Identification of myocardial infarction using consumer smartwatch ECG measurement

Master of Science (Tech) Thesis
University of Turku
Department of Computing
Software Engineering
2023
Samuel Kivi

UNIVERSITY OF TURKU
Department of Computing

Samuel Kivi: Identification of myocardial infarction using consumer smartwatch ECG measurement

Master of Science (Tech) Thesis, 62 p.
Software Engineering
May 2023

The goal of this thesis is to detect and classify acute myocardial infarctions from smartwatch ECG data. As the smartwatches have been increasing in numbers, and many of new smartwatch models have capability to detect ECG data. This study aims to answer to the question whether or not the ECG data from smartwatches can be used to detect acute myocardial infarctions.
To answer to this question, and existing database has been used in tandem with smartwatch ECG data gathered from two different smartwatches. Five different machine learning models have been used to detect and classify ECG data. The best performing machine learning model was Extra Trees, which achieved accuracy of 90.84% with using Leave-One-Out Cross-Validation.
These results show that ECG data from smartwatches could be used to detect infarctions. Measuring ECG with smartwatch is much easier than using clinical ECG measurement devices, meaning that ECG measuring could reach much wider audience that it has prior to this been able to reach.
Further research could include gathering larger database from smartwatch ECG, and the data ownership of smartwatch, and other medical and biological data that companies collect.

Keywords: electrocardiogram, machine learning, acute myocardial infarction, infarction, smartwatch, smart wearables

Table of contents

1 Introduction **1**

1.1 Research motivation . . . 1

1.2 Research questions . . . 4

1.3 Research method . . . 4

2 Background **6**

2.1 Electrocardiogram . . . 6

2.2 Acute myocardial infarction . . . 8

2.3 Machine learning . . . 9

2.4 Consumer smartwatch . . . 11

3 State of the art **13**

3.1 Performance metrics . . . 13

3.2 Detecting myocardial infarction with artificial intelligence . . . 15

3.2.1 K-nearest neighbors . . . 15

3.2.2 Support vector machine . . . 18

3.2.3 Decision trees . . . 21

3.2.4 Random forest . . . 22

3.2.5 Extra trees . . . 24

3.3 Summary of different machine learning methods . . . 25

3.4 Leave-One-Out Cross-Validation . . . 27

4 Data from smartwatches **29**
4.1 Devices . . . 29
4.1.1 Samsung Galaxy Watch 4 . . . 29
4.1.2 Apple Watch Series 6 . . . 31
4.2 Data gathering . . . 32
4.3 Data preprocessing . . . 33
4.3.1 Normalisation . . . 36
4.4 Peak detection . . . 37
4.4.1 R-peak detection . . . 37
4.4.2 P-, Q-, S- and T-wave detection . . . 37
4.5 Feature Extraction . . . 40

5 STAFF III –database **43**
5.1 Database . . . 43

6 Results **46**
6.1 Results of the STAFF database . . . 46
6.2 Results of the smartwatch data . . . 50
6.3 Results of the STAFF database and smartwatch data . . . 54

7 Conclusion **60**
7.1 Further research . . . 61

References **63**

List of acronyms

Bagging Bootstrap aggregating

DT Decision trees

ECG Electrocardiography

ET Extremely randomized trees

HF-ECG High-frequency electrocardiography

HRV Heart Rate Variability

KNN K-nearest neighbors

LOOCV Leave-One-Out Cross-Validation

ML Machine Learning

PDF Portable Document Format

PNG Portable Network Graphics

RF Random forest

S-ECG Serial electrocardiography

SVM Support vector machine

1 Introduction

In this chapter we will look at the motivation of reading this thesis, and how important acute myocardial infarctions are to detect and to treat. We will take a look in the history of studying cardiovascular diseases, and how detection methods of hearts and heart signals irregularities have been improved in the last century. We will also look at the research questions that this thesis has, and general structure of the thesis.

1.1 Research motivation

In the beginning of the 20th century the cardiovascular diseases were the fourth highest cause of deaths in the United States, only behind pneumonia, tuberculosis, and diarrhea. Only ten years later cardiovascular diseases had risen to the top of the causes of death. By the half of the century cardiovascular diseases were the reason for the half of the deaths, not only in the United States but also in the rest of the industrialised world [1].

As the mortality rate from cardiovascular diseases rose, so did the deaths from acute myocardial infarctions. By the mid 20th century, acute myocardial infarctions were the leading cause of death in the United States. Luckily with the major developments in the care of the acute myocardial infarctions, death from the infarctions have been on the descent [1].

Myocardial infarction is defined as the cell death due to ischemia. This means that the cells in the heart are dying due to lack of blood flow to them [2]. This means that myocardial infarctions can cause major damage to the heart muscle and to the blood vessels supplying oxygen and nutrients to the heart.

Even as the deaths from acute myocardial infarctions have been on the downturn and many methods for detecting and treating infarctions have been invented, these changes have been received perhaps with too much optimism, as despite the inventions and overall progress in prevention, detection, curing, and treating cardiovascular diseases still remain the leading cause for death in developed nations [1].

These diseases of the heart can be detected with multiple ways, and one of the most used methods is an electrocardiogram (ECG). ECG was a series of anatomical and technological in the field of heart diseases and technology. ECG has even been called one of the most valuable tools in medicine. ECG can be used to detect different abnormalities from a heart rate, such as acute myocardial infarctions. The electrocardiogram works by showing the electrical signals produced by the heart. These signals are then transmitted by the body through different tissues in the human body, until they reach the skin of the patient. Then the signal is captured by the leads. Usually there are 12 leads that record the humans' signals, but ECG can be recorded with smaller amount of leads. For this thesis we are interested in the electrocardiograms' ability to detect acute myocardial infarctions, especially with smart wearable devices [3] [4].

Electrocardiogram is not perfect, even if it is extremely powerful tool to diagnose heart problems. The accuracy is not perfection and there is always human variation as human doctor is examining the graph created by ECG. There is also variation between different patients. This variation between patients can have a confusing effect that may lead to wrong diagnosis on the root condition of the heart. This then may lead to decreased accuracy of the ECG diagnosis [4] [5].

There are multiple methods of tackling downsides that the ECG has. One is serial ECG (S-ECG). With conventional ECG performed by emergency department correct diagnosis for acute myocardial infarction was only between 24 and 60 percent [6] [7].

There is also a high-frequency electrocardiography (HF-ECG) method for detecting hearts irregularities and acute myocardial infarctions, but in this, thesis we will concentrate on the serial electrocardiography method. This HF-ECG has shown ability to detect and diagnose acute cardiac diseases [8].

The increase of the consumer smart wearables especially the increase in the consumer smartwatches and the ability of these smartwatches to measure simple electrocardiographs, gives us the ability to monitor and track the patients ECG with greatly bigger frequency compared to clinical ECG. This increase in the sheer number of smartwatches that have the ability to detect ECG, provides more ways to diagnose and even prevent acute heart problems and diseases. Smartwatches are also more approachable than devices that are specially designed to detect ECG. This kind of medical devices are not usually marketed towards greater population, but among people, who are already buying smartwatches, the ability to monitor ECG comes as a bonus feature [9] [10].

The last decade of 19th century and the first decade of 20th century, were special in the advancement of identifying heart problems, as both chest x-rays and electrocardiograms were invented [11]. For detecting irregularities in hearts rhythm, electrocardiogram was crucial invention. This invention was made by Dutch physiologist Willem Einthoven in 1902. This invention gave physicians a capable tool, when trying to diagnose problems in the patient's heart, such as arrythmias and acute myocardial infarctions [12].

1.2 Research questions

The main research questions for thesis are:

1. Can machine learning be used to detect and classify acute myocardial infarctions?
2. Can wearable smartwatches be used in detecting acute myocardial infarction?

The question about can machine learning used to detect acute myocardial infarctions, will be answered with machine learning methods using electrocardiogram data from STAFF III database [13]. These machine learning models will be then used with data collected from smart watches, and compared how well will the data from smart watches perform in detecting acute myocardial infarctions.

1.3 Research method

The thesis will consist of literacy overview about the usage of machine learning and artificial intelligence in the field of detecting acute myocardial infarctions. The other part of this thesis will tackle the usage of smartwatch ECG.

In the literacy overview chapter of this thesis, we will immerse ourselves into the field of machine learning and using machine learning to detect heart problems, especially acute myocardial infarctions. This literature overview will be done by focusing om research papers and studies done on how well machine learning models can detect and classify heart's rhythm irregularities.

In the analysis of the smart watches chapter, we will measure ECG data with multiple different smart watches and try to use machine learning methods to detect if there are any irregularities or sign of acute myocardial infarctions. These different manufacturer's smart watches are then going to be compared to each other to find if there are any differences in the capability that machine learning models can detect hearts irregularities.

2 Background

In this chapter, we will take a deeper look on the thesis' theme. We will familiarise ourselves with the electrocardiogram, hearts irregularities, mainly acute myocardial infarctions, and at the end of the chapter we will take a closer look in the machine learning methods that we will be using later in the thesis.

2.1 Electrocardiogram

Electrocardiography has been a critical tool for detecting and diagnosing heart problems for some while now. ECG can be used to detect from a heart rate to severe irregularities in heart such as cardiac arrhythmias, myocardial ischemia, that may lead to acute myocardial infarctions, and even changes in QT interval [14]. Heart diseases have been one of the largest reasons of death for over a half century now and show little to no signs of decreasing. As seen in the Figure 2.1 the cardiovascular diseases have had a quite rapid growth in the last century. These critical and often quite acute diseases can be detected with ECG, and such ECG has become a staple tool in the field of cardiovascular science [1]. ECG has an extremely important part in diagnostics when patient is suspected of having acute myocardial infarction. It is usually the first test to be done to the patient in an emergency setting [15].

Classical ECG consists of 12 leads. These leads are electrodes placed on the skin of the patient, which detect the signals transmitted by the patient's heart. These electrodes are placed on specific parts on the patients skin, and are used to

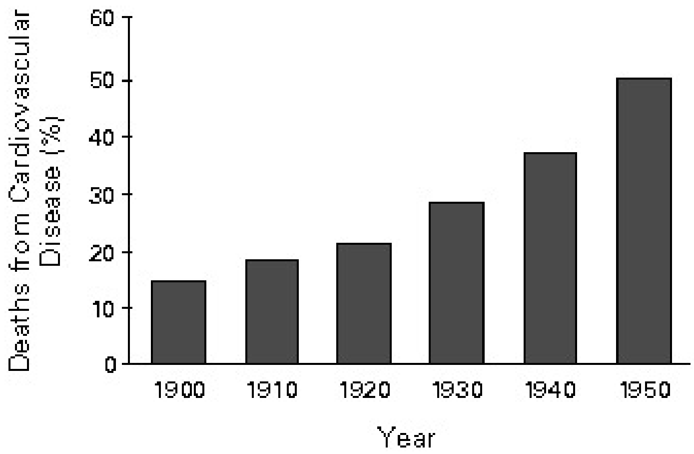

Figure 2.1: Percentage of Deaths from Cardiovascular Disease in the United States from 1900 to 1950 [1]

.

create the ECG graph. The 12 leads that are usually used in a clinical ECG include three bipolar limb leads (leads I, II, III), also known as Einthoven's Triangle and Einthoven's leads [16], six unipolar precordial leads (leads from V_1 to V_6), and three modified unipolar leads (augmented limb leads aV_r, aV_l, aV_f) [4].

Normal heartbeat consists of five different peaks or waves. These peaks are from left to right P-, Q-, R-, S-, And T-peaks. In the Figure 2.2 we can see these different peaks. PQ interval is the time and amplitude between P-wave and Q-peak. QRS complex is the complex that is created by Q-, R-, and S-peaks. ST interval is the section between S-peak and T-wave [17]. Acute myocardial infarctions are usually divided to ST elevation(STEMI), and non-ST elevation(NSTEMI) infarctions by the ST interval [18].

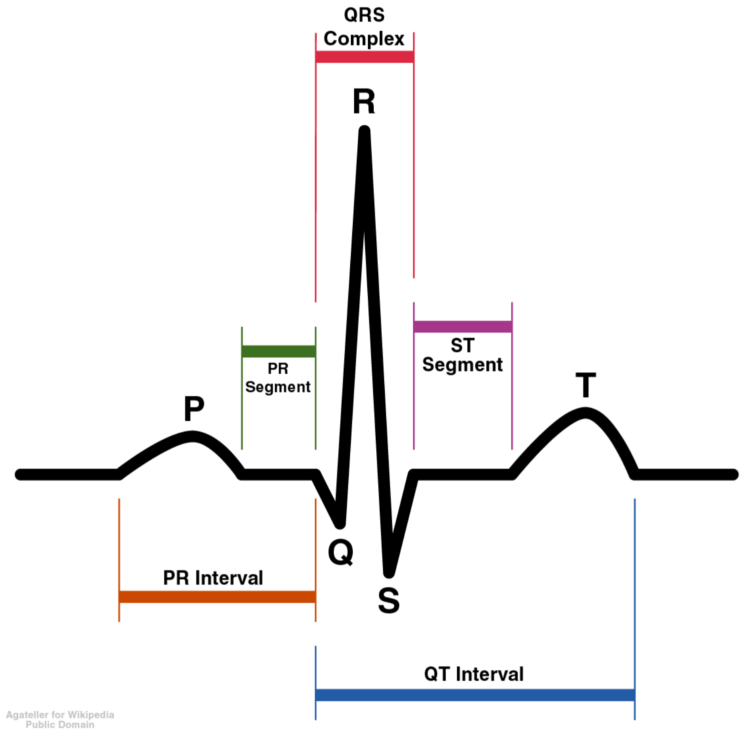

Figure 2.2: ECG from a normal sinus rhythm [19]

.

2.2 Acute myocardial infarction

Myocardial infarction is defined as myocardial cell death due to prolonged ischemia, meaning prolonged blood loss to the specific area of the heart, in this case to the heart. The onset of ischemia is not immediate, but it does not take long to develop. Ischemia can be developed as fast as in 20 minutes from the loss of the correct blood flow. The diagnosis of the ischemia caused myocardial necrosis can take multiple hours, using microscopic examination means. Complete necrosis of the myocardial cells happens from two to four hours after the cut of blood flow, but the onset of necrosis can take longer if the collateral circulation is present to supply blood to the ischemia affected parts [2].

In the beginning stages of myocardial infarction is myocardial ischemia. This ischemia happens from the imbalance of supply and demand of the blood around the heart. This myocardial ischemia can be detected from the patient with ECG and with patient's history. The ischemia can cause different symptoms to the patients, which include pain or discomfort in the upper body, arms, or upper abdomen during both exercise and even rest. Often this pain and discomfort is dispersed and scattered around the body and is not affected by the movement of the painful area. Unluckily these symptoms are not contained only by myocardial infarctions and can be produced by plethora of other conditions. Myocardial infarction can be also occurred without symptoms. If myocardial ischemia is detected by changes in the patients ECG together with myocardial injury, and manifested by rising or falling pattern of cTn (cardiac troponin I and cardiac troponin T [20]) values, an acute myocardial infarction diagnosis can be given. [2].

2.3 Machine learning

Automated learning or more popularly known as machine learning (ML) is "teaching" computers to "learn" from input given to the machines. Learning is this case is the ability to use past experiences to create new knowledge or expertise. Usually in ML, training data given to the algorithm is the "past experiences" and the output that the algorithm creates from input and the training data is the ability to use the past experiences to create domain specific knowledge. For example, nature is full of these ML algorithms. Animals can have these learning experiences that can be regarded as ML algorithms. For example, a rat that eats a new kind of food, after which it develops some symptoms, for example nausea. If the rat comes across the same food again, it will not eat it as it can learn from the past experiences. This example can be used to mimic similar behaviour in simple ML task. Task can be that we want a simple spam filter. The filter can use ML by collecting data and linking that data to some categories. If message is spam, that message can be tagged as spam, and the ML algorithm can try to learn what makes that specific message a spam message. Next time ML algorithm sees similar message it can deduct that the message is likely spam message, and the algorithm can filter it out [21].

This spam message filtering approach can be quite effective, but it has one fatal flaw that learning system should have. It cannot categorize messages that it has not seen before. A good ML algorithm should be able to move from remembering past messages to be able to generalise. The rat in our previous example can use the previously gathered experiences to new unseen food. If new food has same smell, look, and taste as the food that previously gave it nausea, it can induct that the new food can also cause nausea, and the rat will not eat the food. The rat can learn from previous experiences, but it can also use this experience to create new knowledge. We can try to use this example from nature to improve our spam message filtering algorithm. Our algorithm can take into account suspicious words inside the spam messages, and when it comes across these suspicious words in incoming messages, it can induct that the message might be spam [21].

