ACADEMIC CITY UNIVERSITY COLLEGE

DESIGN AND CONSTRUCTION OF A DIGITAL STETHOSCOPE WITH THE APPLICATION OF MACHINE LEARNING

NATHANIEL AWONTIIRIM ASIAK

2022

ACADEMIC CITY UNIVERSITY COLLEGE

DESIGN AND CONSTRUCTION OF A DIGITAL STETHOSCOPE WITH THE APPLICATION OF MACHINE LEARNING

NATHANIEL AWONTIIRIM ASIAK

10180200243

PROJECT SUBMITTED TO THE FACULTY OF ENGINEERING, ACADEMIC CITY UNIVERSITY COLLEGE, IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF BACHELOR OF ENGINEERING DEGREE IN ELECTRICAL AND ELECTRONICS ENGINEERING.

November 6, 2023

Declaration

This is to declare that, the research work documented in this Thesis has been carried out by the under-mentioned student under the supervision of the under-mentioned supervisor. The student and supervisor certify that the work documented in this Thesis is the output of the research conducted by the student as part of final year project work in partial fulfillment of the requirements for the Bachelor of Engineering (BSc) in Electrical and Electronics Engineering degree.

Nathaniel Awontiirim Asiak (Student)	**November 6, 2023**
Dr. Stephen K. Armah (Supervisor)	**November 6, 2023**

Abstract

Auscultation is an important diagnostic process in medicine. In Ghana, almost all the stethoscopes used for auscultation are traditional and acoustic. Electronic stethoscopes on the market are in the price range of $500 which makes them inaccessible to medical professionals in third-world countries. Two main challenges with using the acoustic stethoscope are the corruption of the faint sounds picked through the chest piece of the stethoscope by background noise and the subjectivity in offering diagnoses to a patient. This research showed the design and construction of a digital stethoscope that is able to amplify faint heart sounds and is coupled with a classification algorithm to differentiate normal and abnormal heart sounds using sounds recorded from the electronic stethoscope. An electronic stethoscope is designed with amplification capabilities and from this, a digital acquisition system obtains the sound for classification. Employing digital processing techniques such as spike removal, filtering, zero padding, and other techniques such as feature transformation, an Adaboost classifier was built to discriminate between normal and abnormal heart sounds with an accuracy of 93.95%.

Acknowledgments

I would like to acknowledge the inspirational and instructional guidance, selfless devotion, leadership, and coaching skills of my supervisor, Dr. Stephen K. Armah of Academic City University College-Ghana. There is no doubt in my mind that without his continued support and counsel, I could not have completed this project.

I would also like to thank Mr. Paul for his guidance in building the hardware of the project.

Finally, I thank all my family and friends for their continuous encouragement and moral support.

Dedication

I dedicate this project to my parents Mr. Samuel Asiak and Mrs. Faustina Asiak for their love and support for me throughout this program. This project is also dedicated to Dr. Fred Mcbagonluri for his instrumental guidance in my college journey here at Academic City. They are the real heroes in my story!

TABLE OF CONTENTS

COVER i

TITLE ii

DECLARATION iii

ABSTRACT iv

ACKNOWLEDGEMENT v

DEDICATION vi

TABLE OF CONTENTS vii

LIST OF FIGURES xi

LIST OF TABLES xii

LIST OF ABBREVIATIONS xiii

1 CHAPTER ONE - INTRODUCTION 1

1.1 Background 1

1.2 Problem Statement 4

1.3 Objectives 4

1.3.1 General Objective 4

1.3.2 Specific Objectives 5

1.4 Scope 5

1.5 Layout of Thesis 5

2 CHAPTER TWO - LITERATURE REVIEW 7

2.1 Theoretical Review 7

2.2 Related Work 9

2.2.1 The Electronic Stethoscope 9

2.2.2 Machine Learning with the Electronic Stethoscope 13

2.3 Technical Component Review 16

2.3.1 Microphones 16

2.3.2 Power 18

2.4 Audio Preprocessing 19

2.4.1 Zero Padding 19

2.4.2 Filtering 19

2.4.2.1 Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) Filters 19

2.4.2.2 Butterworth Filters 20

2.4.3 Amplification 21

2.4.4 Feature extraction 21

2.4.4.1 Mel Frequency Cepstral Coefficients 22

2.4.4.2 Spectral Flux 24

2.4.4.3 Spectral Rollof 25

2.4.5 Feature Selection 25

2.4.6 Hyperparameter Tuning 26

2.5 Windowing 26

2.6 MATLAB 27

3 CHAPTER THREE - METHODOLOGY 28

3.1 Block Diagram 28
3.2 Hardware Details 28
3.2.1 Electret Condenser Microphone 29
3.2.2 MAX4466 Pre-amplifier 30
3.2.3 LM386 Audio Amplifier 31
3.3 Software- Machine Learning Techniques 32
3.3.1 Datasets 32
3.3.2 Pre-processing 33
3.3.2.1 Downsampling 33
3.3.2.2 Filtering 34
3.3.3 Feature Transformation 35
3.3.3.1 Feature Extraction 35
3.3.3.2 Feature Selection : Neighborhood Component Analysis 35
3.3.4 Heart Sound Classification Algorithms 36
3.3.4.1 Ensemble Classifiers 36
3.3.4.2 Hyperparameter Optimization Method 36
3.4 Implementation 37
3.4.1 System Architecture 37
3.4.2 Hardware Subsystem 37
3.4.2.1 Hardware Implementation resources 37
3.4.2.1.1 EasyEDA 38
3.4.2.1.2 Ender-3 3D Printer 38
3.4.2.2 Design and Construction of hardware subsystem . . . 39
3.4.2.2.1 Power Unit 39

3.4.2.2.2 Amplification Unit 40

3.4.3 Software subsystem 40

3.4.3.1 Implementation Resources 40

3.4.3.1.1 Matlab Statistics and Machine Learning Toolbox 41

3.4.3.1.2 Matlab Application Designer 42

3.4.3.2 Training 42

3.4.3.3 Binary Classification Using Ensemble 43

4 CHAPTER FOUR - RESULTS AND DISCUSSION 44

4.1 Complete Assembly of Electronic Stethoscope 44

4.2 Software Results 46

4.2.1 Binary Classification using AdaBoost 46

4.3 Testing Results 47

4.4 Discussion 48

5 CHAPTER FIVE- CONCLUSIONS AND RECOMMENDATIONS 51

5.1 Conclusion 51

5.2 Limitations 52

5.3 Recommendations 52

REFERENCES 55

APPENDIX 60

List of Figures

1.1 Pie chart of responses from 11 medical professionals 3

2.1 Frequency response of nth-order Butterworth filter 21

3.1 Block diagram of digital stethoscope 29

3.2 Electret Condenser Microphone 30

3.3 MAX4466 Pre-amplification module 30

3.4 Circuit diagram of Preamplification 31

3.5 LM386 amplifier 31

3.6 Amplification circuit using LM386 32

3.7 Comparison of Mel spectrogram of a downsampled audio signal vs an original audio signa 34

3.8 Ender-3 3D printer 38

3.9 3D printed model of packaging unit 39

3.10 Battery power connection 39

3.11 Amplification unit connection 40

3.12 Flowchart of the software subsystem 41

4.1 Complete assembly of electronic stethoscope 44

4.2 Input signal to Oscilloscope 45

4.3 Amplified signals output from LM386N-1 45

4.4 Interactive GUI Desktop application for stethoscope 46

4.5 Misclassification error plot 47

4.6 Hyperparameters of Optimized AdaBoost ensemble 48

4.7 Confusion matrix of training observations using the AdaBoost ensemble 49

4.8 Confusion matrix of test observations using the AdaBoost ensemble . . 50

List of Tables

3.1 Summary of Audio Dataset 33

4.1 Classification results of Adaboost classifier on training data 47

List of Abbreviations

MFCC Mel Frequency Cepstrum Coefficient

MEMS Micro Electro-mechanical

PCG phonocardiogram

STFT Short Time Fourier Transform

CNN Convoluted Neural Network

RF Random Forest

LB LogitBoost

CSC Cost-Sensitive Classifier

SVM Support Vector Machine

FIR Finite Impulse Response

IIR Infinite Impulse Response

FFT Fast Fourier Transform

ECM Electret Condenser Microphone

NCA Neighborhood Component Analysis

BO Bayesian optimization

1 CHAPTER ONE - INTRODUCTION

1.1 Background

Nearly 40% of all deaths globally are caused by cardiovascular complications and lower respiratory infections. In 2019, 2.6 million people died from diseases like pneumonia, bronchitis and tuberculosis. In sub-Saharan Africa, 9.9% of all deaths were caused by lower respiratory tract infections [1]. These statistics are surprising given the advances in modern medicine. The stethoscope is a critical instrument used by health professionals to diagnose respiratory and circulatory diseases. The stethoscope enables doctors to listen to internal sounds of the body and from these sounds give a diagnosis of what might be wrong with a patient. Despite the very important use of this instrument, very little innovation has been made to it. The first monaural stethoscopes were developed by French physician Dr.Rene Laennec in 1816 [2]. This model of the stethoscope was essentially a cylindrical piece of wood opened at both ends and it brought a change to the method of direct auscultation that was being used at the time.

