Heartbeat Sound Feature Extraction and Classification

by

Fariha Chowdhury Bibrity
19101677
Farhana Jahan
18341012
Md. Shahriar Khan
14301033

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
December,2019

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Fariha Chowdhury Bibrity
19101677

Farhana Jahan
18341012

Md. Shahriar Khan
14301033

Approval

The Thesis titled " Heart Beat Sound Feature Extraction and Classification" Submitted by

1. Fariha Chowdhury Bibrity (19101677)

2. Farhana Jahan (18341012)

3. Md. Shahriar Khan (14301033)

Of Summer, 2019 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on Decembe 24,2019.

Examining Committee:

Supervisor:
(Member)

Dr. Jia Uddin
Associate Professor
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Dr. Md Golam Rabiul Alam
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department: (Chair)

Dr. Mahbubul Alam Majumdar
Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

This research work is solely done by us. We haven't done any unfair means throughout our work.

Abstract

Heart diseases has ranked top as the cause of death globally. The harsh truth is, in this time it is hard to get proper medical treatment in proper time and still it is costly. Now the only light of hope is coming from technology. Heart sound is one of the oldest ways to judge the condition of the heart. This paper shows the outcomes from a set of extracted features of Heartbeat sound by applying the classifier Naïve Bayes, Neural Network, Decision Tree, SVM, Logistic Regression and Nearest Neighbor. Experimental results show that SVM carried the highest accuracy (i.e., 76%) for normal and abnormal heartbeat classification, ANN (i.e., 83%) for normal and murmur classification and Nearest Neighbor (i.e., 73%) for normal and extrasystole classification compared to other machine learning algorithms .This research includes comparing the results from all this algorithms and finding the best possible set of data and algorithms. This machine learning technique contributes to the development of heart disease related researches and developing more efficient machines to detect heart diseases accurately in short time.

Keywords: Classification; Features Extraction; SVM; Heart Beat; Decision tree;Logistic Regression ;Naïve Bayes;PCG;Neural Network;Nearest Neighbor.

Dedication

Our beloved parents and honorable Supervisor Dr.Jia Uddin for their endless support and patience.

Acknowledgement

Firstly, we are grateful to Almighty for our good health and well being that were necessary to complete this thesis and always guiding us to the right path. We wish to express our sincere gratitude to Dr. Jia Uddin, our thesis supervisor for providing us with all the necessary facilities that was required for our research. We would also like to thank Dr. Jia Uddin for his sincere and valuable guidance and encouragement. Moreover, we would also like to thank all the Faculty of BRAC University for guiding us throughout the study period. Finally, we would like to thank our loving parents for their constant support and help to make this research successful.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

ANN Artificial Neural Network

CV Cross Validation

$MFFC$ Mel Frequency Cepstral Coefficient

PCG Phonocardiogram

PPV Precision

$STFT$ Short-Term Fourier Transform

SVM Support Vector Machine

TPR True Positive Rate

Chapter 1

Introduction

One of the major kind of diseases is Cardiovascular diseases and have a significant contribution to health problems. The heart is the organ that keeps individual alive.Thus it is really important to have a healthy heart. It is not easy for a non-medical person to understand his or her heart condition. For the developing countries the scenario is worst. Sometimes patient get diagnosed lately for lack of physician or diagnostic devices.Cardiovascular diseases are critical and should be detected without making any delay [17].In this situation technology can be a very big solution provider. This alarming problem can be reduced with the help of phonocardiogram signals. A phonocardiogram or PCG is a sound and murmuring recording of the heart by phonocardiogram. In medical heart sound is very important entity. For the diagnosis of valvular heart disease, analysis and characterization of the PCG signal are important. The PCG records heart sounds, sounds and other sounds [23].The normal heart sounds are classified in the order of the first, second, third and fourth heart sound by Phonocardiogram test. [20]. The PCG signals can be used in digital signal processing to do research with the heartbeats. There is a big difference between normal cardiac audio signal and irregular cardiac audio signal as their PCG signal differs. By elucidating heart sound data, heart condition is known.However, early detection can be really effective for the treatment as well as for the patient's health. For this reason, the advantages of putting the primary discovery capacity in their grasp of the individual are tremendous. Our study is to make our idea into reality where we have extracted 193 features from the heartbeat sounds and applying machine learning algorithms to identify and classify those audio files of audio that were collected from both way digital stethoscopes and mobile devices.

1.1 Motivation

According to World Health Organization 85% of global deaths are due to heart attack and stroke . In Bangladesh every year there are 5,80,000 deaths occurs from these non-communicable diseases among them heart disease is a major cause of death [32]. More than 3/4 of CVD deaths occur in low- and middle-income countries. An effective method is needed to diagnose heart disease prematurely so people can avail early diagnosis and treatment. The heart test, ectrocardiogram (ECG) and echocardiogram, are two effectively investigated [27]. While both are effective both methods are fairly costly. The other method for examining heart condition is a balancing procedure requiring specialists from medicine and cardiology. So, in this study numerous research on developing an automatic system for detecting abnormal heart sound already conducted to provide a system to any common people [32]. Main thought of this thesis work is to detect normal and abnormal condition of human heart from heartbeat sound based on different features of audio signals using different algorithms. In this study, to predict cardiovascular disease, the comparison of assorted machine learning techniques like support vector machine, Decision Tree, Naive Bayes, Nearest Neighbor, Neural Network and Logistic regression. These algorithms show different accuracy. Therefore we decided to work on classifying the condition of the heart and which algorithm performs well in this system.

1.2 Thesis Orientation

All methods and process used for thesis work are organized as follows:

1. Chapter 2 includes the previous works done by the other researchers.

2. Chapter 3 represents the background analysis on the heart sound, basic heart sound types and abnormal heart sound.

3. Chapter 4 presents the dataset information that used for our study, proposed model and workflow. This chapter also describes feature extraction and algorithms used for classification purpose.

4. Chapter 5 demonstrate the overall result analysis and the discussion in terms of the obtained result.

5. Chapter 6 concludes the thesis work with conclusion and includes the future research direction.

1.3 Objective

The main objectives of this study are briefly given bellow:

1. To design a model that extract the features of sounds and classify normal, abnormal sounds by using different algorithm

2. To extract features STFT is used and extracted features are MFFCs, Chroma, Tonnetz, Mel-spectrogram, Contrast were used.

3. To classify and calculate the accuracy SVM, Nearest Neighbour, Naïve Bayes, Neural Network, Decision tree and Logistic Regression algorithms are used

Chapter 2

Literature Review

In recent years, researcher have made significant progress on identifying different diseases. There are several studies have been conducted for classifying various heartbeat sound to detect the abnormal heart sound. In a model, Guraksin and Uguz used Least Squares Support Vector Machine to classify the sounds [13]. The study was carried on a add up to 120 heart sounds and the sounds were inspected in three categories: metral stenosis, pulmonary stenosis and normal sounds. They used DWT to obtain features from heart sound .These sound signals obtained were separated into sub bands by utilizing DWT. The aim was to select fewer features to represent the information set instead of large number of features gotten by DWT. Then the dimension reduction stage , the entropy of each sub-band was computed by utilizing shannon entropy algorithm to diminish the dimensionality of the include vectors through DWT. In the stage of classification, LS-SVM method was used.
Ryu et al. established a CNN based classification model that can identify normal and abnormal heart sounds [25]. This model states whether the heart sound recordings are normal or abnormal as cardiac diagnosis. By utilizing windows -sinc Hamming filter algorithm, filtering process removed signals regarded as noise from the sound that was recoded. Besides the filtered sound s are recorded. The scaled sound recordings are at that point separated into numerous fragments. Then in the segment classification process firstly extracts feature from each segment and then extracted features are used to decide whether each signal section is normal or abnormal by utilizing CNN. Finally recording classification prepare decides whether the total sound recording is normal or abnormal based on classification comes from the sectional recordings. By using the proposed demonstrate, they have accomplished an in general score of 79.5 with a sensitivity of 70.8 and a specialty of 88.2 in “CinC Challenge 2016” official stage.
Sing et al. in another study proposed a model to see the efficacy of features and different classifiers were used to differentiate normal from the abnormal heart sounds [15]. A phonocardiograph dataset of 60 signals where 30 are normal and 30 murmur signals are used. This research has drawn together up to of 23 features from time, frequency, cepstrum and statistical features that can be used to separate the normal and murmur sounds. They usually assumed that systole would be shorter than diastole. In order to deal with the systolic and diastolic murmuring condition, features of the entire signal, as well as independently were extracted in systolic and diastolic locations. During the feature selection phase, 5 optimal features were choosen and used with the accuracy of 93.33 for the classification. Bayes Net, Naïve Bayes,

SGD, Logit Boost were used as classifier and among all of the classifier Naïve Bayes selected as the best classifier for this model.