This inductive reasoning raises a problem, which is that it can flag real messages as spam, also known as false positive. In the rat example the rat can skip a perfectly fine piece of food, because it smelled similar to food that previously gave it nausea. Other good example in for this reinforced leaning is Pigeon Superstition where pigeons are in a cage. They are fed in regular intervals and are not stimulated with anything else. When these pigeons are fed, they are doing seemingly randoms activities. But the delivery of food is reinforcing these random activities that pigeons are doing, and they are more likely to do these activities, when they are fed next time, and the reinforcement loop is created. Pigeons are more likely to do specific activities when fed, and they induct that the specific activity means that they are getting food [22]. This pigeon example highlights a problem. How to distinguish a useful learning mechanism from a pigeon superstition? For human this ability to distinguish is trivially easy with some common sense, but ML algorithms do not have this "common sense". So when creating ML algorithms, we need to present principles that protect the algorithm from conclusions that are not common sense [21].

2.4 Consumer smartwatch

Smartwatches are small computers that can be worn on the wrist instead of a conventional watch. Smartwatch can be categorised as smart wearable computers. Similar to conventional watches smartwatches can show time, take time, be used to set an alarm, but instead of conventional watches smartwatches can be used to install and run different digital applications. Many of these smartwatches can be connected to users' smartphones, and with that have the ability to receive incoming messages, calls, e-mails, and other notifications. The newer smartwatches even have the ability to track users sleep, measure heart rate, and the most interesting in the case of this study, some smartwatches can be used to measure users ECG [23] [24].

The number of smartwatches in circulation has been steadily increasing in the few latest years, as stated by the Fortune Business Insight's study [10]. The same study also is projecting that the market share of the smartwatches is not decreasing, instead it is going to increase, and quite rapidly as well. The global smartwatch market size was 18.62 billion US dollars in 2020 and was forecasted to increase to 58.21 billion US dollars [10].

With the increase in the usage of smartwatches, also comes the newest models' ability to record ECG. At the current time smartwatches that have the US Food and Drug Administration approval are Apple Watch Series 8, Apple Watch Ultra, Apple Watch Series 7, Google Pixel Watch, Samsung Galaxy Watch 4, Samsung Galaxy Watch 4 Classic, and Fitbit Sense [9]. With Samsung and Apple dominating a lion's share of the whole smartwatch market value with almost half of the shipment share of the new smartwatches in first quarter of 2022 [24]. As seen on the Figure 2.3

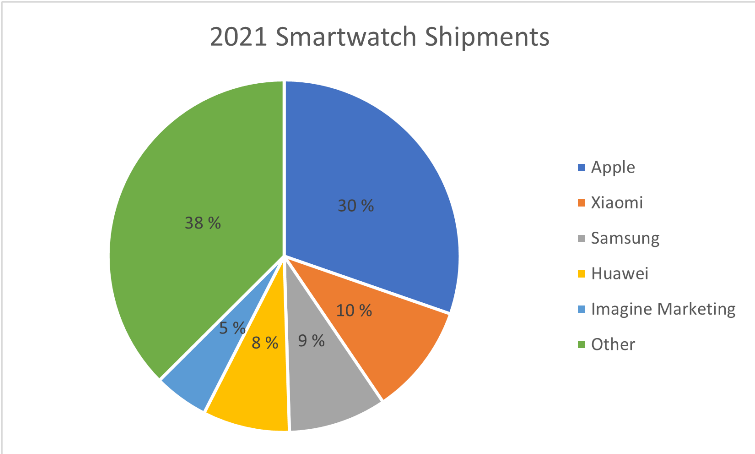

Figure 2.3: Smart wearable shipment volume in 2021 [25]

.

Apple and Samsung, along with Huawei have been able to conglomerate a sizable shipment volume of all smart wearables. Apple has the market share tightly in their grip as Apple has over half of the market share in 2020 [10].

3 State of the art

In this chapter we will take a deeper look on how and how well previous studies have been able to detect ECG with smartwatches. We will also take a look how previous studies have been able to analyze ECG with machine learning methods and machine learning, especially how well these machine learning methods are able to detect myocardial infarctions.

3.1 Performance metrics

When comparing different machine learning methods and models it is important to have some kind of a way to compare models. For this purpose, we introduce confusion matrix and accuracy, sensitivity, and specificity. Confusion matrix is a matrix helps with the performance evaluation of machine learning model in use. Below is a simple binary confusion matrix 3.1. This confusion matrix consists of true positives, true negatives, false positives, and false negatives. True positive is a guess or classification that machine learning model estimated correctly. For example, it labeled ECG data instance as normal, and the data was from normal patient. True negative is a classification model estimated also correctly. For example, it labeled ECG data instance as abnormal, and the data was from abnormal patient. False positive is when model classified data incorrectly. For example, model labeled ECG data instance as normal, and the data was from abnormal patient. False negative is when model classified data incorrectly. For example, model labeled ECG data instance as abnormal, and the data was from normal patient [26].

Actual values

		p	**n**	**total**
Prediction values	**p′**	True positive	False positive	P′
	n′	False negative	True negative	N′
	total	P	N	

Accuracy can then be concluded from the confusion matrix with a simple fraction:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

Sensitivity and recall can be calculated with fraction:

$$Sensitivity or recall = \frac{TP}{TP + FN}$$

Specificity can be calculated with fraction:

$$Specificity = \frac{TN}{TN + FP}$$

Precision can be calculated with fraction:

$$Precision = \frac{TP}{TP + FN}$$

F-score can be calculated with fraction:

$$F - score = 2 \times \frac{precision \times recall}{precision + recall} = \frac{2 \times TP}{2 \times TP + FP + FN}$$

Accuracy is simply the model's ability to classify data correctly. True positives and true negatives are summed up and divided by all the classifications, including false classifications. This score tells us how well model can classify data correctly. Sensitivity is the model's ability to detect positive instances. This is calculated by dividing true positives with true positives and false negatives. Sensitivity is also known as true positive rate. Specificity is model's ability to classify negatives correctly. Specificity also known as true negative rate is calculated by dividing true negatives with true negatives and false positives. Precision is calculated by dividing true positives with sum of true positives and false negatives. F-score is created with precision and recall [27].

3.2 Detecting myocardial infarction with artificial intelligence

The parameters of each study differ greatly, in this chapter I have tried to summarize most important parts and the conclusion of the studies chosen for this literacy overview. Studies use different amount of leads for the ECG, studies also use different machine learning methods.

I chose to include more classical and less computationally demanding algorithms. I also had more experience using these machine learning algorithms.

- K-nearest neighbors
- Support vector machine
- Decision trees
- Random forest
- Extra trees

To conduct the literacy overview I chose to use Google Scholar.

3.2.1 K-nearest neighbors

K-nearest neighbors (KNN) algorithm was created by Fix and Hodges for the US Air Force School of Aviation Medicine [28]. KNN algorithm is a quite straightforward and uncomplicated algorithm. It can solve both classification and regression problems. While it can perform regression prediction, but it is mostly used for classification predictions. KNN works by creating clusters from input data, and when algorithm meets new unseen data, it classifies it by assigning it to the cluster that has its nearest neighbors, meaning that the cluster has most similar data as the new data point. While KNN algorithm is effective, it has weaknesses [29].

KNN can be defined as supervised learning algorithm as it classifies new data with previous training data. It needs a training set that has data points which are labeled into corresponding classes. This classification is done by calculating Euclidean distance to its neighbors, and the closest neighbors are chosen. The new data is then classified according to the classes of the nearest neighbors. This simple Euclidean calculation means that the computational strain is low. One of the biggest weaknesses of the KNN algorithm is, that the algorithm is lazy. This means that the calculations are done when new data is introduced to the algorithm. This means that the training of the algorithm is mainly just storing and memorizing the training data [29].

Let's look at a simple example of a KNN algorithm. In Figure 3.1 we can see a data set which includes green and yellow points. These express two different classes. Blue point is a new unseen data for the algorithm. Algorithm then calculates nearest neighbors for the blue point, and in this case most of the neighbors are green, so the algorithm classifies the new data point as a green class. In our simple example the "K" of KNN algorithm was a five. This "K" tells us how many neighbors are taken into account.

First, I searched for articles and studies that used K-nearest neighbors method. This was done with search keyphrase "Detecting myocardial infarction in electrocardiogram AND nearest neighbor". When searching for latest articles, only seven articles were found, from which not one was using K-nearest neighbors machine

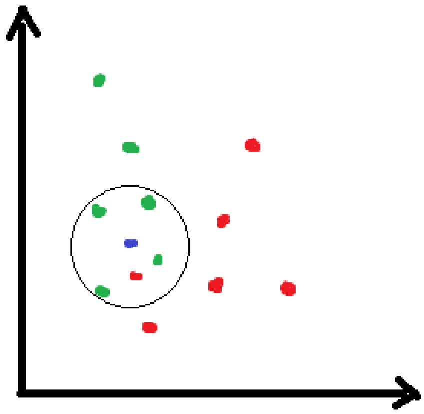

Figure 3.1: KNN graph

.

learning method as the any of the main methods to classify myocardial infarctions. Some studies, such as Elmannai, Saleh, D.Algrani, Mashal, Kwak, El-Sappagh, and Mostafa study [30] used k-nearest neighbor as meta-learner models, which were a part of a bigger convolutional neural network model.

Especially when using Google Scholars "sort by relevance", which is on by default, the studies were quite good. These studies achieved high ninety percent classification accuracies. Some of the articles that were using k-nearest neighbor algorithm as their classification method, were quite old, such as Arif, Malagore, and Afsar article from 2012 [31]. I chose to keep this in the in the comparison as it was able to get quite good classification accuracy. Articles that did not use k-nearest neighbor algorithm were removed from the articles chosen for the comparison.

Other articles that I chose from Google Scholar search were Acharya, Fujita, Sudarshan, et al. article [32], Acharya, Fujita, Adam, et al. article [33], Savostin, Ritter, and Savostina article [34], Sridhar, Lih, Jahmunah, et al. article [35], and Fatimah, Singh, Singhal, Pramanick, Pranav, and Pachori article [36].

In the Arif's article Arif, Malagore, and Afsar were able to achieve overall accuray of 98.3%, with sensitivity of 99.97%, and specificity of 99.9%. They used 20,160

Article	Acc	Sen/Recall	Spe
[31]	0.983	-	-
[32]	0.9880	0.9955	-
[33]	0.985	0.997	0.985
[34]	0.9703	0.9734	-
[35]	0.974	0.984	0.928
[36]	0.9996	0.9996	-

Table 3.1: Accuracy, sensitivity, and specificity of algorithms in the articles achieved. KNN

heartbeats in their method. They used wavelet transform to extract features from the electrocardiogram, and then used k-nearest neighbor classifier to classify heart beats. K-nearest neighbor was given 36 features, and cross validation was done with ten folds [31].

In Acharya's article Acharya, Fujita, Sudarshan, et al. were able to achieve a accuracy of 98.8%, with a sensitivity of 99.45%, and specificity of 96.27%. Acharya and et al. used 47 features. The data set included 61,1405 heart beats. Acharya, and et al. used ten folds for the cross validation, similarly to Arif, and et al. Acharya, and et al. were able to slightly increase the classification accuracy compared to Arif, and et al [32].

3.2.2 Support vector machine

Next machine learning algorithm, from which I searched previous studies and articles, was support vector machines. Support vector machine (SVM) is also similarly to KNN, an algorithm which main function is to classify and label objects. SVM can similarly to KNN classify for example handwritten digits, or in our case can classify if ECG data is normal or the hearts beats are irregular [37].

SVM algorithms work by trying to move the data into a higher dimension, in which the SVM tries to find a Support Vector Classifier, which is then used to classify the data points in the data set. And when new data is introduced to the algorithm, classifying is trivial, by looking at the Support Vector Classifier.

Let's look at an example on how SVM works on a 1-dimensional data set. In Figure 3.2 we can see on how the two different classes stand on a line. SVM cannot create a Support Vector Classifier in this case, so it will change the dimension to two dimensions, as seen on the Figure 3.3. In this dimension SVM can create a Support Vector Classifier, show as a blue line. Now when introducing new data points to the algorithm, it can easily classify them to either red or green classes, depending on which side of the Support Vector Classifier new points land on. SVM algorithm works similarly with data sets have higher dimensions.

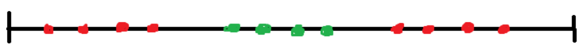

Figure 3.2: SVM 1d

.

The keyphrase used for this search was: "Detecting myocardial infarction in electrocardiogram AND support vector machine". This gave quite good articles and studies, similarly, to search for studies that used K-nearest neighbors algorithm.

Similarly, to K-nearest neighbors search, "sort by relevance" seemed to give the best results. The results were not sorted by the year, so it was critical the check the year, that specific article was published. For example, Dhawan, Wenzel, George, et al. article [38] from 2012, might not have the latest methods that machine learning field might have to offer.

Articles for support vector machine that I chose from Google Scholar search were A. Diker, Z. Cömert, E. Avci, and S. Velappan [39], A. K. Dohare, V. Kumar, and R. Kumar [40], Sridhar, Lih, Jahmunah, et al. article [35], and Fatimah, Singh,

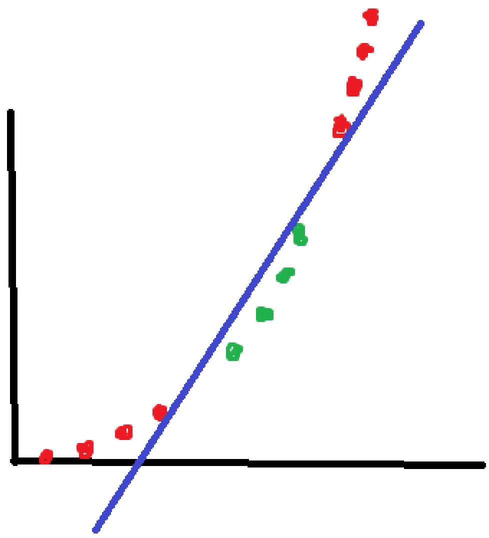

Figure 3.3: SVM 2d

.

Singhal, Pramanick, Pranav, and Pachori article [36]. Sridhar's [35] and Fatimah's [36] articles both had K-nearest neighbors algorithm, as well as support verctor machine, so both studies were in K-nearest neighbors and support vector machine previous studies subchapters.

A. Diker, Z. Cömert, E. Avci, and S. Velappan in their article [39] started with 23 features, that they used in their supporty vector machine. These 23 features helped them to achieve accuracy of 86.44%, sensitivity of 88.33% and specificity of 81.12%. They then fed the features to a generic algorithm. This generic algorithm chose best 9 features for their support vector machine. These better features were able to achieve accuracy of 87.80%, sensitivity of 86.97% and specificity of 88.67%. Both accuracy and specificity increased, but decreasing the features caused sensitivity to decrease. The accuracy of this support vector machine compared to k-nearest neighbor algorithms is not great. Next support vector algorithms luckily achieved results more comparable to k-nearest neighbor algorithms, from sub chapter 3.2.1.

Article	Acc	Sen	Spe
[39]	0.878	0.8697	0.8867
[40]	0.9666	0.9666	0.9666
[35]	0.979585	0.988946	0.938093
[36]	0.9987	0.9961	-

Table 3.2: Accuracy, sensitivity, and specificity of algorithms in the articles achieved. SVM

3.2.3 Decision trees

Next machine learning algorithm, from which I searched previous studies and articles, was decision trees. Similarly, to KNN and SVM, decision trees (DT) is an algorithm able to solve regression and classification problems. This means that DT is a good algorithm to for problems like classifying regular and irregular ECG graphs. DT in the simplest form reduces to a hierarchy or a tree of if/else clauses. DT algorithm can be though as a tree of true or false questions, and at the end of these questions are different classes that DT can classify inputs as [41].

In Figure 3.4 we can see a simple example on how DT algorithm works. This example tries to classify multiple different animals with some questions. In this example each node represents a question or an answer. The depth of this tree is 2, which means that the algorithm asks two questions from the data points. The complexity of the can be increased by adding more layers to the decision tree. This allows us to tackle more complex problems, and in higher dimensions [41].