The current traditional stethoscopes were developed by American doctor George Cammann. Since then, very little innovation has been carried out on this important device until the recent emergence of electronic stethoscopes in 1999. There is a lot of research to develop more efficient stethoscopes and as a result, more electronic stethoscopes are coming on the market. Currently, 3M Limited holds the patent for the development of the Litmann3100 electronic stethoscope which offers 40X amplification, active noise cancellation, and employs artificial intelligence for murmur analysis. This is one of the more advanced applications of artificial intelligence in the stethoscope. The ThinkOne

Stethoscope developed by ThinkLabs is another electronic stethoscope that offers a redesign of the diaphragm [3]. ThinkLabs developed a patented Electromagnetic Diaphragms (EMDs) that uses a capacitive plate to detect vibrations on the diaphragm. The capacitive plate has a high charge voltage which produces a high-intensity electric field. As the diaphragm vibrates back and forth, the distance between the diaphragm and the capacitive varies which varies the intensity of the electric field. The varying voltage change is translated into an analog audio signal which is transmitted for further processing. Though both of these stethoscope designs from ThinkLabs and 3M Limited are impressive, they are rather expensive to afford, costing an average of $500.

Unfortunately, due to the cost of these modern electronic stethoscopes, doctors from low-income countries are not able to access them. Also, with the emergence of telemedicine, and the potential it holds for propelling healthcare in Africa by creating more access for people in low-income countries[4][5], there is the need to rethink the use and design of the stethoscope and other diagnostic tools to make them more ubiquitous.

In Ghana, traditional acoustic stethoscopes are the dominant category of stethoscopes being used. In a survey of 11 medical professionals, 100% of them confirmed the traditional acoustic stethoscope as the only one have seen and used in Ghana as shown in figure 1.1.

Acoustic Stethoscopes rely on the principles of sound propagation in their operation. Sound is picked from the patient using the chest piece and transmitted through an air-filled hollow tube which splits into two at the closer end to the headset. The length of the tubes varies according to manufacturer specification, and though there have been attempts to standardize stethoscopes, there has not been a single agreed standard. According to a comparative study conducted on several stethoscopes from various manufacturers, it was shown that each one had different acoustic properties and hence different performances making it difficult to standardize [6].

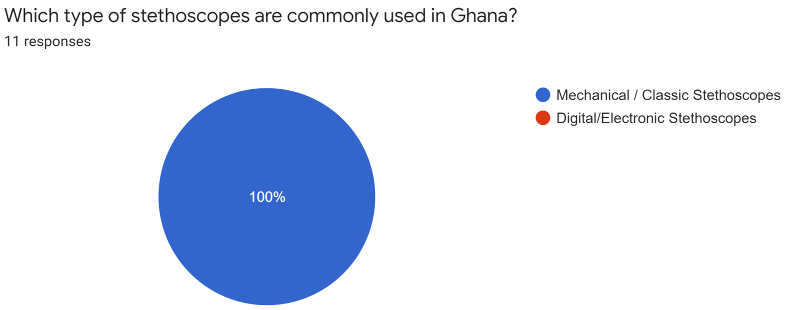


Figure 1.1: Pie chart of responses from 11 medical professionals

Based on the operation of the acoustic stethoscope, there are challenges with noise which interferes with the clarity of the lung or heart sounds being transmitted as well and as such makes the lung sounds very faint to hear clearly. As a result, diagnoses with the acoustic model of the stethoscope have become unsatisfactory. Generally, it would take a very experienced medical professional or cardiologist to use this model effectively; professionals who are in short supply in Africa.

Furthermore, in low-income countries, it is important to augment the diagnostic process due to the unavailability of medical specialists(cardiologists), and even in the case where experienced professionals are present, it is still important to offer objective diagnoses as auscultation contains a lot of subjectivity [7].

The challenges outlined above necessitate the development of the electronic stethoscope that is capable of offering objective predictive diagnoses based on clearly amplified and recorded sounds from patients at a cheaper cost which is what this project seeks to achieve.

1.2 Problem Statement

Traditional acoustic stethoscopes lack the technology to improve the process of auscultating patients. Auscultation involves listening to the heart sounds of patients to determine whether a patient's heart or lungs are healthy or not considering other factors during the medical examination. However, this activity requires the trained and experienced ears of a cardiologist or relevant medical professionals who are in short supply to detect abnormalities based on the sounds produced by the lungs. Unfortunately, this is not achieved due to signal corruption by noise from the background as well as the sounds from the internal body organs during auscultation, and low intensity of the lung sounds which in most instances leave the medical doctor unable to determine with certainty whether a heart or lung sound is normal. As a consequence, cardiovascular diseases and diseases of the lower pulmonary tract account for the majority of deaths in Africa and the Ghana region.

Therefore, there is a need to design an electronic stethoscope that is capable of suppressing noise, amplifying heart or lung sounds, and classifying whether a heart sound is normal or abnormal by using machine learning.

1.3 Objectives

1.3.1 General Objective

The purpose of this study is to design and construct an electronic stethoscope from an already existing acoustic stethoscope to provide a superior experience to medical professionals during auscultation with technology that can amplify heart sounds and classify the heart sounds as normal or abnormal.

1.3.2 Specific Objectives

The specific aims of this project are to:

- Design and construct the circuits for the electronic stethoscope
- Amplify heart sounds from the lung taken using the chest piece of the stethoscope
- Develop an application interface to record lung/heart sounds in MATLAB
- Develop Machine Learning algorithms to classify recorded heart sounds

1.4 Scope

This thesis covers the building of an electronic stethoscope that is able to amplify heart sounds. The electronic stethoscope is designed to be used with 3.5mm headphones or loudspeakers. This thesis does not focus on the wireless transmission of the heart sound nor the analog physical realization of filters. The second component of this project involves using supervised machine learning to classify heart sounds as normal or abnormal. As part of this second component, digital signal processing techniques such as filtering, are performed in Matlab. An application interface is designed to allow end users to record, process, playback, and receive classification results.

1.5 Layout of Thesis

The first chapter introduces the project briefly. This chapter describes the background of the project, the problem statement, the objectives, and the scope of the project as they relate to the development of the electronic stethoscope for auscultation. The second chapter discusses the research of related fields to the project which is mainly about the different design implementations of the electronic stethoscope hardware, different machine learning implementations of the heart sound classification, microphones, and digital and analog signal processing techniques used in the projects. Chapter three discusses the various methods employed to design the hardware of the electronic stethoscope as well as the machine learning algorithms. The specifications of the hardware components used, a block diagram of the system layout, and various machine learning techniques utilized as well as how the methods described are utilized. This chapter also focuses on the actual execution of the project objectives specified in Chapter One. Chapter four discusses the experimentation and the results of these experiments after implementing the project. It covers the experiments and results on the performance of the hardware electronic stethoscope and the performance of the machine learning algorithms in classifying heart sounds. Chapter five contains a summary of all the work that has been done in this thesis and discusses the limitations faced during the project with recommendations for future work.

2 CHAPTER TWO - LITERATURE REVIEW

This chapter discusses previous work, and important theories that are relevant to understanding the two fundamentally important but not mutually independent stages for the classification of heart sounds; the data acquisition system which involves the hardware built to record heart audio sounds, and the use of computer-aided software tools to classify the audio signals.

2.1 Theoretical Review

Auscultation involves listening to the internal sounds to assess airflow through the trachea and the bronchial tree. With stethoscopes, the bell is generally used to detect high-pitched sounds at the apex of the lungs above the clavicle; its diaphragm is used to detect low-pitched sounds in the rest of the chest. Familiarity with the normal vesicular breath sounds found at specific locations on the chest enables health professionals to identify abnormal sounds, which are often referred to as adventitious. Doctors use the nature of sounds from the internal organs to determine whether there is a case of a respiratory infection or not. The characteristics of sounds such as frequency, pitch, and intensity aid in the identification of the sounds as normal or abnormal. Normal sounds from the trachea are measured from tracheal auscultation (though not performed frequently) when heard at the suprasternal notch or the lateral neck, normal tracheal sounds characteristically contain a large amount of sound energy and are easily heard during the two phases of the respiratory cycle. The frequencies of these sounds range from 100 Hz to almost 5000 Hz, with a sharp drop in power at a frequency of approximately 800 Hz and little energy beyond 1500 Hz [8]. In lung sounds, normal breathing sounds occur in the frequency range 100-1000Hz with an energy drop at 200Hz. Lung sounds are characteristically low-pitched and can be heard during inspiration and a little part of the expiration cycle[8].

