Virtual Diagnosis using the heart sound taken through a digital stethoscope to help medical devoid areas.

MSc Research Project
MSc. in Cloud Computing

Anshul Bharadwaj
Student ID: 21197911

School of Computing
National College of Ireland

Supervisor: Dr. Punit Gupta

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Anshul Bharadwaj
Student ID:	21197911
Programme:	MSc. in Cloud Computing
Year:	2023
Module:	MSc Research Project
Supervisor:	Dr. Punit Gupta
Submission Due Date:	14/08/2023
Project Title:	Virtual Diagnosis using the heart sound taken through a digital stethoscope to help medical devoid areas.
Word Count:	7370
Page Count:	19

I hereby certify that the information contained in this (my submission) is informationpertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at therear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	14th August 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placedinto the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Virtual Diagnosis using the heart sound taken through a digital stethoscope to help medical devoid areas.

Anshul Bharadwaj
21197911

Abstract

Cardiovascular illnesses (commonly known as CVDs) are the largest cause of mortality worldwide, with coronary heart disease accounting for 7.2 million deaths in 2004. About 29% of fatalities in 2004 were attributed to cardiovascular disease. If heart disease could be diagnosed earlier, it might have a major impact on global health. Therefore, it is important to create tools that can aid in this process, tools that may be utilised in both clinical settings (by digital stethoscopes) and everyday life (via mobile devices). Researchers in the field of machine learning are captivated by the difficulty of categorising audio samples and distinguishing between different cardiac states in the presence of environmental noise. Since even little changes in heart sounds may be indicative of a broad variety of illnesses, reliable classifiers are necessary for making accurate diagnoses. Despite the obvious benefits that machine learning might bring to the medical industry, its application in this area is currently underutilised. The proposed device, dubbed AI-Doctor, is said to be able to identify alterations in cardiac sound signals using a number of machine learning algorithms for early diagnosis using just acoustic data. The Mel Frequency Cepstral Coefficients (MFCC) serve as the basis for feature extraction in the extensive study that makes use of several machine learning approaches. The results showed that SVM-RBF had the highest accuracy (74.36 %), followed closely by Logistic Regression (73.5 %). The enhanced SVM-RBF model is deployed to Amazon Web Services (AWS), and an API interface is built with Flask. This makes it possible for the model to function as an effective sound signal collection model that can be easily integrated with digital stethoscopes. This groundbreaking finding has the potential to revolutionise how cardiac conditions are diagnosed. The researchers believe that by harnessing the potential of machine learning, they may dramatically advance early detection skills and therefore, medical treatments.

Keywords: MFCC, , Machine Learning, Diagnostic Systems, Sound Signals , LSTM, Random Forest, XgBoost, SVM, Decision Trees, Logistic Regression.

1 Introduction

The heart is primarily responsible for its well-known role as the body's primary pump. It may be found behind the breastbone, to the left of the centre, between the lungs. The diaphragm, a tissues wall that divides the human chest and abdominal cavities, provides structural support (Adhikari, N.C.D., Alka, A., & Garg, R., 2017; Adhikari, N.C.D., 2020). When the heart contracts and relaxes, it makes a lub-dub sound. Medical professionals may learn a lot about a patient's condition just by listening to the heart's rhythms and sounds. In this investigation, we aim to learn whether or not heart illness (or other health issues) may be detected by these acoustic cues.

1.1 Motivation

Murmurs are a sign of a significant heart issue since they are brought on by blood spilling out of a valve that isn't shutting completely. Clinicians and researchers in the medical sector utilise stethoscopes to diagnose heart murmurs and their causes. Then, these sounds might be categorised in relation to the many illnesses that could be impacting the living thing. The researchers hypothesise that by analysing these sounds, they would be able to detect a wide range of problems with a person's (or any other living thing's) health.

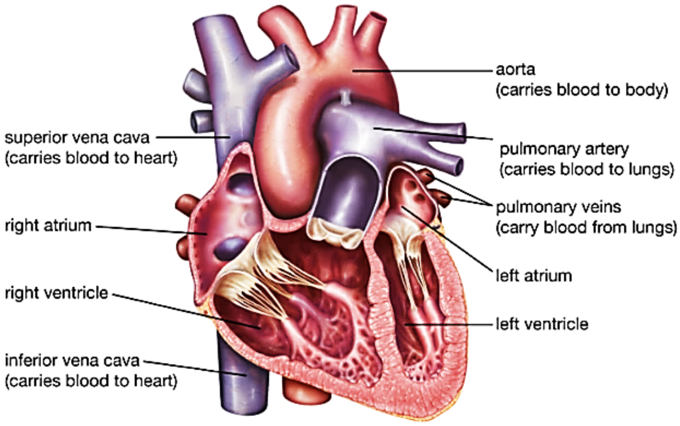

Figure 1. The Location of the Heart and connection to the different organs

1.2 Research Outcome

Using only the heart's sound signals and those from adjacent of the human body, AI-Doctor may diagnose a variety of health conditions with a high degree of accuracy. Doctors can provide comments via a mobile app when machine learning algorithms that incorporate feature drawing from these sound signals have been trained on a small quantity of data in the cloud. In areas with fewer medical facilities or during a pandemic (like the current Covid), this complete system as a service might be helpful.

1.3 Research Questions

In this research, following are the hypothesis that needs to be solved for,

RQ1: What is the optimized pipeline that can be used for deploying the trained model into the diagnostic system?

RQ2: Which data features need to be used for a better result?

RQ3: What are the different algorithms that are used for the diagnosis?

RQ4: How are the different sound signals used for identifying heart problems? How cost effective will this system be in terms of computation, space/storage of the data in real time and feedback system?

In the next part, we'll take a close look at the various sound synthesis methods and machine learning models employed in past papers. This chapter is then followed by an in-depth analysis of the various approaches, methodologies, and strategies, as well as the suggested pipeline for implementing AI-Doctor. The findings section will detail the evaluation of several machine learning models and the extraction of features.

2 Related Work

Cardiac auscultation, sometimes known as listening to one's heart sounds, was evaluated in the studies conducted by Randhawa and Singh using phonocardiogram (PCG) signals (Randhawa, S.K., and Singh, M. (2015)). Digital recordings of cardiac sounds were used to create the PCG signals, which were then sorted into normal, systolic murmur, and diastolic murmur categories. The majority of signals were of the normal variety. Normal cardiac murmurs were present in 60 of the 144 samples tested, whereas diastolic murmurs were present in 45 and systolic murmurs were present in 39. The researchers started with a large dataset of 28 traits and used a technique called feature reduction to narrow it down to only seven critically relevant characteristics. The k-NN method, the fuzzy k-NN technique, and an Artificial Neural Network (ANN) were only some of the classifiers used. It was shown that classifiers based on k-nearest neighbours (k-NN) and fuzzy k-nearest neighbours (fuzzy k-NN) were both capable of accurately identifying heart sound data, with a combined accuracy of 99.6%.
Liang et al. introduced a segmentation method for dividing phonocardiogram (PCG) data into four parts. The initial heart sound, followed by systole, the second heart sound, and finally diastole, make up these sections. This technique was crucial in automating the diagnosis of heart sounds, as it relied on the normalised average Shannon energy of PCG signals to detect interference. The system's efficacy was evaluated using normal and abnormal PCG data totaling 515 periods from 37 users. As the initial and most important stage of heart sound analysis, the results showed a high accuracy rate of 93% in recognizing and segmenting heart sound components (Liang, Lukkarinen, and Hartimo, 1997).
They discussed the limitations of subjective cardiac auscultation in detecting heart issues caused by faulty heart valves. They highlighted the potential for unpredictability resulting from reliance on the interpretation offered by highly skilled doctors. A computer-aided automated or semi-automatic heart sound detection system was advocated as a means of overcoming this problem and arriving at more reliable diagnostic findings (Wu, H., S. Kim, and K. Bae, 2010. August). When viewed together, these materials illuminate the development of cardiac auscultation methods. Methods like signal categorization and automated segmentation algorithms might be used to improve the objectivity of heart sound diagnosis.

