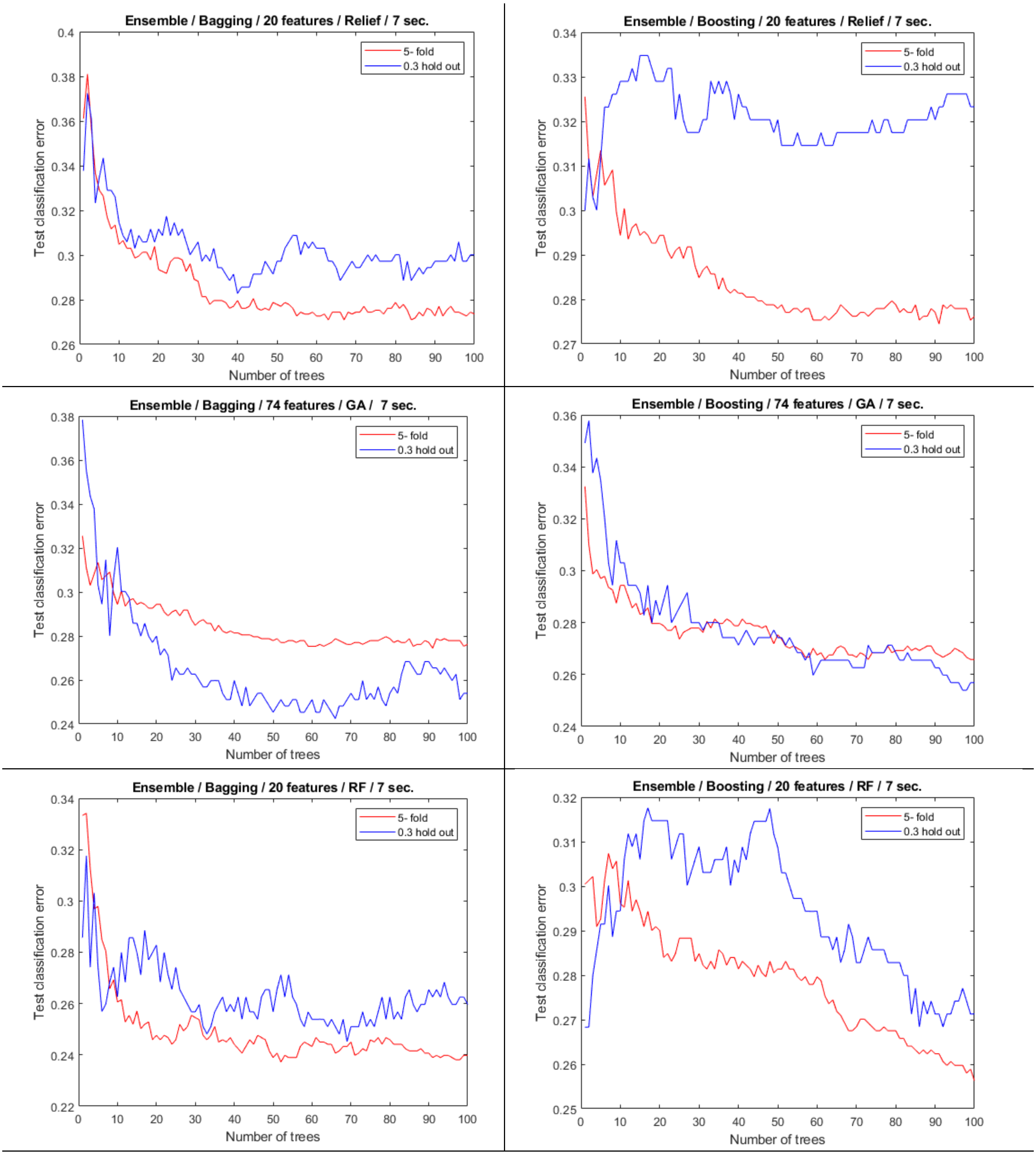
(f) RusBoosting ensemble /20- features RF

Appendix G. 24 Confusion Matrices for **0.3 hold-out** validation method of **DataSet3_all**, FOSW segmentation method over **5** $sec.window$.

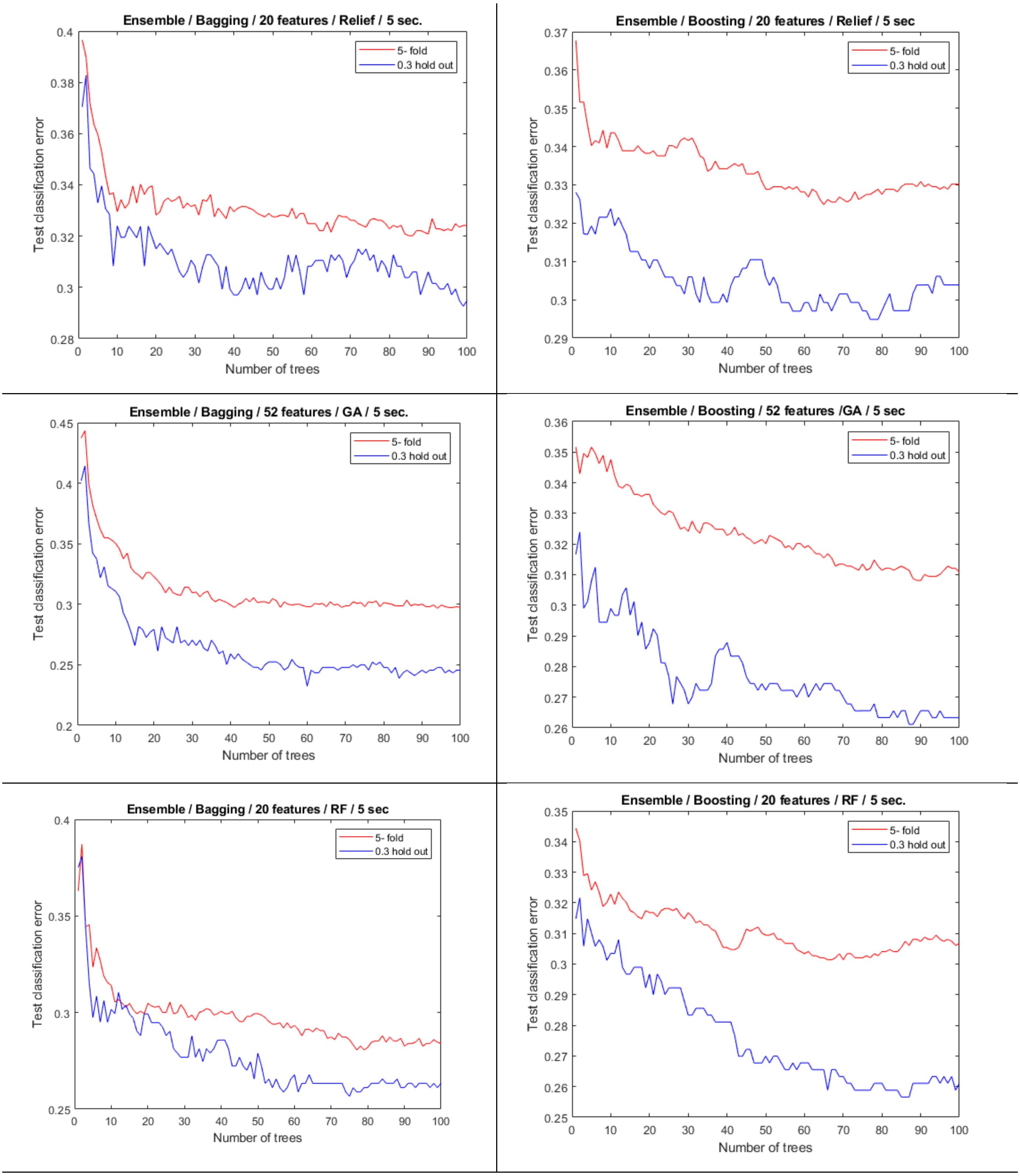
Appendix H. Comparison of Validation Techniques of Ensemble Classifiers