Stridors, wheezes, crackles, squawks, and pleural friction rubs are the most common abnormal noise sounds known and detected in most auscultations. As with normal sounds, they are identified by their characteristics of pitch, intensity, and duration of inspiration and expiration. Stridors are characterized by a musical sound with a high pitch produced as a result of turbulent flow passing through a narrow segment of the upper respiratory tract. Stridors are usually intense with a frequency of 500Hz. Wheezes are also musical abnormal sounds with sound energy in the ranges of 10Hz to 1000Hz. It has a relatively longer duration of 100ms [8]. Crackles are short bursts of sounds heard during inspiration and sometimes during expiration. There are two types of crackles namely; fine crackles and coarse crackles. Each of the two has its own characteristics and is caused by certain diseases. Fine crackles are usually heard during mid-inspiration and are caused by diseases such as idiopathic pulmonary fibrosis. On the other hand, coarse crackles occur at early inspiratory stages and are caused by underlying conditions such as bronchiectasis, and asthma. Finally, pleural friction rubs are crackling sounds produced as a result of the visceral pleura rubbing over the parietal pleura. In healthy persons, the parietal and visceral pleura slide over each other silently but lung diseases could make the visceral pleura rough [8]. Essentially, auscultation is a major part of the diagnostic process due to it relatively being a cheaper option than other options such as electrocardiography and radiology. It is a faster means of diagnosis and remains a preferred choice among doctors.

2.2 Related Work

2.2.1 The Electronic Stethoscope

In earlier models of the electronic stethoscope, just sound amplification was considered. Those designs replaced the acoustic sensing piece by introducing microphones that pick up the internal body sounds. Most of the designs in existence then amplified the sounds picked up by an electret condenser microphone. These basic electronic stethoscopes at best increase the volume of the sounds but do nothing to solve the challenges with noise. Different models of the electronic stethoscope that have evolved have incorporated noise cancellation techniques to filter and remove noise from the sounds recorded. Adaptive noise filtering and active noise cancellation are two of the most employed noise cancellation techniques.

There have been a variety of designs for the electronic stethoscope and it would be practically impossible to provide a comprehensive review of all the available systems, thus only a few publications on the design of the electronic stethoscope have been considered in the succeeding paragraphs.

In the work of Derek et al [9], the electronic stethoscope was modeled and built to wirelessly transmit audio sounds from the chest piece to a mobile phone which provides a database for keeping recorded sounds as well as a playback option to listen to recorded signals. This design does an initial pre-amplification and filtering using a second-order Butterworth filter. Consequently, this design has the potential to eliminate the struggles of doctors concerning the length of the tube and the dangers of getting infections through the air-filled tubes. On the other hand, a challenge with this design is the fact that this design does not take away the subjectivity of diagnosis since doctors have to still base their on their own experience proffer diagnosis.

In Chowdhury et al [10], they also investigated the use of an electret microphone coupled to the chest piece to pick up acoustic signals. The signals are pre-amplified and passed to a low pass filter which attenuates the signals to the range of (20-600Hz) before being passed to an anti-aliasing filter before a digital controller. A Bluetooth module is connected to the digital controller to transmit signals to remote processing workstations for further processing using machine learning to detect whether a heart sounds normal or abnormal. The challenge with such a configuration is the audio quality and fidelity transmitted using Bluetooth as compared to using an aux cable. As it concerns auscultation, every audio detail is relevant to proffer an efficient diagnosis thus, this configuration leads to a loss of data.

After comparing the output of the carbon microphone and the electret microphone, Fattah et al [11] decided to use a piezo disc vibration sensor to counter the challenges with sound attenuation and noise levels experienced with the carbon microphone and electret microphone. The signals from the piezo disc sensor obtained as a result of the stress from the acoustic sounds from the body are further processed using machine learning algorithms. Inasmuch as there is probably noise and attenuation of sounds with the use of the electret microphone, [11] did not offer to perform any audio processing such as amplification, among others to compare the signal output from the electret microphone as compared to the signals from the piezo vibration sensor thus leaving this subject inconclusive. This design also has a major setback due to the fact that medical professionals can not have access to the sounds coming from the body because of the use of the piezo disc vibration sensor.

With the application of the finite element method and theoretical analysis, Weidong et al[12] developed a bat-shaped electronic stethoscope based on the Micro Electro-mechanical (MEMS) technology to detect and perform visual signal processing. The core of the research was to improve the sensitivity and the signal-to-noise ratio of the sensing transducers used to pick up heart sounds. In their work, they explored sound signals being transmitted to a microstructure that converts the acoustic signals to electric signals. The microstructure center mass detects the heart sound signal and the vibration of the sound signal causes a deformation in a single cantilever beam. The stress concentration areas of the single cantilever beam are equipped with piezo resistors whose values vary with the deformation in the single cantilever beam. The changes in the piezoresistors which are coupled as a Wheatstone bridge help in converting the acoustic signal to an electrical signal. The electric signal is filtered, amplified, and then transmitted via Bluetooth to a mobile phone or a computer to view the signal spectrograph of the heart sound. In this study as well, Weidong et al failed to make accommodations to offer playback and a recording feature for use by the doctors.

Denoising in electronic stethoscopes is particularly relevant for efficient diagnoses. The real-time denoising electronic stethoscope [13] proposes a filter design based on adaptive line enhancements to actively filter out noise and audio sounds of higher frequency. This is a much more complex but better denoising technique, however, it was developed on an experimental setup with relatively expensive and non-standalone audio codec systems making it unsuitable for use in rural Ghanaian communities.

The electronic stethoscope developed by Jorge et al [14] employed both analog and digital signal processing techniques to detect and maintain the audio fidelity of wireless transmitted recorded signals. This research presents a mobile phone to visualize the results as a phonocardiogram (PCG). Despite the ease of sound visualization, the power consumption of the Bluetooth technology as well as the inability to process and classify pathological heart sounds in real-time makes it unsuitable for my application.

Adaptive line enhancement coupled with bandpass filtering techniques have been used by Lakhe et al [15] in their development of a digital stethoscope. In their research, a first-order high pass filter and then a low pass filter were responsible for preprocessing the recorded sound signals from the body and were employed with further processing using adaptive line enhancement done in Matlab. Through this technique, it proved the advantages of using adaptive line enhancement for electronic stethoscopes but a clear drawback is the inability to process audio signals in real-time and to make classifications on these signals.

Researchers from John Hopkins University have developed a smart Artificial Intelligence (AI)-enabled stethoscope that employs adaptive noise cancellation with the use of another external microphone, and other digital signal processing techniques to mask out all irrelevant noise including sounds from the beating of the heart [16]. This family of devices uses machine learning algorithms to classify whether a patient has pneumonia or not. Preliminary results revealed an 87% success with health workers determining pneumonia cases correctly through the devices from this research. Despite this success, the use of an external microphone introduces more complexity and might not be the optimal adaptive noise-cancellation technique.

2.2.2 Machine Learning with the Electronic Stethoscope

Audio processing and speech recognition have been at the center of research into audio applications. As an extension of the research done in speech recognition and other purely audio recognition applications, predictive diagnoses of heart sounds have been gaining increasing momentum with several researchers trying to accurately predict different types of heart sounds based on features determined from research of other general audio applications.

Due to the highly specialized nature of the field of cardiology, the accuracy ratings of a lot of machine learning models have been due to the presence of labeled data for training the algorithms. However, it does take a lot of work to get this data or audio signals in good condition for the experts to accurately identify. A first step called preprocessing is usually done to ensure all the important features of the audio are kept in place.

Preprocessing in audio processing is very vital to the success of the learning algorithm. Audio signals from the heart are faint and are unfortunately mixed with the sounds from other parts of the body, thus preprocessing helps to mask the unwanted audio sounds through the use of filters and amplifiers. Other applications of preprocessing are data cleaning, audio re-sampling, and feature extraction. Preprocessing involves ensuring all the audio files have the same characteristics in terms of dimensions. Most times, audio signals have different sampling rates and thus have to be re-sampled to a common sampling frequency in order to have uniformity in the data and just generally, ensure a good representation of the data. Finally, preprocessing allows for extracting important features of audio signals from recorded heart sounds. Mel Frequency Cepstrum Coefficients (MFCCs), spectral centroid, spectral entropy, zero-crossing rate, and energy density are important primary characteristics of audio signals that are used by the machine learning algorithms, thus it is essential to apply good feature engineering techniques through preprocessing to acquire the best features for the machine learning algorithm. However, this is only used in traditional supervised machine learning algorithms.

Due to the large number of studies on heart sound classification using machine learning that have appeared in literature, it is practically impossible to reference all the studies. Thus, only a few studies using various machine learning algorithms have been discussed in succeeding paragraphs.

Convoluted Neural Networks have been used in classifying heart sounds of COVID-19 patients [17]. Using features extracted from Short Time Fourier Transforms (STFTs), MFCCs and fused features of the two from a dataset of recorded audio sounds from 126 patients, a Convoluted Neural Network (CNN) was built to classify heart sounds. However, CNNs are computationally taxing [18] making it unsuitable for use in low-powered and less powerful devices. The features used in this approach are limiting and do not account for some low-level properties of heart audio signals.