Ghiasi et al. proposed a nonlinear cardiac sound evaluation of heart disturbances with recurrence quantification analysis in another model[28]. The work has proposed an effective technique for extracting features and classification to enhance identification with cardiac sounds by Mitral Valve Prolapse (MVP) and Coronary Artery Disease (CAD) of two specific types of cardiac rhythm disease. They used butterworth filtering for preprocessing of PCG signals. Then, heartbeats are separated with state of the art segmental algorithms from preprocessed signals. Beats that are less noisy are chosen using a time and frequency algorithm combination. Recurrence Quantification Analysis is used to describe the dynamics of every abnormality. Then the resulting characteristics are converted into a new space using Fisher's discrimination analysis (FDA), which is a technique for dimensional reduction. For classification new dimension-reduced feature vector is used to determine the normal and abnormal class of the input of the fuzzy C-means clustering (FCM). Then 10 dimensional RQA feature vector is fed into the pattern recognition artificial neural network to classify the CAD and MVP recording. In the test data set for the classification of CAD from MVP recordings the sensitivity of 0.853, specificity of 0.844 and accuracy of 0.848 are obtained.

Wibawa et al. used Conventional Neural Networks to detect anomalous cardiac rhythm based on a cardiac sound spectrogram [32]. 351 normal sound data is used and 129 abnormal sound data are transformed into two dimensional image in this research. Heart beat sound data in two-dimensional forms are represented by a spectrogram. For Classification of abnormal sound CNN is used. Here CNN is set up with 2 convolution layers and 2 pooling layers. The values of 25, 50, 75, 100, 125, and 150 were checked in various periods on CNN. At the 50th epoch the highest results were obtained. The accuracy for training was 82.75 and testing accuracy rate was 82.83.

Yaseen et al. used PCG signals to classify heart sound using multiple features [34]. In this experiment,few steps are followed to complete the whole process.These steps are collection of heart sound, remove the noise, sampling the PCG signals, extract features, training data and classification. MFCCs and DWT, two different types of features were extracted from the heart sounds. Support Vector Machine, centroid displacement based k-nearest neighbors and DNN were used to classify the features as classifiers. To evaluate the performance of features by using these classifiers, the average F1 score and accuracy is calculated for all of the three features that are MFFCs, DWT , combination of MFFCs and DWT by utilizing the classifiers SVM, DNN and centroid based KNN. A normal data set and another abnormal data set are used to extract the features. This experiment can be categorized into nine types. First three experiments were performed by using SVM classifier and three types of features MFCCs, DWT and combination of MFCCs and DWT, another 3 experiments by utilizing DNN. After that 3 experiments conducted by utilizing centroid displacement based KNN. Then all of the features are classified by using classifier.

Chapter 3

Background Analysis

3.1 Heart

Heart is mainly the central part of the circulatory system. It constantly pumps blood and distribute oxygen as well as nutrients throughout the body. These functions are so much important that heart is considered as the most important organ. Visualization of anatomy of the human heart and representation of the spatial characteristics of heart in the chest is beneficial for pathology research, presurgical planning and clinical approach [18]. Heart is a vital organ of our body. It is a muscular organ consisted with special type of muscle. The regarding muscle is called 'cardiac muscle'.

The heart consists of four chambers . This chambers are two atria and two ventricles. Blood is collected from the lungs by the right atrium and pumped into the right ventricle. Blood is collected from the right atrium by the right ventricles and then it pumps them into the lungs where oxygen is charged. The left atrium receives oxygenated blood from the lungs, afterwards pumps it to the left ventricle. Lastly comes the left ventricle which is the strongest chamber. It pumps oxygen-rich blood to the rest of the body. The left ventricles' vigorous contractions create our blood pressure. The preponderant methods of inspecting the heart condition are the ECG and the ultrasound; they are though compared to auscultation, less cost-effective and with more hardware demands [17].

3.2 Heart Sound

The sound of heart beat is caused by the opening and closing of the heart valves as they pump blood for the body. Blood can move in one direction only when the heart is working properly. The valves make this possible by opening and closing in exact coordination with the heart's pumping action .The valves works like doors as they open to allow blood to move in one direction and close to keep blood from backing up. Heart sounds are produced from this specific cardiac event such as closure of a valve or tensing of the chordae tendineae. When a stethoscope is placed on the chest over different regions of the heart, there are four basic heart sounds that can be heard. Listening to heart sounds is called cardiac auscultation. The sounds waves responsible for heart sounds (including abnormal sounds such as murmurs) are generated by vibrations induced by valve closure, abnormal valve opening, vibrations in the ventricular chambers, tensing of the chordae tendineae, and by turbulent or abnormal blood flow across valves or between cardiac chambers. Heart sound is one of the key signals from which the mechanical functional state of the heart can be assessed. Moreover, the heart beat sound directly relates to the variations in pressure during the heart cycles. It also relates to the operation of the heart valves as well as the elasticity of the heart tissues [5]. In children the ratio between heart and respiratory sounds is still greater for which heart sounds are particularly difficult to manage correctly for them [4].

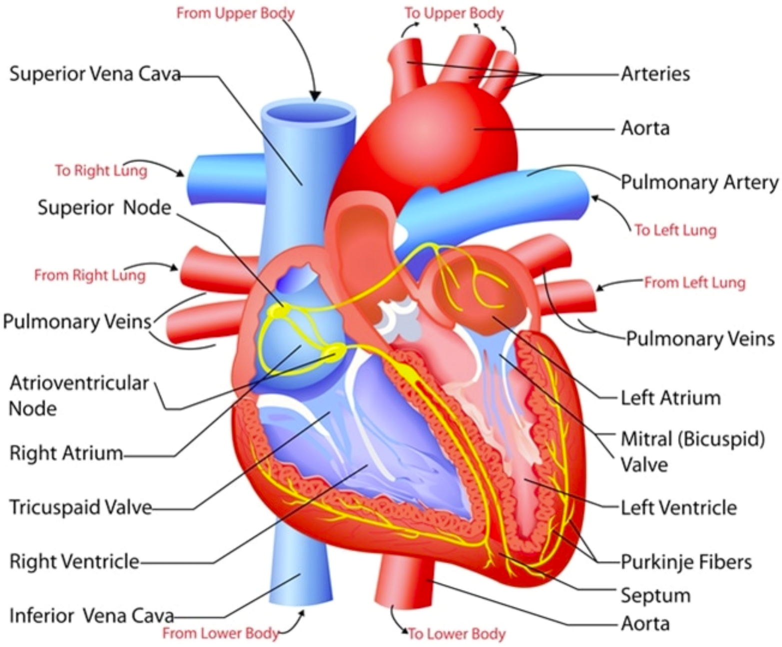

Figure 3.1: Cardiovascular system

3.3 Basic Heart sound Types

The most basic heart sounds are first and second sounds which are normally called as S1 and S2. A normal heart sound has a “lub dub, lub dub” arrangement [29]. The “lub” is the first heart sound, commonly termed S1. S1 is caused by turbulence which is caused by the closure of mitral and tricuspid valves at the start of systole. The second sound,” dub” or S2. This is caused by the closure of aortic and pulmonary valves, marking the end of systole. The heart sounds like lub-dub” [22].

S1 sound is related with the closure of the atrioventricular valves during systole. When the right and left ventricles contract (systole), the tricuspid and mitral valves are forced closed producing S1. Normally it is slightly split. The heart sounds which are produced by the closure of the mitral valve is named as M1. On the other hand which sounds are produced by closure of the tricuspid valve is named as T1. The M1 sound is much louder than the T1 sound. The difference causes for higher pressures in the left side of the heart. M1 radiates to all cardiac listening posts. Whereas T1 is usually only heard at the left lower sternal border part. This is the reason behind making the M1 sound louder the main element of S1

S2 (dub) corresponds with the closure of the semilunar valves during diastole. It is created with the closing of both aortic and pulmonic valves in diastole. It is physiologically split.S2 can provide a lot of useful clinical information which helps to find and diagnose cardiovascular diseases. In this case of S2 the regarding sound produced by the closure of the aortic valve is termed A2. And the sound produced by the closure of the pulmonic valve is named as P2.