2.1 Studies on exploiting heart- or lung-sound signals to diagnose health

With the pipeline shown in study (Panah, D.S., Hines, A., McKeever, J.A., & McKeever, S., 2022), the typical set of cardiac sounds captured by remote devices may be analysed and enhanced. The study's authors, Hines and McKeever, are credited. The quality of a heart sound may be evaluated by the pipeline to see if it is suitable for diagnostic purposes, and the pipeline can implement quality enhancement procedures as needed. The research was done to inform and validate the design decisions for the proposed pipeline, and It showed a significant amount of congruence with the poll findings because of the participation of medical professionals. The suggested pipeline has the potential to decrease the demand for highly skilled medical personnel during the acquisition and processing of diagnostic heart sounds.
In (Aziz, S., Khan, M.U., Shakeel, M., Mushtaq, Z., and Khan A.Z., 2019), the original signal is redesigned with the use of only trustworthy IMFs that can distinguish between healthy and ill individuals in the case of pneumonia. Mel frequency cepstral coefficients (MFCC) and temporal sector information are used for feature extraction. Then, using 5-fold cross validation, a Support Vector Machines (SVM) classifier is trained. A quadratic kernel SVM implementation produced classification accuracy of 99.7% when used on a dataset of 480 percussion signals.

The acoustic stethoscope in (Roy, J.K., Roy, T.S., and Mukhopadhyay, S.C., 2019) uses a microphone and preamplifier module to boost the input audio signal. By computing the frequency, amplitude, and other characteristics of an uninterrupted flow of heart sounds, the MATLAB software analyzes the data. A harmonic distribution is used to display and analyze the heart's filtered sound signal. This inquiry makes use of the frequency domain. A harmonic distribution is utilized to present and evaluate the heart's filtered sound signal. The frequency domain is used for this investigation. The Amplitude Distribution of Harmonics is a useful diagnostic tool for identifying aberrant heart sounds and cardiac murmurs. Matlab's spectrogram, periodogram, and Kalman filtered response are examples of these tools. Also included are root-mean-square (RMS) value, mean value, average energy, and average power.

Figure 2. An example of Digital Stethoscope (**Roy, J.K., Roy, T.S. and Mukhopadhyay, S.C., 2019**)

Automatic auscultations are presented as a tool for the diagnosis of pneumonia (Haider, N.S., 2021). Preprocessing of signals is done using Empirical mode decomposition (EMD), and for feature extraction the mel frequency cepstral coefficients (MFCC) are merged with temporal sector information. Support Vector Machines (SVMs) are the most effective classifiers, with a success rate of 99.7 percent. By merging Hurst analysis, EMD, and spectrum subtraction, the authors of a 2021 study (Shaikh Salleh, S.H. et al.) describe a novel method for denoising breathing sounds for application in computerised diagnostics. Computerised diagnoses based on respiratory sounds are also the subject of this study's investigation. The newly discovered denoising sound technology is expected to aid medical personnel in establishing accurate diagnoses based on respiratory sounds. The study examines the challenges associated with computer-based segmentation and classification methods for phonocardiograms (PCGs) and makes the hypothesis that feature-free approaches, underpinned by strong statistical models, may hold the key to decreasing the overall computational cost of the segmentation strategy.

2.2 An overview of the studies on exploiting heart and lung sound signals to diagnose health issues

An evaluation and enhancement pipeline for mobile heart sound acquisition is proposed in the first study. The need for highly trained doctors to make diagnoses might be reduced if this pipeline is put into action. The second study employs trustworthy MFCC,SVM classifier and IMFs to precisely categorise healthy individuals in contrast to those suffering from pneumonia.By analyzing the frequency, amplitude, and other characteristics of cardiac sounds in MATLAB, the sounds are analyzed for diagnostic purposes.. In the research paper (E. O. Olaniyi and O. K. Oyedotun,2015), an algorithm was presented for utilising auscultations to diagnose pneumonia. Impressive precision was attained by this system using EMD, MFCC, and an SVM classifier. The study (Cheng, X., Ma, Y., Liu, C., Zhang, X. and Guo, Y., 2012) presents a novel approach for denoising breathing sounds by combining EMD, Hurst analysis, and spectrum subtraction. If put into practise, this technique might help doctors more accurately identify respiratory problems. The findings imply that feature-free techniques supported by strong statistical models may be helpful for decreasing the computing expense of PCG segmentation.

2.3 Examining machine learning and deep learning could be applied to diagnose cardiac disease

A method is demonstrated in (Gjoreski, M. et al. 2020) that employs deep learning in conjunction with Spectro-temporal representation of cardiac sounds to guide standard machine learning in order to differentiate between healthy person and those with congestive heart failure (CHF). Identifying CHF is the goal of this strategy. This method's comprehensive accuracy was 92.9%, that is extremely close to the proportion of recordings that experts classify as "unknown." In order to help with the development of models that can differentiate between the decompensated and recompensated stages of CHF, the study also identified 15 expert cues that might be incorporated to ML models. The results of the suggested approach for figuring out the phases of CHF are encouraging, and they could inspire the creation of home-based CHF monitoring systems and better ways to spot people with CHF.

This study (Liu, J., Wang, H., Yang, Z., Quan, J., Liu, L., and Tian, 2022) created a unique algorithm model to identify congenital heart disease (CHD) in children with left-to-right shunts. In terms of accuracy, the system—known as a residual convolution recurrent neural network (RCRnet) classification model—performed better than professional auscultation. The most productive regions for diagnostic testing were found to be the auscultation zones 4, 5, 2, and 3. All four metrics have fairly high ranges (0.888-1.000) for sensitivity, specificity, precision, and accuracy. There was shown to be no statistically significant relationship between LVEF, RVSP, and anomaly severity (Pearson's r 0.3). This held true across all four types of congenital heart disease.

This study (Ge, B. et al. 2023) used a single-cycle, multi-variable method to establish a non-intrusive computer-aided diagnosis strategy for pulmonary hypertension caused by congenital heart disease (CHD-PAH). The time-frequency domain measurements for each cardiac cycle were different, and the segmented heart sounds were used to extract wavelet packet energy attributes. These traits were combined into a single feature vector using XGBoost in order to categorize the subjects as normal, CHD, or CHD-PAH. The method's classification precision increased with its previous iteration to 88.61 percent. The research of (Jabari, M., Rezaee, K., and Zakeri, 2023) offers a hybrid classification technique for automatically classifying various cardiac auscultation classes from a phonocardiogram signal. This technique combines deep learning and human-created characteristics. Phonocardiogram readings are used in this method. This method use L-spectrograms to construct a power spectrum of the signal, which is then fed into a stretched and modified version of the ResNet architecture to extract deep features. In addition, the best windowed signals are used to manually generate frequency domain characteristics. Support vector machines (SVM) with radial basis function kernels are utilised to categorise the proposals, and the altered drone squadron optimisation technique is implemented to avoid overfitting. The precision of the suggested method is 99.38 percent in contrast to the PhysioNet/CinC 2016 Challenge and 94.16 percent when evaluated to the Michigan Heart Sound and Murmur Database (MHSDB). The hybrid model beats other cutting-edge techniques in terms of precision and computing efficiency, and it has the potential to be employed in the implementation of automated cardiac screening systems in less developed and culturally diverse places.