The Optimized ensemble algorithm [10] was used to classify heart signals of patients based on twenty-seven(27) time-frequency extracted features. This study involved imbalanced data of normal heart sounds against abnormal heart sounds which made the classification inaccurate to an extent. Again, the test data was just audio signals from six patients which was not enough to fully ascertain the accuracy of the results of this research.

Reinforcement Learning and Convolutional Recurrent Neural Networks [19] have also been used for the classification of pathological heart sounds recorded in 500 auscultation procedures. This study looked at building a fully interactive auscultation process that determines the most effective parts of the body to perform auscultation and then classify the sounds as needing attention or not. This novel method provides a good way to promote telemedicine through at-home auscultation but as with neural networks, it is computationally taxing to use images in classification especially when it is to be deployed to resource-constrained locations.

Utilizing an ensemble of neural networks, Zabihi et al [20] developed a classification algorithm for detecting normal and abnormal heart sounds based on phonocardiograms. Using audio recordings from one of the largest public data, their algorithms achieved an accuracy of 91.5%. However, the audio sounds were already pre-recorded thus, their work involved only the classification of the heart sounds.

Potes et al [21] in addressing the physioNet challenge [22] used an Adaboost-abstain classifier and a CNN to classify heart sound phonocardiograms. This method allowed a comparison of the results from the Adaboost-abstain classifier as well as the Convoluted neural network in a final decision rule algorithm to determine if a heart sound was normal or abnormal. This method however did not include any signal acquisition which makes it incomplete for our use case.

Homsi et al utilized an ensemble of algorithms to classify normal and abnormal heart sounds. By nesting Random Forests (RFs), LogitBoosts (LBs), and a Cost-Sensitive Classifiers (CSCs), Homsi et al yielded an 84.4% accuracy in their classifier [23].

Support Vector Machines (SVMs) were used by Goda et al [24] in their heart classification method during the physioNet challenge [22]. By extracting both time and frequency domain features of the heart audio sounds, the SVMs obtained a score of at least 86.6% on the various training sets.

2.3 Technical Component Review

2.3.1 Microphones

A microphone is a transducer that converts sound waves into electrical signals. Microphones can be classified according to the type of transducer used in the construction or based on the directionality properties of picking up sound [25].

Based on construction, microphones can be classified into the following :

- Condenser
- Dynamic
- Ribbon
- Piezoelectric
- Fiber Optic
- Laser Microphone
- Micro Electro-mechanical system

Based on the directionality properties of picking up sound :

- Unidirectional

- Omnidirectional
- Bi-directional

In medical applications, the most commonly used microphones are the condenser and the MEMS. The frequencies of lung sounds are very small ranging from 20Hz to 650Hz. Thus, this requires a microphone that has sensitivity within this range, a low signal-to-noise ratio, and can measure little variations in sound pressure.

The condenser satisfies these criteria and would thus be used in this project. The condenser microphone has a very thin fixed plate called the diaphragm that has two sides that are positively and negatively charged respectively. The pressure from the sound waves causes the vibration of the diaphragm. This back-and-forth movement of the diaphragm results in a change in the distance between the two oppositely charged sides of the thin polymer plate which results in a change in the capacitance of the circuit. Hence, an electrical signal is generated from the pressure of the sound waves [25]. Condenser microphones have a uniform frequency response and can respond with clarity to transient sounds, and as such are a good choice to use in the electronic stethoscope because of the varying frequencies at which sound energy drops abruptly in some normal lung sounds and some adventitious sounds. Lastly, the low diaphragm design of the condenser microphone gives it an extended high-frequency while complementing it with the ability to pick low-frequency sounds which is very applicable to pickup low-frequency lung sounds [25].

Other considerations for a choice of microphone for use in the electronic stethoscope include directionality, sensitivity, and power consumption of the microphone. On these criteria, an omnidirectional condenser microphone offers a better and more reliable acoustic sound than a directional sound. In a more in-depth explanation, the directionality describes how the sensitivity of the microphone changes when the sound source changes position in space. An omnidirectional microphone means that the microphone is equally sensitive in all directions, unlike the directional microphones where the sensitivity is unequal in different locations of the perceived sound's location. The choice of an omnidirectional microphone allows for the stethoscope to be positioned at different places in order to obtain the best sound without compromising the quality of the sound [26].

The sensitivity of a microphone describes the ability of the microphone to pick up faint sounds. A note here would be that the sensitivity of the microphone does not determine the quality of a microphone. It only gives information about the characteristics of the microphone and as such should be coupled with another metric such as the Signal-to-Noise (SNR) ratio of the microphone as it relates to the particular application [27]. In the condenser microphone, the sensitivity values vary from -46dB to -35dB [27] and the SNR values may vary according to the manufacturer. However, a good rule of thumb would be to use microphones that have an SNR greater than 60 dBA.

2.3.2 Power

The microphone can be powered via a 5-12V unregulated external power supply. The power source is automatically selected as the highest voltage source.

2.4 Audio Preprocessing

2.4.1 Zero Padding

Heart audio signals that vary in length need to be smoothened to have uniform length. Zero padding consists of extending a signal (or spectrum) with zeros. The addition of zeros to the end of the time-domain waveform does not improve the underlying frequency resolution associated with the time-domain signal. Thus, zero padding provides us with a way to create uniformity in our signals without introducing any artifacts or unnecessary signals in the original signals. The database of heart sounds described in section 3.1 contains audio signals of lengths from 6 seconds to 30 seconds. As such, all heart audio signals less than 30 seconds were zero-padded to 30 seconds.

2.4.2 Filtering

Audio filters are designed to cut off particular frequencies of sound that are irrelevant to the application being developed. There are several filters which include high-pass filters, low-pass filters, and band-pass filters. In this particular case, the heart sounds are very low-frequency components and as such require a filter that allows only low-frequency components to pass.

2.4.2.1 FIR and IIR Filters

Filters can be classified into finite Impulse Response and Infinite Impulse Response filters based on the nature of the response of the filtering technique.

Finite impulse response filters take an input signal and convolve it with an impulse function to produce an output signal. In principle, the FIR filter has a finite duration and is always a function of the current and past samples of the input signal. FIR filter are described by equation 2.1

$$y(n) = \sum_{k=0}^{N} b(k)x(n-k) = \sum_{k=0}^{N} h(k)x(n-k) \tag{2.1}$$

IIR filters are recursive filters that have internal feedback and may continue to respond indefinitely. A recursive filter is described by equation by equation 2.2. Any non-zero

$$a_k$$

results in a recursive IIR filter.

$$\gamma(n) = \sum_{k=1}^{N} a_k \gamma(n-k) + \sum_{m=0}^{M} b_m x(n-m) \tag{2.2}$$

The two most commonly used low-pass filters in audio processing applications are the Butterworth filters and the Chebysev filters.

2.4.2.2 Butterworth Filters

Butterworth filters are low-pass filters that are characterized by the frequency response represented in equation 2.3

$$|H(J\omega)| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}}} \tag{2.3}$$

The image below shows the frequency response of various nth-order ideal Butterworth filters.

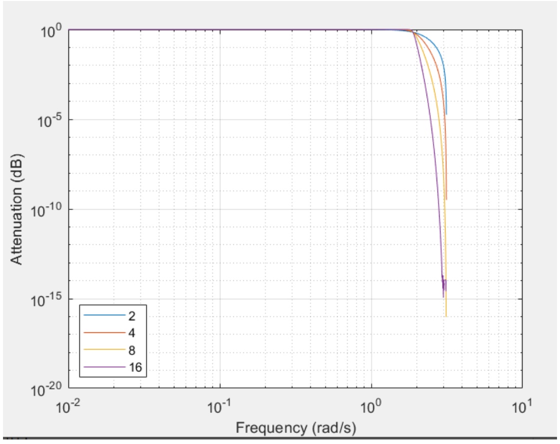


Figure 2.1: Frequency response of nth-order Butterworth filter

2.4.3 Amplification

As mentioned in Chapter One, a key limitation of acoustic stethoscopes is the low intensity of the sound measured. With the introduction of microphones, the conversion of mechanical sound waves to electrical signals allows for magnifying the electrical signals to appropriate sound levels. Amplification is an audio processing technique that allows us to increase sound levels.

2.4.4 Feature extraction

Feature extraction is an important step in audio processing tasks. In general, feature extraction is an essential processing step in pattern recognition and machine learning tasks. The goal of this step is to extract a set of features from the dataset of interest. These features must be informative with respect to the desired properties of the original data. Feature extraction can also be viewed as a data rate reduction procedure because the analysis algorithms are developed based on a relatively small number of features. In this research, the original data, which is the audio signal, is voluminous and as such, it is hard to process directly in any analysis task. Therefore, there is the need to transform the initial data representation to a more suitable one, by extracting audio features that represent the properties of the original signals while reducing the volume of data. In audio and speech recognition or classification tasks [24][17], it is important to consider the time as well as the spectral characteristics of the signal. In that regard, some common features that have been used for various classification or recognition tasks include; zero crossing rate, spectral energy, spectral entropy, spectral centroid, spectral spread, MFCCs, and sample kurtosis. The characteristics of some of these features are discussed in the paragraphs below.