S3 and S4 are the pathologic heart sounds. The origins of the S3 sound are still controversial [4]. The most accepted and popular theory is the ventricular theory. This theory states that it is originated from the ventricular compliance related rapid deceleration of the early transmitral flow and the associated vibration of the entire cardiac-blood pool [22]. In addition with the theory, S3 is associated with volume overload heard during rapid ventricular filling. This heart beat sound is also known as the "ventricular gallop". The S3 sound is mainly produced from the large amount of blood which is striking a very compliant LV. S3 sound is usually a normal finding in children, pregnant females as well as in a well-trained athlete.

The most basic heart sounds are first and second sounds which are normally called as S1 and S2. A normal heart sound has a "lub dub, lub dub" arrangement [28]. The "lub" is the first heart sound, commonly termed S1, and is caused by turbulence caused by the closure of mitral and tricuspid valves at the start of systole. The second sound," dub" or S2, is caused by the closure of aortic and pulmonary valves, marking the end of systole. The heart sounds like "lub-dub" [22].

S1 (lub) is related with the closure of the atrioventricular valves during systole. When the right and left ventricles contract (systole), the tricuspid and mitral valves are forced closed producing S1. Normally it is slightly split. The heart sounds which are produced by the closure of the mitral valve is termed M1. On the other hand which sounds are produced by closure of the tricuspid valve is named as T1. The M1 sound is much louder than the T1 sound. The difference causes for higher pressures in the left side of the heart. M1 radiates to all cardiac listening posts and T1 is usually only heard at the left lower sternal border. This is the reason behind making the M1 sound louder the main component of S1.

S2 (dub) corresponds with the closure of the semilunar valves during diastole. It is produced with the closing of both aortic and pulmonic valves in diastole. It is physiologically split. S2 can provide a lot of useful clinical information which helps to find and diagnose cardiovascular diseases. In this case of S2 the sound produced by the closure of the aortic valve is termed A2 and the sound produced by the closure of thepulmonic valve is termed P2.

S3 and S4 are the pathologic heart sounds. The origins of the S3 sound are still controversial [4].The most accepted theory states is the ventricular theory, which states that it is originated. From the ventricular compliance related rapid deceleration of the early transmitral flow and the associated vibration of the entire cardiac-blood pool [22]. In addition with the theory, S3 is associated with volume overload heard during rapid ventricular filling. This heart sound is also known as the "ventricular gallop". The S3 sound is mainly produced by the large amount of blood striking a very compliant LV. S3 sound usually a normal finding in children, pregnant females as well as in a well-trained athlete.

S4 is associated with a stiff ventricle heard in late diastole with atrial contraction. It also known as the "atrial gallop". This heart sound is almost always abnormal. S4 heart sound occurs during active LV filling when atrial contraction forces blood into a noncompliant LV [4].

Heart sounds can be described by their intensity, pitch, location, quality and timing during the cardiac cycle. The A2 sound is normally much louder than the P2. It is louder because of the higher pressures in the left side of the heart. So A2 sound is the main component of S2.

3.4 Abnormal Heart Sounds

Changes in rhythm can be a signal of abnormality in heart function. Blood can flow abnormally through the heart cause of many reasons including defective valves, congenital heart disorders and anaemia.

There may be gallop where an extra sound causes which involves additional heart sounds S3 and S4. This is not pathological but it can signal heart failure. Other sounds occur from various heart conditions. Clicking and snapping noises are usually generated from abnormally stiff valves opening and shutting. There are the sloshing and rumbling sounds which are called "heart murmurs". A heart murmur is an unusual sound heard between heartbeats. There are two kinds of heart murmurs innocent and abnormal[28].

It is important to remind that a 'noisy' heart is not always a sign of disease or malfunction. Many children have 'innocent' heart murmurs and that don't require any treatment or observation. Heart murmurs are caused by the velocity and turbulence of the blood when it flows through the heart. Heart murmurs are soft sounds. There are systolic and diastolic murmurs. Systolic murmurs occur during the time of contraction. Incompetence in the same valve regarding systole allows some blood to flow back into the ventricle cause diastolic murmur.

Chapter 4

Proposed Model

Most of the early researchers classified normal and abnormal heartbeat sounds using different features of audio signals. In our study we have used distinct features of audio signals for the purpose to distinguish between normal and abnormal heart sounds.

An audio signal is a representation of sound which is vibration. It is created by the movement of source of the sound. Sound is a signal because there is information in these vibrations. The vibrations force the air to form waves of energy, which fly at about 340 metres, each second of the audio signals.

Preprocessing describes any kind of raw data processing to prepare for another treatment procedure [11]. It is basically changing the raw data in an intelligible format. Data pre-processing is an integral step in the machine learning process as the quality of data and the useful data deriving from which our model can directly learn .

Feature extraction is a very important part of the analysis and finding of relationships between various subjects. The audio data supplied by the models can not be directly understood. Features extraction is used to convert them into a comprehensible format. It is a process that explains most information, but in a comprehensible manner. Classification, prevision and recommendation algorithms require feature extraction.

In the classification of machine learning, the computer program learns from the data it receives and uses this learn to classify new observations. This data set can be simply bi-class (as if it is a male or female human, or spam or non-spam), or multi-class, too. There are some examples of issues with classification: speech recognition, writing recognition, bio-metric identification, text classification, etc .

A common practice in applied computer education is the selection of machine learning approaches and a final model. The output of various machine learning algorithms is important to consistently compare. Working on a learning machine, often you have to choose several good models. The different performance features of each model are different. Working with a new data package, it is a good idea to view the data from different perspectives using different techniques. For model selection the same idea applies. It is best practice to use several ways to evaluate the estimated precision of machine learning algorithms to select one or two to finalize.

Figure 4.1: The diagram of the model

Finally using those features we have applied different algorithms to compare the results. However, we have removed noise from the audio signals. Figure 4.1 shows the steps included in this studies.

4.1 Workflow

For this study, we collected raw heartbeat sound. Audio sound format with Wav file are the data used in this study. Python libraries are used in our study for preprocessing, feature extraction and classification. These Python libraries are as follows:
librosa - loading audio files and feature extraction
noisereduce - removing noise from the audio
numpy - Mathematical operations on the extracted features
matplotlib - Plotting
sklearn - used to scale the dataset (feature extracted)
Then utilized six different Machine Learning algorithms, train them with the dataset and calculate accuracy

We separated data into normal and abnormal, murmur and normal, extrasystole and normal and completed three experiment. Figure 4.2, shows a flowchart where the working procedure of our proposed model is given briefly.
FIrstly, we considered all disease heartbeat sounds are abnormal valued 1 and normal heartbeat sounds valued 0. For the purpose of preprocessing, we had to remove all the audio files that have some static noise. Then we extract the features of the data by using STFT. The STFT is one way for time-domain signals to be converted into the time frequency domain. To transform a time domain signal into a time frequency domain and trace the lower energy cycles in the original signal, STFT can be used. Those segments were removed from the original signal to get a newly modified signal with the higher energy cycles[9]. STFT can be used to measure the amplitude of different frequencies that play at a particular time in an audio signal. Instantaneous characteristics are analyzed with the STFT approach since it can provide information that can be lost when evaluating the frequency features as a whole [33]. We scaled and normalized the data, trained them through six different algorithms using cross validation and calculate the average accuracy. Secondly, we took only the Murmur and normal heartbeat sound. We considered all Murmur heartbeat sounds value 1 and normal heartbeat sounds value 0. We remove the static noise and extract features using STFT. Then trained them with those six separate

algorithms and calculate the average accuracy. Lastly, we took the Extrasystole and Normal heartbeat sounds and did the exact same process. We considered all Extrasystole heartbeat sounds value 1 and normal heartbeat sounds value 0.