Researchers suggest two deep learning models in this research (Nguyen, M.T., Lin, W.W., & Huang, J.H., 2023) for categorising heart sounds using log-mel spectrogram data. The models can discriminate between five different types of heart sounds, including normal, mitral regurgitation, systolic heart murmur, aortic stenosis and mitral stenosis. Models with a large short-term storage capacity and a convolution neural network have a 99.67% average accuracy. When compared to the results of their earlier tests, the models' effectiveness increased noticeably.

In (Zhang, M., Li, M., Guo, L., & Liu, J., 2033), the authors detail the development of a portable, low-cost prototype that can diagnose lung and heart sounds at the same time. This prototype may be taken with you and used on the go with the help of a low-cost embedded device. On the Yaseen dataset, the model has a precision of 99.94% and could be run on a Raspberry Pi. Medical personnel may use the $5 digital earpiece and the AI-powered digital eardrum everywhere since they can both automatically diagnose patients and capture audio files for further analysis.

2.4 Overview of the deep learning and machine learning-based virtual cardiac assessment system

In these papers, a number of techniques are put forth for using deep learning and conventional machine learning to diagnose heart conditions like congestive heart failure (CHF), and pulmonary hypertension, left-to-right shunt congenital heart disease (CHD), linked to congenital heart disease (CHD-PAH). Heart sounds are evaluated spectro-temporally, and then various features are retrieved from that analysis using the methods that have been outlined. Studies have shown that accurate, sensitive, specific, and precise diagnosis of cardiac diseases is possible. Some research has also suggested lightweight and inexpensive prototypes that may be used on inexpensive embedded electronics in developing nations. These prototypes can be utilised in areas with limited access to materials. These results may lead to innovative methods for diagnosing cardiac problems and developing home monitoring equipment.

2.5 Research Towards Cloud Computing Methodologies for the deployment of Diagnostic Systems

Integration of cloud computing and IoT technologies is used in intelligent healthcare systems to deliver near real-time applications and enhance patient care. However, there are still obstacles to overcome, like reaction speed, availability, security, and privacy concerns. Interest in edge and fog computing as a solution to these problems has increased recently. With this research, Musa, Ibrahim, and Abd Rahman (2020) hope to provide a succinct evaluation of the widespread application of IoT solutions in health care, from the earliest health monitoring systems to the most recent advancements in fog computing and edge computing for smart health.

Phonocardiogram (PCG) signals are investigated in this study (Belkacem, A.N., Ouhbi, S., Lakas, A., Benkhelifa, E., & Chen, C., 2021). It was found that abnormal PCG signals vary at a higher frequency than normal ones when a device was built for recording them using an Arduino Uno-based microcontroller. Amazon Web Services (AWS) now serves as the safe haven for all of the data, making it easier than ever to retrieve. The author also offers a medical theory for a fully mobile system that can record coughs, convert them into health data for diagnosis, and classify them using machine learning into a variety of respiratory illnesses, including COVID-19. In this part of the research, we provide this medical conjecture. The proposed technology has the potential to play a significant role in the early detection of COVID-19 and in limiting its spread because it is both cheap and simple to use.

2.6 Overview of the server and cloud resources utilised in the installation of the cloud-based diagnostic devices

Each article focuses on a specific discovery related to a certain type of server or deployment method used in cloud, IoT, edge, or fog computing. There is a rising interest in Edge and Fog computing as a way to address problems, and this article examines their application in conjunction with Cloud and Internet of Things architectures in intelligent healthcare systems. The use of Amazon Web Services (AWS) to quickly retrieve stored data from PCG signals is briefly discussed in the paper.

Several promising new avenues might be explored in the quest to create a virtual diagnostic system that uses heart and lung sound signals. Improve the diagnostic system's precision and utility by combining many modalities, employing pre-trained models, heightening openness, personalising treatment recommendations, assessing temporal and behavioural patterns, and discovering novel biomarkers.

3 Research Methodology and Design Specification

3.1 Research Method

AI-doctor is a technology that provides a first medical diagnosis based on the sounds of the cardiovascular system and respiration (or other nearby locations). This system will be housed on the earpiece (digital as illustrated in figure 5(b)) or an app that will capture the heart sound. Details of the AI-Doctor architecture is described below,

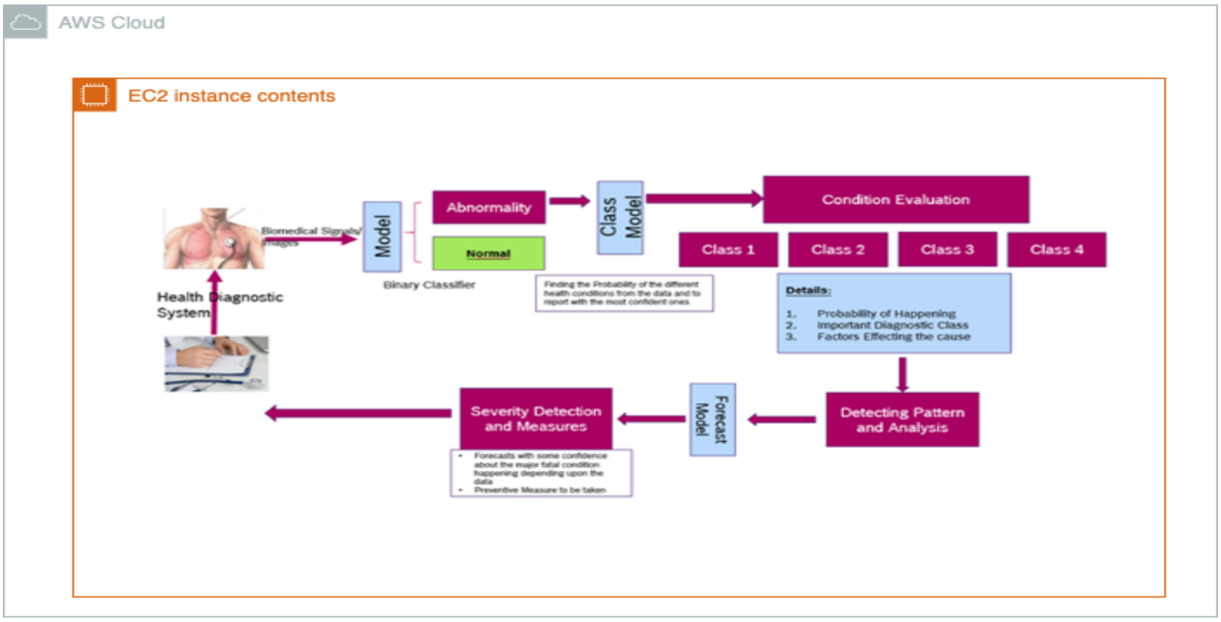

Fig. 3. Architectural Flow of AI-Doctor

Step 1: **Data will be collected** using an edge device, the digital earpiece, to collect heart sound signals.