The keyphrase used for this search was: "Detecting myocardial infarction in electrocardiogram AND decision trees". Second keyphrase for search was: "Detecting myocardial infarction in electrocardiogram AND decision trees -treebagger". "Treebagger" was excluded, because without the exclusion many of the search results were

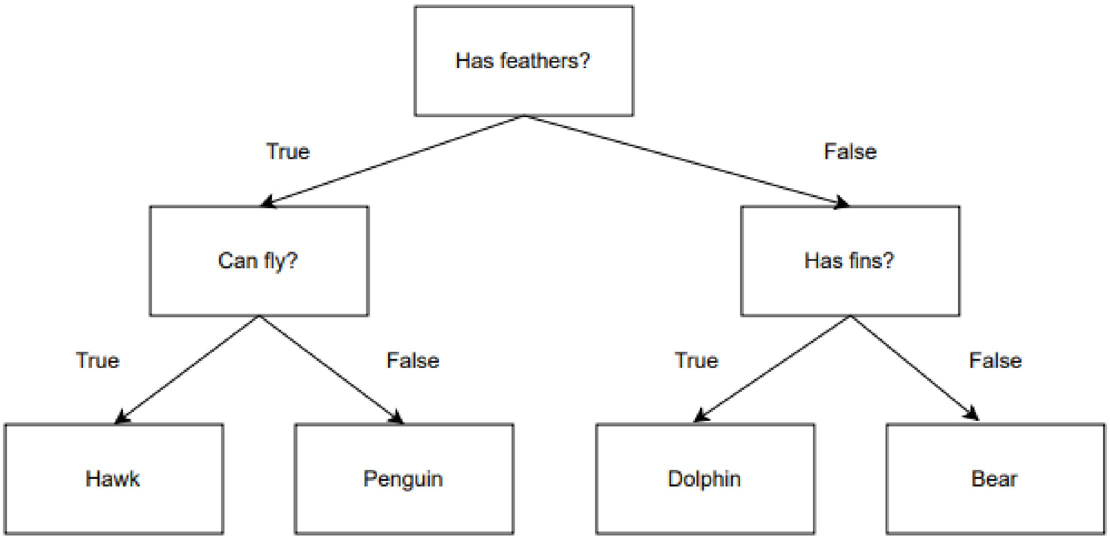

Figure 3.4: DT graph

.

studies done with Random Forest algorithm, specifically using MatLab's TreeBagger method [42]. And excluding TreeBagger results were more likely to use decision trees machine learning algorithm.

Similarly, to two previous searches, "sort by relevance" seemed to give the best results. The results were not sorted by the year, so it was critical the check the year, that specific article was published.

Articles for decision trees, that I chose from Google Scholar search were Liu, Liu, Wang, et al. article [43], Zhang, Lin, Xiong, et al. article [44], Ibrahim, Mesinovic, Yang, and Eid article [45], and Sridhar, Lih, Jahmunah, et al. article [35].

3.2.4 Random forest

Next machine learning algorithm, from which I searched previous studies and articles, was random forest. Random forest (RF) is an ensemble method. This means that it combines multiple machine learning methods to create better and more pow-

Article	Acc	Sen	Spe
[43]	0.895	0.942	0.74
[44]	0.9990	0.9998	0.9952
[45]	0.975	0.935	0.994
[35]	0.955145	0.973553	0.873551

Table 3.3: Accuracy, sensitivity, and specificity of algorithms in the articles achieved.

Decision trees

erful machine learning methods. This means that RF algorithm combines multiple DT algorithms, and uses their combined results to make a classification. Decision trees method has an overfitting problem, meaning that the training data is too closely corresponded. When model is overfitted it cannot predict reliably new unseen data. RF algorithm tries to address this overfitting problem, by having multiple slightly different decision trees. The pretense of RF is that each of the decision trees used can do classification quite well, but are slightly overfitted. And when we use multiple different trees we can get rid of the overfitting by taking an average of all the different trees [41].

The name random forest comes from the "forest" of decision trees, and from including some randomness to each of the decision tree of the forest. This randomness is used to ensure that the decision trees are different enough, to not cause problems with the overfitting. If multiple trees are similar and similarly overfitted, then the random forest method is also overfitted, which then beats the purpose of having multiple decision trees [41].

Random forest also uses thing called Bootstrap aggregating (Bagging) which is a method for making multiple different training data sets from a single bigger data set. This bagging method creates multiple different data sets that then can be used for individual decision trees. Bagging makes slightly different data sets for different

Article	Acc	Sen	Spe
[48]	0.9962	-	-
[47]	0.9971	0.997	0.9973
[49]	0.943	0.9413	0.9798
[46]	0.9988	0.9998	0.9939
[50]	0.9927	-	-

Table 3.4: Accuracy, sensitivity, and specificity of algorithms in the articles achieved. Random forest

decision trees, and in this way helps with the decision tree methods of overfitting.

The keyphrase used for this search was: "Detecting myocardial infarction in electrocardiogram AND random forest".

Articles for random forest, that I chose from Google Scholar search were J. Zhang, M. Liu, P. Xiong, et al. article [46], Z. Wang, L. Qian, C. Han, and L. Shi article [47], S. Nita, S. Bitam, and A. Mellouk article [48], . Kayikcioglu, F. Akdeniz, C. Köse, and T. Kayikciogl article [49], and W. Zhang, R. Li, S. Shen, et al. article [50].

3.2.5 Extra trees

Next machine learning algorithm, from which I searched previous studies and articles, was extra trees. Extra trees (ET) is similar to random forest algorithm. It is also an ensemble method meaning that extra trees consists of multiple different decision trees. Extra trees is similar to random forest, excluding couple of key differences. Where random forest uses bagging to create different training data sets to different decision trees, extra trees uses the complete training data set for each of the decision trees. This means that the decision trees overfitting problem needs to be solved with different methods. Extra trees method uses random values to

Article	Acc	Sen	Spe
[50]	0.9951	-	-
[53]	0.832	0.830	-
[54]	0.96	0.96	-

Table 3.5: Accuracy, sensitivity, and specificity of algorithms in the articles achieved. Extra trees

split the features of the data set. Compared to random forests, which uses greedy algorithm to choose best values to split the features. This random values at which to split the features, means that extra trees is computationally less costly compared random forest [51] [52].

The keyphrase used for this search was: "Detecting myocardial infarction in electrocardiogram AND extra trees". Unlike other machine learning methods ET is not as widely used. This can be attributed to many things, from which can be the popularity of RF method, and as ET and RF are quite similar ensemble methods, RF has shadowed other ensemble methods.

Articles for extra trees, that I chose from Google Scholar search were W. Zhang, R. Li, S. Shen, et al. article [50], M. Shimizu, M. Suzuki, H. Fujii, et al. article [53], and J. Sandelin's article [54].

3.3 Summary of different machine learning methods

All the articles and studies are in a single table 3.6. In the first column we have the article itself, in the second what method the article used. In the third column we have accuracy that writers of the article achieved, fourth column has sensitivity that the writers achieved, and fifth column has specificity. Last column has how many leads was used for the ECG. Every row has a different article, or if the article had multiple different models rows are for different models.

Article	Method	Acc	Sen	Spe	Leads
[31]	KNN	0.983	-	-	12
[32]	KNN	0.9880	0.9955	-	12
[33]	KNN	0.985	0.997	0.985	1
[34]	KNN	0.9703	0.9734	-	12
[35]	KNN	0.974	0.984	0.928	1
[36]	KNN	0.9996	0.9996	-	12
[39]	SVM	0.878	0.8697	0.8867	12
[40]	SVM	0.9666	0.9666	0.9666	12
[35]	SVM	0.979585	0.988946	0.938093	1
[36]	SVM	0.9987	0.9961	-	12
[43]	DT	0.895	0.942	0.74	12
[44]	DT	0.9990	0.9998	0.9952	12
[45]	DT	0.975	0.935	0.994	12
[35]	DT	0.955145	0.973553	0.873551	1
[48]	RF	0.9962	-	-	12
[47]	RF	0.9971	0.997	0.9973	12
[49]	RF	0.943	0.9413	0.9798	5
[46]	RF	0.9988	0.9998	0.9939	12
[50]	RF	0.9927	-	-	12
[50]	ET	0.9951	-	-	12
[53]	ET	0.832	0.830	-	12
[54]	ET	0.96	0.96	-	1

Table 3.6: Accuracy, sensitivity, and specificity of algorithms in the articles achieved.

3.4 Leave-One-Out Cross-Validation

Cross-validation is used to get an estimation on how well will the machine learning model perform. Cross-validation happens with first splitting data set to two parts, one for training the model and other to test the accuracy of the model. With this split we can ensure that the model is not trained with the data that we are using to test the accuracy. If the testing data is same as the training data we cannot get accurate performance of the model, because it can "remember" data from the training phase, meaning that the model will not unnecessarily work with data that it has not seen before [55].

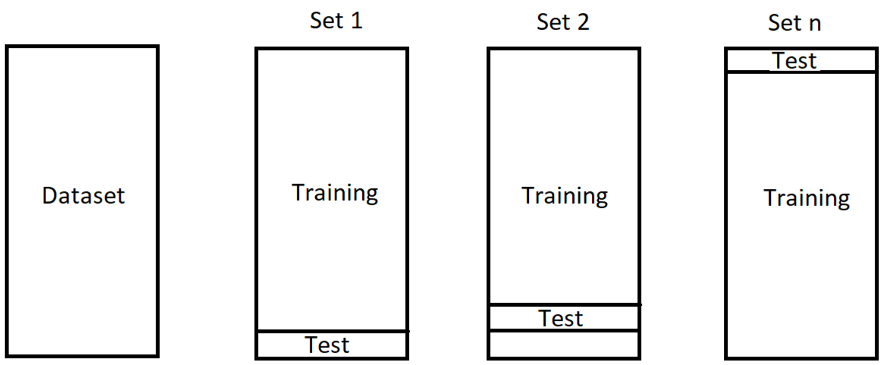

Figure 3.5: Leave-One-Out Cross-Validation

To refine cross-validation even further we can use Leave-One-Out Cross-Validation (LOOCV). In LOOCV the data set in split into as many parts as there are samples in the original data set. Then the model is trained on all the samples, excluding the one sample that was left for testing. After the model is trained, then it is tested with the one test sample. This loop is then repeated for every sample in the data set. In Figure 3.5 we can see how the data set is split into training data, and one test sample, and how this is repeated n-times, for every sample in the data set. LOOCV allows us to use the whole data set, and this is why it makes it important for smaller data sets. LOOCV also gives a more accurate performance estimations, as the model is trained and tested multiple times. This also means that LOOCV is more computationally demanding [55]. The LOOCV method used in this thesis is one from Python's scikit-learn package [56].

4 Data from smartwatches

In this chapter, we will take a look on the how we will get the ECG data from Samsung Galaxy Watch 4 and Apple Watch 6. This can be quite challenging as neither of these devices have been designed for research purposes. After the data acquisition, we will do some preprocessing to the data. Then we will train the machine learning algorithms from previous chapters with the data and classifications that we got out of the Samsung Galaxy Watch 4 and Apple Watch 6. And finally, we will try to use the trained algorithms to try and classify new and unseen data to the algorithms.

4.1 Devices

Both devices that are used for the data gathering are smart wearable watches and have the ability to gather ECG data from the wearer. The watches are from Apple and Samsung, two of the main players in the smartphone and the smartwatch markets.

4.1.1 Samsung Galaxy Watch 4

Samsung Galaxy Watch 4 is a smartwatch made by Samsung. The watch has a single channel ECG measuring capabilities. The ECG data is gathered with the BioActive sensor on the behind of the watches crown, and with the Home Key of the side of the watch as seen on the Figure 4.1. Watch can be used to measure different leads from all over the body, but it is mostly developed to be used on the left wrist, which means that the Einthoven's Lead II is the primary lead to be measured. The watch is capable of 500 Hz sampling size. But as the data is not readily available to users of the watch this sampling size is not the actual sampling size that we can use for the data analysis part of this thesis. This problem is addressed more in the chapter 4.3. In Figure 4.2 we can see the watches front and back, and how it looks on hand. The Samsung Galaxy Watch 4 is not medical device, so when using the data provided by the watch, one must be cautious [57].

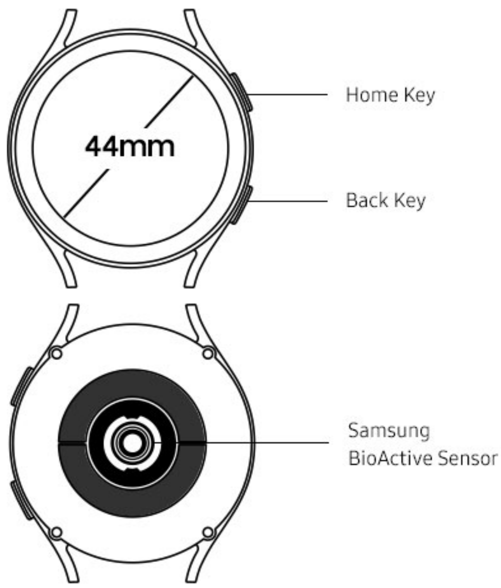

Figure 4.1: Samsung Galaxy Watch 4 blueprint [58]

As Sandelin has argued in his thesis, wearable devices could be used to detect irregularities in the heart's signals. This could be enhanced with weekly or even daily ECG measuring [54]. And if we can use the patient's own smartwatches, without needing new devices, this measuring would be much easier. Samsung Galaxy Watch 4 does not measure ECG automatically, so patient needs to start the measuring themselves, and the measurement is quite sensitivity to movement, meaning that

Figure 4.2: Samsung Galaxy Watch 4

during the measurement, the subject should remain stationary.

4.1.2 Apple Watch Series 6

The other device that was used in this thesis was an Apple Watch Series 6 - smartwatch. Similarly, to Galaxy Watch, Apples watch has a capability to measure single channel ECG signal, and similarly to Galaxy Watch the sensors are on the backside of the crown of the watch and on the upper button on the side of the watch. Watch can be used to measure different leads from all over the body, but it is mostly developed to be used on the left wrist, which means that the Einthoven's Lead II is the primary lead to be measured. The watch is capable of 512 Hz sampling size. But similarly, to Galaxy Watch, this sampling rate is not transferred to the data perfectly as the data must be scraped from the pdf-files (Portable Document Format) that the Apple supplies from the measurement. More on this in the chapter 4.3. In Figure 4.3 we can see the watches front and back, and how it looks on hand.

Figure 4.3: Apple Watch 6

4.2 Data gathering

The smartwatch data was gathered with two different devices, from a single person. Both watches recorded about 30 minutes of ECG data. This data was gathered in a sitting position. The subject was relaxed and had not been exercised before the measurement. For each device the measurements were 30 seconds long and 60 of these measurements were taken in a single sitting. The measurements were taken with the watches tightly on left hand and right hands index finger on the upper button of the watch.

Figure 4.4: Samsung Watch 4 measuring

The measurement produces a single pdf-file. Shown in the Figure 4.5. This pdf-file is then prepossessed to a format that can be used. This preprocessing is explained in more detail in section 4.3.

4.3 Data preprocessing

First problem of the data preprocessing was to extract the data from the wearable smartwatches. Both Apple and Samsung are quite stingy when it comes to the sharing of the data they have collected. In both devices the data comes in pdf-files. This makes it hard for to use this data, as it has to be somehow changed to format, in which data could be used for machine learning applications. With Samsung I had to download data from their Samsung Health app to a pdf-file that had the ECG graph. This graph had to be read with Python. Then the graph had to be extracted, and then the pixels with correct colors had to be read and saved. The x,y-coordinates were the ECG data as raw data points.

At first the quality of the images that were converted from pdf-files, was bad.

This caused issues when reading the graphs from the converted images. This quality issue was fixed with using PyMuPDF [59]. PyMuPDF packages had methods that allowed pdf-files to be converted png-files (Portable Network Graphics) without much data loss.

These png-files were then read. For each read png-file, the file was masked for hues of orange, which was the graph color in the Samsung Health Monitor ECG -pdf files. This mask then changes all the pixels in the image that are not orange, to black, and all the pixels that are hues of orange, are changed to white pixels. Then the coordinates of these white pixels are saved.

Python packages used:

1. Matplotlib [60]
2. Numpy [61]
3. OpenCV [62]
4. PyMuPDF [59]
5. WFDB [63]
6. pyHRV [64]
7. BioSPPy [65]
8. scikit-learn [66]
9. pandas [67]

In Figure 4.5 we can see how Samsung Galaxy Watch 4 outputs ECG data. From this pdf-file we need to extract the raw ECG data.

Figure 4.5: ECG pdf, Samsung Galaxy Watch 4

And in Figure 4.6 we can see when the data has been extracted as raw ECG data from the pdf-file.