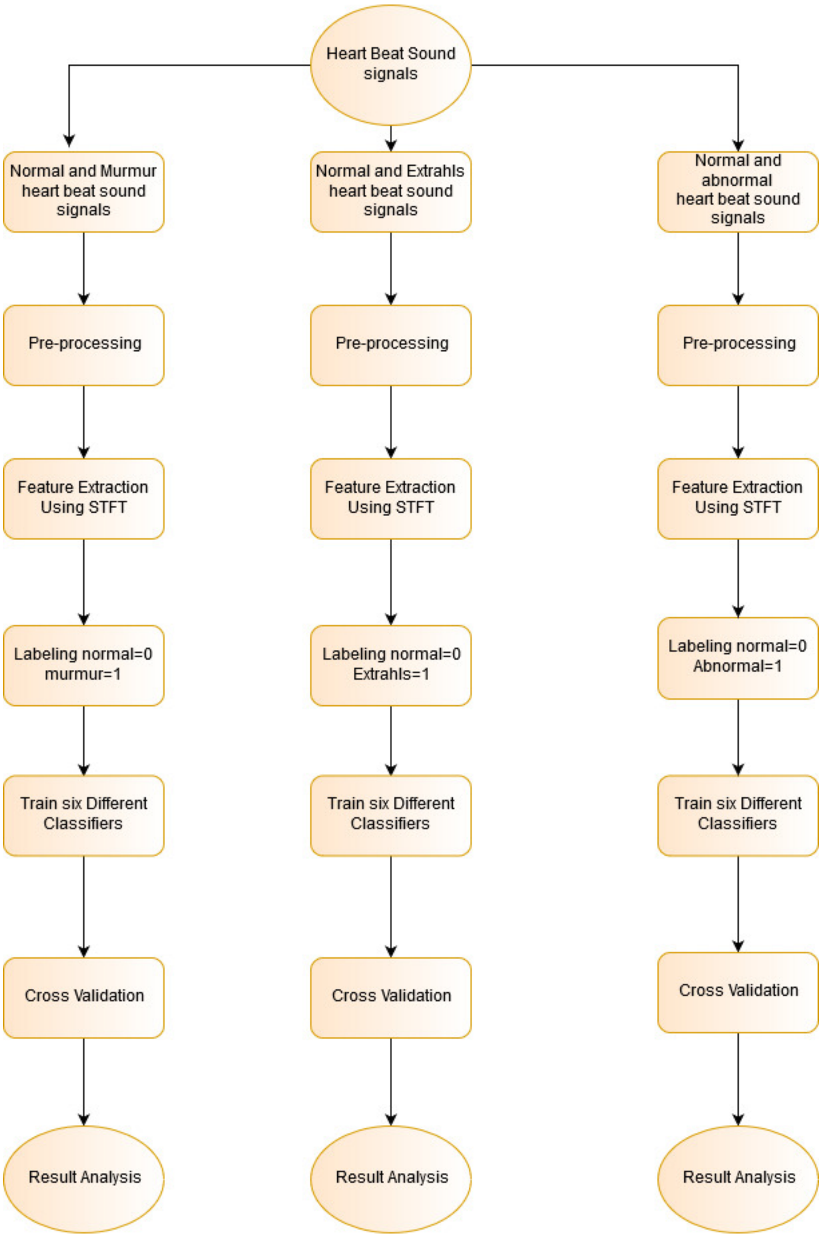

Figure 4.2: Diagram of the workflow

4.2 Data set Description

For our research we got the data set from the website www.Kaggle.com and the data set was originally for a machine learning challenge to classify heart beat sounds

[37]. To classify heartbeat sounds, two different kinds of data sets are used. The information was gathered from two sources. Information from the general public was collected in Dataset A using the iStethoscope Pro iPhone program, and in Dataset B, information were collected from a clinical trial using the Digi Scope wireless stethoscope. Data Set A comprises 176 wav files, divided into four different classes. These classes are normal, murmur, extra heart sound and artifact. There are three classes for dataset B with 656 samples in wav format. Three classes are normal, murmur, Extrasystole. The sound recordings vary between 1 second and 30 seconds of different lengths. These two different datasets are used to feature extraction and classification purpose.

4.3 Preprocessing

The primary work before feature extraction was removing the audio signals which contain no heart beat sounds. In dataset A , 84 audio samples and for the dataset B, 455 audio signals were used. Then few samples were cut from the frequency range in the last from 3500 hz data frame because those were containing the noises. We have used spectral gating for overall noise reduction for the audio clips. The spectral gating follows the following steps-

1. The noise audio clip is an FFT calculated.
2. The statistics on the FFT of the noise (frequency) are calculated
3.Based on the sound statistics (and the desired algorithm sensitivity), a threshold is calculated
4. An FFT is determined by the signal
5.Through comparing the signal FFT to the threshold, a mask is calculated
6. The mask is filtered over time and frequency
7. The mask is applied to the signal's FFT and reversed

In Figure 4.3, shows the signal with the noise. Then the Figure 4.4, shows the noise spectrogram of the heart beat sound and Figure 4.5,4.6,4.7,4.8 and 4.9 are the steps of frequency series,Mask, Masked signal,Spectrogram respectively.

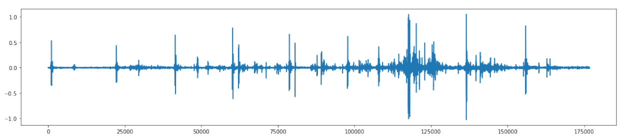

Figure 4.3: Heart beat sound signal with noise

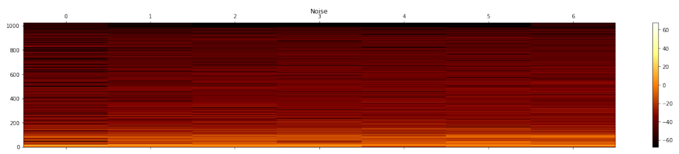

Figure 4.4: Noise spectrogram of heart beat sounds

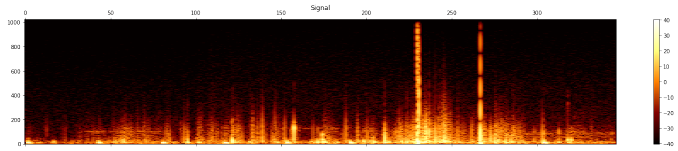

Figure 4.5: Frequency Series

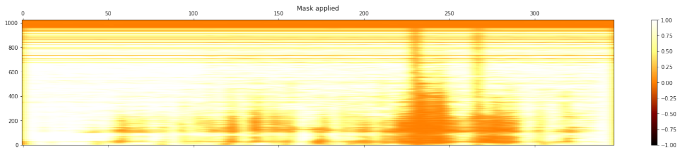

Figure 4.6: Mask (used for noise)

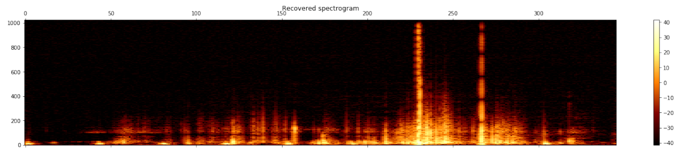

Figure 4.7: Masked signal

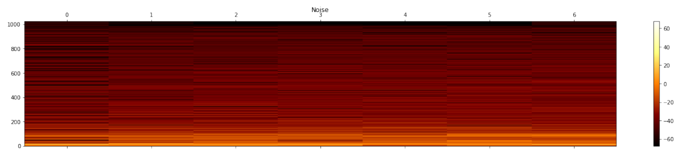

Figure 4.8: Spectrogram

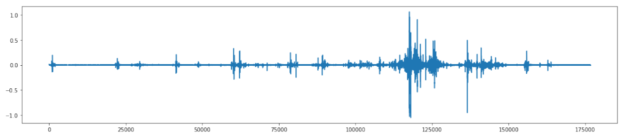

Figure 4.9: Signal after noise reduction

4.4 Feature Extraction

Feature extraction is an essential phase in analyzing and discovering relation between various things. The audio signal is not immediately understood by the models, and an understandable structural function extraction is applied over them. Feature extraction is needed for the purpose of classification,

In this study, total of 193 features were extracted by using STFT. The features extracted are MFCCs, Chroma, Mel-spectrogram, Contrast, Tonnetz. For MFFCCs 40 data points, Chroma 12 data points, Mel-spectrogram 128 data points, contrast 7 data points, Tonnetz 6 data points were found. Then by joining all the columns together, make a 193 column data frame for each audio file. Exactly same operation was done for each file and concatenates together and make a (x, 193) matrix. All data features were extracted by using a python library, librosa.