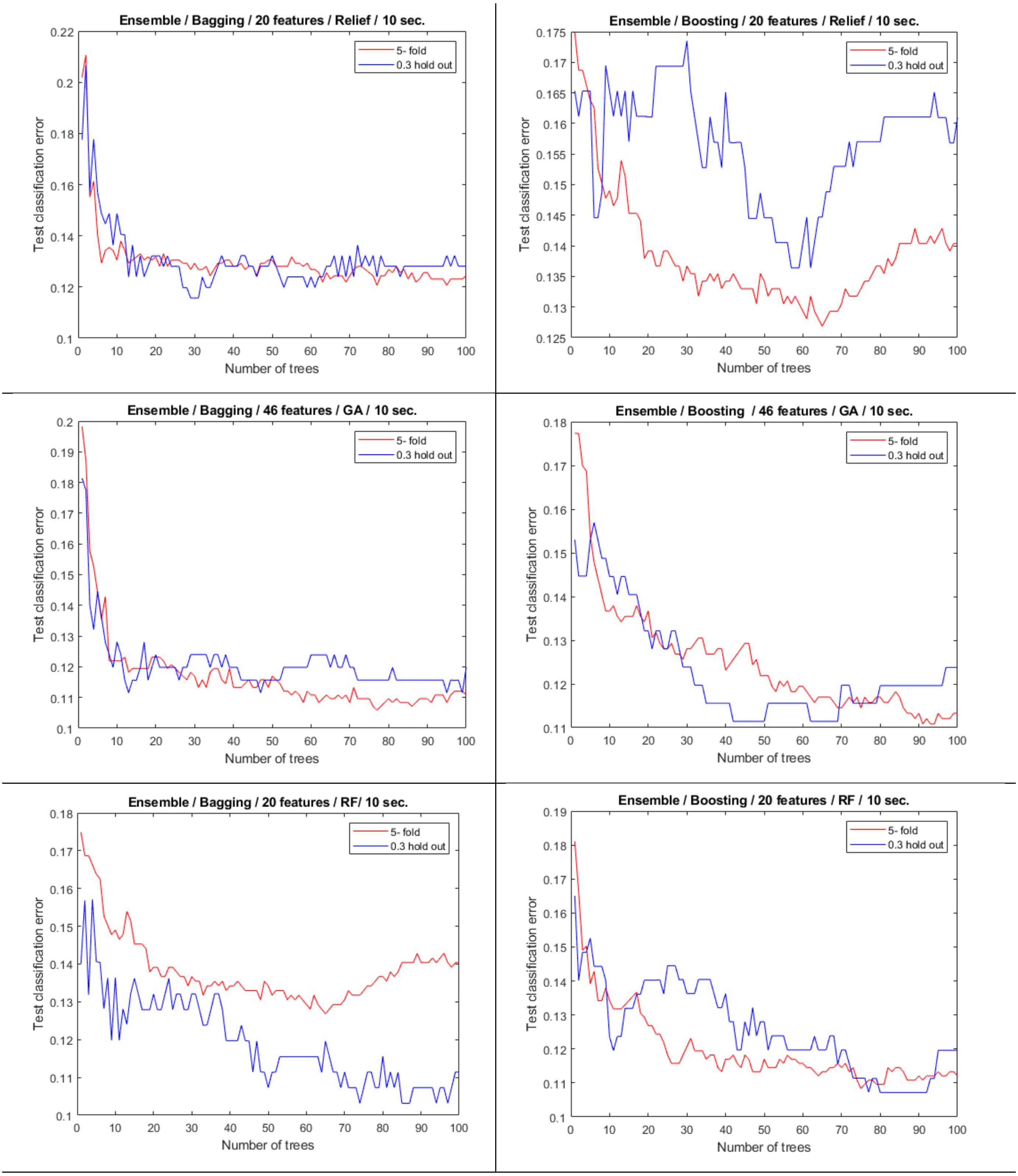
Appendix H. 1 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet1_all** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **10** $sec. window$.



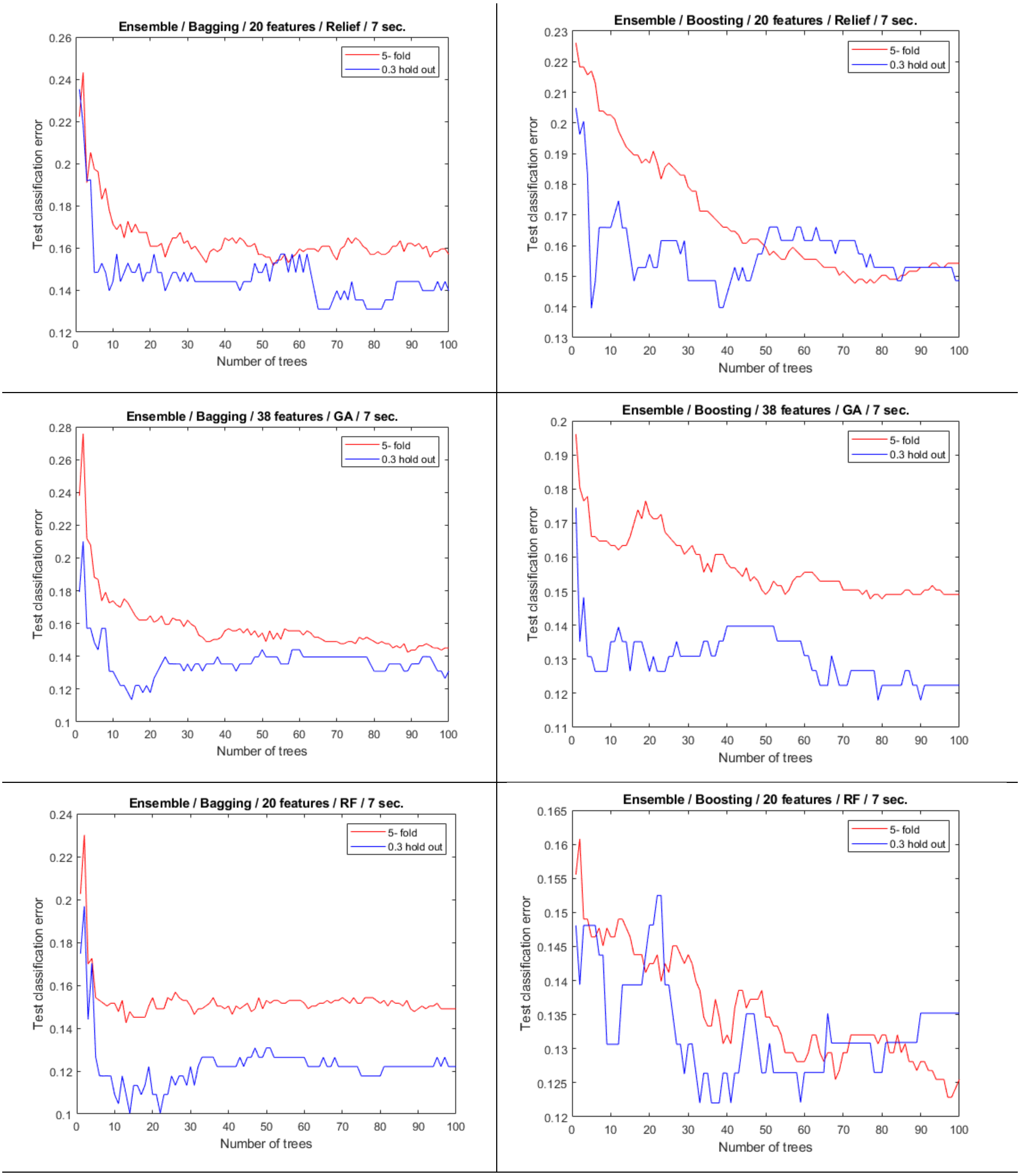
Appendix H. 2 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet1_all** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **7** $sec. window$.



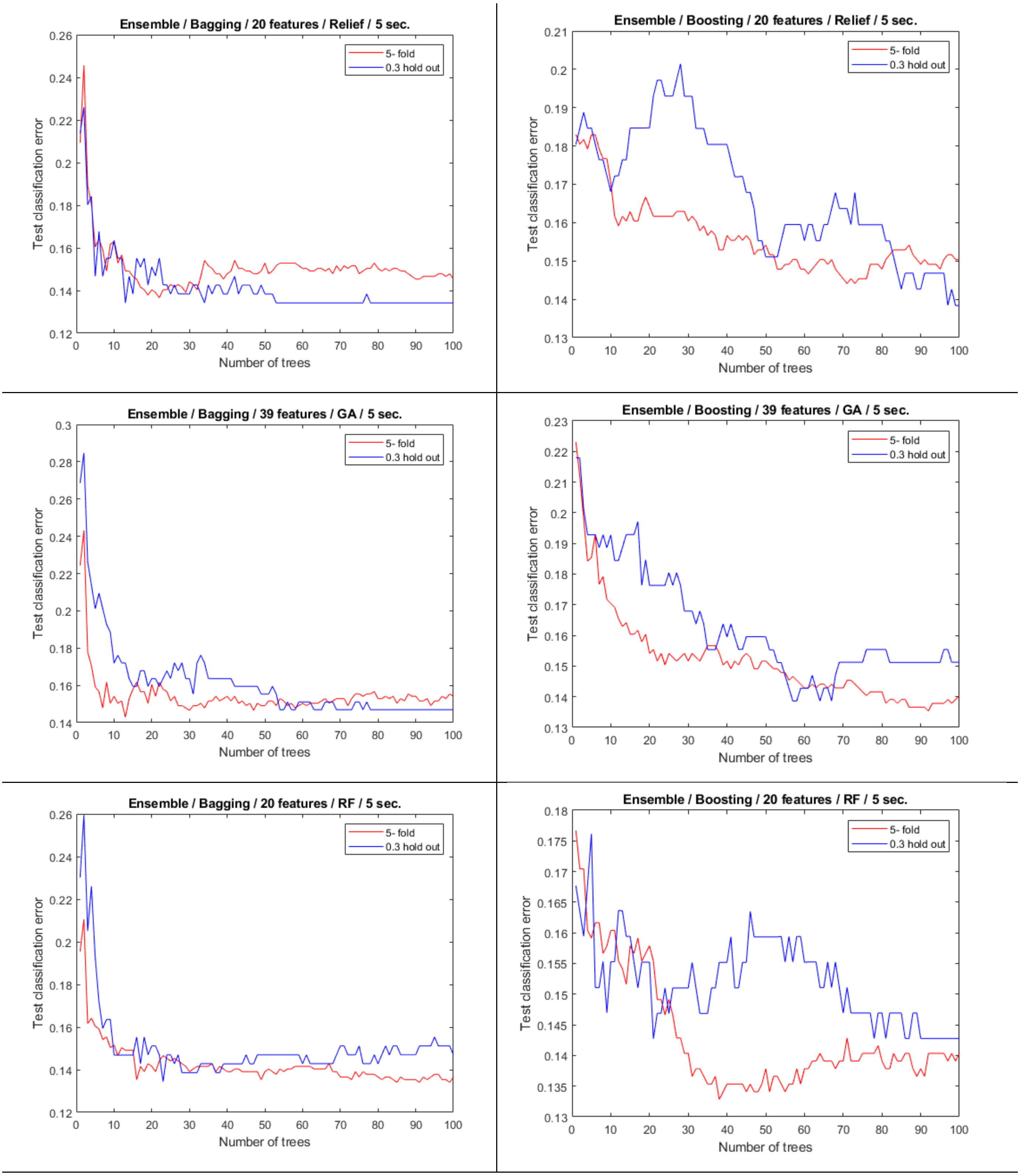
Appendix H. 3 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet1_all** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **5** $sec. window$.