Step 2: **Signals Pre-processing and Features Extraction** – MFCC (Mel Frequency Cepstral Coefficients) [18] is used as the feature separation technique.

$$M(f) = 1125\, ln\left(1 + \frac{f}{700}\right), f \text{ is the frequency of the signal} \tag{1}$$

Step 3: **The Machine Learning Classifier** – Here we have used stratification techniques to deal with the class imbalance [15][16].

$$K_{selected} = \max_{\forall k}\{precision_k\}, \mathrm{k} \in \mathrm{ML_{models}} \tag{2}$$

$$precision_k = \frac{\sum_{j=1}^{n} 1, \forall\ y_p = y_a = 1}{\left(\sum_{j=1}^{n} 1, \forall\ y_p = y_a = 1\right) + \left(\sum_{j=1}^{n} 1, \forall\ y_p = 1\ and\ y_a = 0\right)} \tag{3}$$

Where, y_p, y_a are output predicted and actuals of the signal respectively and n is the total validation sample set.

Step 4: **Determining the Health Conditions** – Higher likelihood scores will be applied for a class determination that is more definite.

3.2 Signal Processing with Signal Analysis and Features Processing

Speech Recognition is a job that must be learned under supervision. In the voice recognition task, the audio signal will be the input, and we will have to anticipate the text from the audio signal as the output. As a result, we cannot utilise the raw audio data as an input to our model since there will be a significant amount of disruption in the signal. As a result, it has been discovered that separating features from the sound data and utilising them as a given data to the basic model produces significantly higher performance than just taking the raw audio data as input. In audio signal processing, the MFCC approach is the most extensively used technique for separating the characteristics from the signal. In order to get the desired outcome, we will convert our audio stream from analogue to digital format in this step, utilising either an 8kHz or 16kHz sampling rate. Pre-emphasis enhances the amount of energy generated at higher frequencies. Vowels and other voiced segments can be identified in an audio signal's frequency domain by observing that the energy at greater frequencies is considerably less compared to the energy at lower frequencies. The efficiency with which the phone can be identified will improve with increased energy available at higher frequencies, improving the model's overall performance. The objective of the MFCC technique is to extract features from an audio signal that may be used to identify phones in a speech stream.

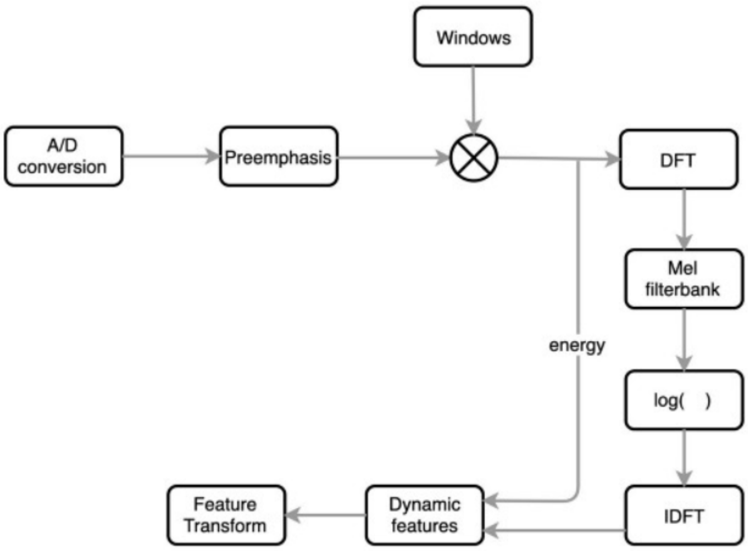

Figure 4: MFCC complete steps

Because there will be multiple phones in the supplied audio stream, we will divide the audio stream into separate segments, each segment have a 25ms width and the signals being spaced out by 10ms, as seen in the following picture. Assuming that a person says three alphabet per second while using four phones and that each phone has three states, this results in 36 states per second or 28 milliseconds per state, which is close to our 25-millisecond window. We will extract 39 different attributes from each section. As a bonus, while we are breaking up a signal, if we immediately slice it off at its edges, the sharp drop in amplitude at the edges will result in noise in the high-frequency domain. We will utilise Hamming windows to cut the signal instead of a rectangular window to avoid producing noise in the high-frequency area, which will be more efficient. By applying the DFT transform to the signal, we will be able to translate it from the time domain to the frequency domain. When it comes to audio signals, studying them in the frequency sector is far easier than analysing them in the time sector . There is a difference between how our ears sense sound and how machines perceive sound, and our ears will experience sound differently from the machines. Low-frequency hearing is more detailed than high-frequency hearing because our ears have a greater resolution at lower frequencies. When we hear noises at 200 Hz and 300 Hz, we can readily distinguish them from sounds at 1500 Hz and 1600 Hz, even though they have a difference of only 100 Hz in frequency between them. The resolution of the machine, on the other hand, remains constant across all frequencies. It has been observed that including the human hearing attribute in the feature extraction stage would increase the overall performance of the model's performance. As a result, we will utilise the mel scale to convert the real frequency to a frequency that humans would perceive as pleasing. The mapping formula is presented in the next section.

$$mel = 1127\, Ln(1 + \frac{f}{700})$$

Compared to when the transmission energy is lower, humans are less sensitive to variations in auditory signal strength when the signal frequency is higher. The curve of the log function will be higher at low input x values and lower at high input values, which is similar to how the log function behaves. To replicate the human hearing system, the log function is used to the Mel-filter's output. The result from the preceding step must first undergo an inverse transform in order to go forward. We must first understand how human-produced sound is created in order to understand why it's necessary to do a reverse transform. The following is the formula for calculating the energy of the sample.

$$Energy = \sum_{t=t_1}^{t_2} x^2[t]$$

Additionally, the MFCC approach will examine the first-order derivative and second-order derivatives of the features, which together account for a total of 26 features in addition to the original 13. To comprehend how the transition is taking place, derivatives are calculated by calculating the difference between the coefficients between the samples of an audio signal. For each audio signal sample, the MFCC approach will yield 39 characteristics, which will be utilised as input for the speech recognition model in the end.

3.3 Machine Prediction

The algorithms that we have taken into the account are Logistic Regression, Support Vector Machines, Decision Trees Classifier, XgBoost, Random Forest Classifier, KNN and LSTM.

3.4 Performance Evaluation

Machine learning techniques are employed in statistical classification challenges to predict the target category for a data set. A categorization model calculates the probability that every single instance relates to a certain class or class of cases. In order to use these models in real-world problem-solving situations, the classifications model's performance must be assessed. Performance metrics are used to assess the effectiveness of classification algorithms for machine learning in a certain context when it comes to these models. Among the performance measures offered are reliability, precision, recollection, and F1-score. Model performance is crucial for machine learning since it enables us to recognise the advantages and disadvantages of these models when making predictions under novel conditions.