4.4.1 STFT

STFT is a sequence of Fourier transforms of a windowed signal. As STFT overcomes the lackings of Fourier Transform it is being used for audio signals feature extraction. STFT overcomes the shortfall in time by taking an application window with a particular time-frequency resolution property. In STFT, a window that could be an operate being zero-valued outside of some chosen interval, is utilized for extracting time info. The STFT increases the symbol by operating a window and the product outside the window is thus valued at nul. As a results of this product, a sub-signal that is found beneath the window is left and FFT of this sub-signal is calculated. Later, the windows Center is shifted and a brand new portion of first signal is processed leading to the subsignal frequency content left below the shifted window. This shifting method is applied until the top of original signal and by this fashion the entire FFT illustration of the signal is obtained [30].

The Short-Time Fourier remodel (STFT) is aimed at interrupting the signal into possibly overlapping frames using a moving window technique and measuring the DFT on each frame. Consequently, the STFT falls within the short-run process techniques class.The size of the moving frame, as already mentioned, plays a very important role as a result.

4.4.2 MFCCs

In automatic speech and speaker recognition or processing the mel-frequency cepstrum (MFC), based on a linear transformation of log power spectrum on a nonlinear mel frequency scale, is an illustration of a short-term power spectrum. The features can be extracted using the MFCCs method and the features extracted with that method show the short time spectrum. These features are widely used for analyzing sound processing tasks. In this method, frequency bands on the Mel scale are equal. It is an accurate approximation of the human voice [29]

Normally, MFCCs are extracted accordingly[14]:

1. Take the transformation of the Fourier signal (windowed extract of)
2. In order to map the above power of the spectrum in the mel scale, triangular overlapping windows are used.
3. Take the power log for each mel frequency.
4. Take as if it were a signal the discrete cosine transformation from the mel log power list.
5. The MFCCs are the amplifications of the result.

4.4.3 Spectral Contrast

Spectral peak, spectral valley and their distinction is considered by spectral contrast in each frequency sub band. Spectral contrast feature can differentiate between the different sound signals. The spectrogram S frame is divided into the subbands for the measurement of spectral contrast [3]. The energy comparison is calculated by comparing the mean energy for each subband in the highest quantile (peak energy) with the lowest quantile (valley energy). High contrast values usually lead to clear signals with small bands and low contrast values to broadband noise.

4.4.4 Chroma

Chroma based features is used for audio signals where the pitch class are used for audio classification.A powerful music analysis method whose pitches can be classified meaningfully and its the frequency interval between each equivalent pair of notes is the same.One way of extracting chroma features by using short-term Fourier transforms in conjunction with binning techniques can be done.The term chroma feature closely relates to the 12 different classes of pitch.One of the chromagram's main characteristics is that they capture theelogi candharmonic characteristics of music 14 while being robust and agile for instrumentation and timbre changes [8].

4.4.5 Mel Spectogram

A spectrogram is a visual representation of the spectrum of sound or other signal frequencies as they vary over time. When using a nonlinear mel frequency scale, we get the Mel Spectrogram. A mel scale is a pitch scale to be at a distance equal to each other. The acoustic time-frequency representation of the sounds is an object of type Mel Spectrogram: the power spectral density P(f, t).It is sampled at several points around the same time ti spaced and frequencies fj [36]

The mel frequency scale is defined as:

$$mel = 2595 * log10(1 + hertz/700), \tag{4.1}$$

4.4.6 Tonnetz

Tonnetz is a conceptual lattice diagram in musical tuning and harmony. The diagram represents space. It was Leonhard Euler who first described the concept of Tonnetz in 1739. There are various representation of Tonnetz. It is a very handy and versatile tool to help musicians. It helps them to understand and visualize the relations between different musical notes or pitches.

Tonnetz can be useful in quickly looking up what notes are there in a specific key or scale. It can be helpful in finding the relatively small or major key. It has amazing effect on analyze and making sense of code progression. Harmonic motion between chords can easily be accomplished by moving around the Tonnetz in three different ways. This ways are P, R, and L. The P transformation is the one which exchanges a triad for its Parallel. The next one is R transformation. It exchanges a triad for its relative. The last one is L transformation. Lastly it is the one which exchanges a triad for its Leading-Tone Exchange.

4.5 Classification

The algorithms of learning the machine are intended to learn and predict data in a different way from static algorithms requiring express human instruction [19]. These algorithms are widely used for the classification purpose by the researchers. In similar classification tasks, the choice of individual classifier was based in their performance. In this study, we have used six different machine learning algorithms for classification. These algorithms are Nearest Neighbor, Neural Network, SVM, Logistic Regression, Decision Tree.

4.5.1 Nearest Neighbor

Among the different methods of classification, one of the easiest solutions is the nearest neighbor and accomplishes reliably high performance. It classified new sample by calculating the distance to the nearest neighbor. K-NN is a simple measurement that saves the data on a similar basis, then classifies all accessible cases.

The k-nearest neighbor algorithm is a non-parametric method used in master learning for classification and regression. The input consists of the k closest examples of training within the functional area, and the output belongs to a class. An object is classified as the object by a majority of its neighbours, the class most common among its k neighbors being assigned the object. In the nearest neighbor algorithm k there is a positive, usually small integer[24].

4.5.2 Neural Network

Neural networks are an integral part of machine learning where the algorithm motivated by the structure of the human brain. They take data, train them to identify the pattern in these data and predict a new set of similar data for performance. Neural network used for cluster and classification. For analyzing and learning data, they comprise of different layers. Every layer tries to distinguish patterns of the data and whenever a pattern is recognized the following hidden later is enacted and

so on.
The different number of neurons may consist of ANN. The sizes of ANNs, namely the number of neurons, range in chemical applications from tens of thousands to less than ten (1-3). The neurons in ANNs can all be put into one or two layers, neurons can form three or even more layers [1].

4.5.3 Logistic regression

Logistic regression is another method borrowed from the statistics field through machine learning. Logistic algorithm provides the outcome provides only two values, either a 0 or 1 [26].Like other forms of regression analysis, logistic regression makes use of one or more predictor variables that may be either continuous or categorical. Statisticians developed the logistical function, often referred to as the sigmoid rule, to describe growth properties in ecology, rising rapidly and maxing out at the setting's carrying capacity. It is a s-shaped generated curve that can take any real-evaluated range and map it to a price between zero and one, but at certain limits in particular. A logistical regression explanation may begin with an explanation of the standard logistical method. The logistical function is a sigmoid function that takes every real value from zero to zero.

4.5.4 Decision Tree

Decision Tree is an algorithm that works like the human brain. The algorithms used in machine learning are wide spread. Sometimes it is also used for regression problems and sometimes for classification. Decision tree works like human brain. We can call it as smart algorithm. Every time we make a question before taking any decision. For example: Is it healthy for the health? If no it will take a decision and if not it will go for another decision. Decision tree is simple to understand, interpret and visualize. Moreover the effort of data preparation is little. It can handle both numerical and categorical data. And nonlinear parameters don't affect its performance. This advantage makes decision tree a very useful algorithm. But it has some disadvantages too. When this algorithm captures noise in the data over fitting occurs. Sometimes the model can get unstable due to small variation in data. In terms of discussing about decision tree it is a must to know about classification. Classification is a technique of categorizing the observation into different category. So basically it is about taking the data, analyzing it and on basis of some condition finally divided them into various categories. Classification is done to perform predictive analysis on data. One use case of classification is classifying fruits on the base of size, color, taste.

4.5.5 SVM

For the audio signal classification svm is also used. SVM works is based upon Statistics learning theory and structure risk minimization Principle which is known as a unique method in machine learning for signal processing [21]. SVM is a very good tool in both image processing and audio signal processing. SVM finds the optimum hyperplane for separation to maximize the distance between training data and the decision limit [10]. We choose linear kernel model for our experiment. The svm algorithm to use a hyperplane with a maximum margin to separate the data points of two classes in N-dimensional space where N is the number of the features we are considering[21].The hyperplane will decode in which category the data will be considered. Again the hyperplane dimension depends on the amount of number the features. The hyperplane is a line if the work is with two features when there are three features the hyperplane is two dimensional. Actually the number of dimensions is equal to number of features -1.

4.5.6 Naive Bayes

Naive Bayes is a classification technique which is based on Bayes' theorem from probability theory. The naive Bayes classifier can make learning simple. This can be done by assuming that features are given class and independent. Although independence is generally thought as a poor assumption. In practice naive Bayes often competes well with more sophisticated classifiers [2].
It has two main advantages. First one is, it is easy to use and another is only one scan of the training data is required for the probability generation [31]. It is easy for a naive Bayes classifier to handle the missing attribute values by simply omitting the corresponding probabilities for those attributes when calculating the likelihood of membership for each class. It also requires related class conditional independence.

