Characterising sheep vocals using a machine learning algorithm

A thesis

submitted in partial fulfilment

of the requirements for the Degree of

Master of Applied Science

at

Lincoln University

by

Bilal Nawaz Kayani

Lincoln University

2021

Abstract

New Zealand's economy is mainly dependent on the farming sector and the sheep sector is one of the most important farming sectors, playing a backbone role to the agricultural industry and placing New Zealand among the top five sheep exporter countries in the world. International consumer trends show concerns over the well-being of animals before slaughter and research also indicates potential negative effects on meat quality of stressed animals. Indicators for sheep well-being have largely been limited to physical weight gain and visually observable behaviour and appearance. There has been recent interest but little substantive research on sheep vocalisation as a means of monitoring sheep well-being. This assumes that sheep vocalisation can be classified as representing different states of well-being. Therefore, this thesis investigated the potential to be able to classify sheep vocalisations in a way that would enable automated assessment of the well-being of New Zealand sheep using recorded vocalisations.

A supervised machine learning approach was used to classify the sheep vocals into happy and unhappy classes. Sheep sounds were collected from a New Zealand Ryeland sheep stud farm and online databases. After collection, these sounds were labelled by an expert, pre-processed to make them clean from unwanted background sound noises and features were extracted and selected for classification. Models were built and trained and tested.

Model use in this research shows that sheep sounds were classified into happy and unhappy classes with an accuracy of 87.5%, for the sheep vocals used in this research. Through demonstrating the ability for automated classification of sheep vocalisations this research opens the door for further study on the well-being of sheep through their vocalisations. Future researchers could also collect larger vocal data sets across different breeds to test for breed-related variance in vocalisations. This may enable future sheep well-being certification systems to be established to assure consumers of the well-being of pre-slaughter sheep life.

Key Words: Sheep well-being, farming, machine learning, vocals, economy, well-being, New Zealand.

Acknowledgements

There are many individuals that I would like to thank for their support, encouragement, and care. I feel blessed am thankful to work under the kind and visionary supervisors here at Lincoln University. First, I am deeply indebted and would like to express my gratitude to Dr Hamish Rennie for his robust, dynamic, and kind personality. It was his support, invaluable guidance, and full cooperation at various steps of my research journey that helped me to achieve this target. He nourished my intellect and encouraged me to become an active and invaluable member of academic society.

I am also truly thankful to my associate supervisor, Dr. Shannon Page for his cordial and kind support during my research. His contribution towards the methodology and technical expertise were of great importance. I am grateful for his scholarly guidance and sympathetic attitude. I always enjoyed working under kind supervision of my supervisors and wish my both supervisors a very happy and healthy life.

No words can explain the respect I have for my parents. Their precious and priceless prayers helped me to attain this milestone. They raised me, supported me, and loved me at each phase of my life. They have sacrificed their comforts for my better future and allowed me to pursue my dream of attaining higher education abroad. My mother is a model for me – she is a successful housewife. I also learnt from my father success lies in hard work, dedication, and honesty.

I also want to express my gratitude for my lovely and caring brothers, sisters, and friends (Dr. Farrukh, Dr. Umar, Dr Zeeshan, Sumaira, Kalsoom, Ahmed, Ibrahim, Ayesha, Mustafa and Noor Fatima) for their unconditional support and prayers.

Finally, I wish to thank my ‘Kiwi’ university friends for their warm hospitality and help. I would like to thank all the staff and students in our Faculty, Environment, Society and Design for kind support and motivation. I consider New Zealand as a great country and hope country stays prosperous and strong.

Table of Contents

Abstract.....i

Acknowledgements.....ii

Table of Contents.....iii

List of Tables.....v

List of Figures.....vi

Glossary of Terms.....vii

Chapter 1 Introduction.....**8**

1.1 Introduction.....8

1.2 Study significance and research gap.....10

1.3 Research Objective.....11

1.4 Thesis Organisation.....11

Chapter 2 Literature Review.....**12**

2.1 Well-Being.....12

2.2 Internationally Recognised Animal Welfare Standards.....13

2.3 New Zealand's Animal Welfare Act 1999.....15

2.4 Three Broad categories of Well-being in Sheep.....15

2.4.1 High level of biological functioning.....16

2.4.2 Freedom from negative sufferings.....16

2.4.3 Feelings of positive experiences.....17

2.5 Understanding animal and sheep communication.....17

2.6 Bioacoustics.....19

2.7 Relationship between vocalisation and welfare in animals.....20

2.8 Review of Speech Recognition.....20

2.9 Past approaches to speech recognition.....21

2.10 Feature extraction.....22

2.11 Feature Selection........23
2.12 Models, methods, and classification algorithms........24
2.13 Chapter summary........26
Chapter 3 Research methodology and data collection........28
3.1 Data Collection........28
3.1.1 Farm data collection........28
3.1.2 Online Acoustics Databases........30
3.2 Data Cleaning and Categorising........30
3.3 Classification system........33
3.4 Audio pre-processing........33
3.5 Feature Extraction........34
3.6 Feature Selection and Reduction........36
3.6.1 Feature Selection........36
3.6.2 Feature Reduction........37
3.7 Classification........38
3.7.1 Linear Discriminant Analysis........38
3.7.2 Support Vector Machine........38
3.7.3 Software and relevant libraries used for data classification........39
3.8 Performance Evaluation........40
3.9 Chapter summary........41
Chapter 4 Results analysis and discussion........42
4.1 Modelling of data........42
4.2 Results and Discussion........43
4.3 Chapter summary........45
Chapter 5 Conclusion........47
5.1 Summary of research findings........47
5.2 Implications........48
5.3 Research limitations and future research work avenues........48

List of Tables

Table 3.1 Confusion (contingency) matrix of a binary-class problem……………………………………………………..40

Table 4.1 Accuracy of models……………………………………………………………………………………………………..43

Table 4.2 Confusion Matrix Table of SVM Model…………………………………………………………………………..43

Table 4.3 Confusion Matrix Table of LDA Model…………………………………………………………………………..43

Table 4.4 Three Basic Measures of Performances…………………………………………………………………………44

List of Figures

Figure 3.1 Flock of sheep used for data recordings...29

Figure 3.2 Sheep sound with background noise...30

Figure 3.3 Cleaned sheep sound signal after removing background noises...31

Figure 3.4 Block diagram of the sheep classification system...32

Figure 3.5 Support vector machine...39

Figure 3.6 K-fold (10-fold) cross-validation...39

Glossary of Terms

Terms	Definition
Fisher score-based feature	Fisher score is one of the most widely used supervised feature selection methods. Fisher score algorithm selects each feature independently in accordance with their scores.
Multi-layer perceptron network.	A multi-layer perceptron (MLP) has the same structure of a single layer perceptron with one or more hidden layers.
Mel-frequency cepstral coefficient	In sound processing, the mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel-scale frequency.
Single-layer perceptron network.	A single layer perceptron (SLP) is a feed-forward network based on a threshold transfer function. SLP is the simplest type of artificial neural networks and can only classify linearly separable cases with a binary target (1, 0).
Supervised Machine learning classifier	Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence. It is defined by its use of labelled datasets to train algorithms to classify data or predict outcomes accurately.

Chapter 1

Introduction

1.1 Introduction

With the value of animals to the growing human population, the growing demand for animal protein, and growing concern in markets for animal welfare (Montossi et al., 2013), it is very important that animal well-being should be given priority in order to have healthy animal life cycle. Concerns over the welfare of animals grown in captivity for consumption was initially couched in terms of animal health while there were no international agreements on animal welfare. The need to fight animal diseases and ensure their better well-being led to an international agreement signed on January 25th, 1924 for the creation of the Office International des Epizooties which later changed to the world Organisation for Animal Health (but the acronym OIE was kept for the original name) in May 2003. OIE being an intergovernmental organisation is responsible for improving animal health worldwide and has a total of 182 Member Countries, indicative of the importance placed on animal welfare generally at an international level. In addition to the OIE, the Food and Agriculture Organisation of the United Nations (FAO) is also playing a pivotal role globally for the well-being of animals. The presence of both these globally recognised bodies demonstrates the global intergovernmental interest in animal health and their well-being (Epizooties, 2020). Therefore, it is very important that animal health be taken seriously, and better well-being steps taken for their overall health.

Domestic laws have also been promulgated addressing animal well-being. This reflects a rising concern for animal rights and more generally for animal well-being for example United Kingdom's Animal Welfare Act 2006 and United States Animal Welfare Act 1966 (Government of the United Kingdom Animal welfare Act 2006; United States Government Animal Welfare Act 1966). While people are happy to eat animal products, surveys have shown that there is an increasing interest in knowing that the animals lived in good conditions while being raised, before slaughter (Terlouw et al., 2008). Also, there is an increased environmental, social and consumer pressure on pastoral livestock production to reduce environmental impacts and to enhance animal welfare and biodiversity (Gregorini, Villalba, Chilibroste, & Provenza, 2017). From a taste perspective, research has also indicated that less stressed animals result in better tasting animal products (Thorslund, Sandøe, Aaslyng, & Lassen, 2016).

A concurrent concern for the maintenance of healthy food chains and secure access to food to meet growing food demands has focussed attention on the animals' health and productivity, which can be seen from the rise of certification systems provided by different organisations (e.g. global animal welfare certification by global animal partnership of the need to address the wide range of ethical issues and rising demand from retailers and the food industry for seeking higher animal welfare standards from the producers of food of animal origin. Montossi et al. (2013) have suggested that precision agriculture enabling better measurement and monitoring of individual animals (e.g. sheep) promises a way to ensure multiple demands are met. Precision agriculture relies on advanced technology assessment and monitoring and rapid information analysis to enable responses. It is within this context of an increasing interest in the well-being of an animal on the farm and advanced technological data analysis that the current research is situated.

There are more than 1.2 billion sheep in the world (FAO, 2002) and understanding the behaviour and well-being of sheep is an interesting research prospect in itself. It might also have potential commercial value through consumer-friendly certification labelling. This is methods of monitoring their well-being. If sheep are 'happy' it may be assumed that they are more likely to be enjoying reasonable well-being. The concept of "happiness" might fit well with the consumer market.

Indicators of sheep well-being have predominantly depended on visual or biophysical/biochemical observations by farmers and veterinarians. Sheep vocalisations offer a means of remote collection of another form of data, but the usefulness of vocalisation data for large numbers of sheep depends on our ability to accurately and meaningfully classify such data. Machine learning algorithms offer a means to achieve this, but until the current research there have been no reported attempts to apply machine learning to sheep vocalisations.

Therefore, this research study aims to classify the sheep vocal data into 'happy' and 'unhappy' categories using machine learning.