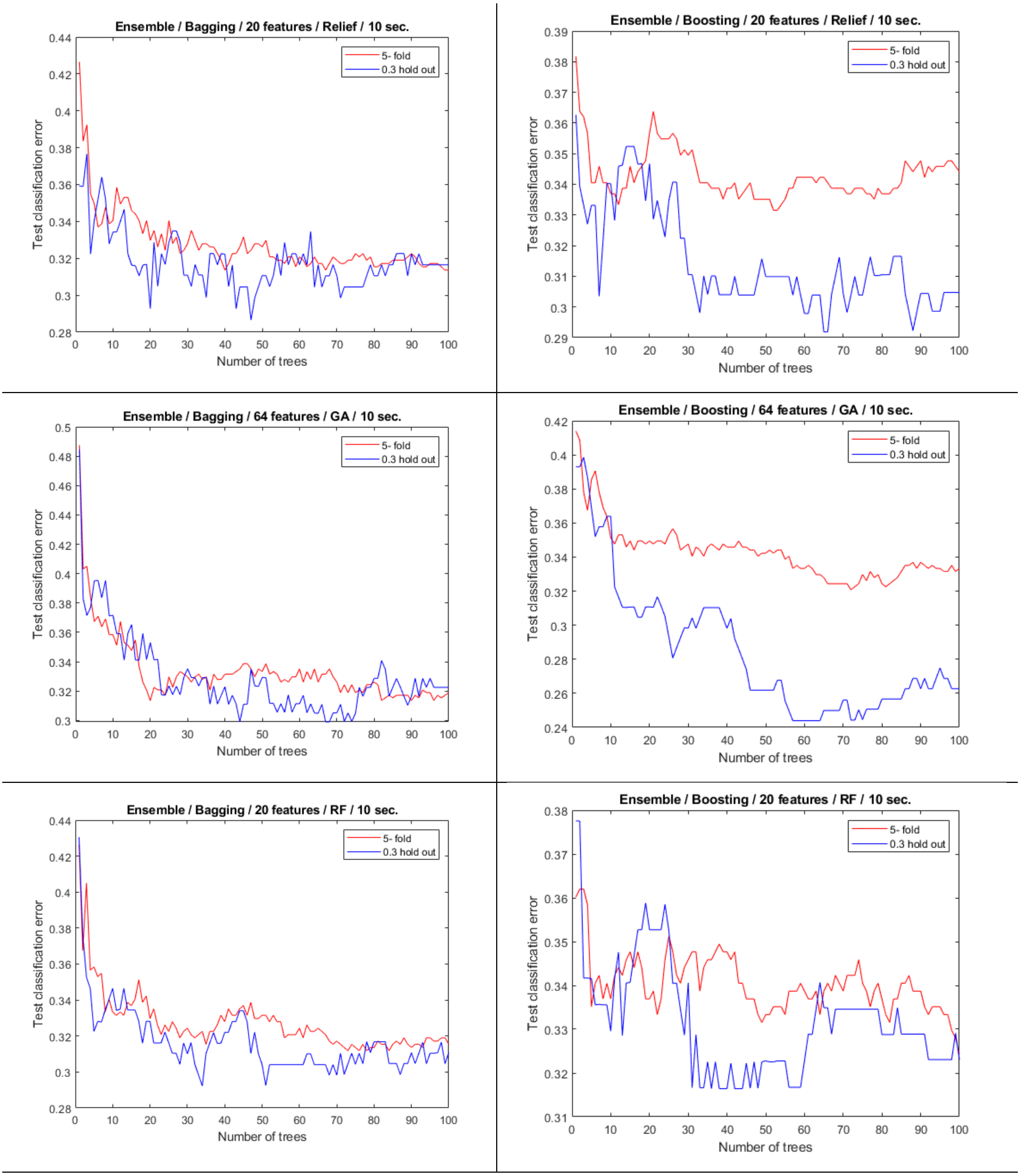
Appendix H. 4 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet2_ac** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **10** $sec. window$.



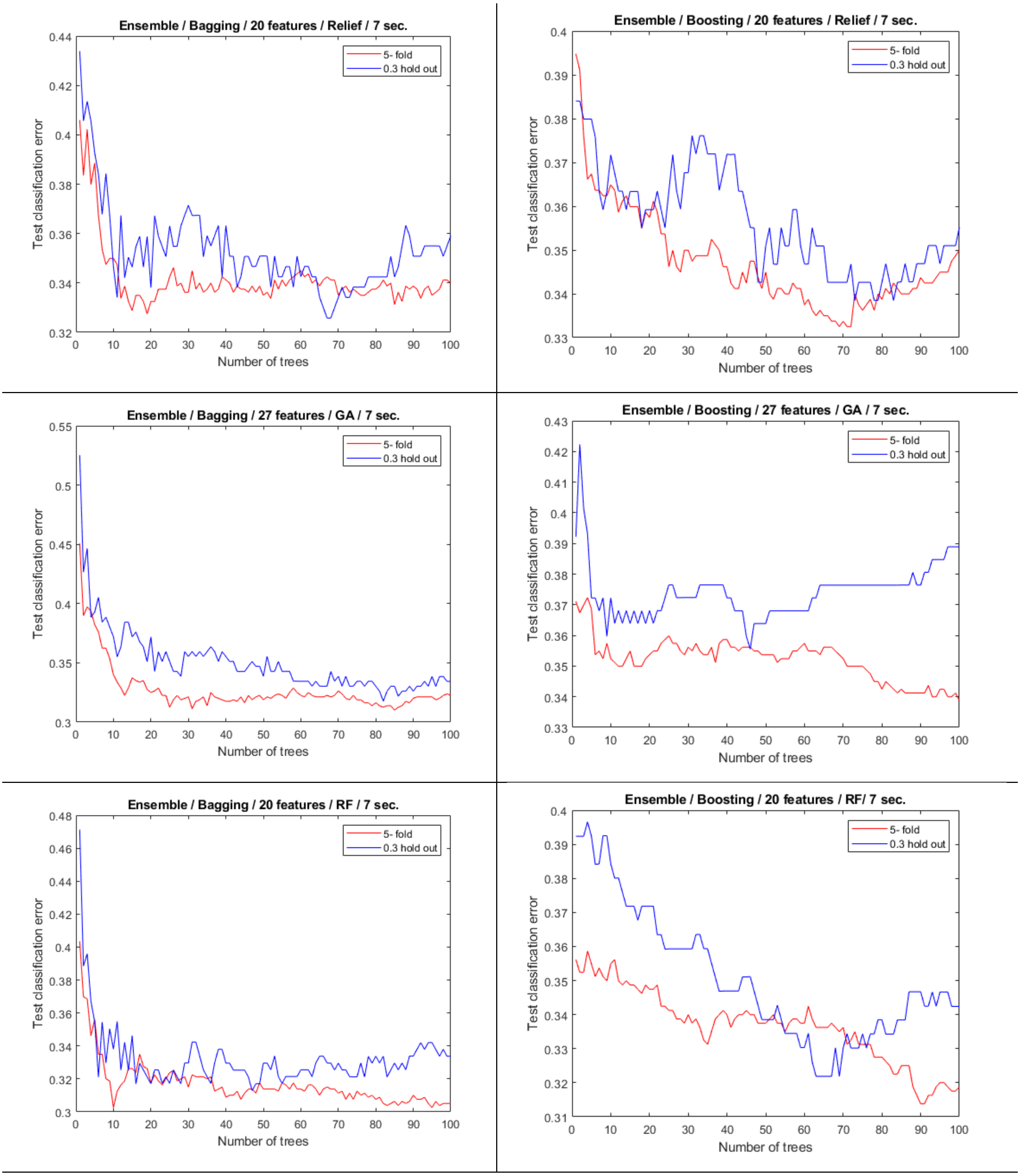
Appendix H. 5 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet2_ac** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **7** $sec.window$.