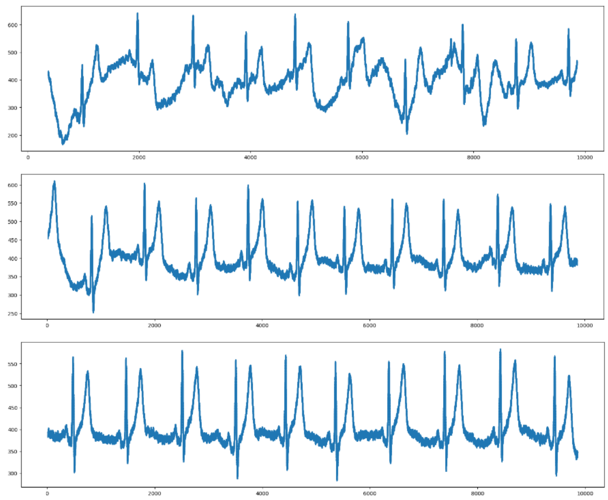

Figure 4.6: ECG pdf to data, Samsung Galaxy Watch 4

Similarly to Samsung, Apple also only gives the ECG data as pdf-files, and we need to extract the data to more usable form. In Figure 4.7 how the data looks from Apple Watch5.

Figure 4.7: ECG pdf, Apple Watch5

In Figure 4.8 we can see how the data has been extracted to raw ECG data points.

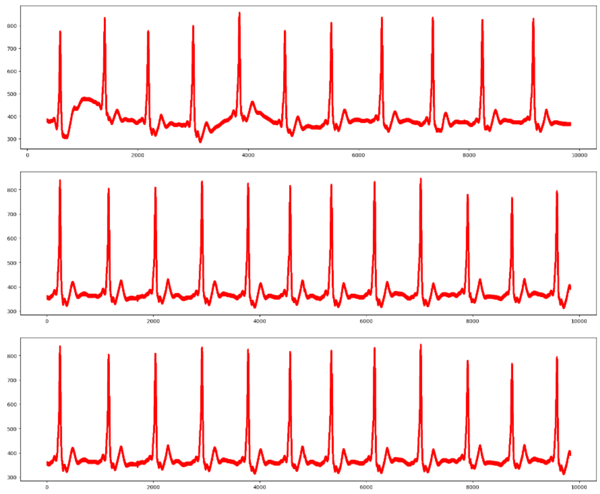

Figure 4.8: ECG pdf to data, Apple Watch5

4.3.1 Normalisation

The ECG samples were normalised between 0 and 1. This was done to uniform the values, as STAFF III data could be negative, and smartwatch data had values between ranges 100 and 800.

$$n = \sum_{0}^{i} \frac{x_i - min(x)}{max(x) - min(x)} \tag{4.1}$$

4.4 Peak detection

Peak detection is important part, when studying the patients ECG graph. Filtering is important to implement before the peak detection as, noise can warp the peak detection [68]. The points of interests that we want to detect from ECG graphs are QRS-complex, and P- and T-waves. With these we can calculate other important values such as ST- and PQ-segments.

4.4.1 R-peak detection

R-peak detection was made with pyHRV Python package [64]. pyHRV is a toolbox for measuring Heart Rate Variability (HRV) from ECG data. pyHRV has tools for gathering Time Domain Frequency Domain, and non linear HRV parameters from the ECG signal data. pyHRV uses BioSPPy package to detect the R-peaks [65]. BioSSPy is a Python package that is used to to analyse and process physiological signals, such as electrocardiograms and electroencephalograms. Were are not interested other than ECG signal analysing and processing functions. The function that we used from BioSSPy was 'biosppy.signals.ecg.ecg'. This function takes signal and sampling rate as inputs, and returns signals time series, filtered ECG signal, and R-peaks of the signal. In Figure 4.9 we can see how R-peaks are detected from signal.

4.4.2 P-, Q-, S- and T-wave detection

After the R-peaks are detected the other important landmarks from the ECG signal can be extracted. P-, Q-, S- and T-waves all relate to R-peak.

The P-peak and the P-wave is detected with the R-peak. Similarly, to Sandelin's thesis, I chose the start of the detection window to be at $1/4^{\text{th}}$, before the R-peak, of the mean interval between the R-peaks of the signal. The end of the detection

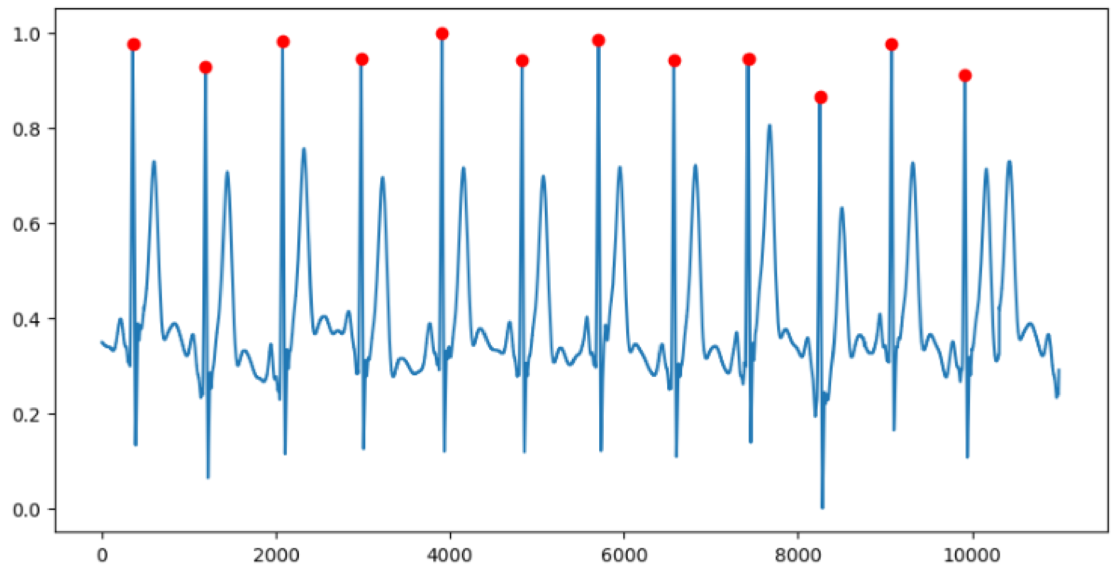

Figure 4.9: R-peak detection

window was at $1/9^{\text{th}}$, before the R-peak. With the start and the end values, the P-peak detection window is created, and the P-peak is the maximum value in the detection window [54]. This P-peak detection window can be seen in Figure 4.10. In the figure we can also see the P-peak itself, and how the window and P-peak relates to R-peak of the heartbeat.

Q-wave was detected with similarly as the P-wave. I chose Q-peak detection window start to be at $1/12^{\text{th}}$ of RR-interval before the R-peak. The end of the detection window was the location of R-peak. The Q-peak is the minimum value between this detection window. This is shown in Figure 4.11.

Q-wave was detected with similarly as the P-wave. I chose Q-peak detection window start to be at the location of the R-peak, and the end of the detection window at $1/12^{\text{th}}$ of RR-interval after the R-peak. The S-peak is the minimum value between this detection window. This is shown in Figure 4.12.

T-peak detection is done similarly to P-peak detection. I chose the start of the detection window $1/9^{\text{th}}$ RR-interval of the R-peak, and the end of the detection

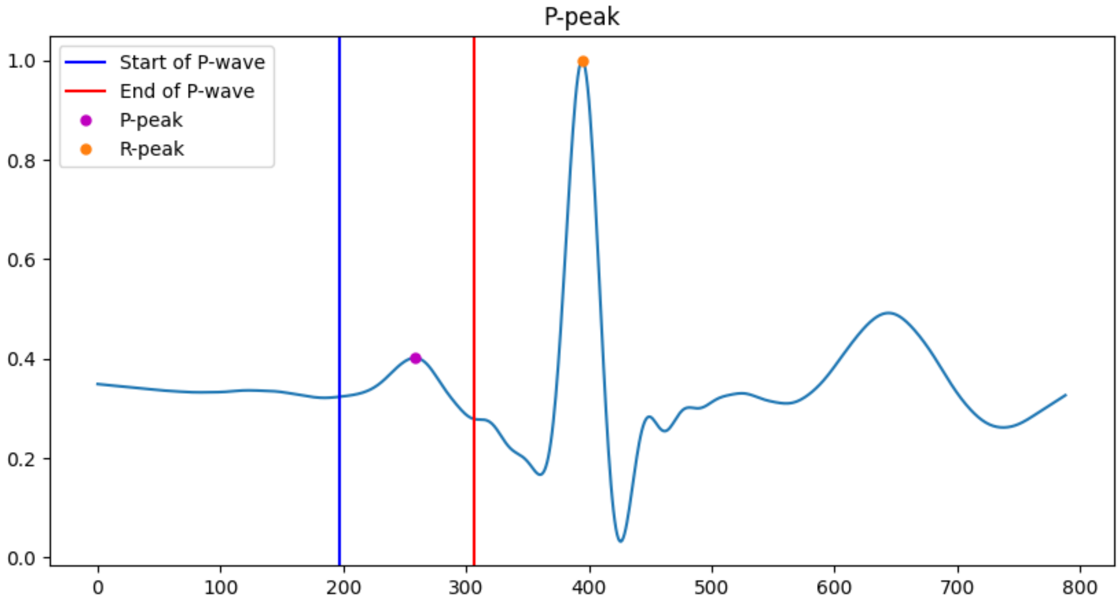

Figure 4.10: P-peak detection window

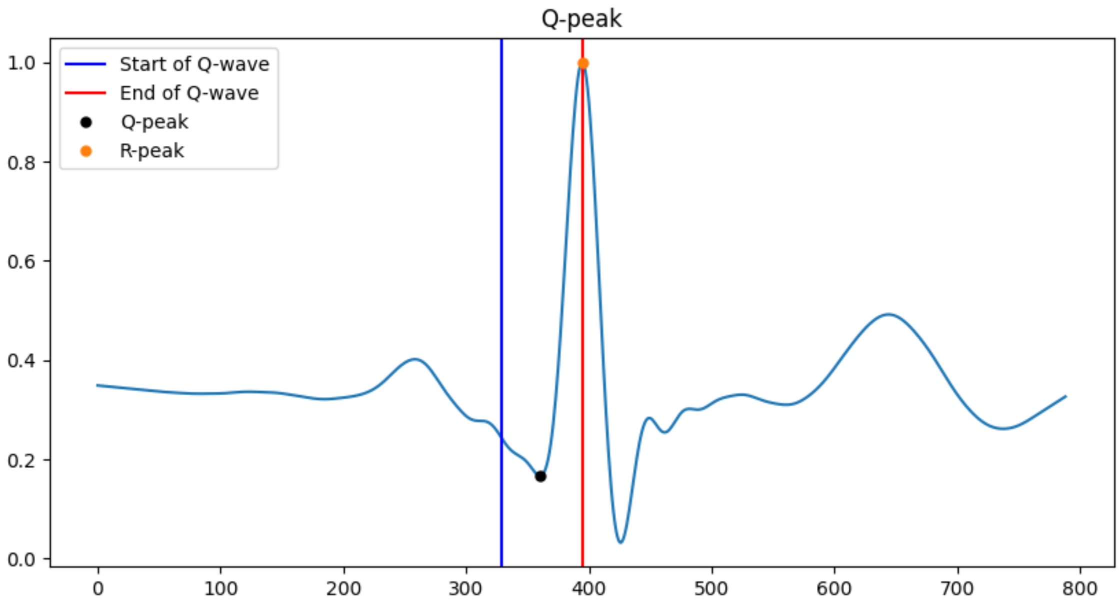

Figure 4.11: Q-peak detection window

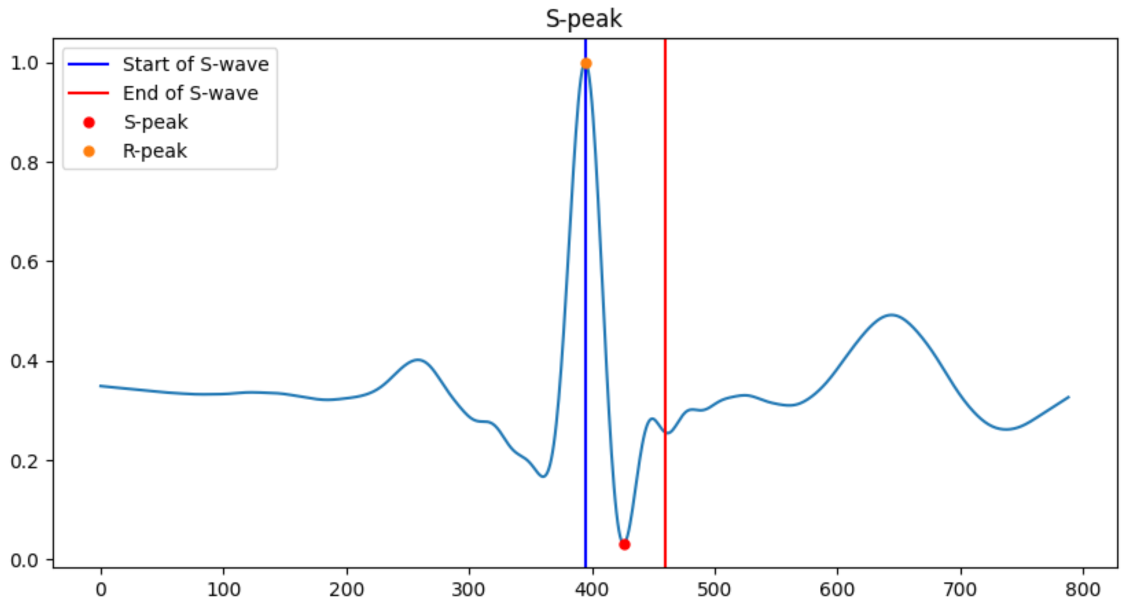

Figure 4.12: S-peak detection window

window to be at the $4/9^{\text{th}}$ RR-interval after the R-peak. The T-peak is the maximum value in this detection window. The detection window, R- and R-peaks can be seen in the Figure 4.13.

In the Figure 4.14 we can see all the heartbeats peaks detected.

4.5 Feature Extraction

To perform the classification of the patients, we first need to extract features from the ECG data. In previous section 4.4 we detected and extracted peaks from the ECG data. These peaks can be used in some features, but we still need more extracted features, to achieve good detection accuracy. These features ideally should give insight on how normal the specific snippet of data is. The features should lessen the burden of computation, in other words they should compress the ECG data. This compression should prevent data loss, meaning that the important features from

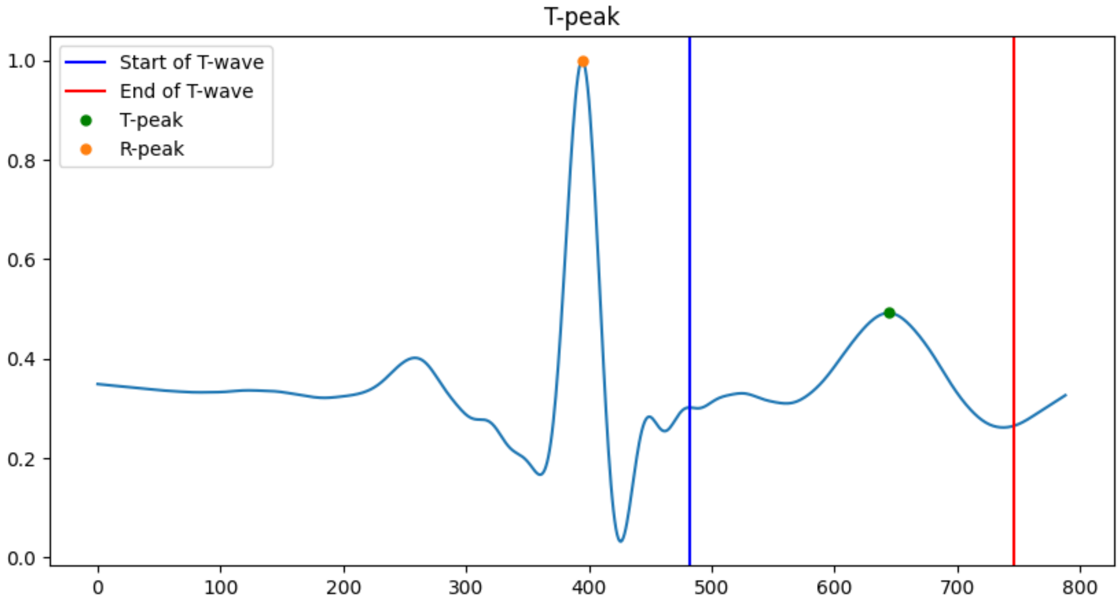

Figure 4.13: T-peak detection window

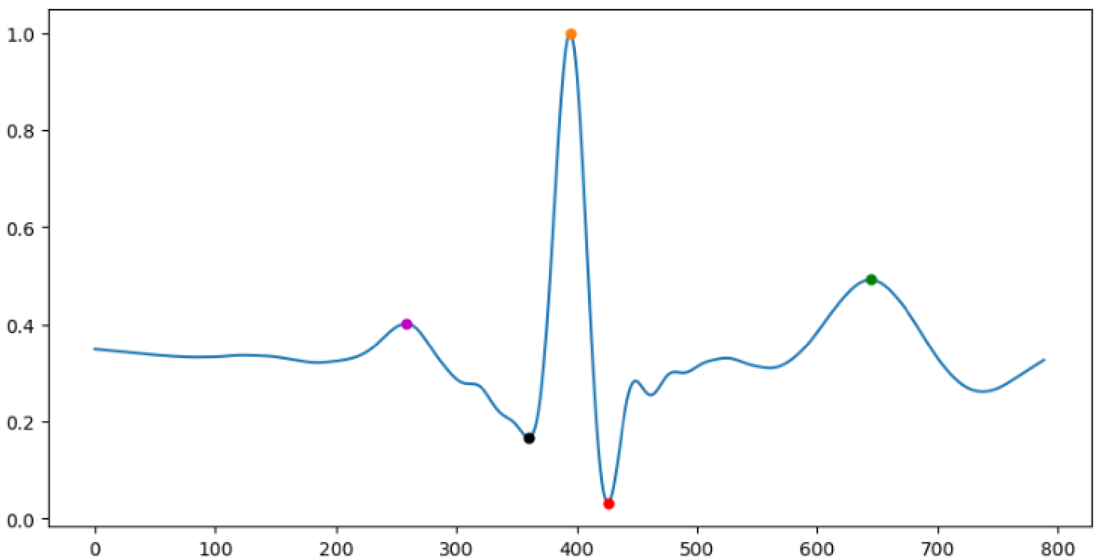

Figure 4.14: All peaks of one heartbeat

the original data would not be lost. These features are then used in the machine learning methods to create models, that can classify if signal is predicted as normal or predicted as having inflation.