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

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

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

4 Implementation and Design Specification

To preserve patient privacy and adhere to applicable rules, it is crucial to adopt best practises for protection, such as employing HTTPS for transmitting information and encrypting personal information. The function of AI Doctor is represented as a flow diagram as below,

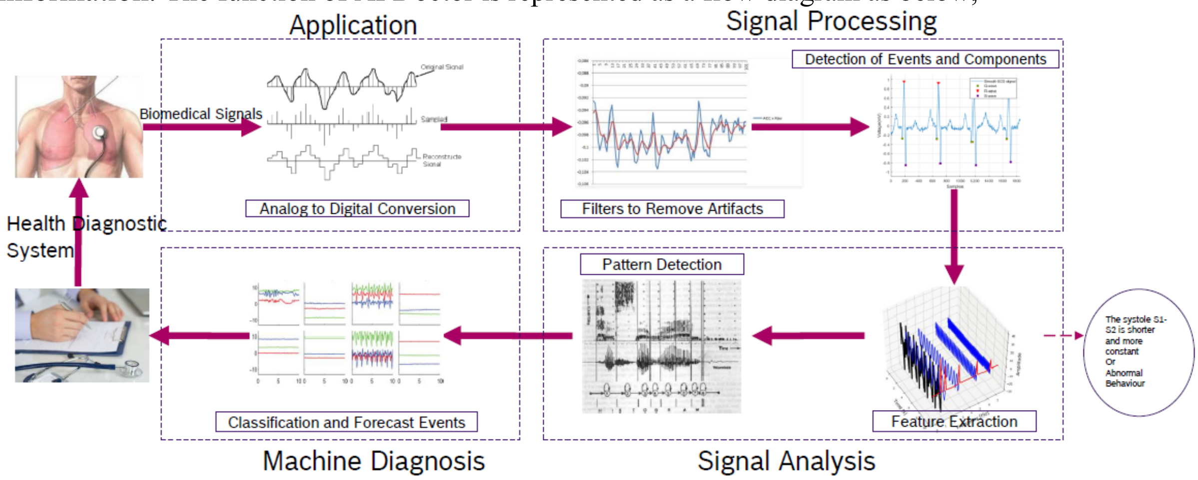

Figure 5: Architectural Flow of the whole process

To carry out the implementation of the AI-Doctor which will work as a virtual diagnosis doctor, the architecture is shown above. This AI-Doctor will be a response system that will be residing on the stethoscope which will acquire the lungs or heart sounds and then feed them to the AI-Doctor response network. This whole setup can be divided into two major stages. The first stage is an edge platform that will take into account the sound data collection and then do the health monitoring using the trained and saved model on the edge. The cloud platform will be taking into account the training and the generation of the saved model which will be sent to the edge platform via software updating. In addition to these stages, there will also be a user interface whether a mobile app or a web app. Let's discuss in detail the response system that will be done to achieve the objectives of the AI Doctor.

Step 1: Bio-Medical Signals Acquisition – The Data will be acquired by the Stethoscope which will act as an edge device over here. The data that will be acquired will then be converted and transformed which will be then used by the different models that are present in the bag of algorithms. In this case, the feature extraction technique that will be taken into the consideration is the MFCC. Mel Frequency Cepstral Coefficients are extracted from the features which are then fed into the modes for the prediction of the classes. Below is the code snippet for the same.

Step 2: Binary Classifier – The next step in the production will be a binary classifier to check for the Normal behaviour and separate them. The whole idea here is to take into account the data imbalance that can be present in the different abnormal classes. Since the different disease cases will be very rare and getting the dataset will be a difficult task, the similar trend of getting the identification of the important features will also be a tough task for the model.

Step 3: Conditions Determination – Once the above step is done properly, the next stage is to get the disease identification which will then take into account the generation of the probability for the different classes. This will take into account the high probability score for a better confident class determination and the lower confident classes to be recognised as the others category which will be sent to the doctors UI portal in the app for further diagnosis. Here the classes that have been taken into the consideration are represented in the below graph.

Step 4: Identification of the Decision Areas – Since the model has taken a particular decision of predicting a particular class based on the input features to the models, the decisive areas are of primary importance for further analysis. This can be achieved using certain python packages like LIME and GRAD-CAM. This will mark the interesting areas for the analysis for the doctors.

Step 5: Recommendation – Once the whole process is taken into the account the recommendation of certain steps can be shared for a further and better diagnosis.

To compare the different features extracted for the different classes,

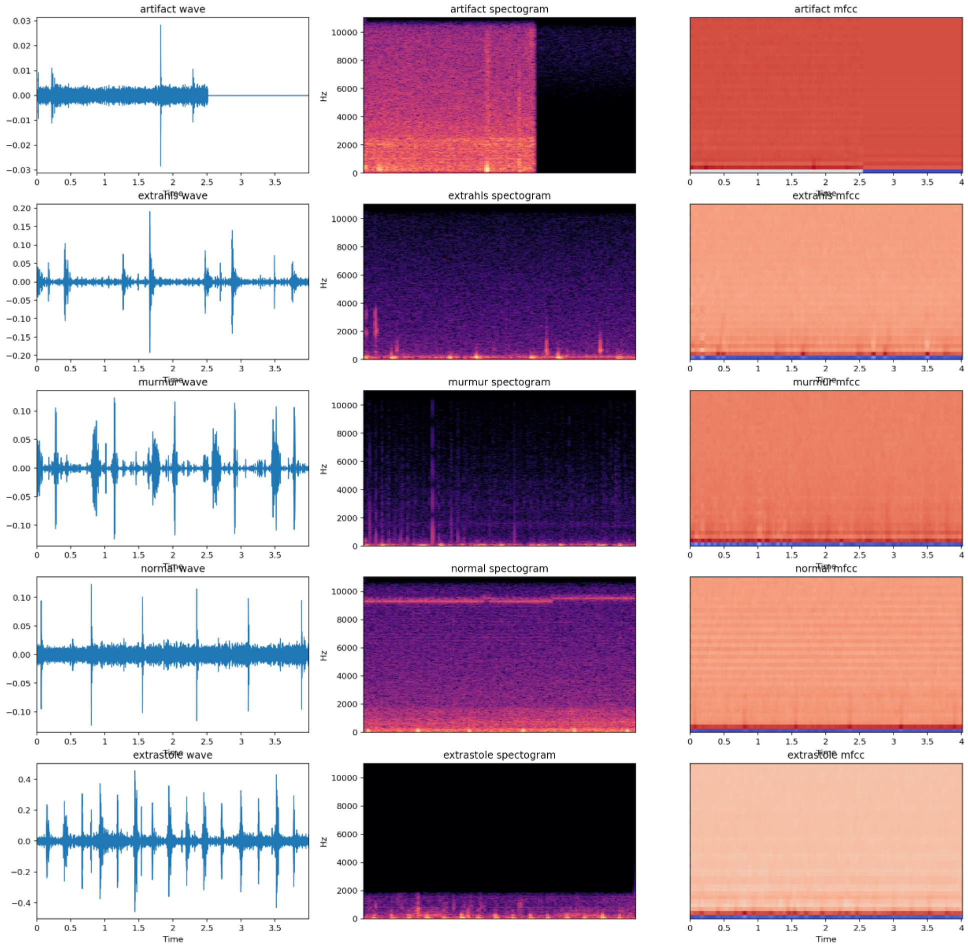

Figure 6: Comparison for the Digital sound signals for the different classes

In this figure, we can see the difference in the features pattern for the different classes. The leftmost column of the figures is the digital reading of sound. The middle column of the images is for the spectral density or spectrogram of the librosa package of the MFCC and the rightmost image is the MFCC.