1.2 Study significance and research gap

This study has chosen New Zealand as research context for characterising sheep vocals into happy and unhappy states, because New Zealand is one of the World's top ten countries in sheep farming (with 26.7 million sheep counted at 30 June 2019) and the sheep farming sector is one of the major contributing sectors to New Zealand's economy with export revenues forecasted to be $3.4 billion for year 2019-2020 (Beef and Lamb, 2019). There are some challenges being faced by this sector and the well-being of sheep is one (Stafford, 2017) which needs to be addressed.

This research project aims to design, implement, and classify sheep vocals into happy and unhappy. As the literature review shows lack of studies were found for charactersing sheep vocals into 'happy' and 'unhappy' states using machine learning.

1.3 Research Objective

The research objective of this research study is to classify the vocalisations of sheep into states of happiness and unhappiness through vocalisation analysis using a machine learning approach.

1.4 Thesis Organisation

This thesis is organised into five chapters. Following this introductory chapter, Chapter 2 provides a review of relevant literature on approaches used to classify animal well-being, and in particular sheep well-being, based on their vocal patterns. This enables the nature of the gap in knowledge to be clearly identified and provides the theoretical context for the research. Chapter 3 provides the methods used, which comprises four parts. The first part discusses the process of data collection, data labelling, followed by feature extraction, feature selection and reduction in section 3.5 and the classification model used for sheep vocals in section 3.6. Chapter 4 presents the results analysis and discusses key findings of this research based on the classification models used for classifying sheep vocals. This is followed by the overall conclusion and recommendations for future researchers in Chapter 5.

Chapter 2

Literature Review

2.1 Well-Being

The concept of well-being varies among academicians, policy makers, psychologists, or interior designers. It is an individual perception related to a specific state of mind in different contexts (White, 2008). Every individual has their own aims, wants/desires, aspirations, motivation levels and preferences so the experience of well-being also varies a lot from an individual to individual. For some it is the level of their own happiness and success, but for others it may be beyond their own individual interest and related to the welfare or the good of society. Therefore, to define the concept of well-being precisely on any uniform standard is extremely difficult (Dolan & Metcalfe, 2011).

The relationship between humans and animals is centuries old and holds much importance in human civilisation in terms of getting various benefits from animals such as meat, milk and wool (Blood & Studdert, 1988). Terms well-being and welfare have been used increasingly by consumers, veterinarians, corporations and others (Hewson, 2003). Similarly, like measurement of well-being in humans, it is a quite difficult and challenging task to provide a uniform definition of animal well-being or welfare (Hewson, 2003).

The various authors have attempted to define animal well-being. It is defined as the strengths/abilities of an animal to cope with physiochemical and social life environment in terms of physiology, behaviour, and cognitive aspects (Scott & MacAngus, 2004). Whereas C.R.W. Spedding (1965, p.3) defined animal welfare as: “Animal welfare is a State of well-being, in which at least basic needs are met and suffering is minimised”. Similarly, D.M. Broom

(1991, p91) defined the welfare as: “The state of an individual in relation to its environment, and this can be measured. Both failures to cope with the environment and difficulty in coping are indicators of poor welfare”.

The two terms animal well-being and animal welfare are used interchangeably in existing research (Broom, 1988). Despite this, Moberg (2013, p.2) defined well-being as: “The state of being happy, healthy or prosperous”. However, (Moberg, 1985) also argued that to understand and define well-being for a particular animal is too vague.

In summary, animal well-being describes how an animal is coping with the conditions in which it lives. The conditions for animal well-being may be named as a good state of welfare if (as indicated by scientific data) it is healthy, happy, having proper food, safety, and able to express natural behaviour. An excellent state of well-being would mean that an animal is free of diseases, has good access to food, areas to shelter in and treatment available on becoming sick.

2.2 Internationally Recognised Animal Welfare Standards

As noted in the Introduction Chapter, animal welfare has gained attention around the globe and has led to an increased public interest (Fidler, 2004; Grandin, 2005). Most of the countries aim to achieve and implement the internationally recognised set of animal welfare standards as they are considered as foundation stones for modern animal welfare. These five standards laid out by Brambell (1965) are known as five freedoms and recommend:-

- *Freedom from hunger and thirst*: availability of clean water and a well-balanced diet would keep animals in good health. Access to clean water and a well-balanced diet will keep animals healthy and strong.
- *Freedom from discomfort:* living area of animals should be calm, spacious, and clean. As providing a neat and clean shelter and environment ensures the animal is healthy and happy.
- *Freedom from pain, injury or disease:* proper care and immediate access to veterinary staff in situations like pain, injury or disease can greatly improve wellbeing and welfare level in animals.
- *Freedom to express normal behaviour*: natural environment should be provided to animals including enough space for moving around and freedom to express their emotions.
- *Freedom from Fear and Distress*: all animals deserve an environment free from fear and distress. Such an environment can help in minimising stress and anxiety and ultimately creates positive and contented feelings, which improves welfare and well-being in their lives.

All these freedoms together ensure greater well-being life cycle of sheep and other animals, therefore it's important to pay more attention towards these above five international animal welfare standards as they affect positively or negatively well-being life cycle of sheep.

This research is for New Zealand sheep as it is an important sector for New Zealand economy, hence it is instructive to discuss some relevant provisions of New Zealand's Animal Welfare Act 1999.

2.3 New Zealand's Animal Welfare Act 1999

The Animal Welfare Act 1999 provides detailed guidelines and standards to be followed for the well-being of animals (Farnworth, Campbell, & Adams, 2011; Mellor & Stafford, 2001). The Act is administered by the Ministry for Primary Industries (MPI) with an objective to create an awareness among New Zealanders about the welfare of animals and lays down overall guidelines for a good and humane treatment of animals for peaceful and healthy living.

Section 4 of Animal Welfare Act 1999 defines physical, health, and behavioural needs in relation to an animal. It includes availability of healthy and abundant food, clean water, secure housing, freedom to exhibit natural behaviour and protection from injuries and disease. Therefore, it is important to first understand the three broader categories of sheep well-being, being a key focus of this study. Based on the existing literature, in the next section we will discuss three main categories of well-being in sheep.

2.4 Three Broad categories of Well-being in Sheep

As per the discussion of (Moberg, 1985), the International standards and the New Zealand Animal Welfare Act indicate three main factors are important in establishing the well-being of sheep. These have also been considered as those factors internal to a sheep and those that are external to a sheep which play vital roles in determining an animal's state of well-being such as physical stimuli, biological responses, heredities, environment (Villalba, Provenza, & Shaw, 2006). Based on these stated factors, well-being of sheep is mainly categorised into three categories, biological functioning, negative and positive experiences (Marai, El-Darawany, Fadiel, & Abdel-Hafez, 2007).

2.4.1 High level of biological functioning

Biological functioning can affect the level of well-being in animals either positively or negatively. Poor resistance to combat disease is an indicator of poor well-being and weak immune system (Broom & Fraser, 2015; Fraser, 1988). Similarly, Dawkins (1988), concludes that symptoms of disease, injury and malnutrition are the indicators of biological suffering, which is broadly acknowledged by the researchers (Broom, 1998; Morton & Griffiths, 1985). Disease and injury are almost by definition indicative of a lack of health and associated well-being. Less immediately apparent but also indicative of lack of well-being is stagnant or low growth. Biological functioning plays a vital role for maintaining well-being in an animal's lifecycle. Presence of any disease directly affects the life of sheep. Sheep which are not meeting their biological needs are more likely to have weak immune systems.

2.4.2 Freedom from negative sufferings

Negative experiences like persistent pain, anxiety, discomfort, lack of food and water are the main reasons behind the poor well-being of animals (Rushen, 1986). Various researchers (Archer, 1988a, 1988b; Broom, 1991; Gendreau & Archer, 2005**)** performed experiment on animals experiencing negative conditions (such as hunger, thirst, insufficient food and living environment) and compared these animals free from those experiences. Their results show that animals experiencing negative conditions remain isolated and did not socialise with other sheep in the flock. Verstegen and Close (1994) found that one of the most likely negative experiences for animals including sheep is related to hunger and thirst. Also, conditions like fear and distress experienced by sheep may have potential impact on their well-being and needs to be minimised for their life cycle (Cresswell, 1960).

2.4.3 Feelings of positive experiences

Positive feelings or experiences play an important role for the sheep well-being (Harrison, 1964). Some of these positive feelings studied by Harrison for examining higher levels of well-being are: comfort, contentment and free will to do activities for pleasure and happiness in their life. Harrison (1964) experiments found that depriving animals of pleasure and wishes in an experiment by separating them from other animals and removing space for grazing were the main reasons for low levels of well-being in animals.

In summary, animal well-being revolves around three main goals, which are: excellent biological functioning, eliminating, or minimising the negative experiences and feelings of positive experiences.

2.5 Understanding animal and sheep communication

Animals communicate with other animals through transferring of information. This information transfer will have a sender (animal transferring information) and a receiver (can be an animal or humans receiving that information). Usually, communication happens between animals within the same species, but can happen between different species as well. Common means of animal communication include:

- Visual,
- Auditory (sound-based), and
- Body gestures. (Laidre, 2012)

Touch-based communication helps animals in finding their mates, defending their territorial boundaries, coordinating for a group behaviour, establishing dominance, and for caring purposes (McGregor, 2005).

It is important to know about sheep behaviour and communication as the focus of this research is on studying sheep vocals for knowing their well-being state. Sheep behaviour (ethology) is little studied in New Zealand (Guadarrama-Maillot & Waas, 2008) but is assumed to rely on communication(Dwyer, 2004). This assumes that sheep communicate their well-being therefore some important sheep communication details are discussed below.

There may be cases where an animal might not be transferring information for communication to other animals but making sounds to show their internal states through which they are passing. Most of the sounds made by sheep include bleats, rumbles, snorts, and grunts (Van Compernolle, 2001; Welch, Leege, Wald, & Kellstedt, 1993). All these sounds have been found to have specific purposes for communication: Bleating is used mostly as a communication for contacting between ewes and lambs, occasionally between other flock members. Bleats of an individual sheep are unique, which enables a ewe and her lambs to recognise each other's vocalisations (Sebe, Aubin, Boue, & Poindron, 2008).

Besides contact communication, bleats may be used for signalling distress, frustration or impatience and isolation. Pregnant ewes grunt, usually when they are in labour (McCune, Vihman, Roug-Hellichius, Delery, & Gogate, 1996). Rumbling sounds are made by rams during courting and by the ewes while with her neonate lambs (Dwyer et al., 1998). A snort is an explosive exhalation through the nostrils and signals aggression and warning (Baotic, Sicks, & Stoeger, 2015). These sounds/calls are often produced or prompted by startled or frightened sheep (Forkman, Boissy, Meunier-Salaün, Canali, & Jones, 2007).