The features that I chose for my machine learning models were:

- Heart rate (BPM)
- Root mean square of successive differences (RMSSD)
- Standard deviation of NN intervals (SDNN)
- Standard deviation of successive differences (SDSD)
- The proportion of RR intervals that have greater deviation than 20ms (pNN20)
- The proportion of RR intervals that have greater deviation than 50ms (pNN50)
- Very Low Frequency peak (VLF Peak)
- Low Frequency peak (LF Peak)
- High Frequency peak (HF Peak)
- Spread of RR intervals on the Poincaré plot (Poincaré plot SD1 and SD2 ratio)
- R-peak amplitude changes
- ST-segment amplitude changes
- ST-segment length changes
- PQ-segment amplitude changes
- PQ-segment length changes
- Slope of average waveform

These features were extracted from the STAFF III databases patient signals and from the smartwatch data that I gathered. The features are paired with classification, that tells if the features are from normal patient or are they from patient undergoing infarction.

5 STAFF III –database

In this chapter we will take a look on what STAFF III database is, and what does the database contain.

5.1 Database

The STAFF III is a database that contains a set of ECG data from patients that are receiving percutaneous transluminal coronary angiography (PTCA). This is a procedure where the blockage of an artery is opened, as seen on Figure 5.1. PTCA is a basic procedure after myocardial infarction, and this procedure has been demonstrated to have ability to normalise blood flow to coronary arteries and it has been able to lower rates of recurrent ischemia, infarction, stroke, and even death [69]. The aim of this database was to get a better understanding about myocardial ischemia in regards of ECG data. And what makes this database special is the ability to gather ECG data from patients that are going through acute myocardial infarctions [70] [13].

STAFF III was collected between 1995 and 1996 at Charleston Area Medical Center in West Virginia, USA. These procedures were single prolonged balloon inflation, to achieve best possible result in PTCA procedures. This single big inflation was used instead of multiple smaller brief inflations. STAFF III database consist of 104 patients. Standard procedure for each of 104 patients was to gather at least one baseline 5-minute-long ECG. Baseline ECGs were gathered either in supine position

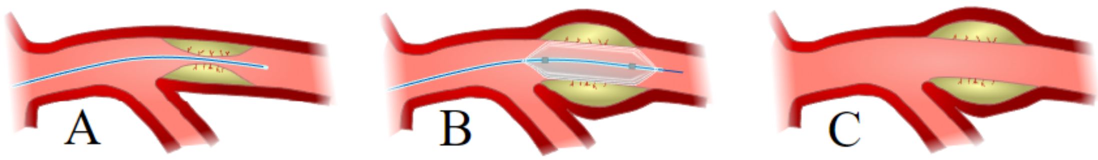

Figure 5.1: Percutaneous transluminal coronary angiography [71]

in relaxing room or in the catheterisation laboratory. For some patients two baseline ECGs were gathered in both places. From each patient up to 5 different inflation ECGs were gathered. These inflation ECGs were differing in lengths, ranging from 90 seconds up to 9 minutes and 54 seconds. Mean length for inflation records were 4 min 23 seconds. Similarly, to baseline ECGs post-inflation were acquired in relaxing room or in catheterisation laboratory, in supine position [13]. In Figure 5.2 we can see patient 102's two baseline ECG graphs, inflation ECG, and post-inflation ECG.

STAFF III database contains *.dat and *.hea files, with sometimes including annotation file *.event. Full database can be found in physionet [13]. All of the patients include multiple data files. Usually with A, and B post position marked files are baseline ECG data. Files with C to G markings are inflation and postinflation data files. The database includes an annotation Excel-file that describes all of the data files that the database contains. This annotations file includes timings of the infarctions, and how long each infarction lasted.

The database contains 152 occlusions, in major coronary arteries. These occlusions are distributed into 58 left anterior descendent (LAD) artery, 59 right coronary artery (RCA), 32 in left circumflex artery (LCX), and 3 in the left main (LM) artery [13]. The annotations file distributes these occlusions even further. For example, left anterior descending artery occlusion are divided into proximal, mid, and diagonal segments. Some patients have had dye injections during the procedure. Some of these injections have been annotated, but not all. Dye injections can cause changes

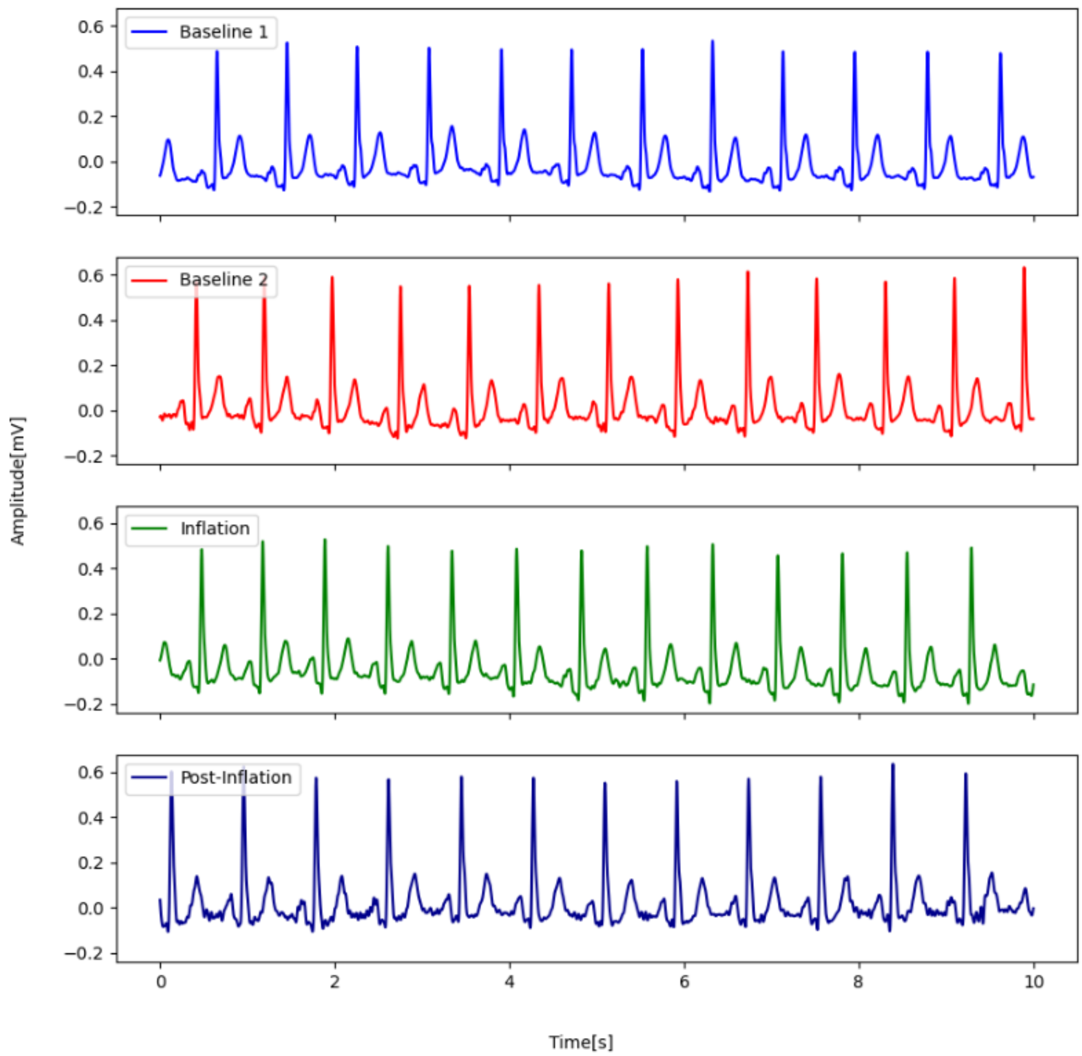

Figure 5.2: Different ECG graphs from patient 102 of STAFF III database [13]

in ECG data, so users are advised to be cautious [13].

The STAFF III database is soon 30 years since it has been gathered, but studies are still being published with the help of the database. This is partly thanks to the uniqueness of the database. The database is great data set when developing new methods for detecting ischemia and myocardial infarctions. STAFF III database does have its limitations, for example it does not have any annotations for QRS occurrence times [70].

6 Results

Results of the machine learning models, using the features that have been gathered from the STAFF III database and from the smartwatch data.

6.1 Results of the STAFF database

From the STAFF III database I was able to gather 271 ECG graphs from patients that were not undergoing infractions, and 130 ECG graphs from patients that were suffering from infarctions. The signals that were from patients suffering from infarctions were split into three different signals. This split was done to get more data sets from irregular ECG signals. The database included 658 different instances of ECG features.

This data set was then used with Leave-One-Out Cross-Validation and KNN, SVM, DT, RF, and ET classification methods. Confusion matrices made by the methods can be seen below.

In the Figure 6.1 we can see how the KNN model performed with only STAFF III data set. 214 ECG graphs were predicted to be healthy and were healthy. 126 graphs were predicted to be healthy but had infarctions. 54 graphs were predicted to have infarctions, but were healthy. 264 were predicted to have infarctions and were having infarctions. The accuracy of KNN model was measly 0.73.

In the Figure 6.2 we can see how the SVM model performed with only STAFF III data set. 184 ECG graphs were predicted to be healthy and were healthy. 64

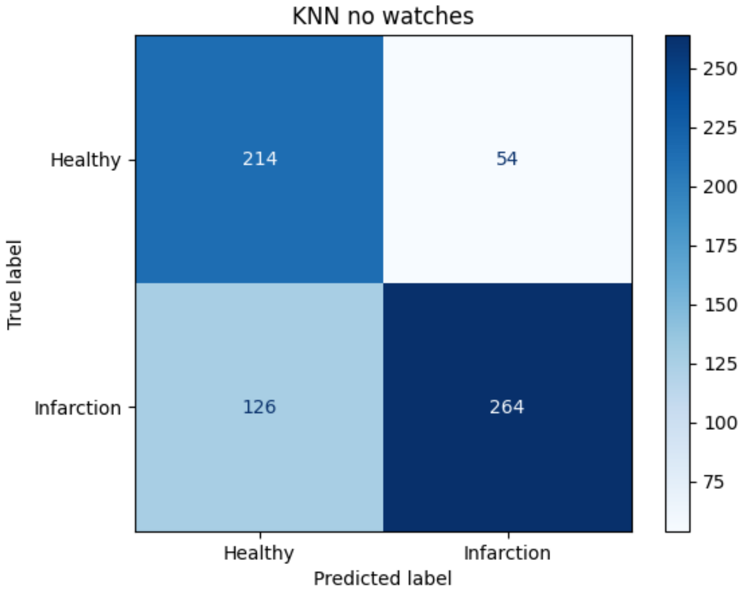

Figure 6.1: KNN confusion matrix without smartwatch data

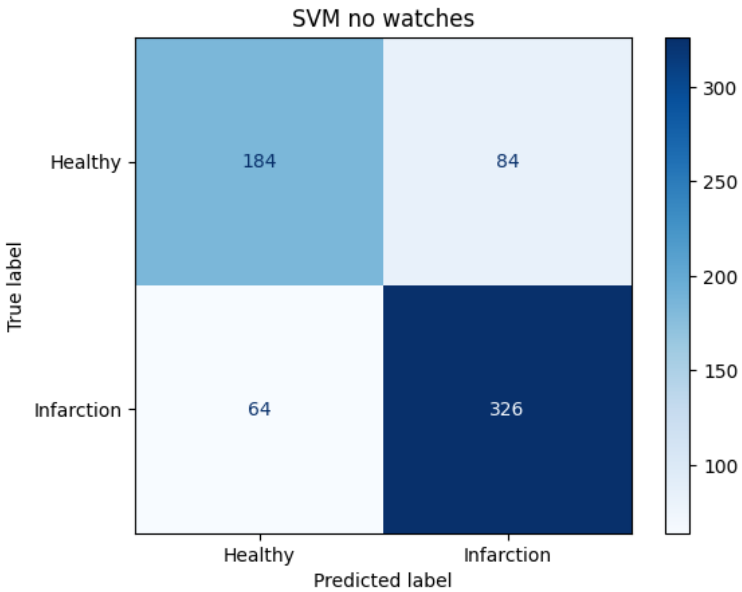

Figure 6.2: SVM confusion matrix without smartwatch data

graphs were predicted to be healthy but had infarctions. 84 graphs were predicted to have infarctions, but were healthy. 326 were predicted to have infarctions and were having infarctions. The accuracy of SVM model was a bit better than KNN, but still disappointing 0.78.

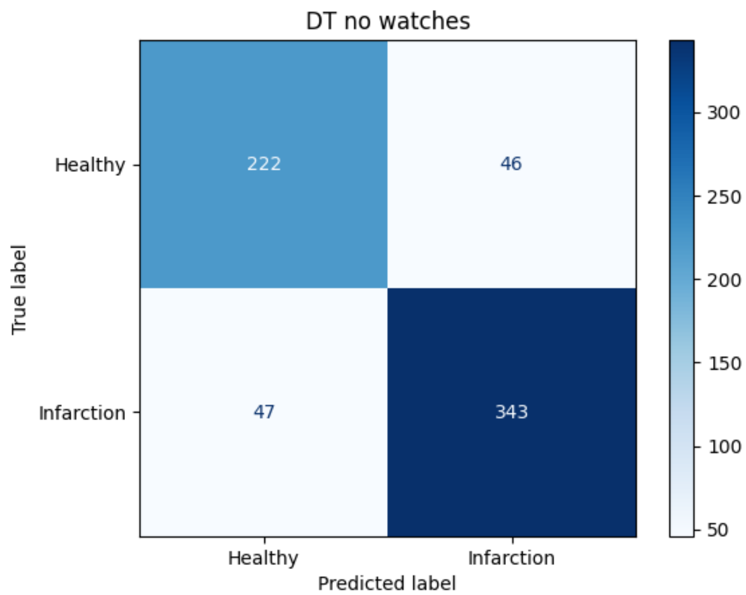

Figure 6.3: DT confusion matrix without smartwatch data

In the Figure 6.3 we can see how the DT model performed with only STAFF III data set. 222 ECG graphs were predicted to be healthy and were healthy. 47 graphs were predicted to be healthy but had infarctions. 46 graphs were predicted to have infarctions, but were healthy. 343 were predicted to have infarctions and were having infarctions. The accuracy of DT was much better than KNN and SVM, a tolerable 0.86.

In the Figure 6.4 we can see how the RF model performed with only STAFF III data set. 193 ECG graphs were predicted to be healthy and were healthy. Only 8 graphs were predicted to be healthy but had infarctions. 75 graphs were predicted to have infarctions, but were healthy.382 were predicted to have infarctions and were having infarctions. The accuracy of RF similar to DT model, with a 0.87. RF model has a better precision at 0.90, than DT model, but at the cost of lower recall at 0.85.