5 Evaluation

In this section, we will be discussing in detail the different results of the machine learning algorithms that have been taken into the account.

5.1 Case Study 1

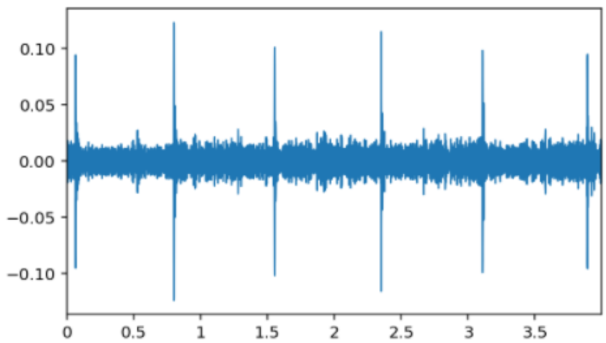

Figure 7: Normal Heard Sound

The lub sound is S1 and the dub sound is S2. At rest, most people's hearts beat between 60 and 100 beats per minute. Because the data may be gathered from toddlers or adults in a relaxed or enthusiastic condition, the heart rates may range from 40 to 140 beats per minute. Dataset B comprises noisy normal data, which is normal data with significant background noise or distortion. You can use it or not, but the test set will include some similarly noisy samples. There is a "whooshing, roaring" or "rumbling" sound in one of two temporal locations: between "lub" and "dub", or between "dub" and "lub". They can indicate significant cardiac conditions. There will be lub and dub. One of the things that confuse non-medical people is that murmurs occur between lub and dub, not on lub or dub.

5.2 Case Study 2

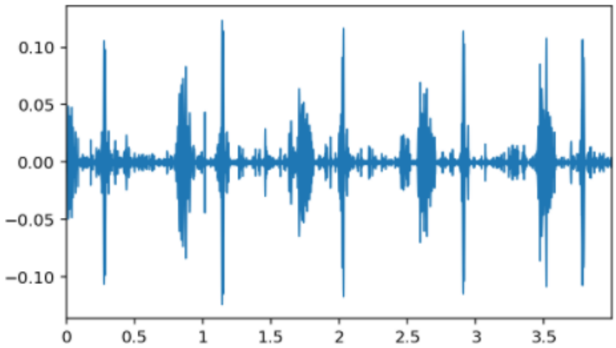

Figure 8: Murmur Heart Sound

Dataset B also contains noisy murmur data, which has significant background noise or distortion. You can use it or not, but the test set will include some similarly noisy samples. A "lub-lub dub" or "lub dub-dub" is an example of an extra heart sound. This may not be the case. In certain cases, it is a warning indication of a condition that may be treated early. The additional heart sound is critical to identify since ultrasonography cannot detect it well. Note the additional heart sounds' temporal description:

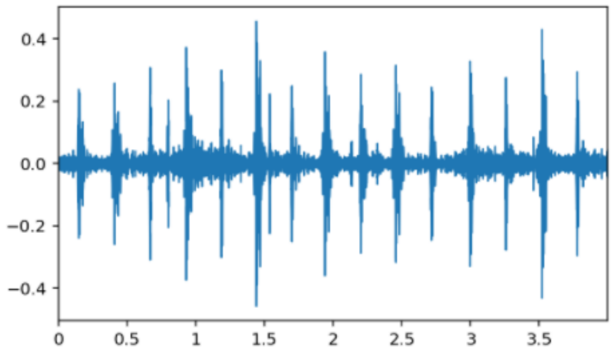

Figure 9: Extra Heart Sound

5.3 Case Study 4

Sounds in the Artifact category include feedback squeals and echoes, speech, music, and noise. At frequencies below 195 Hz, there are no discernible heart sounds, and consequently no temporal periodicity. This category is unique. It is critical to identify this group from the other three so that data collectors can be prompted to attempt again.

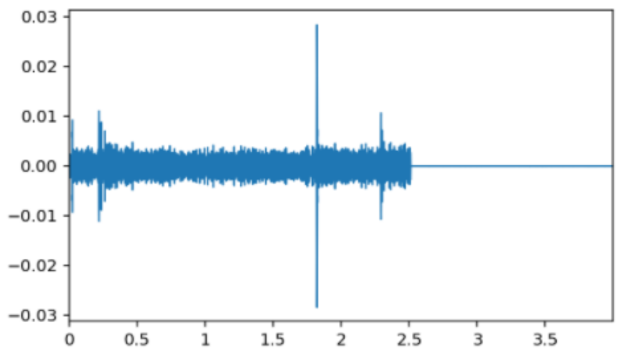

Figure 10: Artifact Heart Sound

5.4 Case Study 4

Extrasystole noises are detected by an out-of-rhythm heart sound containing additional or missed heartbeats, e.g., a "lub-lub dub" or a "lub dub-dub". (Note that this is not the same as an additional heart sound.) This may not be a sickness symptom. It occurs in both adults and children. Heart disorders can induce extrasystoles in rare cases. Treatment will be more successful if these disorders are diagnosed sooner. Note the additional heart sounds' temporal description:

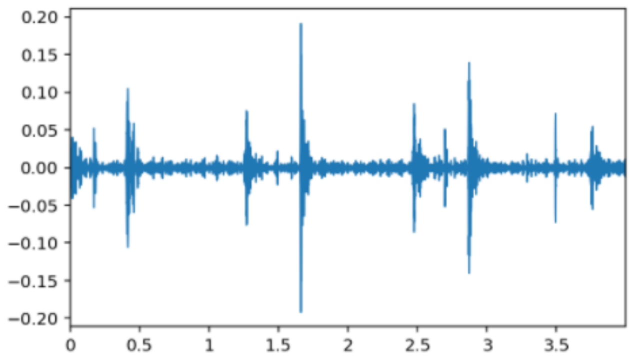

Figure 11: Extrasystole Heart Sound

5.5 Machine Learning Prediction

While analysing the different number of training examples present in the dataset, the below figure talks about the number of samples per class,

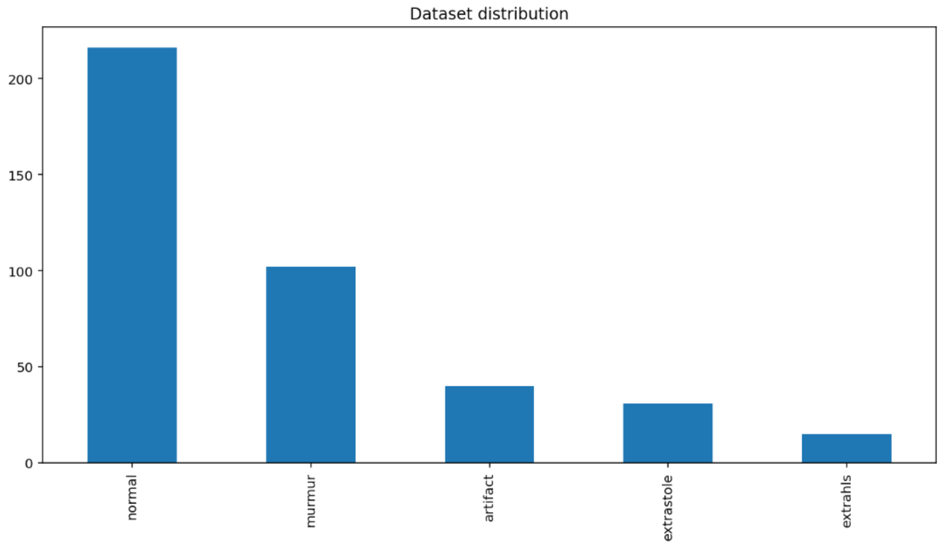

Figure 12: Samples Comparison per class

In this graph, we can see that most of the examples present are for the class Normal which is the exact scenario in the real world. The next class which has the maximum samples are for the class ‘murmur’ followed by the class ‘artefact’ and ‘extrastole’. The class ‘extrahls’ has the least number of samples present. The data is then passed within the MFCC module to get the features extracted and then passed into the machine learning algorithms for the class prediction. For stage 1 of the prediction for the binary classes’ prediction, the SVM-RBF kernel was predicted with 79.8% accuracy.