Chapter 5

Results and Discussion

The chapter's primary goal is to define outcomes and their analysis based on extracted features from audio signals and used machine learning algorithms. In previous Chapters we have discussed the feature extraction and regarding algorithms. Those lead us to different results and understandings. Now we are going to discuss the results we have found by implementing the algorithms.

5.1 Accuracy of Models Applying All Features

The results have been obtained by applying different classification algorithms. Precision is one another way to calculate the performance of the machine learning models.The greater the accuracy, the more accurate the forecast. Thus precision can be used to know the predictive power of the machine learning algorithms.

The Precision, True Positive rate and accuracy of machine learning algorithms are used for performance analysis. We have shown performance of Precision, True Positive rate and accuracy in the Table 5.1,5.2 and 5.3.

$$Precision = \frac{TP}{(TP + FP)}$$

$$TruePositiveRate = \frac{TP}{(TP + FN)}$$

$$Accuarcy = \frac{TP + TN}{TP + TN + FP + FN} * 100\%$$

Where TP means true positive number, TN means true negative number, FP means false positive number, and FN means false negative number.

Table 5.1 displays the overall classification results using six machine learning algorithms where we used the data set with all features and applied Nearest Neighbor, Neural Network, Support Vector machine, Decision Tree, Naive Bayes, and Logistic Regression. In the classification of the normal and abnormal heartbeat sounds the SVM and Nearest Neighbor represents the highest accuracy and the Logistic Regression gives the largest precision number.

Table 5.1: Performance Overview of classification algorithms for normal vs abnormal

Algorithms	Precision (PPV)	Recall (TPR)	Accuracy
SVM	93	59	76
Decision Tree	81	69	63
Logistic Regression	100	59	64
Naive Bayes	96	63	68
Nearest Neighbor	81	61	74
Neural Network	67	64	66

The figure 5.1 provides the accuracy information in the form of bar graph and it shows that the SVM and ANN give better results with the system.

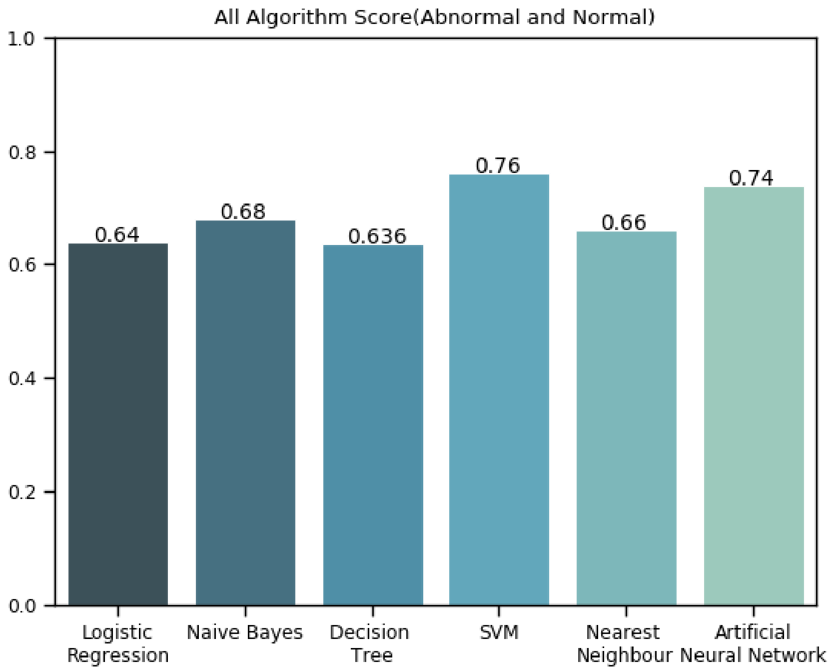

Figure 5.1: Accuracy comparison between Normal and Abnormal

For the classification between the normal and murmur heartbeat sound the largest precision represent by three algorithms Logistic Regression, Naive Bayes and SVM classifier where the accuracy with highest value is the Artificial Neural Network in the Table 5.2. In this case the Naive Bayes, SVM and Nearest Neighbour can be considered as a good classifier for this specific classification.

Table 5.2: Performance Overview of classification algorithms for normal vs murmur

Algorithms	Precision (PPV)	Recall (TPR)	Accuracy
SVM	98	71	76
Decision Tree	87	74	71
Logistic Regression	98	71	71
Naive Bayes	13	64	79
Nearest Neighbor	98	72	76
Neural Network	94	73	83

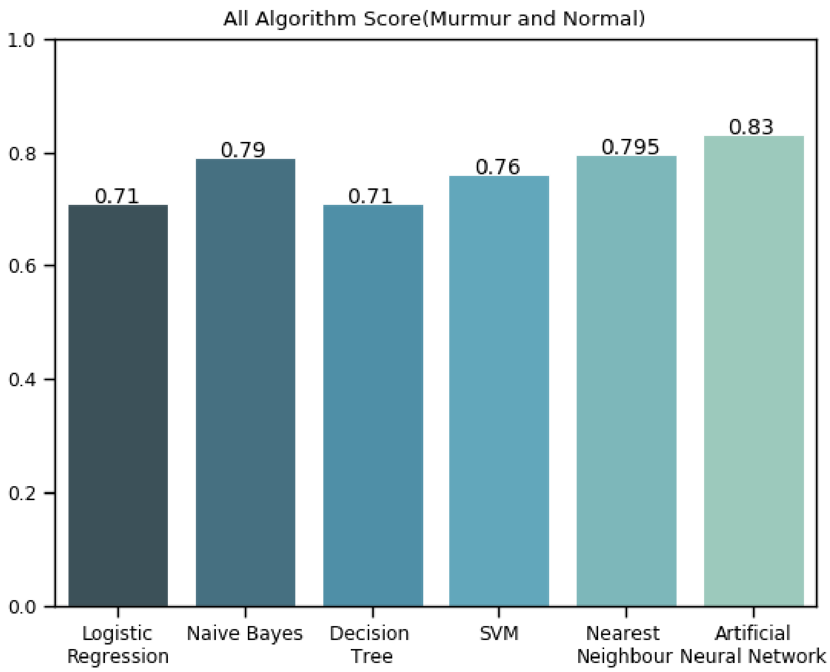

Figure 5.2: Accuracy comparison between Normal and Murmur

Table 5.3 displays classification between the normal and extrasystole heart beat sounds the decision tree and logistic regression performed better than other in case of precision. Again the Nearest Neighbor performed well with largest accuracy.

Table 5.3: Performance Overview of classification algorithms for normal vs extrasystole

Algorithms	Precision (PPV)	Recall (TPR)	Accuracy
SVM	98	71	71
Decision Tree	85	75	68
Logistic Regression	98	71	71
Naive Bayes	13	64	33
Nearest Neighbor	98	72	73
Neural Network	94	73	71

In our model, we divide our data into k subset. Then (k-1) of subset was used for the training purpose and remaining 1 subset was used to test the model. We repeated the process for all combination and we continue until each of the k is used as test the test set. Finally the average of our recorded scores was taken which is performance metric for our model.

5.2 Feature Engineering

Feature engineering is the deciding factor in machine learning whether the model is good or bad. It is an important issue. Choosing fewer features helps to train the dataset faster than expectations. So it's tricky to use the relative features in order to get more accurate outcomes. Some features are linearly related to others. Feature selection is also important as it helps to select the important features which lead to improvement of the feature's accuracy. If quality features aren't provided even the best of breed algorithms will fail and results will be bad [35].

5.3 Feature Importance

Features in machine learning plays vital role. In dataset there may be attributes which reduce accuracy level of a model. On the other hand there may be some attributes which don't even influence the results. So, when working with machine learning it is necessary to choose the appropriate attributes.

5.4 Cross Validation

Cross validation is one of the models for validating how statistical analysis results are generalized to form a single data set. This technique consists of the storage of a specific data set on which the model is not educated. After this, the model is checked on this sample before completion. Steps which take part in cross validation [16]:
1. Store a sample data set.
2. Use the remaining part of the data set to Train the model.
3. Test (validation) set using the stored sample of the data set. This will help to predict the effectiveness of model's performance.