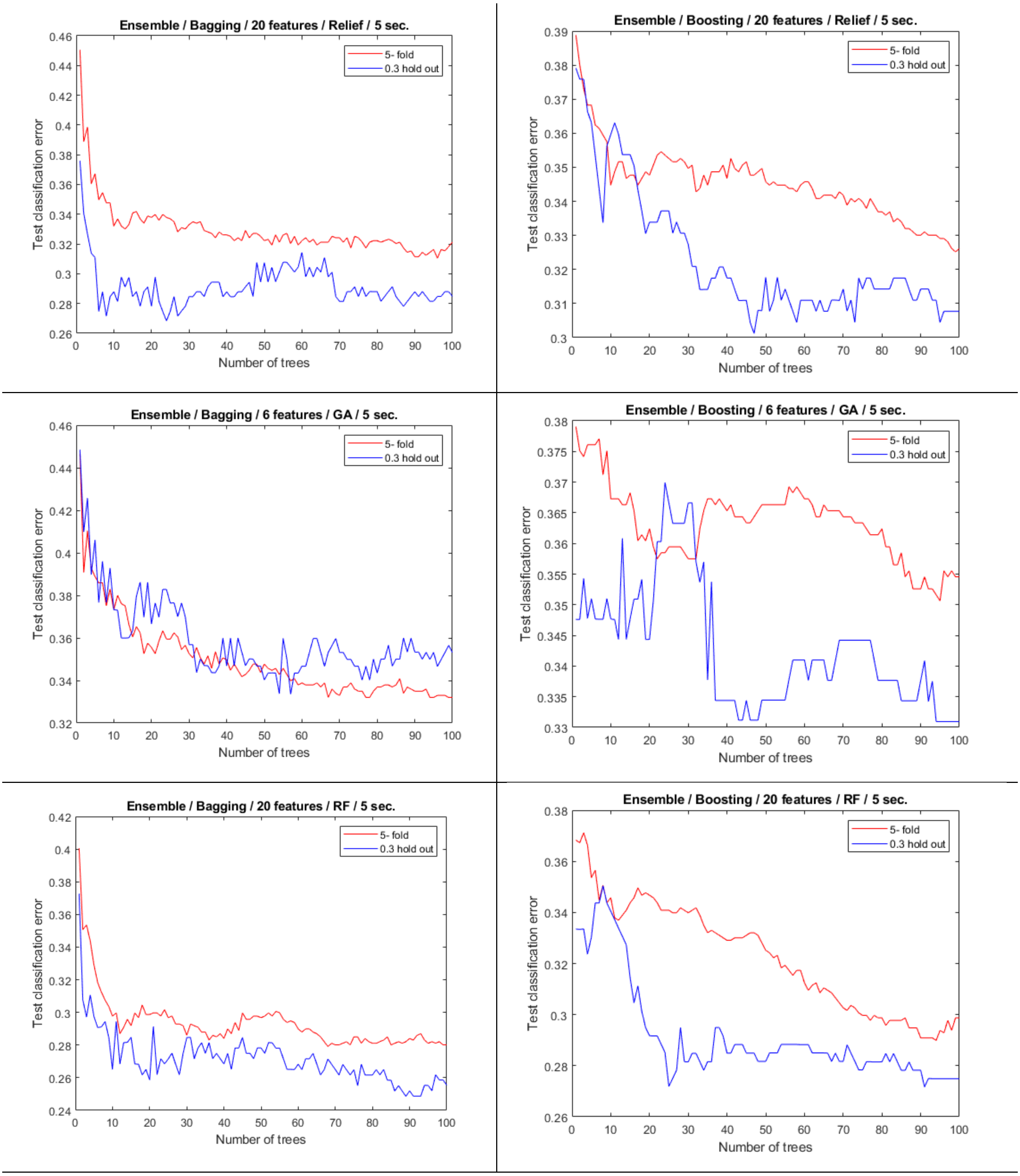
Appendix H. 6 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet2_ac** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **5** $sec. window$.



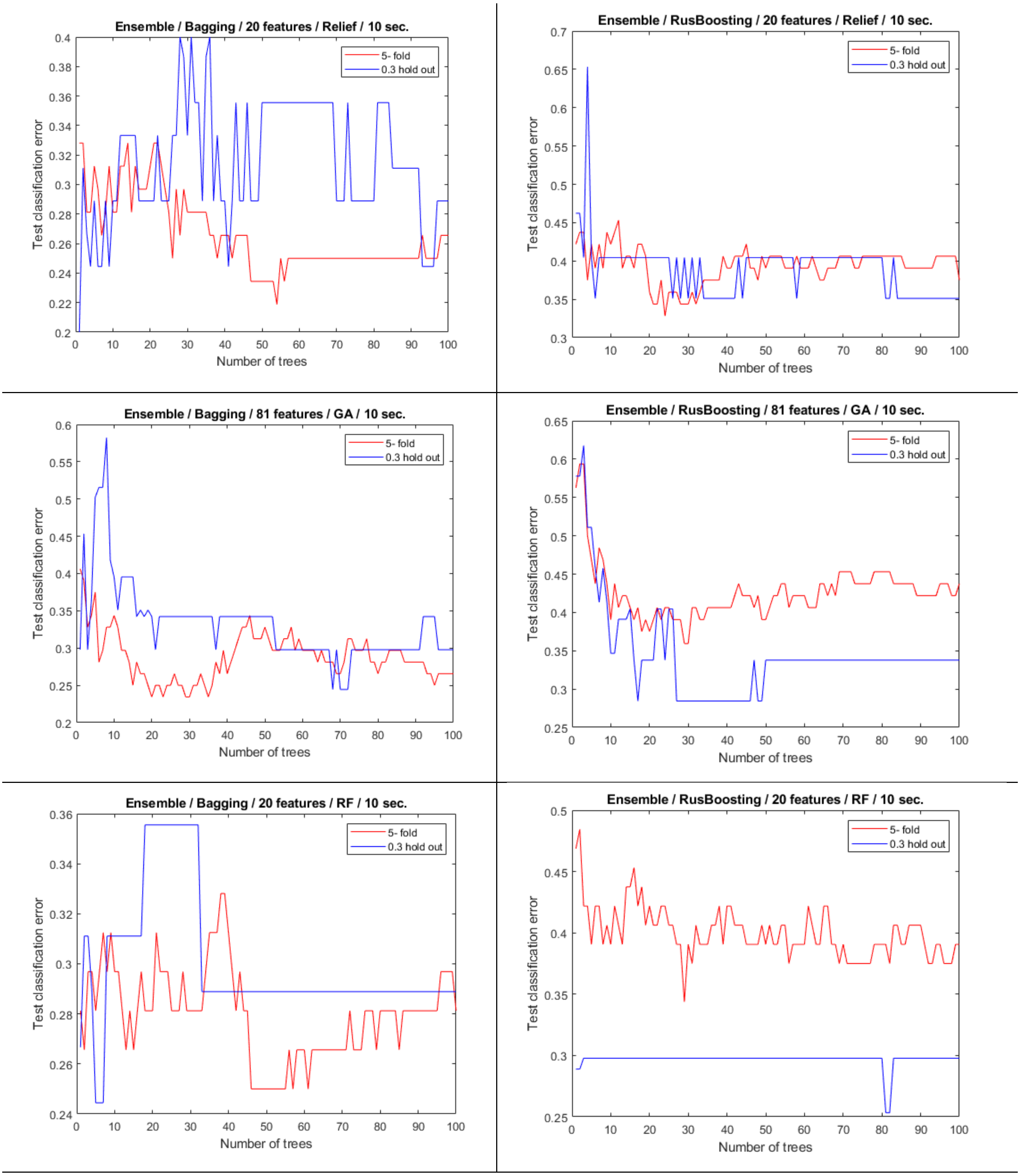
Appendix H. 7 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet2_b** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **10** $sec.window$.