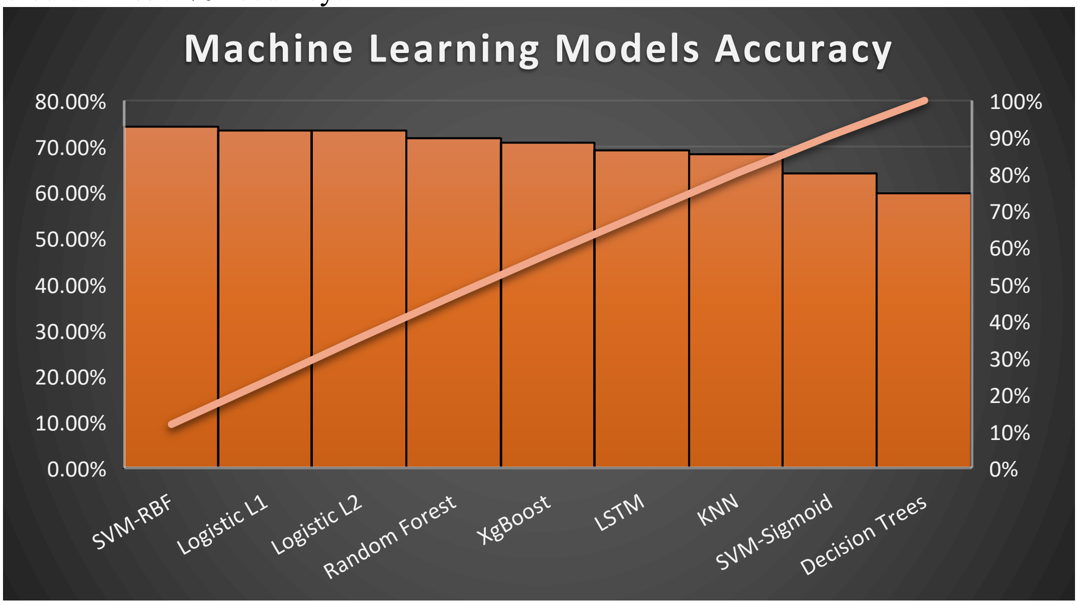

Figure 13: Accuracy comparison for different machine learning models

The above figure is for the complete analysis of the different classes. We find that SVM-RBF here as well performed the best. In this chole of the analysis, the hyperparameters tuning is done with the use of Grid Search and Random Search.

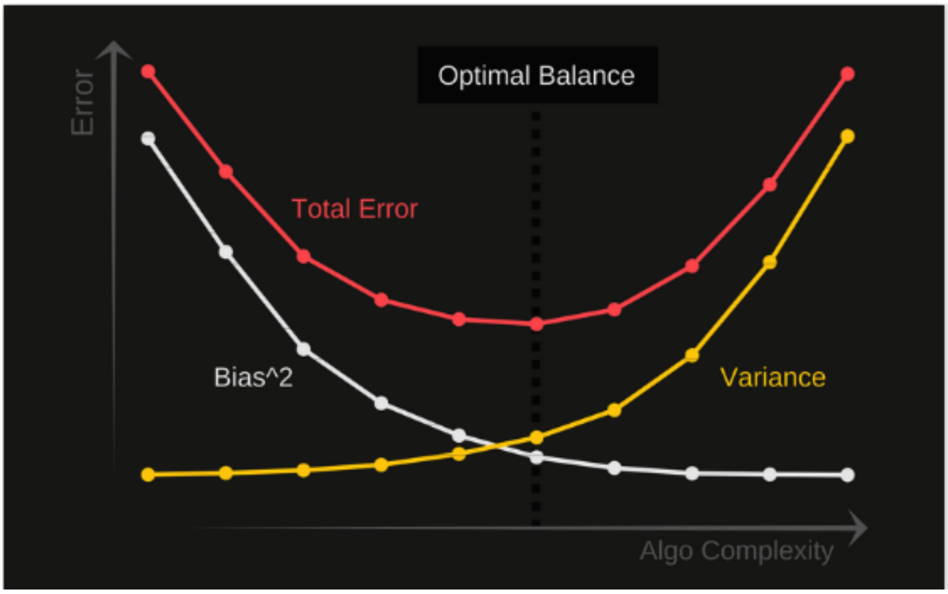

Figure 14: Bias Variance Trade-Off

The model that accurately predicts the correct values is shown in the above image in the middle of the goal. Our forecasts become more and more off as we move away from the target. For several impacts on the target, our model-building approach may be repeated. When a model in supervised learning is unable to recognise the underlying pattern in the data, underfitting occurs. These models frequently exhibit significant levels of bias and low levels of variance.

This issue arises, among other things, when we don't have enough data to establish a reliable model or when we try to fit a linear model to nonlinear data. Additionally, straightforward models like linear and logistic regression are very good at detecting intricate data patterns. When our model successfully captures both the noise and the underlying pattern in the data set, the phenomenon of overfitting occurs in supervised learning. It happens when we train our model on a noisy dataset over a large number of trials. These models have a high amount of variety and little bias. These models, like Decision Trees, are exceedingly complex and prone to overfitting because of this. The accuracy table is as below,

Models	**Accuracy**
XGBoost	70.81%
SVM-Sigmoid	64.10%
SVM-RBF	74.36%
Logistic L1	73.50%
Logistic L2	73.50%
Decision Trees	59.83%
Random Forest	71.79%
KNN	68.38%
LSTM	69.14%

From the above table, we can see that SVM-RBF has an accuracy of 74.36% followed by Logistic Regression with 73.5%.