2.4.4.1 Mel Frequency Cepstral Coefficients

MFCCs are a small set of features that concisely describe the overall shape of a spectral envelope. MFCCs was originally suggested for identifying monosyllabic words in continuously spoken sentences but not for speaker identification. MFCC computation is a replication of the human hearing system intending to artificially implement the ear's working principle with the assumption that the human ear is a reliable speaker recognizer [28]. MFCC features are rooted in the recognized discrepancy of the human ear's critical bandwidths with frequency filters spaced linearly at low frequencies and logarithmically at high frequencies to retain the phonetically vital properties of the speech signal [29]. In other words, the Mel scale is a perceptually motivated scale of frequency intervals, which, if judged by a human listener, are perceived to be equally spaced. Equation 2.4 gives a representation of how the human ear works according to the Mel scale.

$$f_{\mathrm{mel}}(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{2.4}$$

In order to extract the mel frequency cepstral coefficients from a signal frame, the discrete Fourier transform of the signal is taken given by equation 2.5.

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp\left(-j\frac{2\pi}{N}kn\right), \quad k = 0, \ldots, N-1 \tag{2.5}$$

Mel-Spectrum is computed by passing the Fourier-transformed signal through a set of band-pass filters known as mel-filter bank. A mel is a unit of measure based on the human ear's perceived frequency. It does not correspond linearly to the physical frequency of the tone, as the human auditory system apparently does not perceive pitch linearly. The mel scale has approximately a linear frequency spacing below 1kHz and a logarithmic spacing above 1kHz. The approximation of mel from physical frequency can be expressed as given in equation 2.4 where f denotes the physical frequency in Hz, and mel denotes the perceived frequency. For MFCCs computation, filter banks are generally implemented in the frequency domain. The center frequencies of the filters are normally evenly spaced on the frequency axis. However, in order to mimic the human ear perception, the warped axis according to the nonlinear function given in Equation 2.4 is implemented [30][31]. The Mel spectrum of the magnitude spectrum $X(k)$ is computed by multiplying the magnitude spectrum by each of the triangular mel weighting filters (equation 2.6)

$$s(m) = \sum_{k=0}^{N-1} \left[|X(k)|^2 H_m(k)\right]; \quad 0 \leq m \leq M-1 \tag{2.6}$$

where M is the total number of triangular mel weighting filters [32][31]. $H_m(k)$ is the weight given to the kth energy spectrum bin contributing to the mth output band and is expressed as in equation 2.7 with m ranging from 0 to M-1

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \frac{2(k-f(m-1))}{f(m)-f(m-1)}, & f(m-1) \leq k \leq f(m) \\ \frac{2(f(m+1)-k)}{f(m+1)-f(m)}, & f(m) < k \leq f(m+1) \\ 0, & k > f(m+1) \end{cases} \tag{2.7}$$

The sum of the signals from various filters in the filter bank is passed through a non-linear rectifier and the Discrete Cosine Transform (DCT) of the signal is then taken to obtain only real-valued coefficients of the signal. Since most of the signal information is represented by the first few MFCC coefficients, the system can be made robust by extracting only those coefficients ignoring or truncating higher-order DCT components [33]. Finally, MFCCs is calculated as in equation 2.8 where $C(n)$ are the cepstral coefficients and C is the number of MFCCs. Traditional MFCCs systems use only 8–13 cepstral coefficients. The zeroth coefficient is often excluded since it represents the average log energy of the input signal, which only carries little signal-specific information

$$c(n) = \sum_{m=0}^{M-1} \log_{10}(s(m)) \cos\left(\frac{\pi n(m-0.5)}{M}\right); \quad n = 0, 1, 2, \ldots, C-1 \tag{2.8}$$

2.4.4.2 Spectral Flux

Spectral flux measures the spectral change between two successive frames and is computed as the squared difference between the normalized magnitudes of the spectra of the two successive short-term windows:

$$Fl_{(i,i-1)} = \sum_{k=1}^{W_{LL}} \left(EN_i(k) - EN_{i-1}(k)\right)^2 \tag{2.9}$$

where $EN_i(k)$is the kth normalized DFT coefficient at the ith frame.

$$EN_i(k) = \frac{X_i(k)}{\sum_{l=1}^{W_{f_L}} X_i(l)} \tag{2.10}$$

2.4.4.3 Spectral Rollof

Spectral Rollof is defined as the frequency below which a certain percentage (usually around 90%) of the magnitude distribution of the spectrum is concentrated. Therefore, if the mth discrete Fourier transform coefficient corresponds to the spectral roll-off of the ith frame, then it satisfies the following equation 2.11

$$\sum_{k=1}^{m} X_i(k) = C \sum_{k=1}^{Wf_L} X_i(k) \tag{2.11}$$

where C is the adopted percentage (user parameter). The spectral roll-off frequency is usually normalized by dividing it with W_{fL} so that it takes values between 0 and 1. This type of normalization implies that a value of 1 corresponds to the maximum frequency of the signal, i.e. to half the sampling frequency.

2.4.5 Feature Selection

Feature Selection refers to the process of choosing the most relevant features of a signal that are the strongest predictors in an algorithm. Some features are redundant in improving the performance of classification algorithms whilst other features are the dominant predictors of the algorithm and as such it is important to be able to reduce the dimensionality of the feature matrix to allow only the most important features to be fed into the classification algorithm. The main benefits of feature selection are to improve prediction performance, provide faster and more cost-effective predictors, and provide a better understanding of the data generation process [34]. Several dimensionality reduction techniques exist, notably Principal Component Analysis, Ratio of missing values, Backward feature elimination, and Neighborhood Component Analysis.

2.4.6 Hyperparameter Tuning

Machine learning algorithms like ensemble models, decision trees, regression models, and neural networks for regression and classification tasks typically have a number of hyperparameters that need to be set before running them. Hyperparameters determine the performance of the model to a great extent and as such a great emphasis is placed on selecting the most optimal hyperparameters such as the learning rate, and the box level constraint among others for various classification estimators. Several methods have been developed to enable the selection of the most optimal hyperparameters for training machine learning models. The three most popular ones are; Grid search, Random search, and Bayesian Optimization [35].

2.5 Windowing

The Fast Fourier Transform (FFT) is based on the fact that the signal being measured is finite data. The underlying assumption is that the signal is periodic and has an integer number of periods. In cases where this assumption holds, applying the FFT directly works. However, for signals that do not have an integer number of periods, the finiteness of the signal may result in truncation. Windowing is thus the technique to reduce the sharp discontinuities that may be introduced by the FFT [36]. By multiplying the time signals with a finite-length window with an amplitude that varies smoothly and edges to zero, the technique is able to eliminate the undesirable characteristics of the FFT on a non-integer number of period signals. Several windowing functions exist. The most popular ones include; the Hamming window, Hanning window, Blackman-Harris Window, and Kaiser-Bessel window [37].

2.6 MATLAB

MATLAB is a programming platform developed by Mathworks to help engineers and scientists analyze and design systems and products. The platform uses the MATLAB programming language, a matrix-based language for computational mathematics. The MATLAB platform provides a lot of toolboxes and in-built functions for designing digital filters, developing and analyzing machine learning algorithms, and visualizing, plotting, and extracting the features of audio signals. Essentially, the MATLAB platform allows to move the processing to the digital domain [38].

3 CHAPTER THREE - METHODOLOGY

The main hypothesis of this thesis was that recorded heart sounds can be classified using machine learning as normal or abnormal. This chapter is split into hardware and software techniques which were taken to evaluate the above proposition and the implementation techniques that were used to build the system. Before discussing the two-component sections, a broad description of the system block diagram is given.

3.1 Block Diagram

The block diagram represents how the various parts of the system are connected to have a fully operational system. The main components of the hardware section of the electronic stethoscope are the microphone which will convert the sound into electrical impulses, an amplification unit to increase the level of the sound, power unit to provide the power hardware system. In the software section, 2nd order butterworth digital filters have been designed for further preprocessing to ensure sound clarity, machine learning algorithms for classifying heart sounds, and data visualization techniques to enable medical professionals or users to better understand the audio data. Noteworthy to mention is the fact that, the software section allows the audio signals to be recorded which can be accessed later on for further analysis. Figure 3.1 below shows how these moving parts are connected together to give a complete working system.