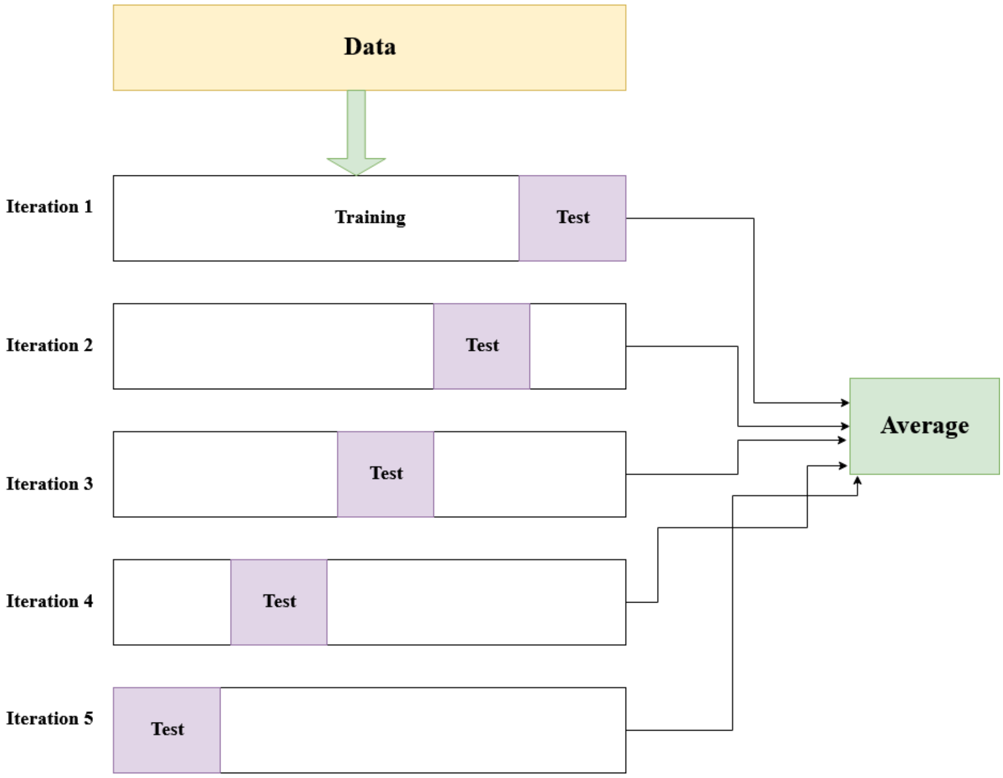

Figure 5.3: Process of calculating of Cross Validation value

5.5 Discussion

Depending on the size of data sets and feature selection the outcome may be observed with variant accuracy. Using the cross validation the results shows in the Tables 5.1,5.2 and 5.3 are classification between the normal with abnormal, normal with murmur and normal with extra systole heart beat respectively. Again from the Figure 5.1 and 5.2 the SVM, ANN and Nearest neighbor are giving acceptable results for the abnormal,murmur and extrasystole classification.

For more observation the ROC curve of SVM classifier is shown in Figure 5.4. The ROC indicates the performance of a binary classier by ploting TRP as y axis and FRP as x axis with different threshold settings.
SVM simultaneously reduces the error of empirical classification and maximizes the geometry . So SVM called Maximum Margin Classifiers [12]. SVM works relatively well when there is clear hyper plane of separation between classes. In Figure 5.5 plot shows that overlapping of multiple attributes. In some attributes overlapping is very low and in some attributes overlapping is much high. But maximum attributes distinct in someway. That is why it is easy for SVM to generate a hyperplane. In table 5.1, shows that SVM gives the highest accuracy of 76% .
Artificial neural network has ability to work with inadequate knowledge because disappearance of a few pieces of information in one place does not restrict the functions of it [6]. After ANN training, the data may produce output even with incomplete information. Murmur and normal heart sounds data is not rich enough but the Neural Network classifier gives the highest accuracy of 83% .

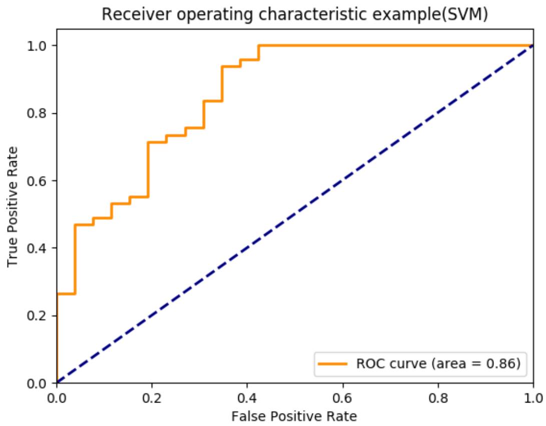

Figure 5.4: ROC curve using SVM classifier for all features

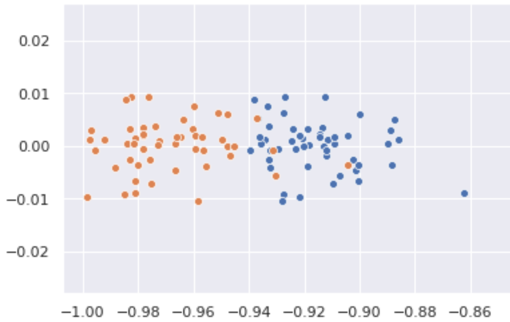

Figure 5.5: SVM classifier for the features

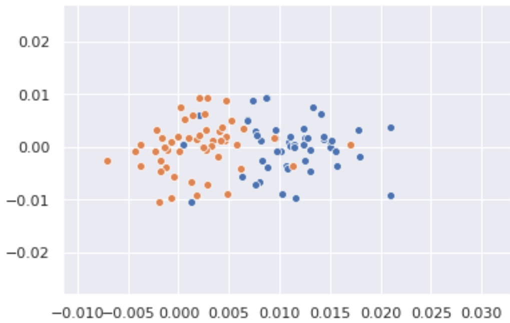

Figure 5.6: KNN classifier for the features

KNN is a classification method based on the object nearest to the object being searched [7]. It works by calculating how far the data points are from each other. Because of the attributes overlapping in our data point is low in Figure 5.6. it is easier for the Nearest Neighbor to calculate the average distance of the data points. For this reason KNN also performed very well with the accuracy of 73% which shows in the figure 5.3.

Chapter 6

Conclusions and Future Work

6.1 Conclusion

In this research, we have extracted different features from heart beat sound and measured the condition of heart by applying different algorithms. We have completed the study in a manner that we can predict if the person has heart disease or not from the calculative results acquired from collected data. In the preprossessing phase we have reduced noise from our training dataset so that we can get the maximum accuracy. Then we extracted features like Tonnetz, MFCCs, Chroma, Contrast, and Mel Spectrogram. We have divided the dataset into two section which are training and testing data sets. Using the extracted features we have calculated result from them by applying six algorithms. From the outcome of the algorithms we have compared the results in order to predict the condition of the heart. When we worked with murmur and normal heartbeat sound artificial neural network gives the best accuracy (i.e.,83%) among all the algorithms. In terms of normal and abnormal heartbeat sound we obtained most accuracy (i.e., 76%) from SVM classifier. And finally for the normal and extrasystole heartbeat sound the KNN gives the highest accuracy (i.e.,73 %). In the three models different algorithms have different performance because of their different way of approach to the features of data set. It is possible that other data sets and algorithms may gives the better accuracy. If we increase the attributes, may be we can obtain more accurate result but it will be more time consuming. The result can be improved by improving feature extraction techniques and managing the data set more technically.

6.2 Future Work

We have used a small and old data set. In terms of this research there is a lot of scope to discover various horizon of information if we work with big amount of data sets collected in valid manner. Moreover we can extract a lot of other features and apply different algorithms to compare them from them in order to get maximum accuracy.

Bibliography

[1] J. Zupan, "Introduction to artificial neural network (ann) methods: What they are and how to use them", *Acta Chimica Slovenica*, vol. 41, Jan. 1994.

[2] I. Rish, "An empirical study of the naïve bayes classifier", *IJCAI 2001 Work Empir Methods Artif Intell*, vol. 3, Jan. 2001.

[3] D.-N. Jiang, L. Lu, H. Zhang, J. Tao, and L. Cai, "Music type classification by spectral contrast feature", *Proceedings. IEEE International Conference on Multimedia and Expo*, vol. 1, 113–116 vol.1, 2002.