6 Conclusion and Future Analysis

The development of non-invasive cardiac sound detection technology has received greater attention recently due to the rise in the prevalence of cardiovascular illnesses. In this paper, the most recent research on computer-aided cardiac sound identification methods has been examined, with a focus on how various signals' properties may be extracted and how different machine learning algorithms can be used to categorise heart sounds. The research findings have been presented in this study. In this research, we find that SVM-RBF has an accuracy of 74.36% followed by Logistic Regression with 73.5%. The features extraction that is used is MFCC (Mel Frequency Cepstral Coefficients).
The following topics of future research are advised to better understand the potential contributions towards technology to promote human health. To enrich the heart sound database, a considerable amount of cardiac sound data will be required. A good source of data is required for finding the hidden characteristics of cardiovascular illnesses is heart sound data. As a result, it is required to put through and improve the heart sound database, as well as the expert annotations that accompany it, to improve model training and provide a more perfect assistant diagnosis. Because large-scaled computer systems are currently accessible in hospitals, it has become viable to develop a complicated deep learning model that will be able to handle heart sound data on a huge scale. The methodologies for data processing and parameter optimization, as a result, require more in-depth investigation. Deep learning modelling necessitates more powerful computer systems with GPU capability, however, compressed deep learning techniques may be run on PCs. Further investigation of the heart sound classification framework based on compacted deep learning algorithms will be helpful for the popularisation and application of portable heart sound detection in the future because it is more precise than the heart's rhythm classification simulation based on traditional algorithms.
In this research as for finalising the best of the models, Accuracy is the performance parameter that has been taken into the consideration. For future studies, moe different performance

metrics need to be taken into account like F1 Score, Precision, Sensitivity, AUR-ROC etc. Different features extraction techniques like Tonnetz, Auto Correlation, PSD (Power Spectral Density) etc. will be taken into then consideration for more enriched features set.

References

A. Methaila, P. Kansal, H. Arya, and P. Kumar, 2014, "Early heart disease prediction using data mining techniques," in Proceedings of Computer Science & Information Technology (CCSIT-2014), vol. 24, pp. 53–59, Sydney, NSW, Australia, 2014.

B. Edmonds, 2005, "Using localised 'Gossip' to structure distributed learning," in Proceedings of AISB symposium on Socially Inspired Computing, pp. 1–12, Hatfield, UK, April 2005.

Cheng, X., Ma, Y., Liu, C., Zhang, X. and Guo, Y., 2012. Research on heart sound identification technology. Science China Information Sciences, 55(2), pp.281-292.

E. O. Olaniyi and O. K. Oyedotun, 2015, "Heart diseases diagnosis using neural networks arbitration," International Journal of Intelligent Systems and Applications, vol. 7, no. 12, pp. 75–82, 2015.

Elamaran, V., Arunkumar, N., Hussein, A.F., Solarte, M. and Ramirez-Gonzalez, G., 2018. Spectral fault recovery analysis revisited with normal and abnormal heart sound signals. IEEE Access, 6, pp.62874-62879.

Gamero, L.G. and Watrous, R., 2003, September. Detection of the first and second heart sound using probabilistic models. In Proceedings of the 25th annual international conference of the ieee engineering in medicine and biology society (ieee cat. no. 03ch37439) (Vol. 3, pp. 2877-2880). IEEE.

H. Kahramanli and N. Allahverdi, 2008, "Design of a hybrid system for the diabetes and heart diseases," Expert Systems with Applications, vol. 35, no. 1-2, pp. 82–89, 2008.

J. López-Sendón, 2011. "The heart failure epidemic," Medicographia, vol. 33, pp. 363–369.

J. Mourão-Miranda, A. L. W. Bokde, C. Born, H. Hampel, and M. Stetter, 2015. "Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data," NeuroImage, vol. 28, no. 4, pp. 980–995.

K. Vanisree and J. Singaraju, 2011. "Decision support system for congenital heart disease diagnosis based on signs and symptoms using neural networks," International Journal of Computer Applications, vol. 19, no. 6, pp. 6–12.

Liang, H., Lukkarinen, S. and Hartimo, I., 1997, September. Heart sound segmentation algorithm based on heart sound envelogram. In Computers in Cardiology 1997 (pp. 105-108). IEEE.

Liu, C., Springer, D., Li, Q., Moody, B., Juan, R.A., Chorro, F.J., Castells, F., Roig, J.M., Silva, I., Johnson, A.E. and Syed, Z., 2016. An open access database for the evaluation of heart sound algorithms. Physiological Measurement, 37(12), p.2181.

M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, 2013, "Classification of heart disease using artificial neural network and feature subset selection," Global Journal of Computer Science and Technology Neural & Artificial Intelligence, vol. 13, no. 11, 2013.

M. Durairaj and N. Ramasamy, 2016. "A comparison of the perceptive approaches for preprocessing the data set for predicting fertility success rate," International Journal of Control Theory and Applications, vol. 9, pp. 256–260.

M. Gudadhe, K. Wankhade, and S. Dongre, 2010, "Decision support system for heart disease based on support vector machine and artificial neural network," in Proceedings of International Conference on Computer and Communication Technology (ICCCT), pp. 741–745, Allahabad, India, September 2010.

O. W. Samuel, G. M. Asogbon, A. K. Sangaiah, P. Fang, and G. Li, 2017, "An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction," Expert Systems with Applications, vol. 68, pp. 163–172, 2017.

Oh, S.L., Jahmunah, V., Ooi, C.P., Tan, R.S., Ciaccio, E.J., Yamakawa, T., Tanabe, M., Kobayashi, M. and Acharya, U.R., 2020. Classification of heart sound signals using a novel deep WaveNet model. Computer Methods and Programs in Biomedicine, 196, p.105604.

P. A. Heidenreich, J. G. Trogdon, O. A. Khavjou et al., 2011. "Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association," Circulation, vol. 123, no. 8, pp. 933–944.

Q. K. Al-Shayea, 2011. "Artificial neural networks in medical diagnosis," International Journal of Computer Science Issues, vol. 8, no. 2, pp. 150–154.

R. Das, I. Turkoglu, and A. Sengur, 2009, “Effective diagnosis of heart disease through neural networks ensembles,” Expert Systems with Applications, vol. 36, no. 4, pp. 7675–7680, 2009.
R. Detrano, A. Janosi, and W. Steinbrunn, 1989, “International application of a new probability algorithm for the diagnosis of coronary artery disease,” American Journal of Cardiology, vol. 64, no. 5, pp. 304–310, 1989.
Randhawa, S.K. and Singh, M., 2015. Classification of heart sound signals using multi-modal features. Procedia Computer Science, 58, pp.165-171.
S. Ghwanmeh, A. Mohammad, and A. Al-Ibrahim, 2013. “Innovative artificial neural networks-based decision support system for heart diseases diagnosis,” Journal of Intelligent Learning Systems and Applications, vol. 5, no. 3, pp. 176–183.
S. Nazir, S. Shahzad, and L. Septem Riza, 2017, “Birthmark-based software classification using rough sets,” Arabian Journal for Science and Engineering, vol. 42, no. 2, pp. 859–871, 2017.
S. Nazir, S. Shahzad, S. Mahfooz, and M. Nazir, 2015 , “Fuzzy logic based decision support system for component security evaluation,” International Arab Journal of Information Technology, vol. 15, pp. 1–9, 2015.
S. Palaniappan and R. Awang, 2008, “Intelligent heart disease prediction system using data mining techniques,” in Proceedings of IEEE/ACS International Conference on Computer Systems and Applications (AICCSA 2008), pp. 108–115, Doha, Qatar, March-April 2008.
Wu, H., Kim, S. and Bae, K., 2010, August. Hidden Markov model with heart sound signals for identification of heart diseases. In Proceedings of 20th International Congress on Acoustics (ICA), Sydney, Australia (pp. 23-27).
Xiefeng, C., Wang, Y., Dai, S., Zhao, P. and Liu, Q., 2019. Heart sound signals can be used for emotion recognition. Scientific reports, 9(1), pp.1-11.