Acoustic utterances play an important role in communication between animals. These sounds can travel long distances and help other animals to keep in contact and share information with them. Vocalisation is an expression of the inner state of an animal, that may occur spontaneously or as a result of some external event or happening (Gregory & Grandin, 1998).

Vocalisations, therefore, are an important indicator of an animal's state of welfare and also indicate the needs of an animal (Dawkins, 1988). Animal vocalisations are considered a useful tool in knowing well-being or stress in an individual animal (Weary & Fraser, 1995).

2.6 Bioacoustics

For the purpose of this research, auditory communication was the focus. This reflected the assumption that sheep sounds might be more readily able to be learnt and used by automata than visual and touch data. For instance, touch and body gesture communication might be too subtle to record unless a camera was sufficiently well targeted, from the right angle and provided sufficient scope to observe the respondent sheep. The field of bioacoustics is a result of multi-disciplinary research that deals with the study of vocalisations other than by human beings. The main goal of bioacoustics revolves around studying and determining the role of animal vocalisations in the communication process (Manteuffel, Puppe, & Schön, 2004). It is therefore ideally suited to draw on for examining auditory communication of well-being of sheep (Bishop, Falzon, Trotter, Kwan, & Meek, 2017, 2019). Recently, the field has gained much importance, because of the availability of better scientific tools (e.g. sound analysis, bioacoustics, and software such as audacity and praat).

Researchers from varying disciplines are contributing to this field (e.g., biologists, animal behaviourists, psychologists). More recently engineers from the signal processing field having communication expertise are also contributing to the field because of better recording technologies and ease in gathering accurate and real-time data related to an animal's environment (Clemins, Johnson, Leong, & Savage, 2005). There is a considerable growth and improvement in the analysis techniques of vocalisations but still there is a technology gap between animal and human vocalisation analysis techniques (Bardeli et al., 2010).

2.7 Relationship between vocalisation and welfare in animals

Social context plays a pivotal role in animal communication. Since early 1970s there has been interest in the vocalisation of animals for instance: cattle vocalisation research (Kiley, 1972) and this has led to an increased interest in the field of animal well-being and vocalisations (Cordeiro, Nääs, Baracho, Jacob, & Moura, 2018; Manteuffel et al., 2004; Smith, Bruner, & Kendall, 2019). Vocalisations in cattle signal their physiological and emotional state (Watts & Stookey, 2000). For example, there is an increase in calling rate between cows and calves after separating from each other (Weary & Chua, 2000). These studies provide evidence that vocalisation in animals can be very useful in assessing their needs and stress levels and thereby providing a measure of welfare.

Searby and Jouventin (2003) studied acoustic recognition between ewes and lambs by spectrographic analysis of their vocalisation. Results of their research revealed that ewes and lambs can recognise each other solely based on their calls.

2.8 Review of Speech Recognition

As the above review reveals, speech is a natural means of communication among mammals, not just humans. Speech recognition is an inter-disciplinary field, incorporating knowledge from linguistics, computer science, and electrical engineering. Speech recognition systems require training for translating speech to an understandable form, so systems that need training are known as speaker dependent systems, while systems that do not require training are called speaker independent systems (Graves, Mohamed, & Hinton, 2013).

Vocal patterns from one animal to another animal vary in pronunciation, delivery, nasality, tone, and swiftness. Background noise and echoes may distort or reduce the quality of vocal patterns during transmission. Therefore, interpreting vocal patterns becomes an even more complex problem because of the variations and distortions in sound patterns (Hirsch & Pearce, 2000).

It is important to highlight difference between voice recognition and speech recognition. The term voice recognition refers to the identification of 'who is saying' regardless of what is being said. Whereas, the term speech recognition refers to the 'recognition of what is said' rather than 'who is saying' (Perrachione, Del Tufo, & Gabrieli, 2011). The focus of this research is on the speech recognition of sheep rather than voice recognition as the intent was to measure the well-being of sheep.

2.9 Past approaches to speech recognition

Automatic speech recognition started gaining popularity among researchers around the 1950s. Past approaches like the template matching approach, knowledge based approach and statistical based modelling approaches were commonly used (Juang & Rabiner, 2005).

The expert-based template approach deals with templates having pre-recorded words, where speech is compared against this template to find the best possible match (Povey et al., 2011). These modelling variations in speech help in differentiating speech but at the same time are challenging because of not easily available experts-based knowledge. Statistical models like Hidden Markov Models (HMMs) are used for bioacoustics analyses (Ranjard et al., 2017). Modern techniques of digital signal processing, power spectral density, linear predictive coding (LPC) and hidden Markov models (HMMs) were used to identify and recognise two psychologically stressed conditions in cows – hungry before feeding and separating from her calf based on the voice signals (Ikeda & Ishii, 2008).

Based on the successful models for human speech and recognition like the hidden Markov model, Clemins et al. (2005) used the features of Mel-frequency cepstral coefficients (MFCCs) features, which was proposed by Noll (1964) and plays a pivotal role in features extraction (explained in next chapter's section 3.4) for classifying African elephant vocalisations. The success in doing so supports classification of vocalisations in nonhuman species.

2.10 Feature extraction

A feature set (audio attributes) is the main component of an audio-based classification (Bishop et al., 2017, 2019; Yongwha Chung et al., 2013; Jahns, 2008), various different features or combinations of multiple feature sets have been used in different animal studies (Ovaskainen, Moliterno de Camargo, & Somervuo, 2018; Schrader & Hammerschmidt, 1997; Souli & Lachiri, 2018; Weninger & Schuller, 2011).

Lee, Chou, Han, and Huang (2006) identified animals from their sounds by using the averaged Mel-frequency cepstral coefficients (explained in next chapter's section 3.4) technique for calculating vocalisations features and linear discriminant analysis (LDA) for classification in his experiment. As noted previously, vocal patterns from one animal to another animal vary in pronunciation, delivery, nasality, tone, and swiftness. Background noise and echoes may distort or reduce the quality of vocal patterns during transmission. Therefore, interpreting vocal patterns becomes even a more complex problem because of the variations and distortions in sound patterns (Hirsch & Pearce, 2000).

Bioacoustics analyses were restricted to the sonograms before the availability of digital signal processing with the discrete Fast-Fourier-Transformation techniques (Cooley & Tukey, 1965). Now, numerical descriptions and statistical examinations of vocal utterances of animals can be done with help of Digital Signal Processing (DSP) (Hopp, Owren, & Evans, 2012). Animal calls are complex, and extraction of useful features (sound characteristics) is still a problem today.

Too few or too many features will lead to incorrect classifiers accuracy. Therefore, well-defined, and widely used methods are required for the feature extraction, because the better the features we have, the more accurate classification results would be. Feature extraction from time domain is the most used available method, which entails extracting information about features such as duration, loudness of calls, means and standard deviation of energy.

2.11 Feature Selection

After feature extraction from data comes the next stage of analysis which is feature selection. This is required as extracted features are large in number and there are many that may be irrelevant and redundant which would increase time and computational needs and makes it difficult for interpreting the classification results. The most common used feature section method for supervised feature section is wrapper and filter-based feature selection methods. A filter-based method is used before classification as a pre-processing step and selects features with high-ranking score, which is used in this research (explained in glossary) method (Gu, Li, & Han, 2012; Guyon & Elisseeff, 2003; Yang & Pedersen, 1997).

2.12 Models, methods, and classification algorithms

Modern statistical-based speech recognition algorithms have two important modelling pillars: acoustic modelling and language modelling. There are many models used for classification based on the research applicability and requirement. Some of the important ones are discussed below.

The first computational based model known as binary threshold model was proposed by McCulloch and Pitts in 1943. This model has a binary output either 0 or 1. The model got much recognition and appreciation for further research especially for learning procedures that would automatically find the values of weights for a network McCulloch and Pitts (1943). Rosenblatt and Gaponoff (1984) later discovered an interactive learning procedure for single-layer perceptron networks.

The following section will discuss the dynamic time wrapping (DTW) algorithm. DTW algorithm enables similarities to be found between two temporal time series sequences and can be used for speech recognition (Itakura, 1975, 1990; Sakoe & Chiba, 1978; Sakoe, Chiba, Waibel, & Lee, 1990; Vintsyuk, 1971).

In late 1980s neural networks emerged as a great approach for acoustic modelling in automatic speech recognition and interest in the neural networks research area revived when Hopfield (1987) studied and proposed that a network can be analysed in terms of an energy function, prompting the development of the Boltzmann Machine in 1988 (Aarts & Korst, 1989). Soon, after that researchers such as (Fodor & Pylyshyn, 1988; McClelland & Elman, 1986; Usher & McClelland, 2001) proposed a much faster learning procedure called backpropagation, which could train a multi-layer perceptron. Neural networks become very popular because of backpropagation research. Their research work gave a hope and motivation to researchers that the goal of achieving machine intelligence was now within their reach (Fogel, 2006).

There is a growing interest among researchers to use Support Vector Machine (SVM) classifier approaches. A support vector machine is a supervised machine learning algorithm which works on the categorised and labelled data for data classification purposes and provides greater flexibility and simplicity for a range of classification problems and is used in research studies where data is scarce and not abundantly available (Pisner & Schnyer, 2020). Abbasi, Derakhshanfar, Abbasi, and Sarbaz (2013) used a SVM in their research for classifying normal and abnormal lung sounds and found that SVM are a very good classifier with classification accuracy about 93.51% for classification of lung sounds while (Bouril, Aleinikava, Guillem, & Mirsky, 2016) classified normal and abnormal heart sounds of 3000 cardio recordings using a support vector machine and found the classification accuracy of 78.64 percent. Similarly in another study by Travieso, Yadav, Singh, and Dutta (2019) for classifying cardiac diseases using normal and abnormal heart diseases found support vector machine classified with an accuracy of 97.78 percent. Therefore, SVM has been used by various researchers for supervised machine learning classification of audio sounds across the varying areas of research work like classifying heart diseases based on heart sounds and bird species recognition (Abbasi et al., 2013; Choi & Jiang, 2010; Ko, Park, & Ko, 2018; Satapathy, Dehuri, Jagadev, & Mishra, 2019; Travieso et al., 2019). In this research SVM will be used for classification based on the prior available research and suitability in this research context.

Linear discriminant analysis is a commonly used supervised classification technique and is used for classifying multi classes or category data. LDA is often used because of its robustness and interpretable classification results (Gaber, Tharwat, Ibrahim, & Hassanien, 2017). Linear discriminant classifier has been used for varying classification purposes such as to detect cough sounds in pediatric wards (Amrulloh, Abeyratne, Swarnkar, Triasih, & Setyati, 2015) identifying respiratory diseases from lung sounds (Fraiwan et al., 2020).