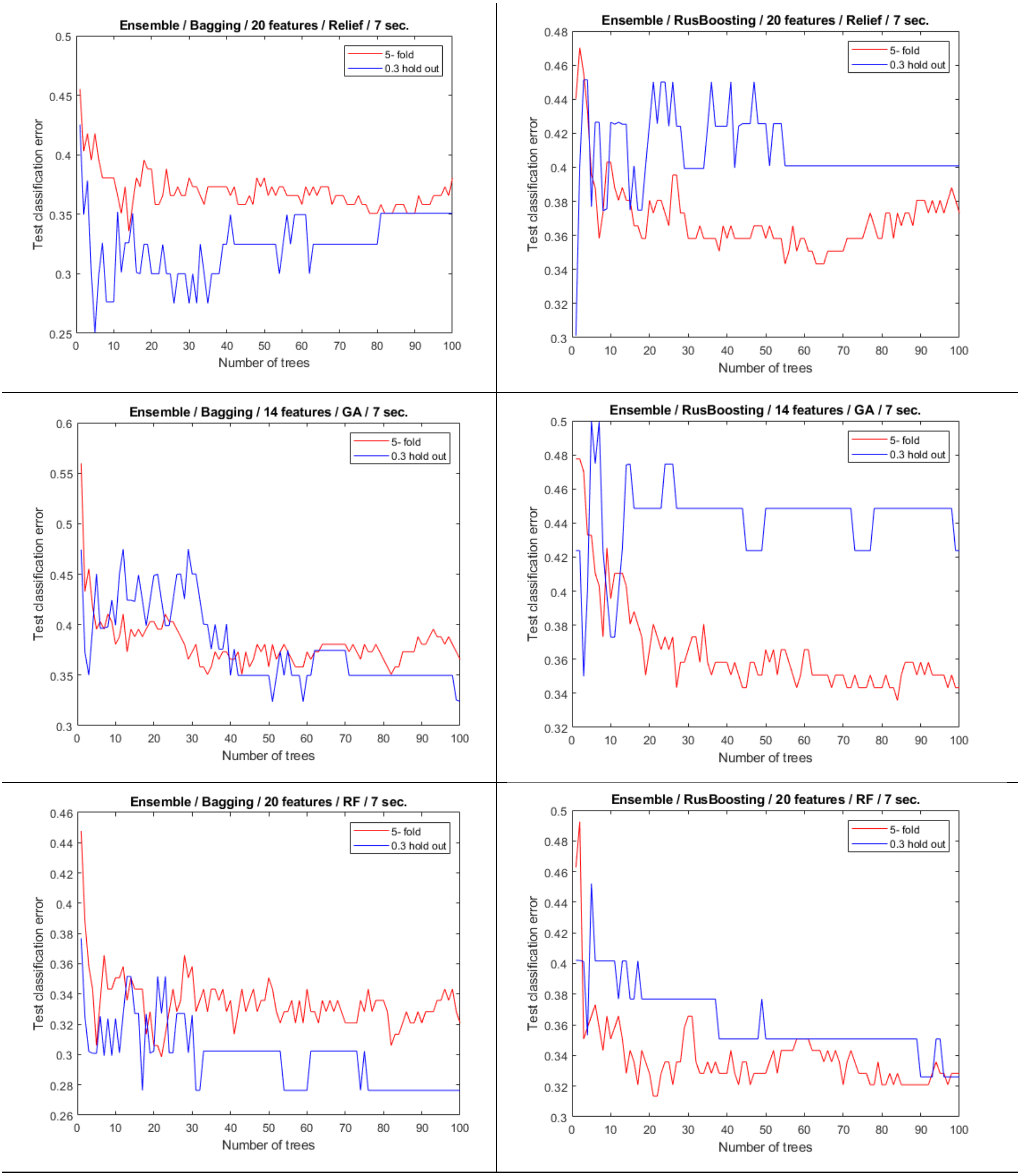
Appendix H. 8 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet2_b** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **7** $sec. window$.



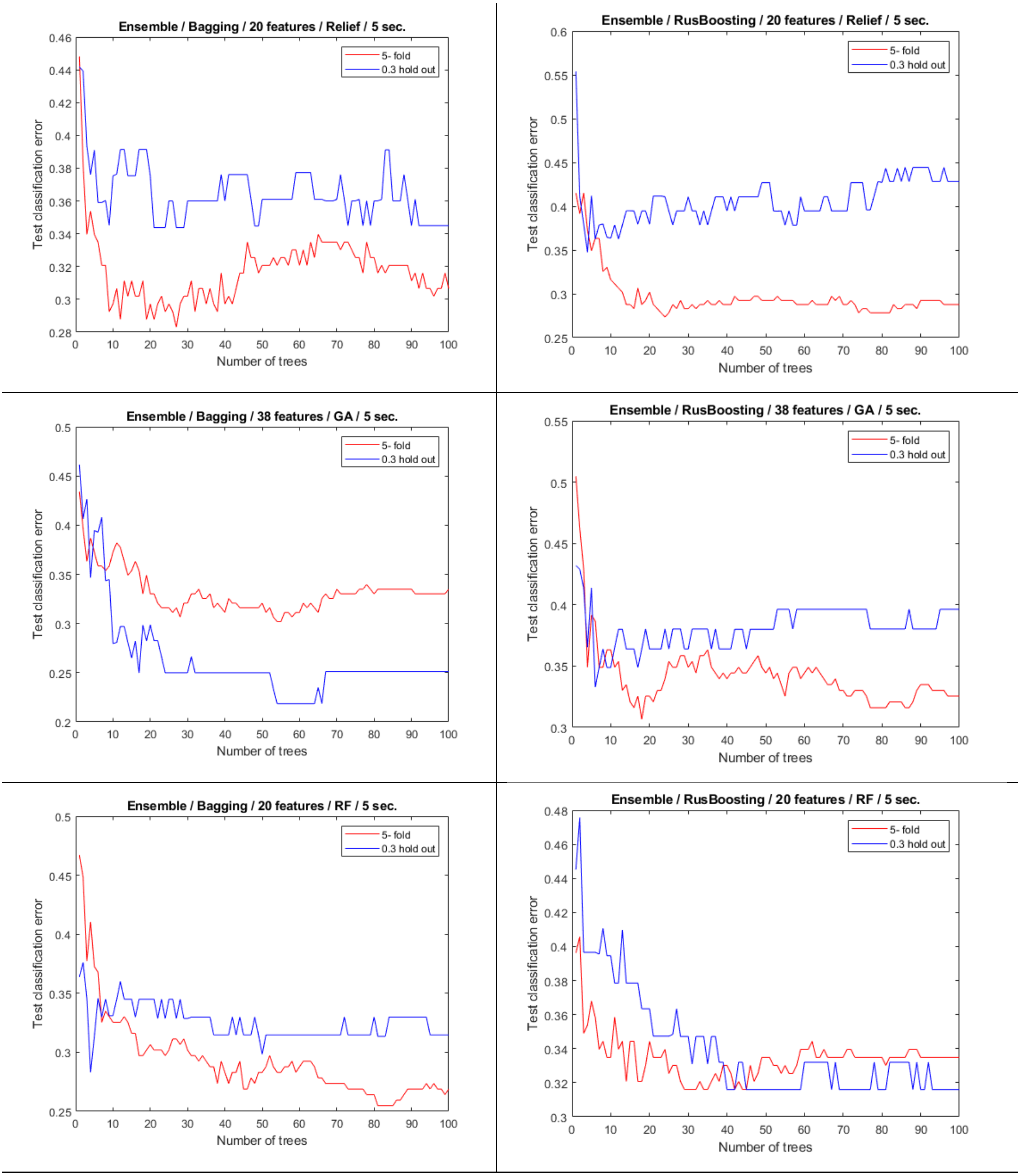
Appendix H. 9 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & Boost) for **DataSet2_b** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **5** $sec.window$.



Appendix H. 10 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & RusBoost) for **DataSet3_all** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **10** $sec. window$.



Appendix H. 11 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & RusBoost) for **DataSet3_all** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **7** $sec. window$.



Appendix H. 12 Validation techniques comparison (5-fold & 0.3 hold-out) of Ensemble classifiers (Bag & RusBoost) for **DataSet3_all** (3 FS: ReliefF, GA, RF), FOSW segmentation method over **5** $sec. window$.

Appendix I. **Publications and Awards Gallery**

University of Northampton The Graduate School

Images of Research 2016-17

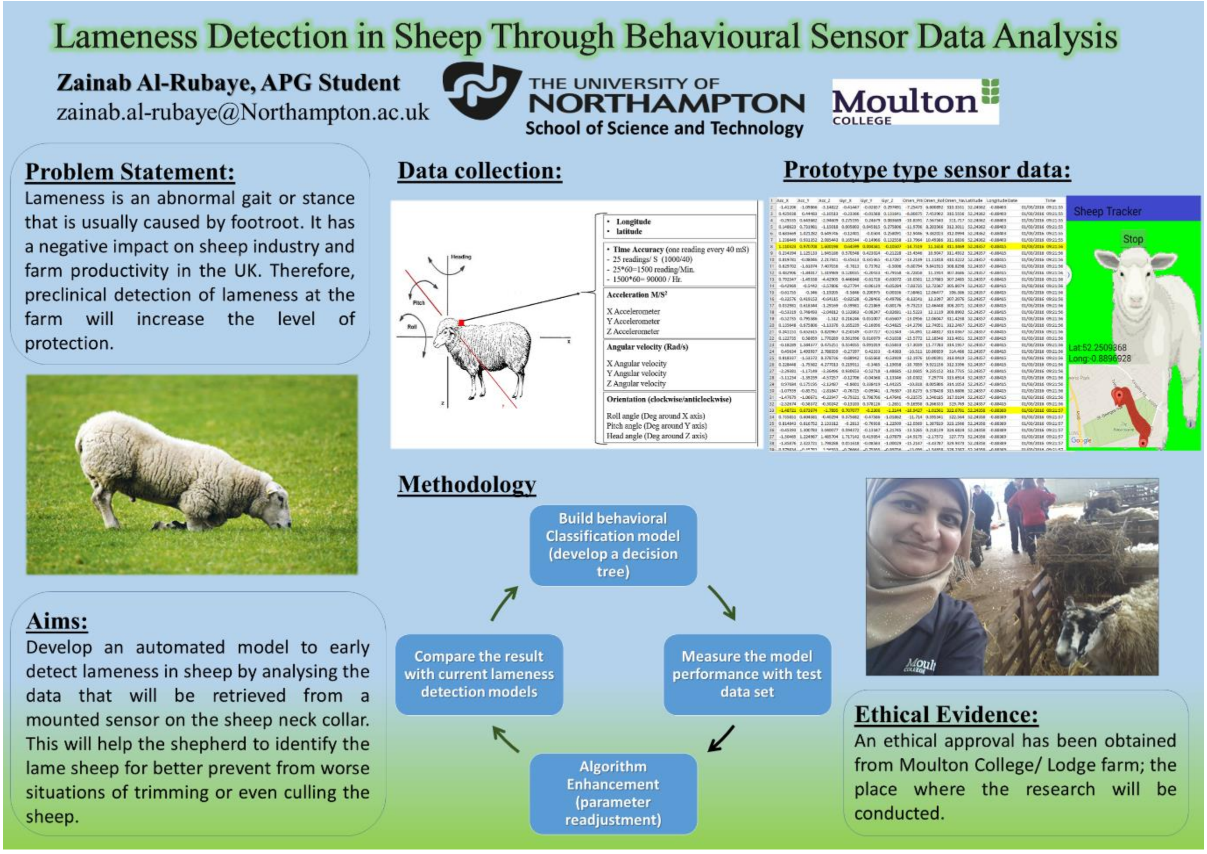
Winners: Chosen by voters at the exhibition and online