In the Figure 6.5 we can see how the ET model performed with only STAFF III data set. 234 ECG graphs were predicted to be healthy and were healthy. 25 graphs were predicted to be healthy but had infarctions. 34 graphs were predicted to have

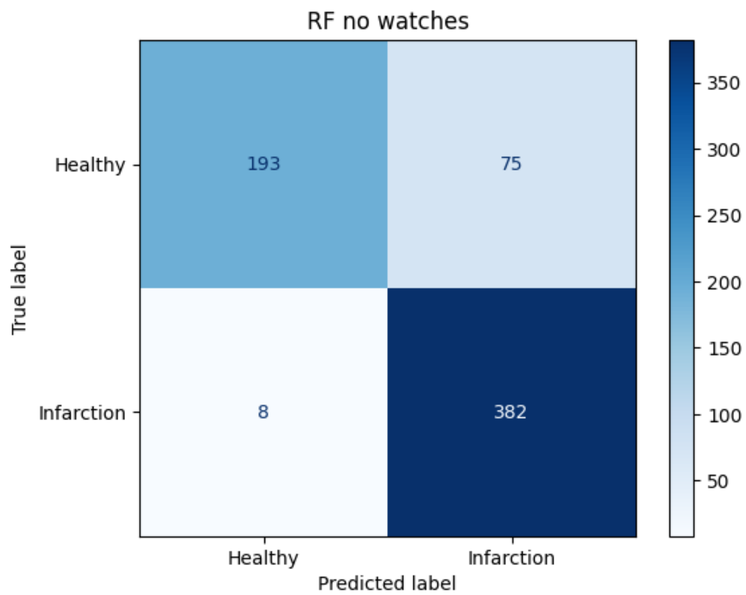

Figure 6.4: RF confusion matrix without smartwatch data

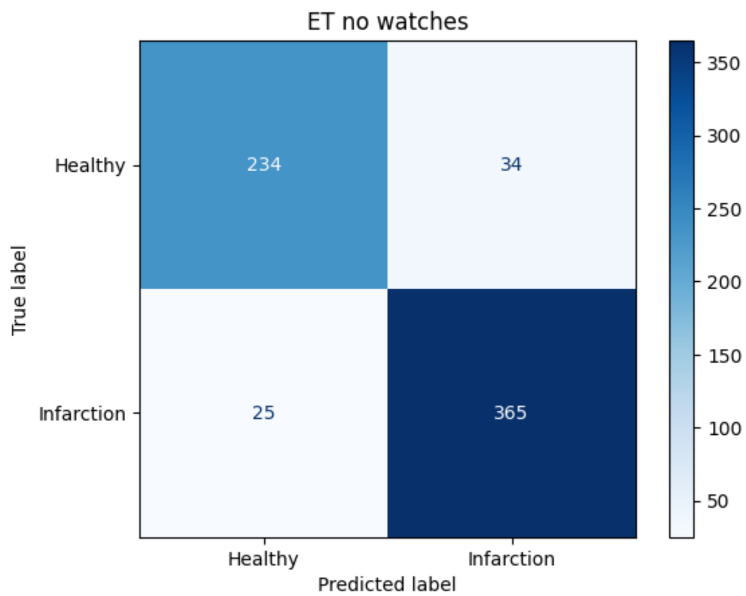

Figure 6.5: ET confusion matrix without smartwatch data

infarctions but were healthy. 365 were predicted to have infarctions and were having infarctions. The accuracy of ET was best one from all five models, with a 0.91. The precision, recall, and F1 score was also best from all of the models, this means that extra trees model is clearly best one from the five chosen models.

Model	**Accuracy**	**Precision**	**Recall**	**F1 score**
KNN	72.64%	72.98%	73.77%	72.49%
SVM	77.51%	76.85%	76.12%	76.41%
DT	86.02%	85.55%	85.46%	85.50%
RF	87.39%	89.80%	84.98%	86.25%
ET	91.03%	90.91%	90.45%	90.66%

Table 6.1: Performance metrics for classification models without smartwatch data

As we can see from table 6.1, the results from different methods differ quite bit. K-nearest neighbors and support vector machine methods are not great at classifying normal and abnormal ECG signals. They both are still much better than guessing the classes, but compared to decision trees, random forest, and extra trees methods, they do not match up. The computationally burdensome methods random forest and extra trees have better performance, with extra trees beating random forest with great margin. The KNN and SVM performance is quite disappointing. All of the methods were using LOOCV as the data set was quite small.

Table 6.1 answers to our first research question. These machine learning methods can be used to detect and classify acute myocardial infractions, from the STAFF III database. Extra trees model was able to classify normal ECG and infarction ECG graphs with 91 percent accuracy.

6.2 Results of the smartwatch data

STAFF III database was used for training set, and smartwatch data was used as testing set. As we can see from table 6.2, STAFF III database data and smartwatch data is not comparable. This is why the data set should combine both STAFF III data and data from smartwatches.

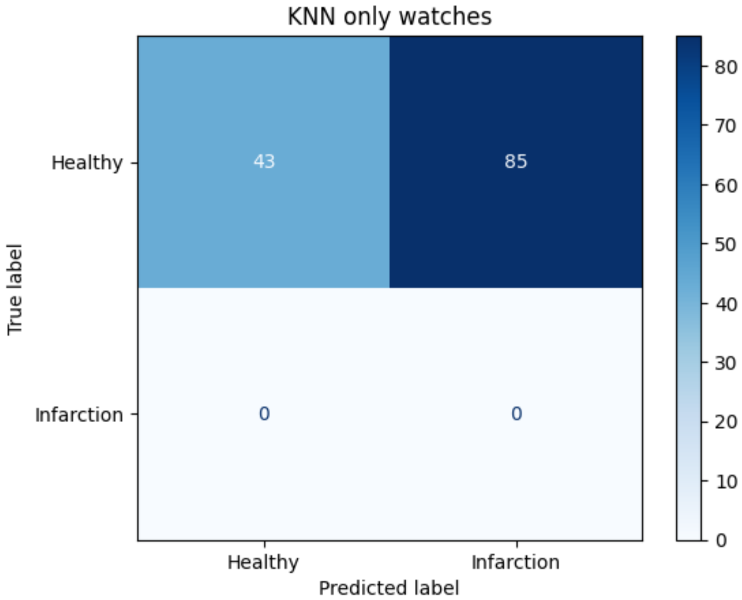

Figure 6.6: KNN confusion matrix only smartwatch data

In the Figure 6.6 we can see how the KNN model performed with training from STAFF III data set, and smartwatch data as testing data. 43 ECG graphs were predicted to be healthy and were healthy. 0 graphs were predicted to be healthy but had infarctions. 85 graphs were predicted to have infarctions, but were healthy. 0 were predicted to have infarctions and were having infarctions. The accuracy of KNN model was measly 0.34.

In the Figure 6.7 we can see how the SVM model performed with training from STAFF III data set, and smartwatch data as testing data. 20 ECG graphs were predicted to be healthy and were healthy. 0 graphs were predicted to be healthy but had infarctions. 108 graphs were predicted to have infarctions, but were healthy. 0 were predicted to have infarctions and were having infarctions. The accuracy of SVM model was only 0.16.

In the Figure 6.8 we can see how the DT model performed with training from STAFF III data set, and smartwatch data as testing data. 9 ECG graphs were predicted to be healthy and were healthy. 0 graphs were predicted to be healthy but

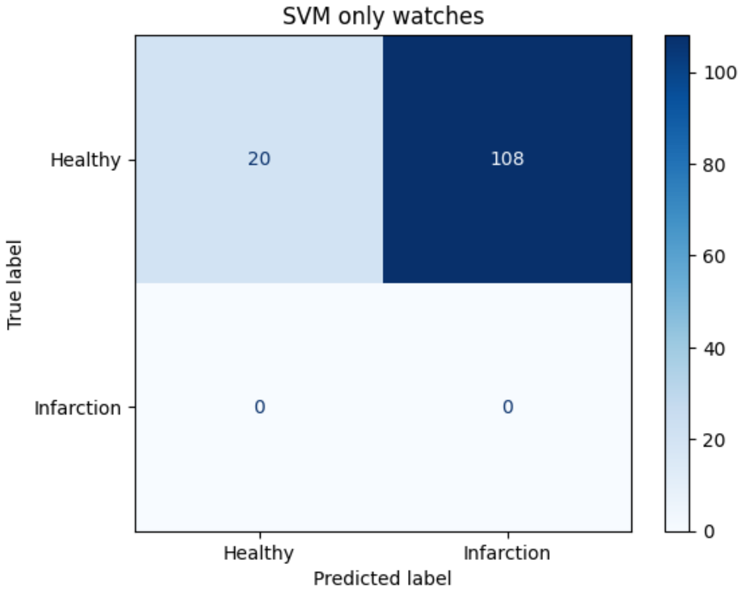

Figure 6.7: SVM confusion matrix only smartwatch data

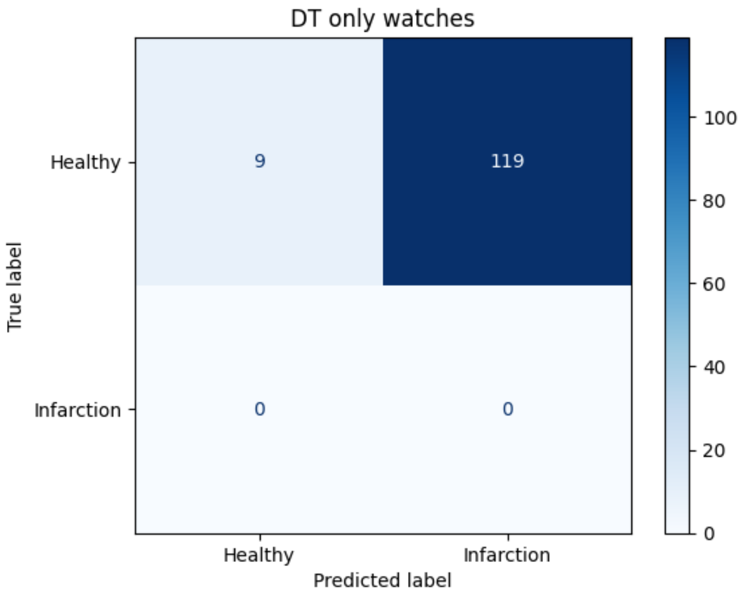

Figure 6.8: DT confusion matrix only smartwatch data

had infarctions. 119 graphs were predicted to have infarctions, but were healthy. 0 were predicted to have infarctions and were having infarctions. The accuracy of DT model was only 0.09.

In the Figure 6.9 we can see how the RF model performed with training from

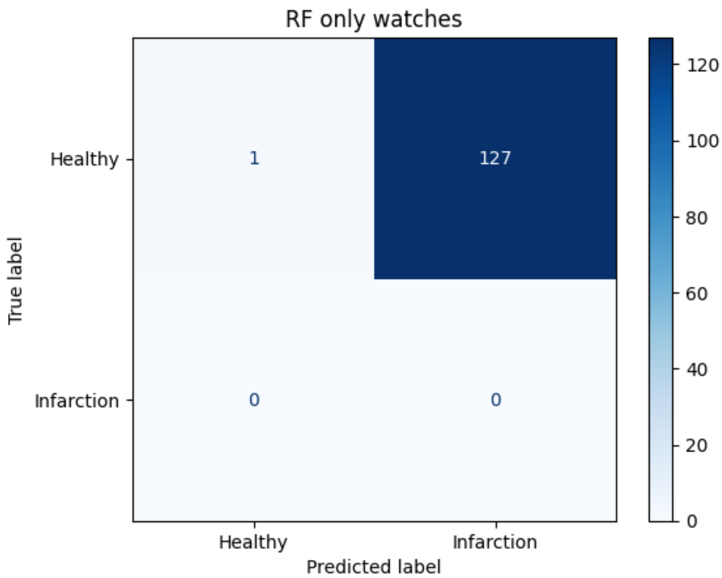

Figure 6.9: RF confusion matrix only smartwatch data

STAFF III data set, and smartwatch data as testing data. 1 ECG graphs were predicted to be healthy and were healthy. 0 graphs were predicted to be healthy but had infarctions. 127 graphs were predicted to have infarctions, but were healthy. 0 were predicted to have infarctions and were having infarctions. The accuracy of RF model was only 0.008.

In the Figure 6.10 we can see how the ET model performed with training from STAFF III data set, and smartwatch data as testing data. 4 ECG graphs were predicted to be healthy and were healthy. 0 graphs were predicted to have infarctions, but were healthy. 124 graphs were predicted to have infarctions, but were healthy. 0 were predicted to have infarctions and were having infarctions. The accuracy of ET model was only 0.03.

From the table 6.2, we can see that smartwatch data and STAFF III database data were so different, that they could not be used interchangeably, meaning that other data set could not be used to train the models, and other data set used to test.

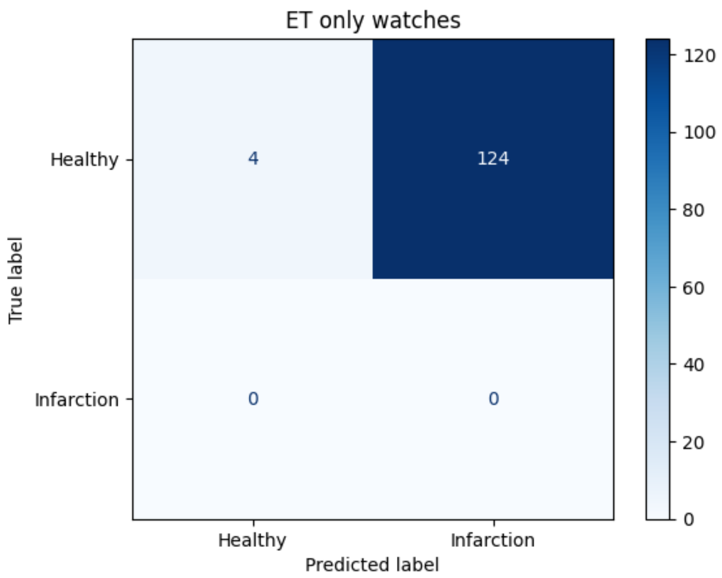

Figure 6.10: ET confusion matrix only smartwatch data

Model	Accuracy	Precision	Recall	F1 score
KNN	33.59%	50.00%	16.80%	25.15%
SVM	15.63%	50.00%	7.81%	13.51%
DT	9.38%	50.00%	4.69%	8.57%
RF	0.78%	50.00%	0.39%	0.78%
ET	3.13%	50.00%	1.56%	3.03%

Table 6.2: Performance metrics for classification models, trained with STAFF III and tested with smartwatch data

6.3 Results of the STAFF database and smartwatch data

Both STAFF database and smartwatch data was merged for this classification, as the smartwatch data set was quite small, containing only 128 entries.

As with only the STAFF III data, this data set was then used with Leave-One-

Out Cross-Validation and KNN, SVM, DT, RF, and ET classification methods. Confusion matrices made by the methods can be seen below.

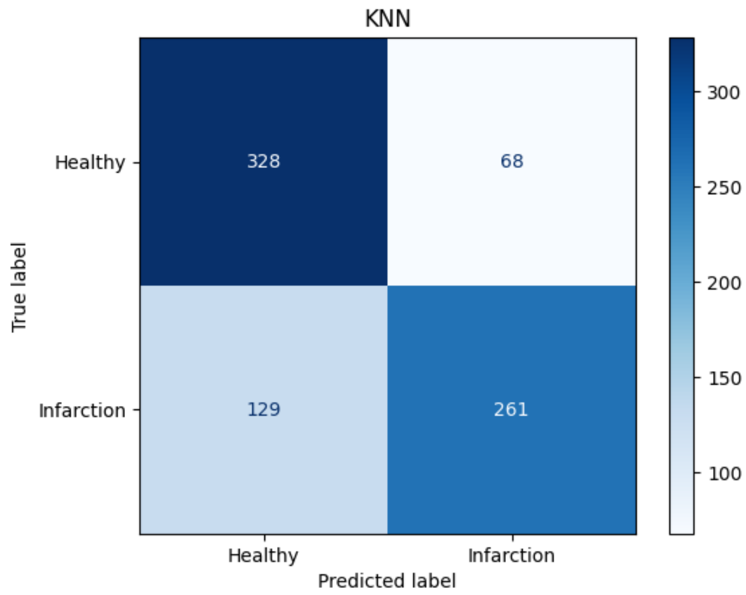

Figure 6.11: KNN confusion matrix

In the Figure 6.11 we can see how the KNN model performed with STAFF III and smartwatch data sets. 328 ECG graphs were predicted to be healthy and were healthy. 129 graphs were predicted to be healthy but had infarctions. 68 graphs were predicted to have infarctions, but were healthy. 261 were predicted to have infarctions and were having infarctions. The accuracy of KNN model was measly 0.75.