2.13 Chapter summary

As per literature review, it is evident that animal well-being is an important research area which needs to be given more attention and vocals may provide an important, means of identifying the well-being of animals. From literature we can conclude that for vocal analysis of the animal sounds, their audio-based features need to be extracted and then used to train and then to test the machine leaning model for gauging the accuracy of the model.

As described in the literature, linear discriminant and support vector machine models have been used to classify animal vocals (Cordeiro et al., 2018; Manteuffel et al., 2004; Smith et al., 2019), but none of them have been used for studying sheep vocals. In this research study besides using a linear discriminant analysis model, a support vector machine classifier seems likely to aid the model's suitability for binary classification of supervised data (i.e., happy, and unhappy sheep vocals).

In summary, based on the available literature review support vector machine and linear discriminant analysis has been used for binary supervised classification in different areas of research. As per the literature review, the author has found no study which has been conducted which developed a model for classifying happy and unhappy sheep sounds using either a support vector machine or linear discriminant analysis classifier. Therefore, this research study is the first of its kind utilising machine learning algorithms (support vector machine and linear discriminant analysis) to attempt to classify sheep vocals into happy and unhappy sheep sounds for the well-being purpose. In the next chapter the complete methodology steps used or this research study are discussed.

Chapter 3

Research methodology and data collection

This chapter describes the data collection process of sheep sounds in section 3.1 followed by pre-processing techniques used to clean the data and fit for training the model in section 3.2 and 3.3. Classification system block diagram is presented in section 3.4, followed by feature extraction discussion in Section 3.5. Feature selection and reduction is discussed in section 3.6 followed by classification models and software used to classify sheep vocals into happy and unhappy classes based on their vocals in section 3.7 and finally the overall performance evaluation of the model is discussed in section 3.8, followed by chapter summary in section 3.9.

3.1 Data Collection

Data related to sheep vocals in different situation like hunger, panic etc. is not easily and amply available because sheep, in most cases, remain quiet, making calls very occasionally unless disturbed. Initially it was intended to use collars with recorders to capture individual sheep sounds, however the funding for this did not eventuate. Similarly funding constraints meant that in field recorders such as those used by (Bishop et al., 2017, 2019) were also not obtainable. The methods used for sheep sounds data collection are discussed below.

3.1.1 Farm data collection

A university staff member was enlisted to record the sounds from his own flock of sheep. The person selected was chosen on the basis that they were someone having enough sheep farming experience, expertise, availability, and willingness to record sheep sounds on the spot when the sound is produced. They also needed ready access to the sheep. The staff member, Associate Professor Hamish Rennie has sheep farming expertise and co-owns a sheep stud[1]. The sounds were recorded from the sheep farm during the time of late October and November, 2018. Recordings made were the recordings of ryeland sheep (rams, ewes and lambs) from the Eclectic Energy owned Eel Bend Ryeland Stud (Flock No.270 in NZ Flock Book) at Irwell, New Zealand, using a Samsung Galaxy J5 cell phone.

Docility is a recognised trait of ryeland sheep and it was a small stud of less than 100 sheep, shown in figure 3.1. The owners treated the sheep as pets and were still building the stud flock, leasing rams rather than selling them, and had not sold any sheep. Consequently, he had a sufficiently close relationship with the sheep that he could move freely around the paddocks and approach close to the sheep (within 2m) on the farm without disturbing them. An opportunistic approach was adopted. When a sheep was vocalising he would start recording. Often this could be anticipated by observed behavioural characteristics, but on many occasions when the recorder was turned on in anticipation sheep bleats were likely nothing was recorded as they had either stopped or did not start when anticipated. A total of 40 sound samples were collected from the farm in this way.

[1] Dr Rennie subsequently became a member of the supervision team during the closing stages of the research due to his familiarity with the research and the unavailability of one of the original supervisors.

Figure 3.1 Flock of sheep used for data recordings

3.1.2 Online Acoustics Databases

The internet was the second data source, online acoustic databases were searched for sheep audio calls and there were several where multiple sheep audio segments are freely available on audio databases as highlighted below. Online databases used for sheep vocals collection purposes were:

http://soundbible.com

https://www.zapsplat.com

http://www.findsounds.com/ISAPI/search.dll?keywords=sheep

http://www.soundarchive.online/lots-of-sheep-baa-ing-and-bleating-near-and-far-3-2-10036-mp3-audio-sound-free-download-voices-mammal/

https://retired.sounddogs.com/results.asp?Type=&CategoryID=1003&SubcategoryID=52

3.2 Data Cleaning and Categorising

To address the aim of the research, the data needed to be cleaned, categorised and labelled in terms of the well-being of the sheep. In other words, identifying what a particular vocalisation indicated the sheep's state of well-being might be at that point in time. It was impractical to both record the sounds and assess the physical or biological functioning of a sheep in a sense of veterinary science for the field collected data. It was assumed that most of the sheep, being on a stud farm and treated largely as pets, had probably had better than bad life experiences and were well fed and watered, unlikely to be in distress. Such assumptions could not be made for the sounds collected from the online datasets although in some cases accompanying videos enabled a sense of the context within which sounds were being made. Data cleaning from background noises e.g. winds, vehicles etc. was a problematic task. Therefore, it a considerable time and a lot of effort went into the cleaning process of the data, which highlights collecting good data is not an easy task e.g. at one instance while collecting a sheep sound sample that was happy grumbling sound while a sheep was grazing but it was unable to be distinguished from the background sound of the sheep biting the grass and masticating. Figures 3.2 and 3.3 represent sheep sound signal recorded with background noises and cleaned sheep sound signal from background noises.

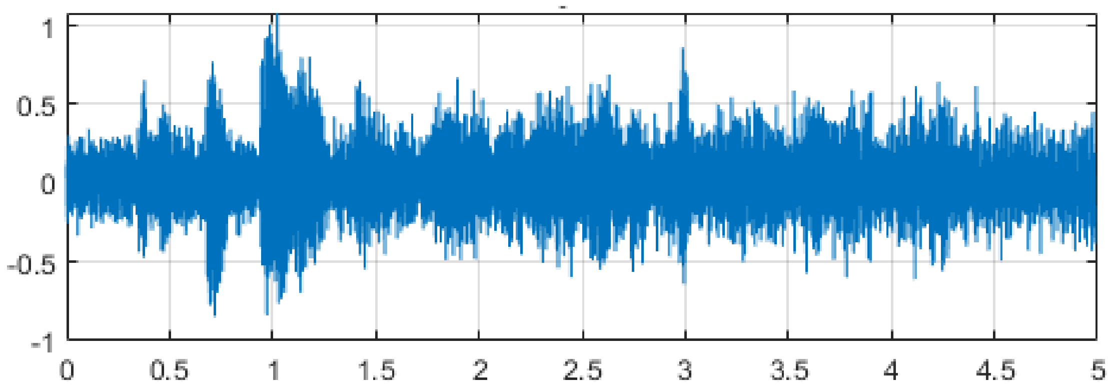

Figure 3.2 Sheep sound with background noises

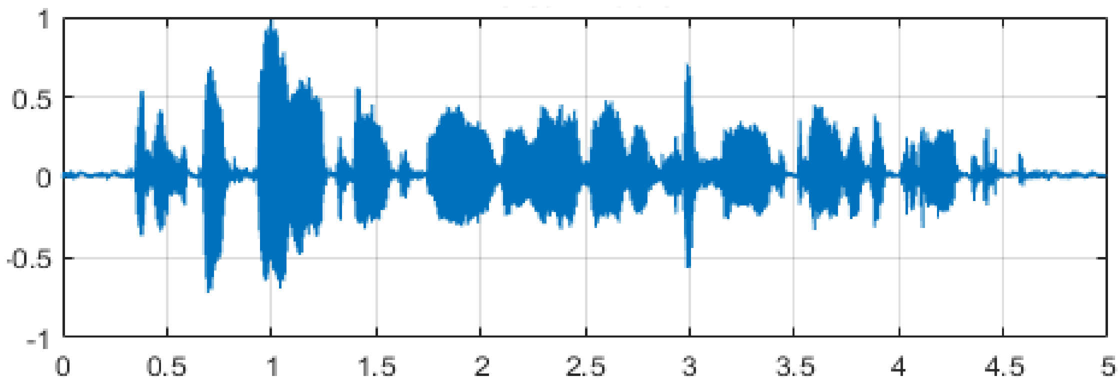

Figure 3.3 Cleaned sheep sound signal after removing background noises

Labelled sheep vocals were not found in any repository or sounds database collection over the internet. So, after collecting the data from both sources – internet and from the farm, it was imperative to label those audio recordings using a supervised classification technique. Dr

Rennie who grew up on a sheep farm and has been involved with stud farming for eight years and is undertaking research on sheep sociology, farming and sheep rearing and was familiar with the data was used for the expert classification process.

The categorising of the data involved his listening to the data he had collected and labelling the data based on his understanding of the meaning of the sounds. This occurred some weeks after the data had been collected, the sound data samples were randomised in their presentation to him and were not connected with any information that would enable him to identify the sheep or circumstances of the recording. The labelling was done in two broad categories of animal well-being: happy and unhappy classes. Due to the data scarcity for machine learning classifiers, sheep audio sounds were labelled happy and unhappy sheep sounds. The total number of sound samples collected were 137 which includes 97 sounds collected from online databases.

Sheep vocals were labelled once and then the same already labelled sheep vocals were represented to Dr Rennie for classification for data validity through the double-blind check and a total of 100 samples were finally selected from 103 sound samples, giving a data labelling margin of an error of 2.91 % and overall data labelling consistency to be 97.08 %. In total 100 labelled sheep vocal samples are used in this research study, which includes 49 'happy' sheep sounds samples and remaining 51 'unhappy' sheep sound samples. 35 happy sheep sound samples out of 49 total happy sheep samples were collected from farm and remaining 14 happy sheep sound samples were collected from online sound databases. Similarly, 46 unhappy sheep sound samples were collected from online databases out of 51 total unhappy sheep sound samples and remaining 5 unhappy sheep samples were collected from farm.

3.3 Classification system

The below figure shows the block diagram of classification system used for sheep vocals classification into happy and unhappy sounds, followed by explanation of how each block of diagram used in this research in next sections.