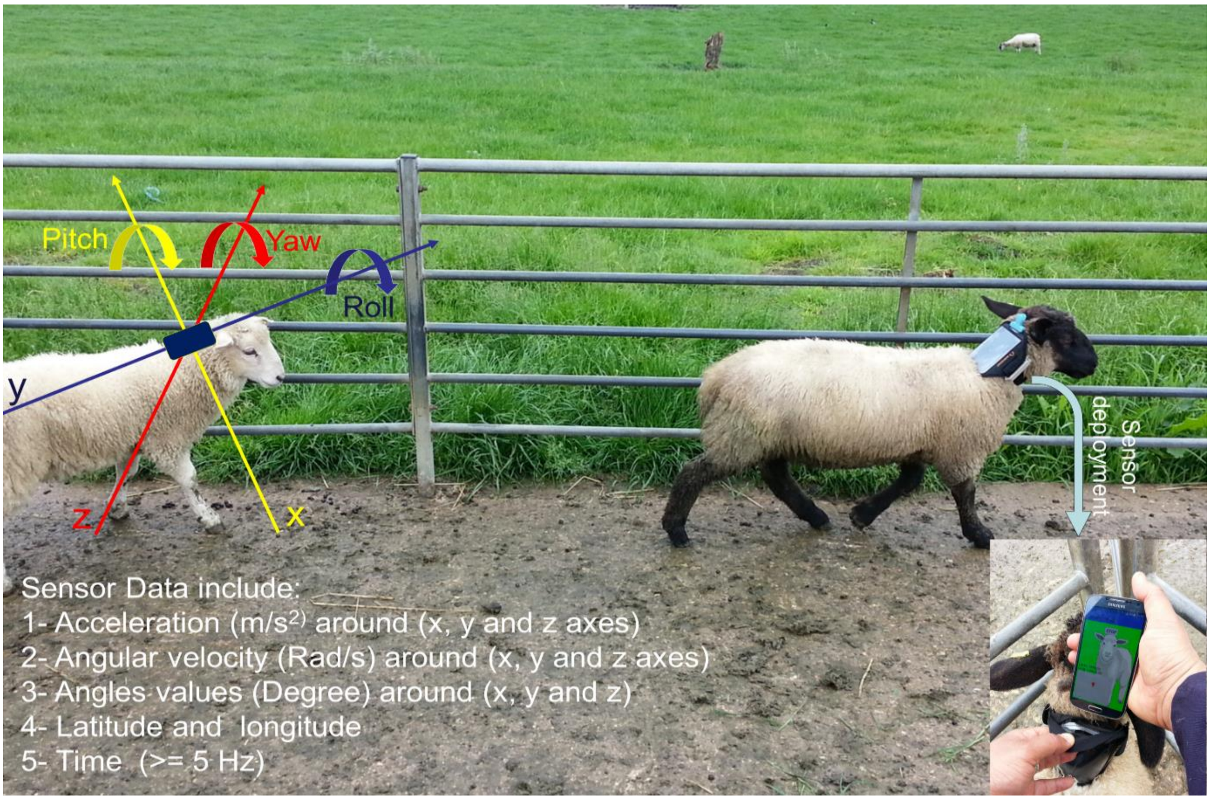
1st

Zainab Al-Rubaye, Postgraduate Research Student, Computing and Immersive Technology, Faculty of Arts, Science and Technology, with Moulton College.

Lameness is a clinical symptom related to movement disorder in the locomotion systems of the animal. It is considered one of the primary welfare concerns in the sheep industry in the UK due to the annual loss, which is estimated to be £10 for each ewe according to a 2016 Agriculture and Horticulture Development Board report. This research aims to develop an automatic model for early detection of lameness in sheep by analysing the data that will be retrieved from a sensor mounted on a collar on the sheep's neck. This extensive spatiotemporal data will be classified to infer the associated behaviour of the lame sheep. The prior detection of the lame sheep will be expected to decrease the prevalence of lameness and enable the shepherd to react quickly with better treatment.

Appendix I. 1 Poster competition 2nd place winner in 2016, and Image of Research 1st place winner in 2017 at the University of Northampton.





Appendix I. 2 1st winner poster in the Poster competition, and 1st place winner image in the Image of Research at the University of Northampton in 2016 and 2017.

Moulton
COLLEGE

Zainab Al-Rubaye
33 Brunswish
Northampton
NN1 4BG

The Principal/jlm
Ref: 2911/7

18th May 2017

Dear Zainab

I just wanted to drop you a quick note to congratulate you on winning first prize in the UoN's Images of Research competition.

I was extremely impressed by the image you created and can clearly see why you won.

Kind regards.

Yours sincerely,

Steve

STEVE DAVIES
PRINCIPAL

Moulton College, West Street, Moulton, Northampton, NN3 7RR
t +44 (0)1604 491131 e enquiries@moulton.ac.uk w www.moulton.ac.uk
Principal and Chief Executive Stephen M Davies MHort (RHS), Dip Hort (Edin), Cert Ed, FCIHort, CMgr FCMI
Deputy Principal and Deputy Chief Executive Gerald Davies BSc (Hons), DMS
A charity providing educational services to the community

Appendix I. 3 Acknowledgement letter from Moulton College Principal for winning 1st place in the Image of Research at the University of Northampton.

Appendix I. 4 Gallery from BBC recording day in October 2017 at Lodge Farm/ Moulton College/ Northampton.

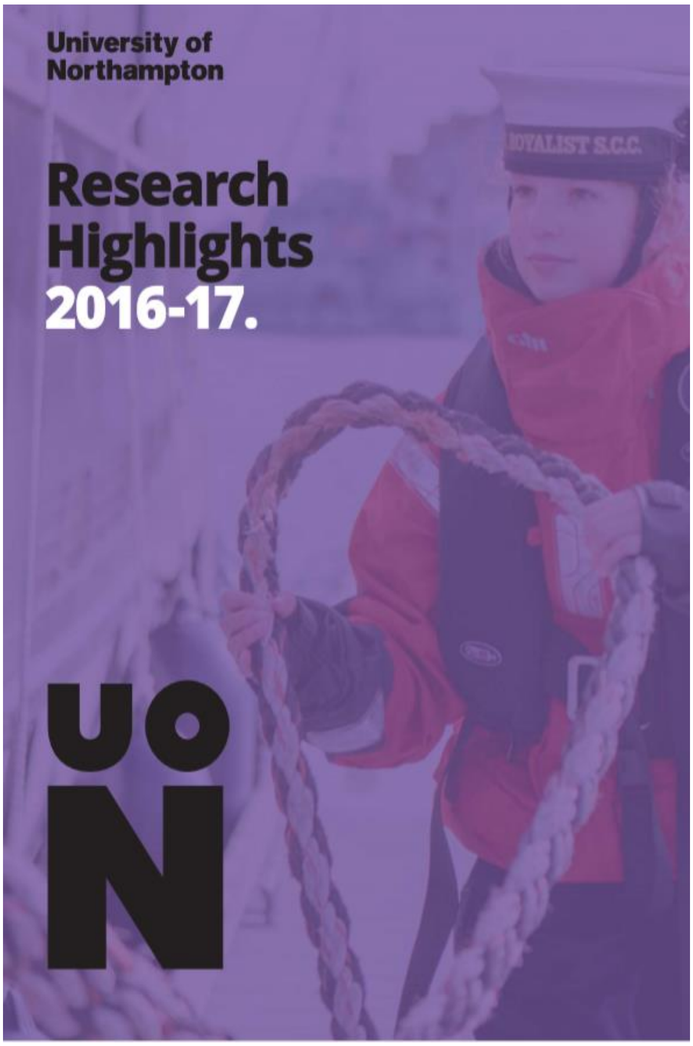



Photo: Wearable sensor

'Fitbit' for sheep: wearable tech monitors for lameness in the flock.