In the Figure 6.12 we can see how the SVM model performed with STAFF III and smartwatch data sets. 298 ECG graphs were predicted to be healthy and were healthy. 78 graphs were predicted to be healthy but had infarctions. 98 graphs were predicted to have infarctions, but were healthy. 312 were predicted to have infarctions and were having infarctions. The accuracy of SVM model was similar to KNN with a measly 0.78.

In the Figure 6.13 we can see how the DT model performed with STAFF III

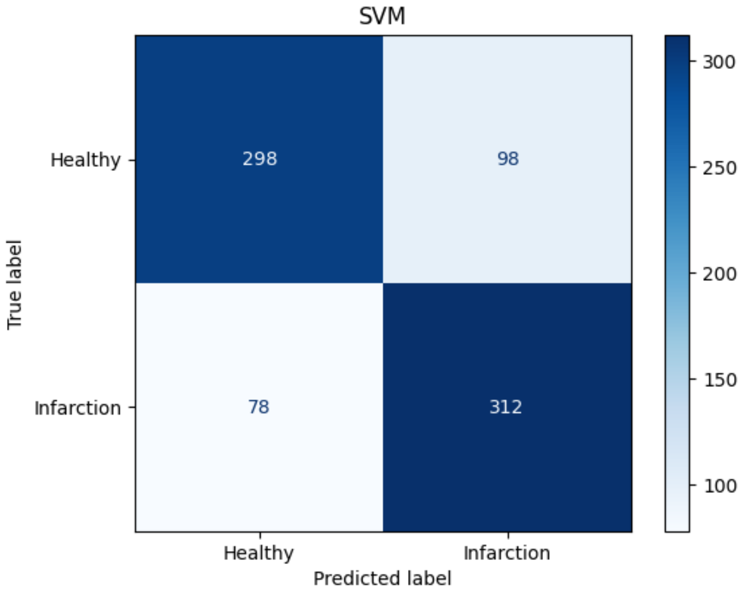

Figure 6.12: SVM confusion matrix

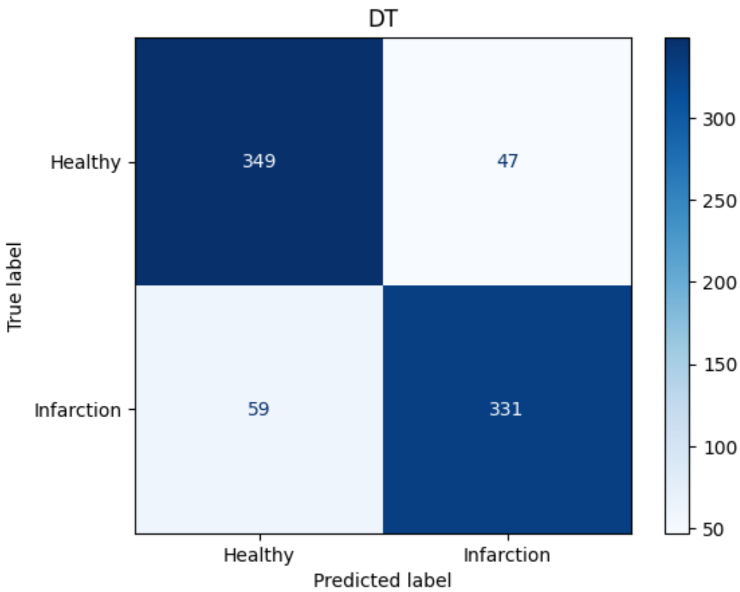

Figure 6.13: DT confusion matrix

and smartwatch data sets. 349 ECG graphs were predicted to be healthy and were healthy. 59 graphs were predicted to be healthy but had infarctions. 47 graphs were predicted to have infarctions, but were healthy. 331 were predicted to have infarctions and were having infarctions. The accuracy of DT was quite a bit better than both KNN and SVM models, with a 0.87.

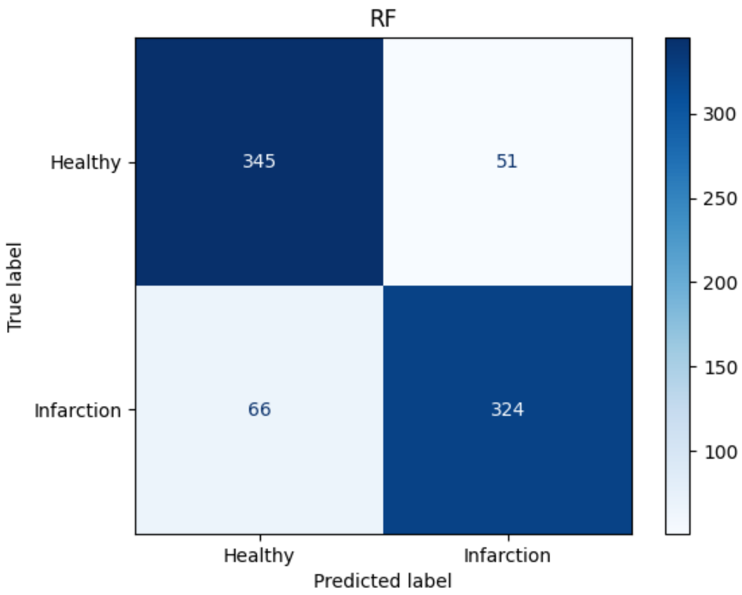

Figure 6.14: RF confusion matrix

In the Figure 6.14 we can see how the RF model performed with STAFF III and smartwatch data sets. 345 ECG graphs were predicted to be healthy and were healthy. 66 graphs were predicted to be healthy but had infarctions. 51 graphs were predicted to have infarctions, but were healthy. 324 were predicted to have infarctions and were having infarctions. The performance of RF was worse than DT model with accuracy of 0.85.

In the Figure 6.15 we can see how the ET model performed with STAFF III and smartwatch data sets. 363 ECG graphs were predicted to be healthy and were healthy. 39 graphs were predicted to be healthy but had infarctions. 33 graphs were predicted to have infarctions, but were healthy. 351 were predicted to have infarctions and were having infarctions. Like with only the STAFF III data, the performance of RF model was the best. Accuracy was 0.91, precision 0.91, recall 0.91, and F1 score 0.91.

As we can see from table 6.3, the more robust methods such as ET, RF, and

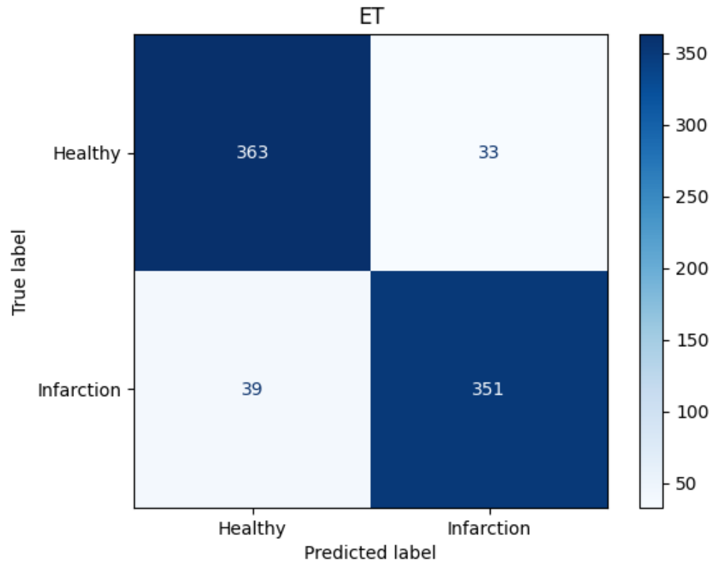

Figure 6.15: ET confusion matrix

Model	Accuracy	Precision	Recall	F1 score
KNN	74.94%	75.55%	74.88%	74.75%
SVM	77.61%	77.68%	77.63%	77.60%
DT	86.77%	86.80%	86.76%	86.76%
RF	85.11%	85.17%	85.10%	85.10%
ET	90.84%	90.85%	90.83%	90.84%

Table 6.3: Performance metrics for classification models

DT perform better than KNN and SVM methods. All the models performed worse with data from both STAFF III and smartwatches. Still DT and ET models were able to classify ECG graphs with good accuracy, with ET even reaching 90 percent classification accuracy. Like the models that were using only STAFF III data, models with both data sets were using LOOCV as the data set was still quite small, even with the smartwatch data.

Table 6.1 answers to our research question. These machine learning methods can be used to detect and classify acute myocardial infractions, from the STAFF

III database and gathered smartwatch data. Extra trees model was able to classify normal ECG and infarction ECG graphs with 91 percent accuracy.

7 Conclusion

In this thesis we were trying to detect and classify acute myocardial infarctions from electrocardiograph data, that was both collected from medical ECG machines, as well as gathered from Samsung Galaxy Watch 4 and from Apple Watch Series 6 watches. These wearable smartwatches could be a more practical way to measure patient's ECG data, as watches do not require complex ECG measuring devices compared to ECG machines used in medical settings. If these everyday carry watches could measure ECG data well enough, that the data could be used to detect acute myocardial infarctions, it would be highly beneficial.

The use of machine learning in detecting infractions, and other abnormalities in ECG graph data has been studied before, some of these studies and articles have been listed and discussed in chapter 3. These studies have been done with clinical ECG data, so a study how commercial use smartwatches that have the ability to measure ECG data would give a different perspective how useful machine learning methods could be. When in clinical settings the data is purer, this means that the studies done only with clinical data cannot be directly used, as ECG data gathered by patient's themselves, with wearable smart devices us not as clean as from clinical devices.

This thesis had two research questions. Answer for the first question can be seen from the table 6.1. The classification accuracy for the best model, which was extra trees, was over 90%. The machine learning models that were listed in that table were made from STAFF III database. This database included healthy and abnormal ECG data. This answers our first research question, and the answer is yes. Machine learning can be used to detect and classify myocardial infarctions, from data that includes healthy and infarction ECG data. The second question was about if wearable smartwatches could be used to detect acute myocardial infarctions. To get answer to this I gathered data from two different smartwatches. Machine learning models were then trained with smartwatch and STAFF III database data. In table 6.3 we can see how well. The best machine learning model was also extra trees which achieved classification accuracy of 91%. This means that the smartwatch data tandem with STAFF III data can be used to detect acute myocardial infarctions.

7.1 Further research

For further research it could be interesting to compare different manufacturers. Would there be differences between different manufacturers that offer smartwatches capable of measuring ECG. The ECG measuring ability between different commercial smartwatches, could be one reason when choosing smartwatches.

There are many reasons why Samsung and Apple are not releasing raw ECG data what their smartwatches measure. Regulatory requirements can be one reason why Samsung and Apple are not willing to release their raw ECG data, as the raw data might not be properly validated. Other reason could be that the ECG data is sensitive information, and if that sensitive information is shared to end users, this could pose a cyber security risks. These smartwatches are not medical equipment, so the raw data is not as reliable as data gathered from medical grade ECG measuring devices. This means that if the data is used as high-quality medical data, users could be harmed.

Lastly perhaps the biggest reason why Apple and Samsung are not willing to release their hard-earned data is intellectual property concerns. They have used their own valuable resources to gather this valuable user data, that they themselves can use to better their own algorithms. This is why they are not keen to release this data to other parties, as they understand the value that this biomedical data has and will have in the future.

References

[1] E. Braunwald, "Cardiovascular Medicine at the Turn of the Millennium: Triumphs, Concerns, and Opportunities", *New England Journal of Medicine*, vol. 337, no. 19, pp. 1360–1369, Nov. 1997, Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJM199711063371906, ISSN: 0028-4793. DOI: 10.1056/NEJM199711063371906. [Online]. Available: https://doi.org/10.1056/NEJM199711063371906 (visited on 11/30/2022).

[2] K. Thygesen, J. S. Alpert, A. S. Jaffe, *et al.*, "Fourth Universal Definition of Myocardial Infarction (2018)", en, *Circulation*, vol. 138, no. 20, Nov. 2018, ISSN: 0009-7322, 1524-4539. DOI: 10.1161/CIR.0000000000000617. [Online]. Available: https://www.ahajournals.org/doi/10.1161/CIR.0000000000000617 (visited on 11/30/2022).

[3] B. J. Drew and M. G. Adams, "Clinical consequences of st-segment changes caused by body position mimicking transient myocardial ischemia: Hazards of st-segment monitoring?", eng, *Journal of electrocardiology*, vol. 34, no. 3, pp. 261–264, 2001, ISSN: 0022-0736.

[4] D. L. Mann, D. P. Zipes, P. Libby, R. O. Bonow, and D. Bhatt, *Braunwald's Heart Disease E-Book: A Textbook of Cardiovascular Medicine*. London, UNITED STATES: Elsevier - Health Sciences Division, 2014, ISBN: 978-0-323-29064-7. [Online]. Available: http://ebookcentral.proquest.com/lib/kutu/detail.action?docID=2010634 (visited on 12/07/2022).

[5] B. J. A. Schijvenaars, G. van Herpen, and J. A. Kors, "Intraindividual variability in electrocardiograms", eng, *Journal of Electrocardiology*, vol. 41, no. 3, pp. 190–196, 2008, ISSN: 1532-8430. DOI: 10.1016/j.jelectrocard.2008.01.012.

[6] F. M. Fesmire, R. F. Percy, J. B. Bardoner, D. R. Wharton, and F. B. Calhoun, "Usefulness of automated serial 12-lead ECG monitoring during the initial emergency department evaluation of patients with chest pain", eng, *Annals of Emergency Medicine*, vol. 31, no. 1, pp. 3–11, Jan. 1998, ISSN: 0196-0644. DOI: 10.1016/s0196-0644(98)70274-4.

[7] J. Velez, W. J. Brady, A. D. Perron, and L. Garvey, "Serial electrocardiography", en, *The American Journal of Emergency Medicine*, vol. 20, no. 1, pp. 43–49, Jan. 2002, ISSN: 0735-6757. DOI: 10.1053/ajem.2002.28335. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0735675702180521 (visited on 12/02/2022).

[8] P. Jurak, J. Halamek, P. Leinveber, *et al.*, "Ultra-high-frequency ECG measurement", vol. 40, Jan. 2013, pp. 783–786, ISBN: 978-1-4799-0884-4.

[9] R. Udvarlaki, *Here are all the smartwatches that can take an ECG*, en-US, Section: Smartwatches, Jul. 2022. [Online]. Available: https://pocketnow.com/best-smartwatches-for-ecg/ (visited on 12/14/2022).

[10] *Smartwatch Market Size, Share, Trends | Growth Analysis [2028]*. [Online]. Available: https://www.fortunebusinessinsights.com/smartwatch-market-106625 (visited on 12/14/2022).

[11] M. AlGhatrif and J. Lindsay, "A brief review: History to understand fundamentals of electrocardiography", *Journal of Community Hospital Internal Medicine Perspectives*, vol. 2, no. 1, p. 14 383, Jan. 2012, Publisher: Taylor & Francis _eprint: https://doi.org/10.3402/jchimp.v2i1.14383, ISSN: null. DOI:

10.3402/jchimp.v2i1.14383. [Online]. Available: https://doi.org/10.3402/jchimp.v2i1.14383 (visited on 01/18/2023).

[12] W. B. Fye, "A History of the origin, evolution, and impact of electrocardiography", en, *The American Journal of Cardiology*, vol. 73, no. 13, pp. 937–949, May 1994, ISSN: 0002-9149. DOI: 10.1016/0002-9149(94)90135-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/000291499490135X (visited on 01/18/2023).

[13] L. Sörnmo, J. P. Martínez, and P. Laguna, *STAFF III Database*, Type: dataset, 2017. DOI: 10.13026/C20P4H. [Online]. Available: https://physionet.org/content/staffiii/ (visited on 12/02/2022).

[14] B. J. Drew, R. M. Califf, M. Funk, *et al.*, "Practice Standards for Electrocardiographic Monitoring in Hospital Settings", *Circulation*, vol. 110, no. 17, pp. 2721–2746, Oct. 2004, Publisher: American Heart Association. DOI: 10.1161/01.CIR.0000145144.56673.59. [Online]. Available: https://www.ahajournals.org/doi/10.1161/01.CIR.0000145144.56673.59 (visited on 12/07/2022).

[15] J. Lehmacher, J. T. Neumann, N. A. Sörensen, *et al.*, "Predictive Value of Serial ECGs in Patients with Suspected Myocardial Infarction", *Journal of Clinical Medicine*, vol. 9, no. 7, p. 2303, Jul. 2020, ISSN: 2077-0383. DOI: 10.3390/jcm9072303. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7408822/ (visited on 12/02/2022).