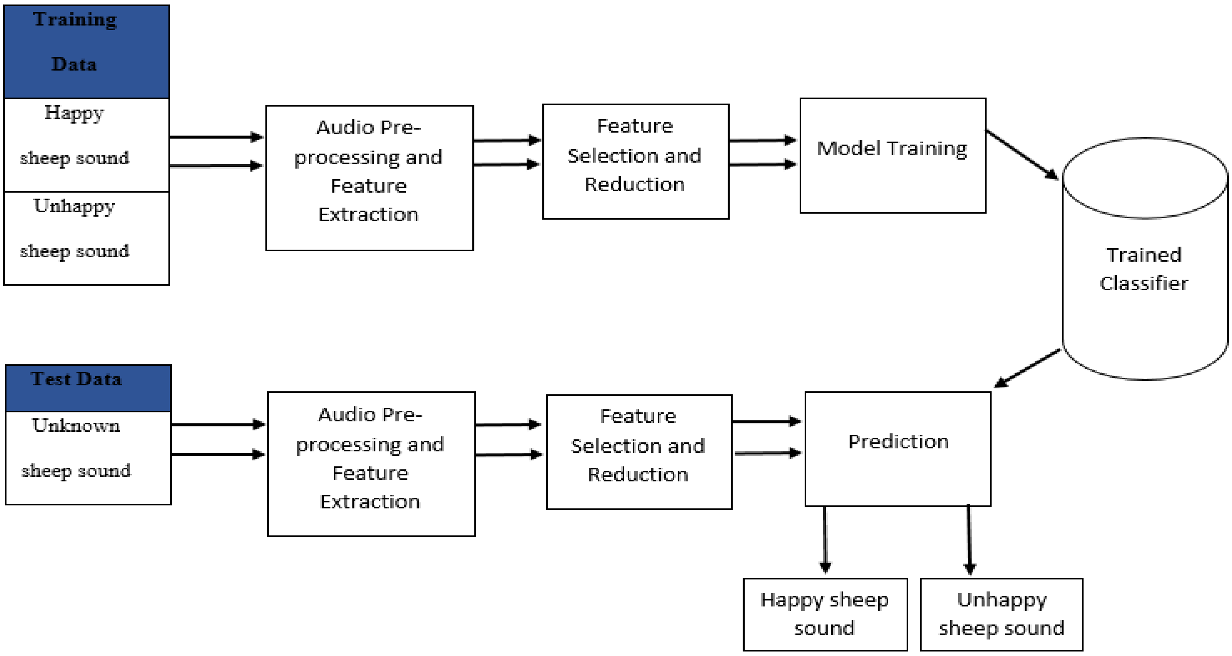

Figure 3.4 Block diagram of the sheep classification system.

3.4 Audio pre-processing

An audio segment of each individual sheep was resampled to 44.10 kHz. Threshold plays a pivotal role in determining the start- and endpoint of an audio segment. Besides duration, other signal parameters (e.g., dominant frequency peak) strongly depend on the onset of a signal. To achieve frequency amplitudes, fast Fourier transform (FFT) was performed on each audio segment. Fourier components, whose absolute value was less than the threshold, were discarded. The remaining components were used to reconstruct the audio segment using the inverse Fourier transform (IFFT) technique.

3.5 Feature Extraction

This is the first audio-based study conducted to classify happy and unhappy sheep in New Zealand (see discussion in 2.9 and 2.10). No audio feature has been reported with 100% accuracy at discriminating different classes. Multiple features (each of them described below) were extracted for higher classification accuracy from each frame of the audio segment and used for happy and unhappy sheep sound classification in this study, these features are given below.

Zero-Crossing Rate

Zero-crossing rate (ZCR) is rate of sign changes along a signal and is the weighted measure of the number of times a discretised time-domain signal represented by x[N] crossed a zero (i.e., changes sign) from positive to zero to negative or from negative to zero to positive, where N represents length of window. For a simple periodic signal, it is roughly related to the frequency of the signal.

Zero crossing rate is defined by equation as

$$ZCR = \frac{1}{N}\sum_{i=1}^{N}|sign(x[i]) - sign(x[i-1])|. \qquad (0.1)$$

$sign$ is 1 for positive argument and -1 for negative argument.

Short-Time Energy

Short-time energy is the sum of the squared time-domain signal $x[n]$, and is defined as

$$STE = \sum_{i=1}^{N}(x[i])^2. \qquad (0.2)$$

This feature indicates how signal amplitude changes over time.

Variance

Variance can be considered as centred signal power and is defined as

$$V_X = \frac{1}{N-1}\sum_{i=1}^{N}(x[i] - \mu_X)^2, \qquad (0.3)$$

where, μ_X is the mean of the signal.

Entropy

The concept of Entropy was introduced by Shannon Claude in 1948 and is defined as the measure of uncertainty in the signal. Given the discrete variable *X*, the Equation used to define entropy of *X* is:

$$H(X) = -\sum_{i=1}^{N} P(xi)\log P(xi), \qquad (0.4)$$

Where Σ denotes the sum and $P(xi)$ is the probability of discrete variable $P(x_1),\ldots\ldots,P(x_n)$.

Spectral Centroid

The spectral centroid indicates the centre of mass of the spectrum. It is related to the brightness of the sound and is defined as

$$SC = \frac{\sum_k f_k |X(k)|}{\sum_k |X(k)|}. \qquad (0.5)$$

f_k and $X(k)$ respectively represent centre frequency and weighted frequency value. Centre frequency is measured between upper and lower cut-off frequencies.

Mel-Frequency Cepstral Coefficients

Mel-frequency cepstral coefficients (MFCC) are calculated by windowing each segment followed by FFT. The frequency scale is then converted to a nonlinear log scale. Finally, the log mel-spectrum is converted back to the time domain using a discrete cosine transform (DCT).

3.6 Feature Selection and Reduction

Features extracted were numerically unbalanced, therefore, to address the numerical unbalances among the training features and to avoid outliers, Z-score normalisation was used, as defined in equation (0.6) below.

$$\hat{F} = \frac{F - \mu}{\sigma} \tag{0.6}$$

where, $\hat{F}, \mu, \text{and}\ \sigma$ respectively indicate normalised features, feature mean and standard deviation.

The total number of features (i.e., dimensions) extracted were 16. The total 16 features include 11 MFCC based features and five other features such as zero-crossing rate, short-time energy, variance, entropy, spectral centroid *(features discussed in section 3.4 above).* The number of features have a direct impact on the complexity of a machine learning model. In addition, many features can computationally be expensive, i.e., require more memory, processing power and can substantially degrade classification performances. Feature selection and reduction techniques are discussed in below sections.

3.6.1 Feature Selection

Most of the supervised feature selection techniques depend on the relevance between the class label and the features. As such techniques require class labels and therefore are referred to as supervised (Saeys, Inza, & Larrañaga, 2007). A Fisher score-based filter (definition in glossary) was found to be the appropriate feature selection method for supervised feature selection (Guyon & Elisseeff, 2003; May, Dandy, & Maier, 2011). Features were then ranked per their Fisher scores and the best feature was selected.

The Fisher score F_s of a feature set $\boldsymbol{X} = [\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n}] \in \mathbb{R}^{d \times n}$ was computed as

$$F_s(\boldsymbol{x}^j) = \frac{\sum_{k=1}^{c} n_k \left(\mu_k^j - \mu^j\right)^2}{v^j}, \qquad (0.7)$$

where $\boldsymbol{x}^j$ denotes the jth row of $\boldsymbol{X}$. μ^j and v^j are the mean and variance of the feature j, μ_k^j is the mean of feature j corresponding to the class k and n_k is the number of sample points belonging to class k (Gu, Li, & Han, 2011).

An LDA-based wrapper was used to sequentially select the ranked features. The area under the curve of the receiver operating characteristic (AUC_{ROC}), being robust under skewed-class distribution (Fawcett, 2006), was used as the objective function. 5-fold cross-validation with the top-ranked feature was performed and the resulting mean AUC_{ROC} was saved. The successive feature was then selected subject to the improvement in the combined mean AUC_{ROC}.

3.6.2 Feature Reduction

In contrast to feature selection, feature reduction or dimensionality reduction techniques do not require class labels and are referred to as unsupervised. In this study, principal component analysis (PCA) – a standard analysis tool for this type of data (Shlens, 2014) – was used to extract the relevant information and consequently, reduce the number of dimensions of the feature set or alternatively minimise the redundancy among different features while preserving structure/pattern in the data (Valletta, Torney, Kings, Thornton, & Madden, 2017).

3.7 Classification

Due to the small size of the data set (number of sound samples), a linear and a nonlinear classifier were used to automatically discriminate happy and unhappy sheep. In this study, LDA was used as a base classifier and support vector machine (SVM) was used to achieve improved accuracies.

3.7.1 Linear Discriminant Analysis

LDA is a widely used classifier and assumes data of all the classes are normally distributed with a common covariance matrix $\sum$. Generally, LDA, due to stable estimates and simple boundaries, works well even for data distributions that deviate from normal. LDA discriminates two classes by a hyperplane. It minimises the inter-class variance and maximises the distance between means of classes. The LDA function for k classes with means μ_k and prior probabilities π_k is defined as (Hastie, R., & J., 2008)

$$F_k(x) = x^T \textstyle\sum^{-1} \mu_k - \frac{1}{2} \mu_k^T \textstyle\sum^{-1} \mu_k + \log(\pi_k). \qquad (0.8)$$

LDA as a base classifier has been used in animal well-being studies (Marais, Le Roux, Wolhuter, & Niesler, 2014).

3.7.2 Support Vector Machine

There can be a multitude of separating hyperplanes for linearly separable data. The linear SVM – a widely used classification algorithm – tries to find an optimal hyperplane that maximally separates the instances of two classes per their distances (margin) as shown in Figure 3.. Margin maximisation on training data leads to better generalisation capability, i.e., better classification performance on test data (Hastie et al., 2008). The nonlinearly separable input data, however, is generally transformed into a high-dimensional feature space using a nonlinear transformation ϕ. The nonlinear SVM classifier was constructed by using radial basis function (RBF) as the kernel, defined in equation (0.9).

$$\mathrm{K}(X,Y) = e^{\left(\frac{-|X-Y|^2}{c}\right)}, \tag{0.9}$$

where c is the kernel width.