3.2 Hardware Details

The hardware section of the electronic stethoscope is important as it is the primary data acquisition source. Without this part, heart sound audio signals will have to be acquired from other sources in other to utilize the system proposed in this thesis. Thus, in the para-

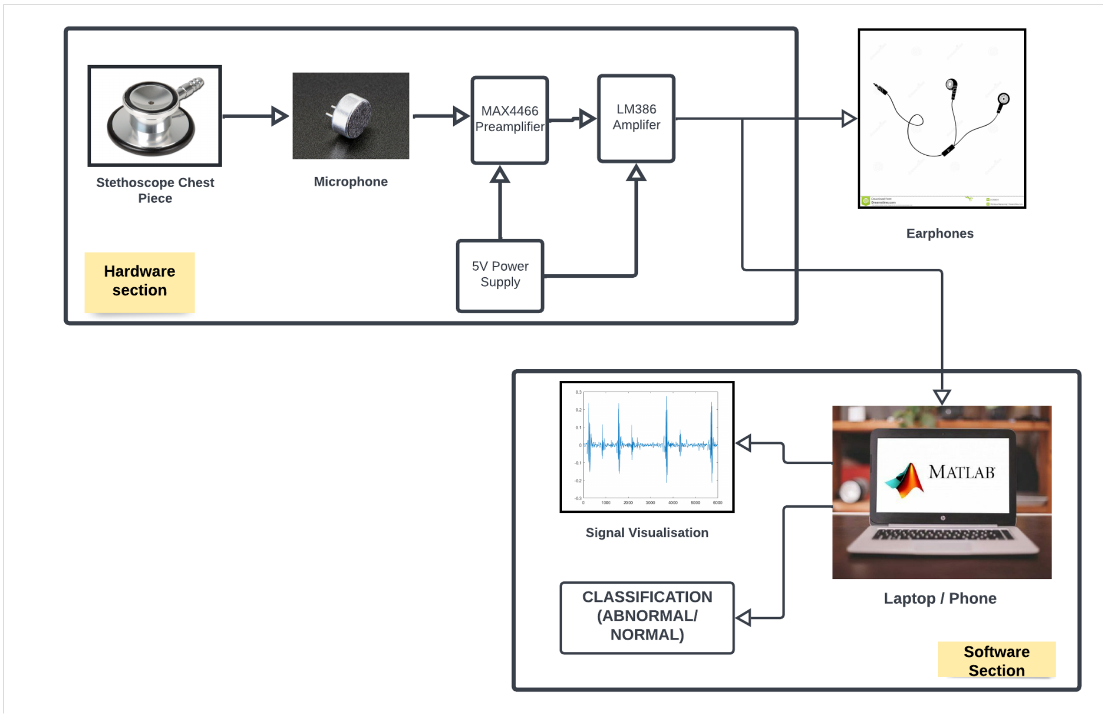


Figure 3.1: Block diagram of digital stethoscope

graphs below, the various hardware components and how they are used in this project will be discussed.

3.2.1 Electret Condenser Microphone

Data acquisition using a sensitive sound sensor with a frequency response in the audio range is the first stage of the stethoscope design. An Electret Condenser Microphone (ECM) transducer was employed to achieve this purpose. Low amplitude values in the 1–5 mV range are typical of heart sounds produced by ECM transducers on stethoscope heads. The ECM is made of the electret which eliminates the need for a polarizing voltage to operate the microphone. In this project, the electret condenser microphone is omnidirectional, has a sensitivity value of -44±2dB, a signal-to-noise ratio of 60 dBA, and a frequency response range of 20-20000Hz.

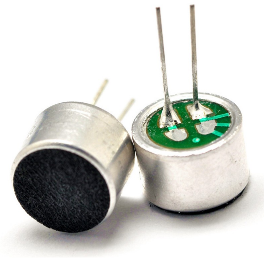


Figure 3.2: Electret Condenser Microphone

3.2.2 MAX4466 Pre-amplifier

The MAX4466 is a micropower operational amplifier designed specifically for microphone preamplifiers. They combine an optimized gain bandwidth product versus supply current with low voltage operation in ultra-small packages to deliver the best of both worlds. The MAX4466 is decompensated to create a 600kHz gain bandwidth product with a minimum stable gain of +5V/V. This amplifier also has Rail-to-Rail® outputs, a high Open Loop Voltage gain (AVOL), an outstanding power supply, and common-mode rejection ratios for use in noisy situations.

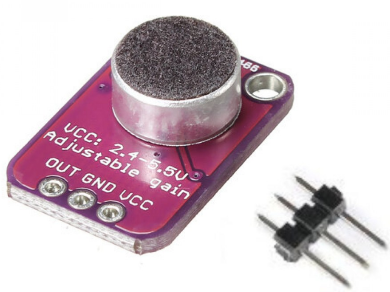


Figure 3.3: MAX4466 Pre-amplification module

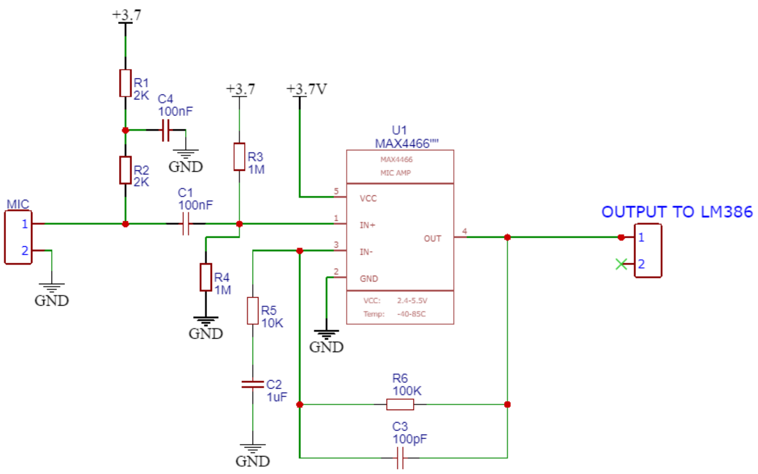


Figure 3.4: Circuit diagram of Preamplification

3.2.3 LM386 Audio Amplifier

The LM386M-1 and LM386MX-1 are power amplifiers designed for use in low-voltage consumer applications. The gain is internally set to 20 to keep the external part count low, but the addition of an external resistor and capacitor between pins 1 and 8 will increase the gain to any value from 20 to 200. The inputs are ground referenced while the output automatically biases to one-half the supply voltage. The quiescent power drain is only 24 mW when operating from a 6-V supply, making the LM386M-1 and LM386MX-1 ideal for battery operation.

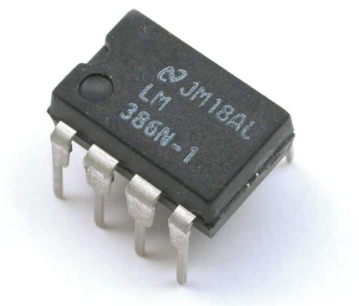


Figure 3.5: LM386 amplifier

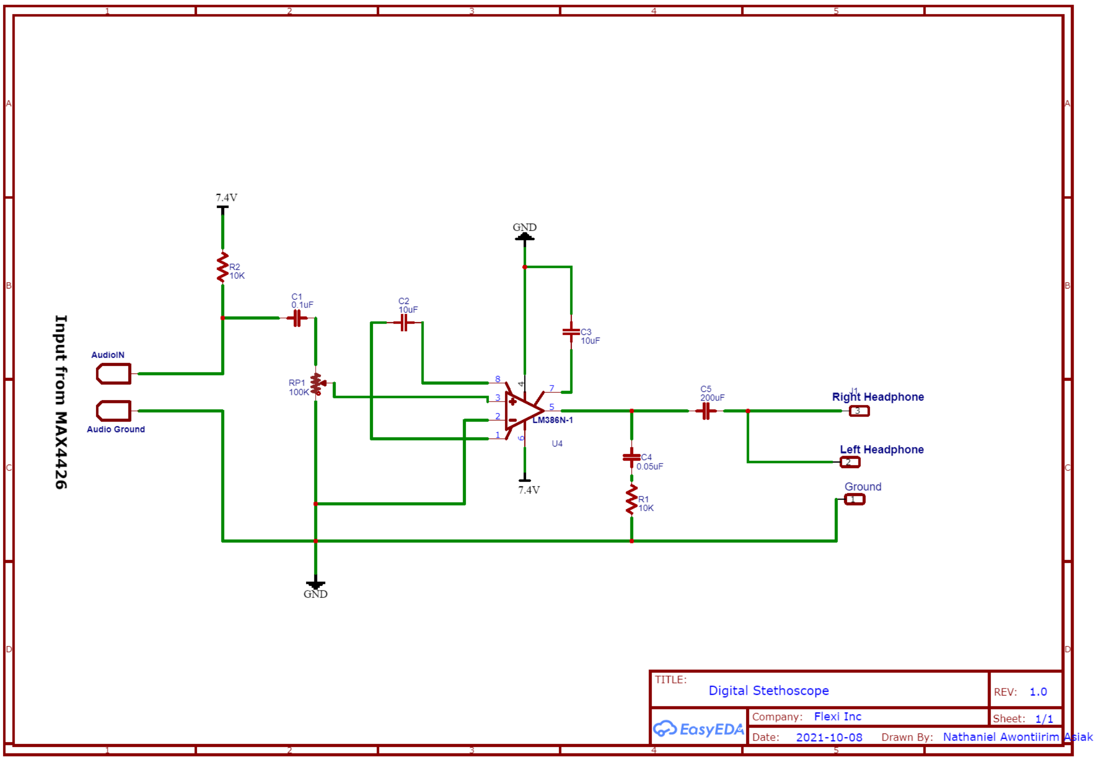


Figure 3.6: Amplification circuit using LM386

3.3 Software- Machine Learning Techniques

The core of the software section of the electronic stethoscope proposed in this thesis is the ability to classify heart sounds as normal or abnormal. In that regard, the succeeding paragraphs will discuss the various processing, data formatting, and machine learning models utilized in this project.