[4] S. Cortes, R. Jane, A. Torres, J. A. Fiz, and J. Morera, "Detection and adaptive cancellation of heart sound interference in tracheal sounds", in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2006, pp. 2860–2863. DOI: 10.1109/IEMBS.2006.260727.

[5] D. Kumar, P. Carvalho, M. Antunes, J. Henriques, A. S. e. Melo, R. Schmidt, and J. Habetha, "Third heart sound detection using wavelet transform-simplicity filter", in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2007, pp. 1277–1281. DOI: 10.1109/IEMBS.2007.4352530.

[6] E. Grossi and M. Buscema, "Introduction to artificial neural networks", *European journal of gastroenterology hepatology*, vol. 19, pp. 1046–54, Jan. 2008. DOI: 10.1097/MEG.0b013e3282f198a0.

[7] H. Parvin, H. Alizadeh, and B. Minaei-Bidgoli, "Mknn: Modified k-nearest neighbor", 2008.

[8] Wei Wang, J. Dong, and T. Tan, "Effective image splicing detection based on image chroma", in *2009 16th IEEE International Conference on Image Processing (ICIP)*, Nov. 2009, pp. 1257–1260. DOI: 10.1109/ICIP.2009.5413549.

[9] S. Abdullah, T. Putra, M. Nuawi, Z. Nopiah, A. Arifin, and L. Abdullah, "Extracting fatigue damage features using stft and cwt", *WSEAS Transactions on Signal Processing*, vol. 6, pp. 91–100, Jul. 2010.

[10] H. Huang, G. Hu, and L. Zhu, "Ensemble of support vector machines for heartbeat classification", in *IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS*, Oct. 2010, pp. 1327–1330. DOI: 10.1109/ICOSP.2010.5657034.

[11] (2010). Preprocessing techniques in character recognition, [Online]. Available: //www.intechopen.com/books/character-recognition/preprocessing-techniques-in-character-recognition/ (visited on 09/17/2010).

[12] D. Srivastava and L. Bhambhu, “Data classification using support vector machine”, *Journal of Theoretical and Applied Information Technology*, vol. 12, pp. 1–7, Feb. 2010.

[13] G. Güraksın and H. Uuz, “Classification of heart sounds based on the least squares support vector machine”, *International journal of innovative computing, information control: IJICIC*, vol. 7, Dec. 2011.

[14] M. Sahidullah and G. Saha, “Design, analysis and experimental evaluation of block based transformation in mfcc computation for speaker recognition”, *Speech Communication*, vol. 54, pp. 543–565, May 2012. DOI: 10.1016/j.specom.2011.11.004.

[15] M. Singh and A. Cheema, “Heart sounds classification using feature extraction of phonocardiography signal”, *International Journal of Computer Applications*, vol. 77, pp. 13–17, Sep. 2013. DOI: 10.5120/13381-1001.

[16] K. Choudhary and S. Wadhwa, “Glaucoma detection using cross validation algorithm”, in *2014 Fourth International Conference on Advanced Computing Communication Technologies*, Feb. 2014, pp. 478–482. DOI: 10.1109/ACCT.2014.29.

[17] M. S. Jahan and A. B. M. Aowlad Hossain, “A low cost stethoscopic system for real time auscultation of heart sound”, in *2014 International Conference on Informatics, Electronics Vision (ICIEV)*, May 2014, pp. 1–4. DOI: 10.1109/ICIEV.2014.6850764.

[18] C. D. Papadaniil and L. J. Hadjileontiadis, “Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features”, *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1138–1152, Jul. 2014, ISSN: 2168-2208. DOI: 10.1109/JBHI.2013.2294399.

[19] V. Bountourakis, L. Vrysis, and G. Papanikolaou, “Machine learning algorithms for environmental sound recognition: Towards soundscape semantics”, Oct. 2015. DOI: 10.1145/2814895.2814905.

[20] G. V. H. Prasad and P. R. Kumar, “Performance analysis of feature selection methods for feature extracted pcg signals”, in *2015 13th International Conference on Electromagnetic Interference and Compatibility (INCEMIC)*, Jul. 2015, pp. 225–231. DOI: 10.1109/INCEMIC.2015.8055885.

[21] Xiaowu Sun, Lizhen Liu, Hanshi Wang, Wei Song, and Jingli Lu, “Image classification via support vector machine”, in *2015 4th International Conference on Computer Science and Network Technology (ICCSNT)*, vol. 01, Dec. 2015, pp. 485–489. DOI: 10.1109/ICCSNT.2015.7490795.

[22] A. Chaudhuri and T. Jayanthi, “Effective s1 s2 detection system with beat track method”, in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Mar. 2016, pp. 714–718. DOI: 10.1109/WiSPNET.2016.7566226.

[23] M. İşcan, F. Yığıt, and C. Yilmaz, “Heartbeat pattern classification algorithm based on gaussian mixture model”, in *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, May 2016, pp. 1–6. DOI: 10.1109/MeMeA.2016.7533715.

[24] J. Novakovic, A. Veljovic, S. Ilic, and M. Papic, "Experimental study of using the k-nearest neighbour classifier with filter methods", Jun. 2016.

[25] H. Ryu, J. Park, and H. Shin, "Classification of heart sound recordings using convolution neural network", in *2016 Computing in Cardiology Conference (CinC)*, Sep. 2016, pp. 1153–1156.

[26] D. B. Springer, L. Tarassenko, and G. D. Clifford, "Logistic regression-hsmm-based heart sound segmentation", *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 822–832, Apr. 2016, ISSN: 1558-2531. DOI: 10.1109/TBME.2015.2475278.

[27] R. Thomas, L. L. Hsi, S. C. Boon, and E. Gunawan, "Heart sound segmentation using fractal decomposition", in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2016, pp. 6234–6237. DOI: 10.1109/EMBC.2016.7592153.

[28] A. R. Jadhav, A. G. Ghontale, and A. Ganesh, "Heart sounds segmentation and classification using adaptive learning neural networks", in *2017 International Conference on Signal Processing and Communication (ICSPC)*, Jul. 2017, pp. 33–38. DOI: 10.1109/CSPC.2017.8305881.

[29] M. Sadeghi and H. Marvi, "Optimal mfcc features extraction by differential evolution algorithm for speaker recognition", in *2017 3rd Iranian Conference on Intelligent Systems and Signal Processing (ICSPIS)*, Dec. 2017, pp. 169–173. DOI: 10.1109/ICSPIS.2017.8311610.

[30] A. Elbir, H. O. İlhan, G. Serbes, and N. Aydın, "Short time fourier transform based music genre classification", in *2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, Apr. 2018, pp. 1–4. DOI: 10.1109/EBBT.2018.8391437.

[31] I. Sarker, A. Kabir, A. Colman, and J. Han, "An improved naive bayes classifier-based noise detection technique for classifying user phone call behavior", in. Apr. 2018, pp. 72–85, ISBN: 978-981-13-0291-6. DOI: 10.1007/978-981-13-0292-3_5.

[32] M. Satria Wibawa, I. M. May Sanjaya, N. Novianti, and P. Crisnapati, "Abnormal heart rhythm detection based on spectrogram of heart sound using convolutional neural network", Aug. 2018, pp. 1–4. DOI: 10.1109/CITSM.2018.8674341.

[33] Y. Seo, B. Jang, and S. Im, "Drone detection using convolutional neural networks with acoustic stft features", in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Nov. 2018, pp. 1–6. DOI: 10.1109/AVSS.2018.8639425.

[34] Yaseen, G.-Y. Son, and S. Kwon, "Classification of heart sound signal using multiple features", *Applied Sciences*, vol. 8, p. 2344, Nov. 2018. DOI: 10.3390/app8122344.

[35] X. Ye, X. Wu, and Y. Guo, "Real-time quality prediction of casting billet based on random forest algorithm", in *2018 IEEE International Conference on Progress in Informatics and Computing (PIC)*, Dec. 2018, pp. 140–143. DOI: 10.1109/PIC.2018.8706306.

[36] (2019). Phonetic sciences, amsterdam, [Online]. Available: //www.fon.hum.uva.nl/ (visited on 01/15/2019).

[37] P. Bentley, G. Nordehn, M. Coimbra, and S. Mannor, *The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results*. [Online]. Available: http://www.peterjbentley.com/heartchallenge/index.html.