Shepherds may well have once watched their flocks by night, but thanks to new University-led research, wearable tech could soon be keeping a virtual eye out for problems 24/7.

Photo: Lameness in sheep

12 University of Northampton Research Highlights 2016-2017


Photo: Dr Wanda McCormick with The One Show team

Lameness in livestock is both a welfare and commercial concern. The main cause of sheep lameness in the UK, the infectious disease foot-rot, costs the farming industry up to £80 million a year through the resulting deaths or infertility of ewes, or the poor growth and survival of lambs.

Sheep that have been infected are usually treated as soon as they are found - but with hundreds of sheep **spread out across their fields, it can be difficult for farmers to catch it** early. When the underlying cause of lameness is also contagious, the need to act quickly is vital.

Hoping to address this problem, University computing experts, Dr Ali Al-Sherbaz and Dr Scott Turner teamed up with animal welfare expert, Dr Wanda McCormick at nearby Moulton College to develop a **lameness 'early warning system'.**

Dr Al-Sherbaz said: "Infected sheep have a characteristic limp, and we decided that the best way to detect it remotely was to develop a wearable sensor."

And whilst the team's choice for a prototype– a mobile phone worn around the neck of the sheep – might surprise many, Dr Al-Sherbaz says it was an obvious choice.
"All smart phones contain gyroscopes, accelerometers and GPS sensors. So as well as being readily available and easy to collect data with, they provided us with everything we needed to measure sheep movements."

To be exact, the phones recorded **movement along nine different axes** at a rate of ten readings a second, and the PhD student attached to the project, Zainab Al-Rubaye, has spent much of the last two years monitoring **sheep during hundreds of five-minute** test periods.

Which is all well and good - but how does that translate into something that can alert farmers to a distressed animal?

"The first thing we needed was lots baseline data, from sheep we knew to be healthy and sheep we knew to be infected," explained Dr Turner. "Then we carefully analysed these numbers **to find where there was divergence** between the two."

Using this approach (which involved data mining and machine learning) the team created an algorithm that could do just that – or to put it more simply: they taught a computer to distinguish between healthy and lame sheep using just their sensor data.

"Now we know we can tell the **difference, the next step is to build the** software into a bespoke sensor that will automatically alert the farm when a sheep starts to limp," added Dr Al-Sherbaz.

Its potential as a commercial product is already starting to attract attention – leading to a prime-time appearance in November 2017 on BBC's The One Show - but there is still work to be done.

The main hurdle to this will be reducing the power consumption of the sensor, which in turn reduces both its size and cost. And, says Dr Turner, the team are already part way there, having now honed in on the few useful measures of movement, allowing them to completely ignore the others.

"What we would really like now is **to find a commercial partner who** can help us miniaturise the sensor and produce them on a scale that is economically viable for sheep farmers."

Biography:
Scott Turner

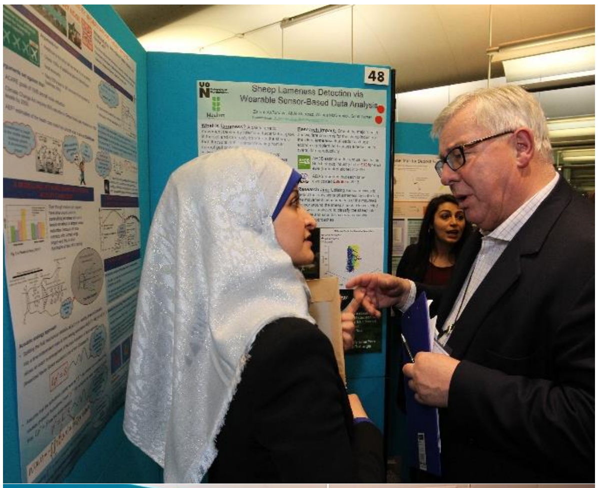
Scott Turner is Associate Professor of Computing and Immersive Technologies at the University of Northampton with an interest in Applied Computing, especially around **Artificial Intelligence, Robotics** and Data Analysis. He is also the Chair of the University's STEAM Steering group, with an interest in engaging the public with Computing and Engineering.

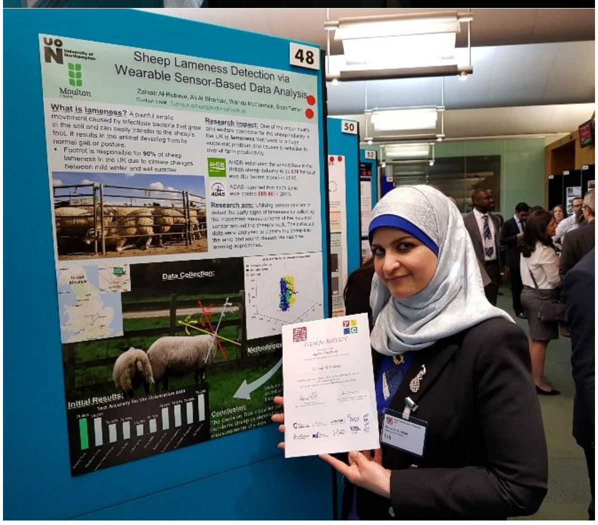
Biography:
Ali Al-Sherbaz

Ali Al-Sherbaz is Associate Professor of Mobile and Wireless Technologies at the University of Northampton with an interest in Connected Vehicles, Sensors, Machine to Machine Communications. He has twenty years theoretical and practical experience in researching and teaching at various universities both in the Middle East and in the UK, with good experience in both academic and industrial projects.

University of Northampton Research Highlights 2016-2017 13

Appendix I. 5 The annual Research Highlights 2016-17 of the University of Northampton includes research story in page 12 and 13.





Appendix I. 6 Gallery from STEM for Britain poster exhibition in March 2018 at the House of Commons/ UK Parliament / London / Westminister.

STEM for BRITAIN

Sponsoring Member

Stephen Metcalfe MP

Chairman, Parliamentary and Scientific Committee

Zainab Al-Rubaye

was selected to present a Poster at the STEM for BRITAIN Exhibition
in the Engineering Section
held at the House of Commons on Monday 12th March 2018

Stephen Metcalfe MP
Chair, Organising Committee, STEM for BRITAIN

Professor Dame Ann Dowling OM DBE FRS FREng
President, Royal Academy of Engineering

Appendix I. 7 Certificate of attendance for the STEM for the Britain Exhibition.

ROYAL ACADEMY OF ENGINEERING

Royal Academy of Engineering
Prince Philip House
3 Carlton House Terrace
London SW1Y 5DG
www.raeng.org.uk

From the President
Professor Dame Ann Dowling OM DBE FREng FRS

Direct tel: +44 (0) 20 7766 0657
president@raeng.org.uk

Mrs Al-Rubaye
University of Northampton

20 April 2018

Dear Zainab,

STEM for BRITAIN: Engineering Exhibition

I would like to congratulate you on behalf of the Academy for being shortlisted to present your poster at the annual STEM for BRITAIN poster competition that took place on 12 March at Portcullis House.