3.3.1 Datasets

Two datasets were used for this research with a collective dataset of over 5000 audio samples recorded in the wav file format. The first dataset which includes 3240 audio samples was obtained from The PhysioNet/Computing in Cardiology Challenge 2016 [39]. These recordings were made with a variety of instruments in both clinical and nonclinical settings (such as in-home visits). The recordings ranged in length from a few seconds to several minutes. Subject demographics (age and gender), recording information (number per patient, body location, and length of recording), synchronously recorded signals (such as ECG), sampling frequency, and sensor type are among the additional data offered. The second dataset was obtained from recordings by Jordanian researchers using 3M™ Littmann® Electronic Stethoscope model 3200 of 112 patients(35 healthy and 77 unhealthy) at the King Abdullah University Hospital in Ramtha, Jordan [40]. Each recording lasts between 5 and 30 seconds, which is long enough to cover at least one breathing cycle. The labeling of the data showed the type of filter mode when the sound was recorded. The letter B is associated with Bell mode filtration, which amplifies sounds in the [20-1000]Hz range while emphasizing low-frequency sounds in the [20-200]Hz region. The letter D is associated with Diaphragm mode filtration, which amplifies sounds in the [20-2000]Hz range while emphasizing frequencies in the [100-500]Hz region. The letter E is associated with extended mode filtration, which enhances sounds in the [20-1000]Hz range while emphasizing frequencies in the [50-500]Hz region. The datasets were split according to table 3.1 for training, validation, and testing of the binary classification algorithm.

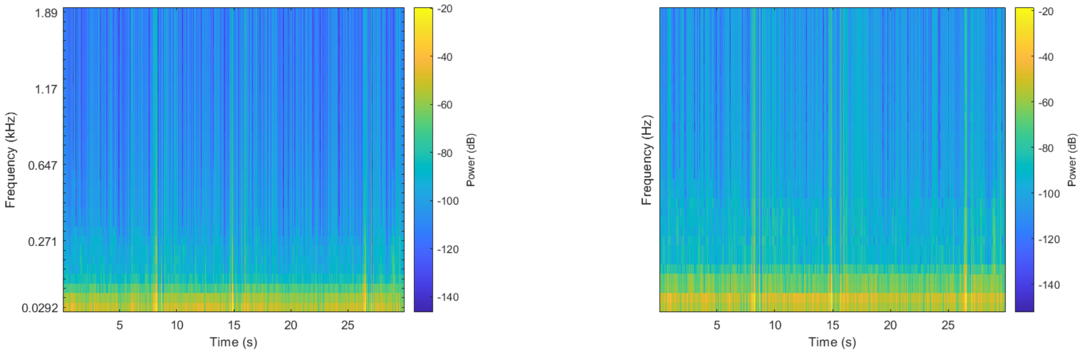
Table 3.1: Summary of Audio Dataset

Database	#Patients	#Recordings	Normal	Abnormal	Unsure
PhysioNet Challenge	764	3240	2365	586	289
Database 2	112	112	35	77	0

3.3.2 Pre-processing

3.3.2.1 Downsampling

In audio classification, large datasets improve the performance of algorithms. However, the size of audio data is limited by the resolution of the audio signals and the amount



(a) Mel spectrogram of an original audio signal

(b) Mel spectrogram of a downsampled audio

Figure 3.7: Comparison of Mel spectrogram of a downsampled audio signal vs an original audio signa

of available memory capacity; The higher the resolution of the audio files, the more memory they require to train the models. Also, audio is highly dimensional and contains redundant and often unnecessary information. As a result, high-resolution audio files are usually downsampled to lower resolutions to reduce the amount of memory consumption. The amount of downsampling usually depends on the problem domain and the original resolution of the audio files. That said, the downsampled audio files should remain meaningful to the problem. Thus, the downsampled audio files should maintain the characteristics of the original images which are relevant to the classification task. The audio files for the binary heart sound classification had sampling rates of 4000Hz (Physionet Database), and 2000Hz(Database 2), a bit resolution of 16Hz, and an average record time of 30 seconds. However, the important details of heart sounds occur in frequencies up to 1000Hz.

3.3.2.2 Filtering

A second-order butterworth filter with a a frequency range between 20-400Hz was designed for this application. The cut-off frequency of 400Hz was utilized to remove high-frequency noise artifacts from the audio signal. As mentioned in the literature review of heart sounds, the highest abnormal heart sounds occur at a maximum frequency of 400Hz and thus it is relevant to ensure frequencies higher than this cut-off frequency are removed. Noteworthy is the fact sudden spikes in the signals were also removed using an algorithm developed by [41] before further processing tasks were carried out.

3.3.3 Feature Transformation

3.3.3.1 Feature Extraction

Feature extraction as described in section 2.4.3 allows to obtain useful information about the audio signals that could be used in modeling the classification algorithms. As described in section 2.4.3, audio features can be classified into time domain, frequency domain, and time-frequency domain features. Statistical features such as the standard deviation, mean value, median value, and other statistical distributions offer a very good representation of the time-varying characteristics of the audio signal. Thus, nine(9) statistical features were extracted as part of the preliminary feature selection process. Further, five(5) frequency domain features were extracted to represent the signal variation in the frequency domain. Finally, the first thirteen(13) MFCCs were selected to represent the time-frequency domain variation of the audio signal. Thus, twenty-six(27) features were extracted in the preliminary feature selection process for each audio file. However, each feature was chosen across 5-second hamming windows, thus introducing the 6x27 dimensionality for each audio signal.

3.3.3.2 Feature Selection : Neighborhood Component Analysis

Neighborhood Component Analysis (NCA) is an automated approach for selecting a small subset of features that carry information most relevant to the classification task while minimizing redundancy among selected features. It is a non-parametric method for selecting features with the goal of maximizing the prediction accuracy of regression and classification algorithms. The NCA algorithm generated 14 features that contributed most to the performance of the algorithm.

3.3.4 Heart Sound Classification Algorithms

3.3.4.1 Ensemble Classifiers

The fundamental principle of ensemble methodology is to combine a number of models, each of which performs the same original problem, to produce a superior composite global model with more precise and dependable estimates or choices than a single model can provide.

3.3.4.2 Hyperparameter Optimization Method

The Bayesian optimization method was chosen to select the hyperparameters of the algorithms. Bayesian optimization is a popular reprocessing algorithm that bases future assessment levels on existing results. Bayesian optimization uses two main components to calculate the hyperparameters: a surrogate model and an acquisition function. The surrogate model seeks to match all of the points in the objective function that can now be viewed. The acquisition function controls how various points are used, balancing exploration and exploitation. The Bayesian optimization model balances the search and use processes to find the optimal location while avoiding losing the best configuration in undeveloped areas.

3.4 Implementation

3.4.1 System Architecture

The application has two major subsystems. The hardware subsystem and the software subsystem. The requirements of the two parts are different as the hardware subsystem deals with the physical realization to acquire the heart sounds and the software subsystem deals with designing digital signal processing methods to put the signals in the appropriate shape and implementing machine learning algorithms to classify the heart sounds. The two subsystems are discussed in 3.5.0.2 and 3.5.0.3 respectively.

3.4.2 Hardware Subsystem

EasyEDA was used to design the circuits for the hardware implementation. The hardware design process involved an initial 3D modeling of the stethoscope in Solidworks to ensure all the parts fit properly. This first step was important in preventing audio leakage and ensuring the electronic components were well packaged. The majority of the hardware system electronics assembly was 3D printed. The packaging unit is separated into two compartments; the power and charging compartment which contains the batteries and the battery chargers, and the amplification unit which contains the microphone assembly, the amplification circuit, and the 3.5mm output jack to the speakers or headphones.

3.4.2.1 Hardware Implementation resources

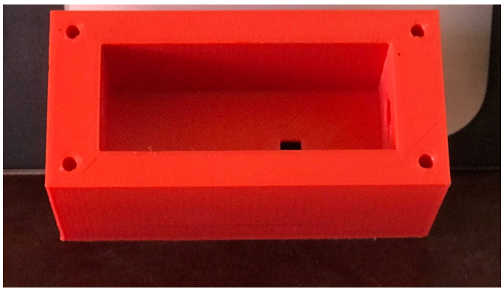
3.4.2.1.1 EasyEDA

EasyEDA is a web-based Electronic Design Automation tool package that allows hardware engineers to create, simulate, exchange, and debate schematics, simulations, and printed circuit boards both publicly and privately. EasyEDA provides for the construction and editing of schematic diagrams, Simulation Program with Integrated Circuit Emphasis(SPICE) simulation of mixed analog and digital circuits, and the design and editing of printed circuit board layouts, as well as the fabrication of printed circuit boards if desired. It has a vast and expanding user base.