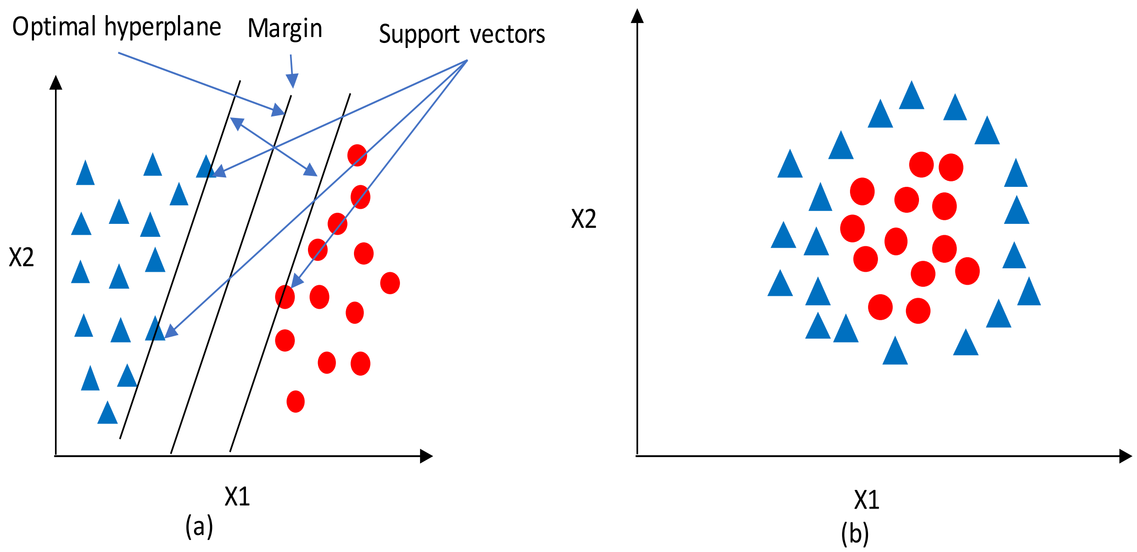

Figure 3.5 Circles and triangles are the instances of two classes (i.e., labels). (a) Linearly separable data. Illustration of margin and support vectors (i.e., instances on the boundary of the slab). (b) Nonlinearly separable classes.

3.7.3 Software and relevant libraries used for data classification

The implementation of the project is done in python programming language because of the different python libraries available for feature scaling and model training testing. The python libraries used in this research project are glob, Librosa, pandas and sklearn. For feature scaling and model training, various functions from scikit learn library and matplotlib were used for analysing and summarising the results.

3.8 Performance Evaluation

The overall performance metrics were the mean value of 10-fold cross-validation. The audio segment of each sheep was independent of another sheep. The 10-fold cross-validation helped to attain a generalised audio-based sheep classification system.

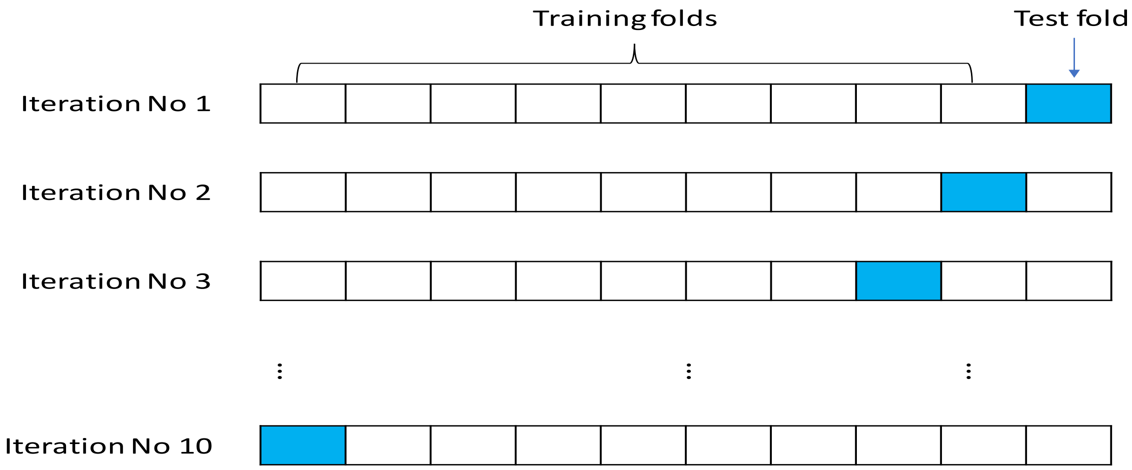

Figure 3.6 Illustration of k-fold (10-fold) cross-validation.

The coloured square in each iteration represents the test data. Various performance metrics, for binary classification, could be formulated from elements of a confusion matrix, shown in Table 3.1. Sensitivity (sn) $\left(\frac{TP}{TP+FN}\right)$, specificity (sp) $\left(\frac{TN}{TN+FP}\right)$, and prediction (pr) $\left(\frac{TP}{TP+FP}\right)$ are three basic measures of performances.

	Predicted class	
Actual positive	True positive (TP)	False negative (FN)
Actual negative	False positive (FP)	True negative (TN)

Table 3.1 Confusion (contingency) matrix of a binary-class problem.

3.9 Chapter summary

This chapter described the data collection process and data labelling, the relevant audio-based feature extraction issues and feature selection and reduction techniques. Classification models used for data classification purpose in this research study were then discussed in section 3.7. In the last section overall performance evaluation matrix and validation for sheep vocals classification system was discussed.

Chapter 4

Results analysis and discussion

This chapter describes the effectiveness of this research followed by technical testing and evaluation of the approach presented in the previous chapter. Previous chapters have discussed the details of the research objectives, audio data analysis and feature testing as well as modelling details within the scope of the study. Details of the implementation of technology by audio classification are described in this section by comparing the models used for this purpose.

4.1 Modelling of data

In the study, different classification models were based on a spectrogram produced from sheep's noise and comparisons were made between models. The features extracted from audio files were converted into a dataset which was later used for training the model.

In the first model, Linear Discriminant Analysis was used on the training and testing data followed Support vector machine (SVM) classification algorithm. After features selection and reduction step, data is split into the training and testing sets for classification. Data is split using built-in scikit learn library functions, which separates data randomly into training and testing sets. Nine (90%) of the ten parts were used for feature selection/reduction and training the classifier, remaining one part (not involved in the training) is used for testing the overall accuracy at classifying (predicting) happy and unhappy sheep. The features that were extracted from the audio files were supplied to these two classical machine learning models.

4.2 Results and Discussion

The majority of happy sheep sound samples (i.e. 35 out of total 49 happy sound) samples were collected from the farm and remaining 14 from online databases, whereas 45 unhappy sound samples were captured from online and only 6 unhappy sound samples from the farm, which indicates that overall sheep on farm are happier than sheep sounds collected through online databases. The difficulty in cleaning the data was significant task as data collected was having background noises which were not required and needs to be cleaned from these background noises to have good and cleaned sheep sounds samples, this took a lot of time and effort and giving an overall sample of 100 samples used for machine learning classification purposes.

The results obtained from training the model using different classifiers are shown in this section. Also, the steps needed to remove the embed features in an audio file (.wav) were also provided. To build a Sheep audio segmentation model with good accuracy a modern solution using machine learning methods is needed. The model was tested on 100 sheep vocals that belonged to two different classes (i.e., happy and unhappy) of sounds. The framework proposed in this study focused on using a machine learning model to classify the happy and unhappy vocals of sheep and classified sheep vocals into happy and unhappy categories and using linear discriminant analysis and support vector machine models for this purpose and chooses the one with best accuracy. Below are the results of each classifier model i.e., LDA and SVM used for classifying sheep vocals in into happy and unhappy classes.

Classifier	**Accuracy**	**Log Loss**
Support Vector Machine	87.50%	0.52
Linear Discriminant Analysis	81.25%	0.39

Table 4.1 Accuracy of classification test models

	Predicted class (On Test Data)	
Actual Happy = 6	True happy (TH) = 5	False unhappy (FUH) = 1
Actual Unhappy = 10	False happy (FH) = 1	True unhappy (TUN) = 9

Table 4.2 Confusion Matrix Table of SVM Model

	Predicted class (On Test Data)	
Actual Happy = 6	True happy (TH) = 5	False unhappy (FUH) = 1
Actual Unhappy = 10	False happy (FH) = 2	True unhappy (TUN) = 8

Table 4.3 Confusion Matrix Table of LDA Model

Performance Measure	Support Vector Machine (SVM)	Linear Discriminant Analysis (LDA)
Sensitivity (sn) $\left(\frac{TH}{TH+FUH}\right)$	83.33%	83.33%
Specificity (sp) $\left(\frac{TUN}{TUN+FH}\right)$	90%	80%
Prediction (pr) $\left(\frac{TH}{TH+FH}\right)$	83.33%	71.42%

Table 4.4 Three Basic Measures of Performance

As can be seen from the results shown in (Table 4.1) that proposed algorithm in this research study classified sheep sounds into happy and unhappy categories with 87.50% accuracy. The sheep sounds data was extracted out from the background noises and cleaned using filter approach using Matlab software, where each audio segment of each individual sheep was resampled to 44.10 kHz and each Fourier components, whose absolute value was less than the threshold were discarded and remaining components were used to reconstruct the audio segment using the inverse Fourier transform (IFFT). The results of this research are difficult to compare with others research in the field, as there appear too few examples of automated animal vocalisation detection studies (Bishop et al., 2017; Y Chung et al., 2013; Yeo, Al-Haddad, & Ng, 2012) but no example of automated sheep vocalisation based classification study found to know happy and unhappy state of sheep for well-being purpose. Therefore, it is difficult to make comparison of obtained results with previous research studies.

In a previous sheep research study (Bishop et al., 2017, 2019) sheep behaviour was not directly observed and data was recorded through a statically placed song meter (sm3) recording unit and collected data was manually segmented and scanned to extract vocals data from the background noise. This may lead to incorrect labelling and researcher's selection bias. In contrast, in my research automated segmentation of sounds is done and direct sheep behaviour was observed while collecting audio sounds from the farm, which helps in more accurately classifying and labelling data. Although sheep sounds collected from online databases were not observed while being recorded they are still auto segmented to minimise the chances of selection bias by missing any specific vocalisation part.

4.3 Chapter summary

The linear discriminant analysis classifier has achieved a classification accuracy of 81.25% regarding sheep sound classification into binary classes of happy and unhappy classes. However, the support vector machine model has showed high results as compared to LDA classifier with an accuracy of 87.50% for classifying sheep sounds into happy and unhappy sounds classes as shown in Table. 4.1. Therefore, SVM-based comprehensive learning model does better classification job than LDA and provides best predictive results.

Chapter 5
Conclusion

This chapter describes and discusses the findings of this research study by drawing together the contents of this research thesis. This research aimed to develop a sound analysis method/process to distinguish the difference between happy sheep sounds and unhappy sheep sounds from their vocals using machine learning algorithm. In this chapter, a summary of the results findings, research contributions will be discussed. Also, the research limitations and future research suggestions will be offered.

5.1 Summary of research findings

Sheep vocalisation data was drawn from field data collected specifically for this research and online recordings. This was achieved by initially categorising the data using a sheep expert and subsequently characterising using extracted acoustic features into the two main categories of happy and unhappy sheep vocals.

The collected data confirmed that the initial assumption that the sheep on the farm were generally assessed as happier than the sheep whose sounds were collected from the online database. This is perhaps explained by the personal care of the sheep on the farm. We do not know the source of the sheep or the circumstances that they were recorded in online and so we do not know if they were being well-cared for or being recorded during a time of hardship, so the result is not unexpected. This does however, provide confidence in the distinction between the happy and unhappy sound categorisation.