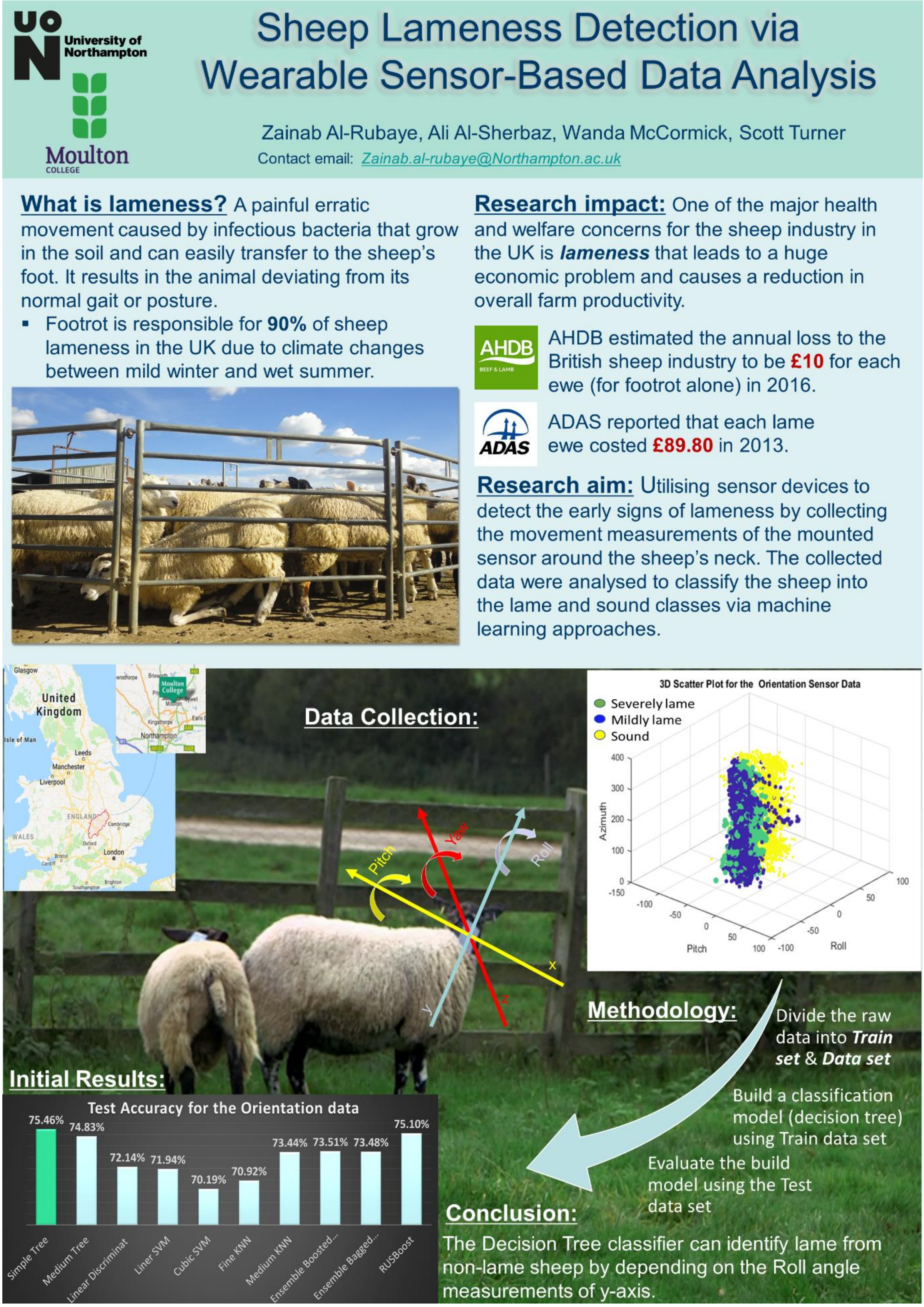
I was delighted to hear from the Fellows of the Academy who judged the competition about all the superb engineering research on display this year from the 45 researchers selected to present their posters. Such talented engineers bringing their work to Parliament demonstrates just how varied the world of engineering is and the many ways in which it impacts our society.

I hope that you enjoyed the evening and many congratulations once again exhibiting your research.

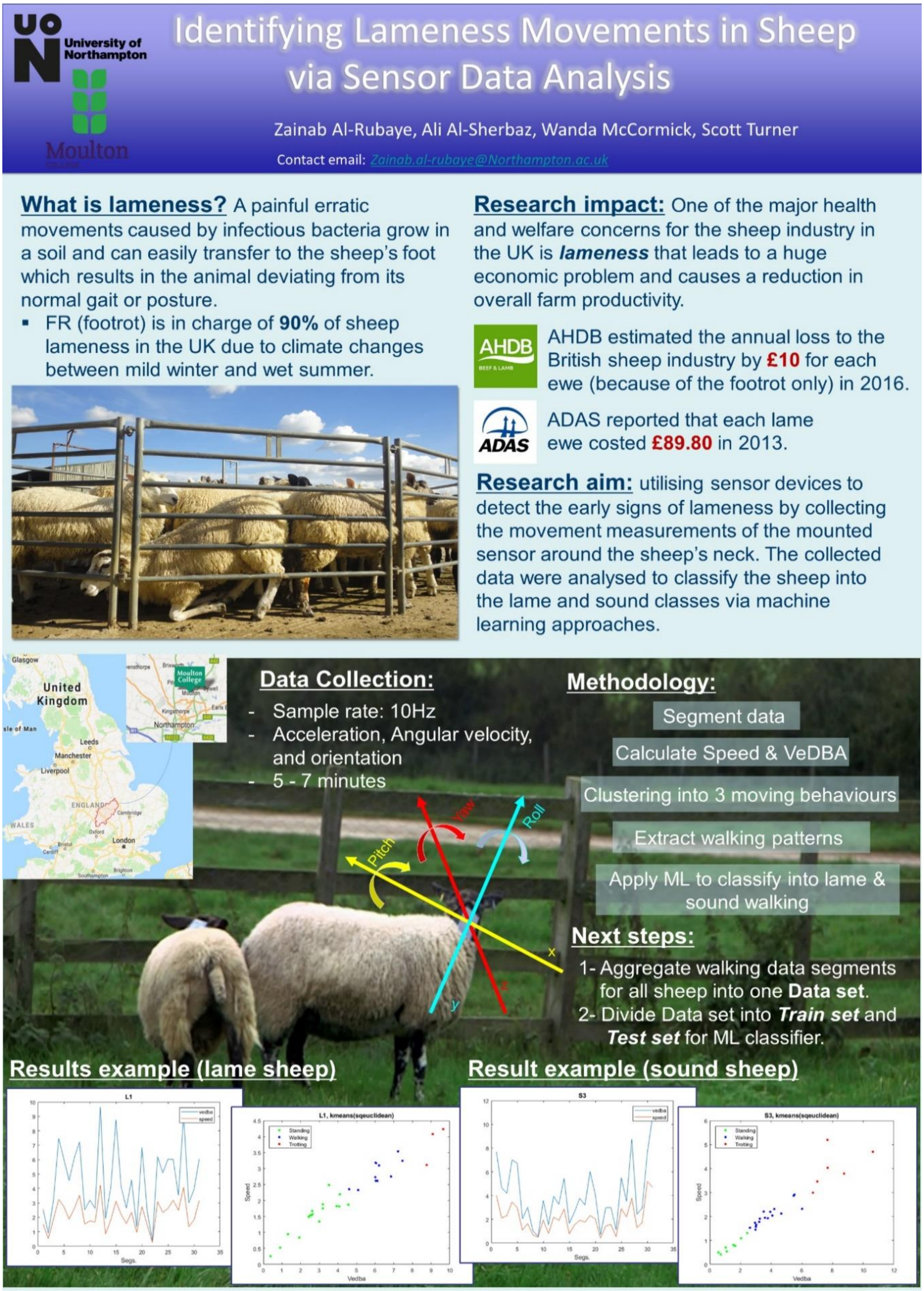
Yours sincerely

Ann Dowling

Appendix I. 8 Acknowledgement letter from Royal Academy of Engineering for being shortlisted to participate in STEM for Britain annual poster competition.



Appendix I. 9 Poster presented at STEM for Britain exhibition in the House of Commons, the UK parliament. 12 March 2018.



Appendix I. 10 Poster presented at Recent advances in animal welfare science VI, UFAW Animal Welfare Conference, Centre for life, Newcastle, UK. 28 June 2018.

Sensor Data Classification for the Indication of Lameness in Sheep

Zainab Al-Rubaye[1,3](✉), Ali Al-Sherbaz[1], Wanda McCormick[2], and Scott Turner[1]

[1] Department of Computing and Immersive Technologies, School of Art, Science and Technology, Northampton NN2 6JD, UK
{zainab.al-rubaye, ali.al-sherbaz, scott.turner}@northampton.ac.uk, zaynebraid@scbaghdad.edu.iq
[2] Department of Biology, Faculty of Science and Technology, Anglia Ruskin University, Cambridge CB1 1PT, UK
wanda.mccormick@anglia.ac.uk
[3] Computer Science Department, College of Science, University of Baghdad, Baghdad, Iraq

Abstract. Lameness is a vital welfare issue in most sheep farming countries, including the UK. The pre-detection at the farm level could prevent the disease from becoming chronic. The development of wearable sensor technologies enables the idea of remotely monitoring the changes in animal movements which relate to lameness. In this study, 3D-acceleration, 3D-orientation, and 3D-linear acceleration sensor data were recorded at ten samples per second via the sensor attached to sheep neck collar. This research aimed at determining the best accuracy among various supervised machine learning techniques which can predict the early signs of lameness while the sheep are walking on a flat field. The most influencing predictors for lameness indication were also addressed here. The experimental results revealed that the Decision Tree classifier has the highest accuracy of 75.46%, and the orientation sensor data (angles) around the neck are the strongest predictors to differentiate among severely lame, mildly lame and sound classes of sheep.

Keywords: Sensor data classification · Machine learning · Decision tree Lameness detection · Sheep

1 Introduction

Lameness is a painful impaired movement disorder, which relates to an animal's locomotion system and causes a deviation from normal gait or posture [1]. In sheep, footrot is the most common cause, resulting in 90% of the sheep lameness cases in the UK [2, 3]. Unfortunately, lameness has a negative impact on the sheep industry and overall farm productivity. Statistics from the Agriculture and Horticulture Development Board (AHDB) estimated the annual UK economic loss to be £10 for each ewe in 2016 [4]. The underlying reasons for the commercial loss in the UK sheep industry can be related to declines in various outcomes, including sheep body condition; lambing

I. Romdhani et al. (Eds.): CollaborateCom 2017, LNICST 252, pp. 309–320, 2018.
https://doi.org/10.1007/978-3-030-00916-8_29

Appendix I. 11 Sensor Data Classification for the Indication of Lameness in Sheep. In *Collaborate Computing: Networking, Applications and Worksharing*. Chapter published in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer International Publishing, pp. 309–320.