[16] F. T. Bennett, K. R. Bennett, and A. K. Markov, "Einthoven's Triangle: Lead Errors and an Algorithm for Solution", en, *The American Journal of the Medical Sciences*, vol. 329, no. 2, pp. 71–77, Feb. 2005, ISSN: 0002-9629. DOI: 10.1097/00000441-200502000-00004. [Online]. Available: https://www.

sciencedirect.com/science/article/pii/S0002962915339100 (visited on 12/07/2022).

[17] Y. Sattar and L. Chhabra, "Electrocardiogram", eng, in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK549803/ (visited on 04/01/2023).

[18] G. W. Reed, J. E. Rossi, and C. P. Cannon, "Acute myocardial infarction", en, *The Lancet*, vol. 389, no. 10065, pp. 197–210, Jan. 2017, ISSN: 0140-6736. DOI: 10.1016/S0140-6736(16)30677-8. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140673616306778 (visited on 12/16/2022).

[19] A. (Atkielski), *File:SinusRhythmLabels.png - Wikipedia*, en. [Online]. Available: https://commons.wikimedia.org/wiki/File:SinusRhythmLabels.png;%20https://commons.wikimedia.org/w/index.php?curid=1560893 (visited on 12/07/2022).

[20] Z. Cao, M. Zhao, C. Xu, *et al.*, "Diagnostic Roles of Postmortem cTn I and cTn T in Cardiac Death with Special Regard to Myocardial Infarction: A Systematic Literature Review and Meta-Analysis", *International Journal of Molecular Sciences*, vol. 20, no. 13, p. 3351, Jul. 2019, ISSN: 1422-0067. DOI: 10.3390/ijms20133351. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6651768/ (visited on 12/16/2022).

[21] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*, en, 1st ed. Cambridge University Press, May 2014, pp. 20–27, ISBN: 978-1-107-05713-5 978-1-107-29801-9. DOI: 10.1017/CBO9781107298019. [Online]. Available: https://www.cambridge.org/core/product/identifier/9781107298019/type/book (visited on 12/21/2022).

[22] B. F. Skinner, *Classics in the History of Psychology – Skinner (1948)*. [Online]. Available: `http://www.yorku.ca/pclassic/Skinner/Pigeon/` (visited on 12/21/2022).

[23] K.-L. Hsiao and C.-C. Chen, "What drives smartwatch purchase intention? Perspectives from hardware, software, design, and value", en, *Telematics and Informatics*, vol. 35, no. 1, pp. 103–113, Apr. 2018, ISSN: 0736-5853. DOI: `10.1016/j.tele.2017.10.002`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0736585317305804` (visited on 12/14/2022).

[24] S. Lim, *Smartwatch Market Grows 13% YoY in Q1 2022; Apple Stays First, Samsung Solidifies Second Place*, en-US, Section: Press Releases, May 2022. [Online]. Available: `https://www.counterpointresearch.com/smartwatch-market-grows-13-yoy-q1-2022-apple-stays-first-samsung-solidifies-second-place/` (visited on 12/14/2022).

[25] J. Ubrani, M. Shirer, and R. Llamas, *Wearables Deliver Double-Digit Growth for Both Q4 and the Full Year 2021, According to IDC*. [Online]. Available: `https://www.idc.com/getdoc.jsp?containerId=prUS48935722` (visited on 12/14/2022).

[26] M. Khanna, *Sensitivity, Specificity and Accuracy - Decoding the Relationship*, en, Jun. 2021. [Online]. Available: `https://www.analyticsvidhya.com/blog/2021/06/classification-problem-relation-between-sensitivity-specificity-and-accuracy/` (visited on 01/25/2023).

[27] A. Kumar, *Machine Learning - Sensitivity vs Specificity Difference*, en-US, Apr. 2022. [Online]. Available: `https://vitalflux.com/ml-metrics-sensitivity-vs-specificity-difference/` (visited on 01/25/2023).

[28] E. Fix and J. J.L. Hodges, "Discriminatory Analysis - Nonparametric Discrimination: Consistency Properties", en, Tech. Rep., Section: Technical Reports.

[Online]. Available: https://apps.dtic.mil/sti/citations/ADA800276 (visited on 02/16/2023).

[29] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification", May 2019, pp. 1255–1260. DOI: 10.1109/ICCS45141.2019.9065747.

[30] H. Elmannai, H. Saleh, A. D. Algarni, *et al.*, "Diagnosis Myocardial Infarction Based on Stacking Ensemble of Convolutional Neural Network", en, *Electronics*, vol. 11, no. 23, p. 3976, Jan. 2022, Number: 23 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 2079-9292. DOI: 10.3390/electronics11233976. [Online]. Available: https://www.mdpi.com/2079-9292/11/23/3976 (visited on 01/04/2023).

[31] M. Arif, I. A. Malagore, and F. A. Afsar, "Detection and Localization of Myocardial Infarction using K-nearest Neighbor Classifier", English, *Journal of Medical Systems*, vol. 36, no. 1, pp. 279–89, Feb. 2012, Num Pages: 279-89 Place: New York, Netherlands Publisher: Springer Nature B.V., ISSN: 0148-5598. DOI: 10.1007/s10916-010-9474-3. [Online]. Available: https://www.proquest.com/docview/921482829/abstract/4E33DFFCDE884868PQ/1 (visited on 01/04/2023).

[32] U. R. Acharya, H. Fujita, V. K. Sudarshan, *et al.*, "Automated detection and localization of myocardial infarction using electrocardiogram: A comparative study of different leads", en, *Knowledge-Based Systems*, vol. 99, pp. 146–156, May 2016, ISSN: 0950-7051. DOI: 10.1016/j.knosys.2016.01.040. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705116000708 (visited on 01/04/2023).

[33] U. R. Acharya, H. Fujita, M. Adam, *et al.*, "Automated characterization and classification of coronary artery disease and myocardial infarction by decom-

position of ECG signals: A comparative study", en, *Information Sciences*, vol. 377, pp. 17–29, Jan. 2017, ISSN: 0020-0255. DOI: 10.1016/j.ins.2016.10.013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025516311938 (visited on 01/04/2023).

[34] A. A. Savostin, D. V. Ritter, and G. V. Savostina, "Using the K-Nearest Neighbors Algorithm for Automated Detection of Myocardial Infarction by Electrocardiogram Data Entries", en, *Pattern Recognition and Image Analysis*, vol. 29, no. 4, pp. 730–737, Oct. 2019, ISSN: 1555-6212. DOI: 10.1134/S1054661819040151. [Online]. Available: https://doi.org/10.1134/S105466181904015 (visited on 01/04/2023).

[35] C. Sridhar, O. S. Lih, V. Jahmunah, *et al.*, "Accurate detection of myocardial infarction using non linear features with ECG signals", en, *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3227–3244, Oct. 2020, ISSN: 1868-5145. DOI: 10.1007/s12652-020-02536-4. [Online]. Available: https://doi.org/10.1007/s12652-020-02536-4 (visited on 01/04/2023).

[36] B. Fatimah, P. Singh, A. Singhal, D. Pramanick, P. S., and R. B. Pachori, "Efficient detection of myocardial infarction from single lead ECG signal", en, *Biomedical Signal Processing and Control*, vol. 68, p. 102 678, Jul. 2021, ISSN: 1746-8094. DOI: 10.1016/j.bspc.2021.102678. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809421002755 (visited on 01/04/2023).

[37] W. S. Noble, "What is a support vector machine?", en, *Nature Biotechnology*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006, Number: 12 Publisher: Nature Publishing Group, ISSN: 1546-1696. DOI: 10.1038/nbt1206-1565. [Online]. Available: https://www.nature.com/articles/nbt1206-1565 (visited on 02/16/2023).

[38] A. Dhawan, B. Wenzel, S. George, I. Gussak, B. Bojovic, and D. Panescu, "Detection of acute myocardial infarction from serial ecg using multilayer support vector machine", in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2012, pp. 2704–2707.

[39] A. Diker, Z. Cömert, E. Avci, and S. Velappan, "Intelligent system based on Genetic Algorithm and support vector machine for detection of myocardial infarction from ECG signals", in *2018 26th Signal Processing and Communications Applications Conference (SIU)*, May 2018, pp. 1–4. DOI: 10.1109/SIU.2018.8404299.

[40] A. K. Dohare, V. Kumar, and R. Kumar, "Detection of myocardial infarction in 12 lead ECG using support vector machine", en, *Applied Soft Computing*, vol. 64, pp. 138–147, Mar. 2018, ISSN: 1568-4946. DOI: 10.1016/j.asoc.2017.12.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494617307159 (visited on 01/06/2023).

[41] A. C. Mu, "Introduction to Machine Learning with Python", en, pp. 70–92, 2016.

[42] MathWorks, *Treebagger*. [Online]. Available: https://www.mathworks.com/help/stats/treebagger.html (visited on 11/01/2023).

[43] B. Liu, J. Liu, G. Wang, *et al.*, "A novel electrocardiogram parameterization algorithm and its application in myocardial infarction detection", en, *Computers in Biology and Medicine*, vol. 61, pp. 178–184, Jun. 2015, ISSN: 0010-4825. DOI: 10.1016/j.compbiomed.2014.08.010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482514002121 (visited on 01/11/2023).

[44] J. Zhang, F. Lin, P. Xiong, *et al.*, "Automated Detection and Localization of Myocardial Infarction With Staked Sparse Autoencoder and TreeBagger",

IEEE Access, vol. 7, pp. 70 634–70 642, 2019, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2919068.

[45] L. Ibrahim, M. Mesinovic, K.-W. Yang, and M. A. Eid, ”Explainable Prediction of Acute Myocardial Infarction Using Machine Learning and Shapley Values”, *IEEE Access*, vol. 8, pp. 210 410–210 417, 2020, Conference Name: IEEE Access, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3040166.

[46] J. Zhang, M. Liu, P. Xiong, *et al.*, ”A multi-dimensional association information analysis approach to automated detection and localization of myocardial infarction”, en, *Engineering Applications of Artificial Intelligence*, vol. 97, p. 104 092, Jan. 2021, ISSN: 0952-1976. DOI: 10.1016/j.engappai.2020.104092. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197620303444 (visited on 01/11/2023).

[47] Z. Wang, L. Qian, C. Han, and L. Shi, ”Application of multi-feature fusion and random forests to the automated detection of myocardial infarction”, en, *Cognitive Systems Research*, vol. 59, pp. 15–26, Jan. 2020, ISSN: 1389-0417. DOI: 10.1016/j.cogsys.2019.09.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S138904171930467X (visited on 01/18/2023).

[48] S. NITA, S. BITAM, and A. MELLOUK, ”An Enhanced Random Forest for Cardiac Diseases Identification based on ECG signal”, in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*, ISSN: 2376-6506, Jun. 2018, pp. 1339–1344. DOI: 10.1109/IWCMC.2018.8450361.

[49] İ. Kayikcioglu, F. Akdeniz, C. Köse, and T. Kayikcioglu, ”Time-frequency approach to ECG classification of myocardial infarction”, en, *Computers & Electrical Engineering*, vol. 84, p. 106 621, Jun. 2020, ISSN: 0045-7906. DOI:

10.1016/j.compeleceng.2020.106621. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790620304766 (visited on 01/18/2023).

[50] W. Zhang, R. Li, S. Shen, *et al.*, "Interpretable Detection and Location of Myocardial Infarction Based on Ventricular Fusion Rule Features", en, *Journal of Healthcare Engineering*, vol. 2021, e4123471, Oct. 2021, Publisher: Hindawi, ISSN: 2040-2295. DOI: 10.1155/2021/4123471. [Online]. Available: https://www.hindawi.com/journals/jhe/2021/4123471/ (visited on 01/18/2023).

[51] K. Thankachan, *What? When? How?: ExtraTrees Classifier*, en, Aug. 2022. [Online]. Available: https://towardsdatascience.com/what-when-how-extratrees-classifier-c939f905851c (visited on 02/17/2023).

[52] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees", en, *Machine Learning*, vol. 63, no. 1, pp. 3–42, Apr. 2006, ISSN: 0885-6125, 1573-0565. DOI: 10.1007/s10994-006-6226-1. [Online]. Available: http://link.springer.com/10.1007/s10994-006-6226-1 (visited on 02/17/2023).

[53] M. Shimizu, M. Suzuki, H. Fujii, S. Kimura, M. Nishizaki, and T. Sasano, "Machine learning of microvolt-level 12-lead electrocardiogram can help distinguish takotsubo syndrome and acute anterior myocardial infarction", en, *Cardiovascular Digital Health Journal*, vol. 3, no. 4, pp. 179–188, Aug. 2022, ISSN: 2666-6936. DOI: 10.1016/j.cvdhj.2022.07.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666693622000445 (visited on 01/19/2023).

[54] J. Sandelin, "Identification of myocardial infarction by high-frequency serial ECG measurement", eng, Apr. 2022, Accepted: 2022-04-19T21:01:30Z. [Online]. Available: https://www.utupub.fi/handle/10024/153691 (visited on 11/20/2022).

[55] A. R. Lemos, *Cross-Validation*, en, Apr. 2022. [Online]. Available: https://towardsdatascience.com/cross-validation-705644663568 (visited on 04/01/2023).

[56] *Sklearn.model_selection.LeaveOneOut*, en. [Online]. Available: https://scikit-learn/stable/modules/generated/sklearn.model_selection.LeaveOneOut.html (visited on 04/01/2023).

[57] *Are Galaxy Watch medical devices?*, en-002. [Online]. Available: https://www.samsung.com/africa_en/support/mobile-devices/are-galaxy-watch-medical-devices/ (visited on 03/08/2023).

[58] *Samsung Galaxy Watch 4*, en. [Online]. Available: https://www.samsung.com/global/galaxy/galaxy-watch4/specs/ (visited on 03/08/2023).

[59] *PyMuPDF 1.21.0*, original-date: 2012-10-06T18:54:25Z, Feb. 2023. [Online]. Available: https://github.com/pymupdf/PyMuPDF (visited on 02/03/2023).

[60] *Matplotlib — Visualization with Python*. [Online]. Available: https://matplotlib.org/ (visited on 02/01/2023).

[61] *NumPy*. [Online]. Available: https://numpy.org/ (visited on 02/01/2023).

[62] *Opencv/opencv-python*, original-date: 2016-04-08T13:36:40Z, Feb. 2023. [Online]. Available: https://github.com/opencv/opencv-python (visited on 02/01/2023).

[63] *The WFDB Python Package*, original-date: 2016-06-06T13:36:38Z, Feb. 2023. [Online]. Available: https://github.com/MIT-LCP/wfdb-python (visited on 02/24/2023).

[64] P. Gomes, *Pyhrv/index.rst at master · PGomes92/pyhrv*, en. [Online]. Available: https://github.com/PGomes92/pyhrv (visited on 03/16/2023).

[65] *Welcome to BioSPPy — BioSPPy 0.6.1 documentation*. [Online]. Available: https://biosppy.readthedocs.io/en/stable/ (visited on 03/28/2023).

[66] *Scikit-learn: Machine learning in Python — scikit-learn 1.2.2 documentation*. [Online]. Available: `https://scikit-learn.org/stable/` (visited on 04/04/2023).

[67] *Pandas - Python Data Analysis Library*. [Online]. Available: `https://pandas.pydata.org/` (visited on 04/04/2023).

[68] I. I. Christov, "Real time electrocardiogram QRS detection using combined adaptive threshold", *BioMedical Engineering OnLine*, vol. 3, no. 1, p. 28, Aug. 2004, ISSN: 1475-925X. DOI: `10.1186/1475-925X-3-28`. [Online]. Available: `https://doi.org/10.1186/1475-925X-3-28` (visited on 03/16/2023).

[69] C. L. Grines, D. A. Cox, G. W. Stone, *et al.*, "Coronary Angioplasty with or without Stent Implantation for Acute Myocardial Infarction", *New England Journal of Medicine*, vol. 341, no. 26, pp. 1949–1956, Dec. 1999, Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJM199912233412601, ISSN: 0028-4793. DOI: `10.1056/NEJM199912233412601`. [Online]. Available: `https://doi.org/10.1056/NEJM199912233412601` (visited on 02/24/2023).

[70] P. Laguna and L. Sörnmo, "The STAFF III ECG database and its significance for methodological development and evaluation", en, *Journal of Electrocardiology*, vol. 47, no. 4, pp. 408–417, Jul. 2014, ISSN: 0022-0736. DOI: `10.1016/j.jelectrocard.2014.04.018`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0022073614001502` (visited on 02/24/2023).

[71] [Online]. Available: `https://upload.wikimedia.org/wikipedia/commons/6/68/Angioplasty-scheme.svg` (visited on 02/24/2023).