Two models (an LDA and an SVM) were then built using suitable machine learning approach for classifying the sheep vocals. The model was tested and found to deliver higher classification accuracy results with SVM model having higher classification results as compared to LDA classification model.

5.2 Implications

The classification model predictability accuracy for classifying sheep vocals into happy and unhappy classes indicates that automated sheep monitoring of well-being based on their vocalisations is possible. This may assist sheep farmers objectively record and demonstrate, for instance through certification systems, the well-being of their farmed sheep. This research therefore contributes to the scientific body of knowledge on applied bioacoustics and provides a useful baseline and direction for future researchers to work on sheep well-being based on sheep vocals.

There is a lack of literature and research available on sheep well-being studies based on their vocals. This is one of the first research studies for studying New Zealand sheep well-being based on their vocals, which contributes to the body of knowledge and practically with laying out a process of studying sheep vocals for knowing their well-being state and using machine learning models for classifying the sheep vocals.

5.3 Research limitations and future research work avenues

As with any research, this research has limitations. The small data sample reflects both the practical difficulties in collecting sheep sounds given financial and time resource limitations. Sheep vocals are not frequently produced because of the natural behaviour of sheep and that leads to data scarcity and indicates that a longer period and greater resources are needed to collect different sheep sounds. Because of this data scarcity and resource issues for collecting varying types of sheep vocals, this research categorised sheep vocals into two major categories i.e., happy and unhappy sounds. A wider range of categories might be able to be demonstrated with a larger sample size. The initial categorisation was based on one expert and future researchers may wish to use multiple experts to ensure the accuracy of the sounds for indicating sheep well-being.

Future researchers may also find it useful in studying sheep well-being based on their vocals by collecting a large sounds data sample with more categories over a longer period. Researchers can record sounds while lambs and mothers are bonding, while in stressed settings (e.g., in shearing shed) and other different conditions. Future researchers may collect sheep sounds from other farms and locations across New Zealand and this may change the classification performance of an algorithm but this is yet to be ascertained. Also, in this research sheep are not numbered to keep track which sheep made happy and which made unhappy sound, so in in future research sheep flock can be tracked using some collar mounted recording device to monitor the particular sheep making specific sounds. This research assumes that sheep sounds collected are not breed-specific and future researchers can study sheep well-being based on different breeds to compare and know more about specific breeds. Future researchers can also take a guide from this research study to do research on the well-being of other New Zealand animals based on their vocals.

Despite its limitations, this research has demonstrated that monitoring and analysis of sheep vocalisations using machine learning technology shows considerable promise as a means of certifying sheep well-being.

References

Aarts, E., & Korst, J. (1989). *Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing*: John Wiley & Sons, Inc.

Abbasi, S., Derakhshanfar, R., Abbasi, A., & Sarbaz, Y. (2013). *Classification of normal and abnormal lung sounds using neural network and support vector machines.* Paper presented at the 2013 21st Iranian Conference on Electrical Engineering (ICEE).

Amrulloh, Y. A., Abeyratne, U. R., Swarnkar, V., Triasih, R., & Setyati, A. (2015). Automatic cough segmentation from non-contact sound recordings in pediatric wards. *Biomedical Signal Processing and Control, 21*, 126-136.

Archer, J. (1988b). *The behavioural biology of aggression* (Vol. 1): Cambridge University Press.

Baotic, A., Sicks, F., & Stoeger, A. S. (2015). Nocturnal "humming" vocalizations: adding a piece to the puzzle of giraffe vocal communication. *BMC research notes, 8*(1), 425.

Bardeli, R., Wolff, D., Kurth, F., Koch, M., Tauchert, K.-H., & Frommolt, K.-H. (2010). Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. *Pattern Recognition Letters, 31*(12), 1524-1534.

Beef and Lamb. (2019). B+LNZ Annual Report 2019. Retrieved from chrome-extension://oemmndcbldboiebfnladdacbdfmadadm/https://beeflambnz.com/sites/default/files/BLNZ-AR-2019.pdf.

Bishop, J. C., Falzon, G., Trotter, M., Kwan, P., & Meek, P. D. (2017). *Sound analysis and detection, and the potential for precision livestock farming - a sheep vocalization case study.* Paper presented at the 1st Asian-Australasian conference on precision pastures and livestock farming.

Bishop, J. C., Falzon, G., Trotter, M., Kwan, P., & Meek, P. D. (2019). Livestock vocalisation classification in farm soundscapes. *Computers and electronics in agriculture, 162*, 531-542. Retrieved from http://www.sciencedirect.com/science/article/pii/S0168169918312870.

Blood, D. C., & Studdert, V. P. (1988). *Baillière's comprehensive veterinary dictionary*: Baillière Tindall.

Bouril, A., Aleinikava, D., Guillem, M. S., & Mirsky, G. M. (2016). *Automated classification of normal and abnormal heart sounds using support vector machines.* Paper presented at the 2016 Computing in Cardiology Conference (CinC).

Brambell, F. (1965). Report on the technical committee to enquire into the welfare of livestock kept under intensive conditions. In: HMSO, London.

Broom, D. M. (1988). The scientific assessment of animal welfare. *Applied Animal Behaviour Science, 20*(1-2), 5-19.

Broom, D. M. (1991). Animal welfare: concepts and measurement. *Journal of animal science, 69*(10), 4167-4175.

Broom, D. M. (1998). Welfare, stress, and the evolution of feelings. *Advances in the Study of Behavior, 27*, 371-403.

Broom, D. M., & Fraser, A. F. (2015). *Domestic animal behaviour and welfare*: Cabi.

Choi, S., & Jiang, Z. (2010). Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique. *Computers in biology and medicine, 40*(1), 8-20.

Chung, Y., Lee, J., Oh, S., Park, D., Chang, H., & Kim, S. (2013). Automatic detection of cow's oestrus in audio surveillance system. *Asian-Australasian journal of animal sciences, 26*(7), 1030.
Chung, Y., Oh, S., Lee, J., Park, D., Chang, H.-H., & Kim, S. (2013). Automatic detection and recognition of pig wasting diseases using sound data in audio surveillance systems. *Sensors, 13*(10), 12929-12942.
Clemins, P. J., Johnson, M. T., Leong, K. M., & Savage, A. (2005). Automatic classification and speaker identification of African elephant (Loxodonta africana) vocalizations. *The Journal of the Acoustical Society of America, 117*(2), 956-963.
Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of computation, 19*(90), 297-301.
Cordeiro, A. F. d. S., Nääs, I. d. A., Baracho, M. d. S., Jacob, F. G., & Moura, D. J. d. (2018). The use of vocalization signals to estimate the level of pain in piglets. *Engenharia Agrícola, 38*(4), 486-490.
Cresswell, E. (1960). Ranging behaviour studies with Romney Marsh and Cheviot sheep in New Zealand. *Animal Behaviour, 8*(1-2), 32-38.
Dawkins, M. S. (1988). Behavioural deprivation: a central problem in animal welfare. *Applied Animal Behaviour Science, 20*(3-4), 209-225.
Dolan, P., & Metcalfe, R. (2011). Measuring subjective wellbeing for public policy: Recommendations on measures.
Dwyer, C. (2004). How has the risk of predation shaped the behavioural responses of sheep to fear and distress? *Animal Welfare, 13*(3), 269-281.
Dwyer, C., McLean, K., Deans, L., Chirnside, J., Calvert, S., & Lawrence, A. (1998). Vocalisations between mother and young in sheep: effects of breed and maternal experience. *Applied Animal Behaviour Science, 58*(1), 105-119.
Epizooties, O. I. d. (2020). The World Organisation for Animal Health (OIE).
Farnworth, M. J., Campbell, J., & Adams, N. J. (2011). What's in a name? Perceptions of stray and feral cat welfare and control in Aotearoa, New Zealand. *Journal of Applied Animal Welfare Science, 14*(1), 59-74.
Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861-874. Retrieved from http://www.sciencedirect.com/science/article/pii/S016786550500303X.
Fidler, D. P. (2004). Global outbreak of avian influenza A (H5N1) and international law. *ASIL Insights (January 2004) The American Society of International Law, Washington*.
Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(1), 3-71.
Fogel, D. B. (2006). *Evolutionary computation: toward a new philosophy of machine intelligence* (Vol. 1): John Wiley & Sons.
Forkman, B., Boissy, A., Meunier-Salaün, M.-C., Canali, E., & Jones, R. (2007). A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiology & Behavior, 92*(3), 340-374.
Fraiwan, L., Hassanin, O., Fraiwan, M., Khassawneh, B., Ibnian, A. M., & Alkhodari, M. (2020). Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers. *Biocybernetics and Biomedical Engineering*.

Fraser, A. (1988). Behavioural needs in relation to livestock maintenance. *Applied Animal Behaviour Science, 19*(3-4), 368-376.
Gaber, T., Tharwat, A., Ibrahim, A., & Hassanien, A. E. (2017). Linear discriminant analysis: A detailed tutorial. *AI communications, 30*(2), 169-190.
Gendreau, P. L., & Archer, J. (2005). Subtypes of aggression in humans and animals. *Developmental origins of aggression*, 25-46.
Government of the United Kingdom Animal welfare Act 2006, A. W. A. Animal Welfare Act 2006. Retrieved from Animal welfare act 2006 UK
Grandin, T. (2005). Maintenance of good animal welfare standards in beef slaughter plants by use of auditing programs. *Journal of the American Veterinary Medical Association, 226*(3), 370-373.
Graves, A., Mohamed, A.-r., & Hinton, G. (2013). *Speech recognition with deep recurrent neural networks.* Paper presented at the Acoustics, speech and signal processing (icassp), 2013 ieee international conference on.
Gregorini, P., Villalba, J. J., Chilibroste, P., & Provenza, F. D. (2017). Grazing management: setting the table, designing the menu and influencing the diner. *Animal Production Science, 57*(7), 1248-1268.
Gregory, N. G., & Grandin, T. (1998). *Animal welfare and meat science*: CABI Pub.
Gu, Q., Li, Z., & Han, J. (2011). *Generalized fisher score for feature selection.* Paper presented at the 27th Conference on uncertainty in artificial intelligence, Barcelona, Spain.
Gu, Q., Li, Z., & Han, J. (2012). Generalized fisher score for feature selection. *arXiv preprint arXiv:1202.3725*.
Guadarrama-Maillot, V., & Waas, J. R. (2008). New Zealand trends in animal behaviour research. *New Zealand Journal of Zoology, 35*(4), 305-321.
Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research, 3*(Mar), 1157-1182.
Harrison, R. (1964). Animal machines: the new factory farming industry. *Animal machines: the new factory farming industry.*
Hastie, T., R., T., & J., F. (2008). *The Elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York, USA: Springer.
Hewson, C. J. (2003). What is animal welfare? Common definitions and their practical consequences. *The Canadian Veterinary Journal, 44*(6), 496.
Hirsch, H.-G., & Pearce, D. (2000). *The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions.* Paper presented at the ASR2000-Automatic Speech Recognition: Challenges for the new Millenium ISCA Tutorial and Research Workshop (ITRW).
Hopfield, J. J. (1987). Neural networks and physical systems with emergent collective computational abilities. In *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* (pp. 411-415): World Scientific.
Hopp, S. L., Owren, M. J., & Evans, C. S. (2012). *Animal acoustic communication: Sound analysis and research methods*: Springer Science & Business Media.
Ikeda, Y., & Ishii, Y. (2008). Recognition of two psychological conditions of a single cow by her voice. *Computers and electronics in agriculture, 62*(1), 67-72.
Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 23*(1), 67-72.

Itakura, F. (1990). Minimum prediction residual principle applied to speech recognition. In *Readings in speech recognition* (pp. 154-158): Elsevier.
Jahns, G. (2008). Call recognition to identify cow conditions—A call-recogniser translating calls to text. *Computers and electronics in agriculture, 62*(1), 54-58. Retrieved from http://www.sciencedirect.com/science/article/pii/S0168169907001937.
Juang, B.-H., & Rabiner, L. R. (2005). Automatic speech recognition–a brief history of the technology development. *Georgia Institute of Technology. Atlanta Rutgers University and the University of California. Santa Barbara, 1*, 67.
Kiley, M. (1972). The vocalizations of ungulates, their causation and function. *Zeitschrift für Tierpsychologie, 31*(2), 171-222.
Ko, K., Park, S., & Ko, H. (2018). *Convolutional feature vectors and support vector machine for animal sound classification.* Paper presented at the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
Laidre, M. E. (2012). Principles of animal communication. In: Academic Press.
Lee, C.-H., Chou, C.-H., Han, C.-C., & Huang, R.-Z. (2006). Automatic recognition of animal vocalizations using averaged mfcc and linear discriminant analysis. *Pattern Recognition Letters, 27*(2), 93-101.
Manteuffel, G., Puppe, B., & Schön, P. C. (2004). Vocalization of farm animals as a measure of welfare. *Applied Animal Behaviour Science, 88*(1-2), 163-182.
Marai, I., El-Darawany, A., Fadiel, A., & Abdel-Hafez, M. (2007). Physiological traits as affected by heat stress in sheep—a review. *Small Ruminant Research, 71*(1), 1-12.
Marais, J., Le Roux, S. P., Wolhuter, R., & Niesler, T. (2014). *Automatic classification of sheep behaviour using 3-axis accelerometer data.* Paper presented at the 25th Annual Symposium of the Pattern Recognition Association of South Africa (PRASA).
May, R., Dandy, G., & Maier, H. (2011). Review of input variable selection methods for artificial neural networks. In *Artificial Neural Networks-Methodological Advances and Biomedical Applications* (Vol. 10): InTech.
McClelland, J. L., & Elman, J. L. (1986). The trace model of speech perception. *Cognitive psychology, 18*(1), 1-86.
McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics, 5*(4), 115-133.
McCune, L., Vihman, M. M., Roug-Hellichius, L., Delery, D. B., & Gogate, L. (1996). Grunt communication in human infants (Homo sapiens). *Journal of Comparative Psychology, 110*(1), 27.
McGregor, P. K. (2005). *Animal communication networks*: Cambridge University Press.
Mellor, D., & Stafford, K. (2001). Integrating practical, regulatory and ethical strategies for enhancing farm animal welfare. *Australian veterinary journal, 79*(11), 762-768.
Moberg, G. P. (1985). *Animal stress*: American Physiological Society.
Montossi, F., Font-i-Furnols, M., Del Campo, M., San Julián, R., Brito, G., & Sañudo, C. (2013). Sustainable sheep production and consumer preference trends: Compatibilities, contradictions, and unresolved dilemmas. *Meat science, 95*(4), 772-789.
Morton, D. a., & Griffiths, P. (1985). Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Vet Rec, 116*(16), 431-436.

Noll, A. M. (1964). Short-time spectrum and “cepstrum” techniques for vocal-pitch detection. *The Journal of the Acoustical Society of America, 36*(2), 296-302.
Ovaskainen, O., Moliterno de Camargo, U., & Somervuo, P. (2018). Animal sound identifier (ASI): software for automated identification of vocal animals. *Ecology Letters, 21*(8), 1244-1254. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.13092. doi:10.1111/ele.13092
Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. (2011). Human voice recognition depends on language ability. *Science, 333*(6042), 595-595.
Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In *Machine Learning* (pp. 101-121): Elsevier.
Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., . . . Schwarz, P. (2011). *The Kaldi speech recognition toolkit.* Paper presented at the IEEE 2011 workshop on automatic speech recognition and understanding.
Ranjard, L., Reed, B. S., Landers, T. J., Rayner, M. J., Friesen, M. R., Sagar, R. L., & Dunphy, B. J. (2017). MatlabHTK: a simple interface for bioacoustic analyses using hidden Markov models. *Methods in Ecology and Evolution, 8*(5), 615-621.
Rosenblatt, R., & Gaponoff, M. (1984). The microcomputer as a vehicle for continuing medical education. *Journal of Family Practice, 18*(4), 629-632.
Rushen, J. (1986). Aversion of sheep to electro-immobilization and physical restraint. *Applied Animal Behaviour Science, 15*(4), 315-324.
Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics, 23*(19), 2507-2517. Retrieved from https://doi.org/10.1093/bioinformatics/btm344. doi:10.1093/bioinformatics/btm344
Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 26*(1), 43-49.
Sakoe, H., Chiba, S., Waibel, A., & Lee, K. (1990). Dynamic programming algorithm optimization for spoken word recognition. *Readings in speech recognition, 159*, 224.
Satapathy, S. K., Dehuri, S., Jagadev, A. K., & Mishra, S. (2019). *EEG Brain Signal Classification for Epileptic Seizure Disorder Detection*: Academic Press.
Schrader, L., & Hammerschmidt, K. (1997). Computer-aided analysis of acoustic parameters in animal vocalisation: A multi-parametric approach. *Bioacoustics, 7*(4), 247-265. Retrieved from https://doi.org/10.1080/09524622.1997.9753338.
Scott, G., & MacAngus, G. (2004). The ability of laying hens to negotiate perches of different materials with clean or dirty surfaces. *Animal Welfare, 13*(3), 361-365.
Searby, A., & Jouventin, P. (2003). Mother-lamb acoustic recognition in sheep: a frequency coding. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*(1526), 1765-1771.
Sebe, F., Aubin, T., Boue, A., & Poindron, P. (2008). Mother–young vocal communication and acoustic recognition promote preferential nursing in sheep. *Journal of Experimental Biology, 211*(22), 3554-3562.
Shlens, J. (2014). A tutorial on principal component analysis. *Computing Research Repository, abs/1404.1100*.

Smith, B. J., Bruner, K. E., & Kendall, L. V. (2019). Female-and Intruder-induced Ultrasonic Vocalizations in C57BL/6J Mice as Proxy Indicators for Animal Wellbeing. *Comparative Medicine, 69*(5), 374-383.

Souli, S., & Lachiri, Z. (2018). Audio sounds classification using scattering features and support vectors machines for medical surveillance. *Applied Acoustics, 130*, 270-282. Retrieved from http://www.sciencedirect.com/science/article/pii/S0003682X17304280.

Stafford, K. (2017). *Livestock production in New Zealand: the complete guide to dairy cattle, beef cattle, sheep, deer, goats, pigs and poultry*: Massey University Press.

Terlouw, E., Arnould, C., Auperin, B., Berri, C., Le Bihan-Duval, E., Deiss, V., . . . Mounier, L. (2008). Pre-slaughter conditions, animal stress and welfare: current status and possible future research. *Animal: an international journal of animal bioscience, 2*(10), 1501.

Thorslund, C. A., Sandøe, P., Aaslyng, M. D., & Lassen, J. (2016). A good taste in the meat, a good taste in the mouth–Animal welfare as an aspect of pork quality in three European countries. *Livestock Science, 193*, 58-65.

Travieso, C. M., Yadav, A., Singh, A., & Dutta, M. K. (2019). Machine learning-based classification of cardiac diseases from PCG recorded heart sounds. *Neural Computing and Applications*, 1-14.

United States Government Animal Welfare Act 1966, A. W. A. Animal Welfare Act. Retrieved from https://www.aphis.usda.gov/aphis/ourfocus/animalwelfare/SA_AWA.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological review, 108*(3), 550.

Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour, 124*, 203-220. Retrieved from http://www.sciencedirect.com/science/article/pii/S0003347216303360.

Van Compernolle, D. (2001). Recognizing speech of goats, wolves, sheep and… non-natives. *Speech Communication, 35*(1-2), 71-79.

Verstegen, M., & Close, W. (1994). The environment of the growing pig.

Villalba, J. J., Provenza, F. D., & Shaw, R. (2006). Sheep self-medicate when challenged with illness-inducing foods. *Animal Behaviour, 71*(5), 1131-1139.

Vintsyuk, T. (1971). Element-wise recognition of continuous speech composed of words from a specified dictionary. *Cybernetics and Systems Analysis, 7*(2), 361-372.

Watts, J. M., & Stookey, J. M. (2000). Vocal behaviour in cattle: the animal's commentary on its biological processes and welfare. *Applied Animal Behaviour Science, 67*(1-2), 15-33.

Weary, D. M., & Chua, B. (2000). Effects of early separation on the dairy cow and calf: 1. Separation at 6 h, 1 day and 4 days after birth. *Applied Animal Behaviour Science, 69*(3), 177-188.

Weary, D. M., & Fraser, D. (1995). Calling by domestic piglets: reliable signals of need? *Animal Behaviour, 50*(4), 1047-1055.

Welch, M. R., Leege, D. C., Wald, K. D., & Kellstedt, L. A. (1993). Are the sheep hearing the shepherds? Cue perceptions, congregational responses, and political communication processes. *Rediscovering the religious factor in American politics*, 235-254.

Weninger, F., & Schuller, B. (2011). *Audio recognition in the wild: Static and dynamic classification on a real-world database of animal vocalizations.* Paper presented at the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

White, S. C. (2008). *But what is well-being? A framework for analysis in social and development policy and practice.* Paper presented at the Conference on regeneration and wellbeing: research into practice, University of Bradford.

Yang, Y., & Pedersen, J. O. (1997). *A comparative study on feature selection in text categorization.* Paper presented at the Icml.

Yeo, C. Y., Al-Haddad, S., & Ng, C. K. (2012). *Dog voice identification (ID) for detection system.* Paper presented at the 2012 Second International Conference on Digital Information Processing and Communications (ICDIPC).