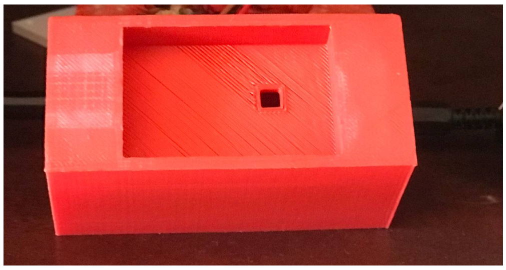
3.4.2.1.2 Ender-3 3D Printer

The Ender-3 3D printer was used to print the casing for the assembly of the electronics for the hardware. This 3D printer is a resource provided by the school in the CNC room of the engineering workshop. The models were designed in Solidworks and then sliced before printing using the ender-3 3D printer.

Figure 3.8: Ender-3 3D printer



(a) Amplification compartment of 3D printed packaging unit



(b) Power compartment of 3D printed packaging unit

Figure 3.9: 3D printed model of packaging unit

3.4.2.2 Design and Construction of hardware subsystem

3.4.2.2.1 Power Unit

The power unit of the electronic stethoscope was designed to utilize two lithium polymer batteries rated at 3.7 volts each. The polymer batteries were connected in a series of parallel configurations to be able to meet the different power requirements of the amplifier and pre-amplifier circuits. Figure 3.10 shows the construction of the power circuit of the device.

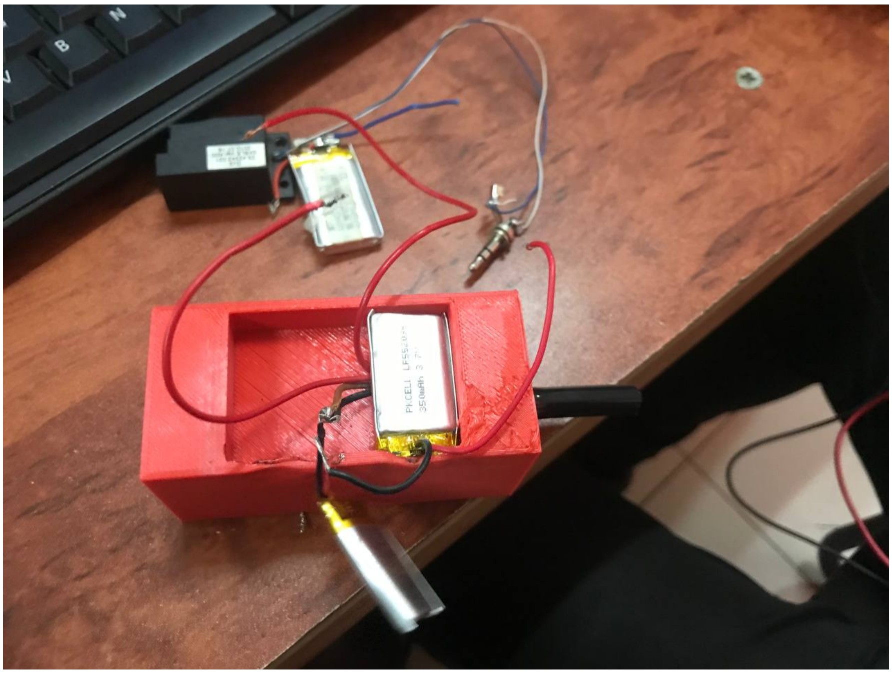


Figure 3.10: Battery power connection

3.4.2.2.2 Amplification Unit

The amplification compartment contains the amplifier, pre-amplifier, and the 3.5mm audio output jack to the speaker or headphones. The microphone assembly is carefully fitted into a hole that has a protruding end that fits into the chest piece of the stethoscope. Figure 3.11 shows the assembling stage of the amplification unit.

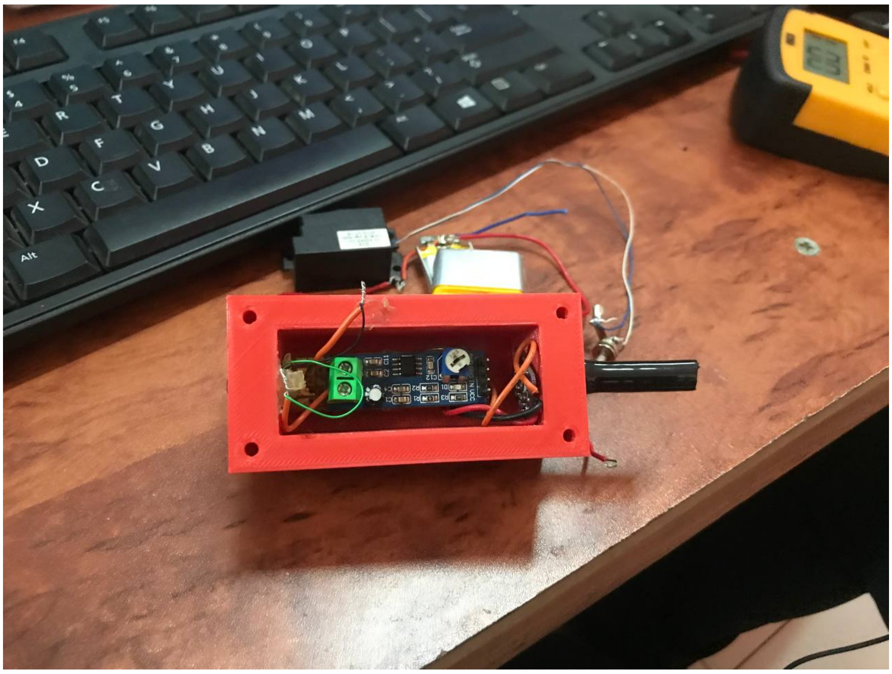


Figure 3.11: Amplification unit connection

3.4.3 Software subsystem

The software application has two major modules : Training and Testing modules. These modules are different in the functions they play in determining the classification of heart sounds but are not mutually exclusive entirely. However, different behavior is expected for the application under the two modules. A flowchart of the application is shown in Figure 3.12.

3.4.3.1 Implementation Resources

The software system was designed primarily with 3 toolboxes from the Matlab appli-

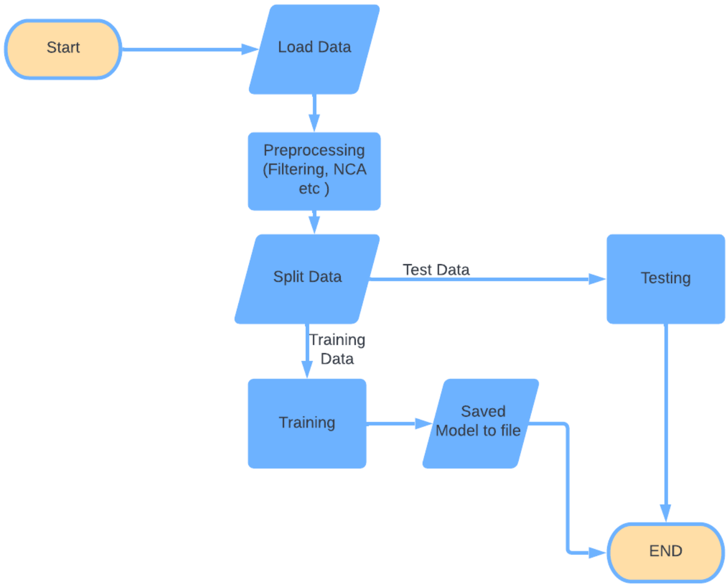


Figure 3.12: Flowchart of the software subsystem

cation developed by Mathworks Inc. These toolboxes provide intuitive ways of understanding concepts and allow for tuning the models until the desired results are obtained. Adopting these toolboxes described below reduced the boilerplate code required to preprocess signals, training, and testing algorithms.

3.4.3.1.1 Matlab Statistics and Machine Learning Toolbox

The Statistics and Machine Learning Toolbox is a collection of functions and programs for describing, analyzing, and modeling data. For exploratory data analysis, descriptive statistics, visualization, and clustering may be used to fit probability distributions to data. In addition, the Classification and Regression Learner application enables inferences to be drawn from data and develop prediction models interactively or programmatically using regression and classification techniques.

The Classification Learner software teaches data classification to models. This program allows you to experiment with supervised machine learning using a variety of classifiers. It allows us to explore the data, choose features, set up validation schemes, train models, and evaluate the outcomes. Decision trees, discriminant analysis, support vector machines, logistic regression, closest neighbors, naive Bayes, kernel approximation, ensemble, and neural network classification may all be automated to find the optimum classification model type.

3.4.3.1.2 Matlab Application Designer

The Matlab App Designer is a development environment that allows for the creation of an interactive graphical user interface application's appearance and program functionality. It includes a fully integrated MATLAB Editor as well as a huge number of interactive User Interface components. It also includes a grid layout manager to help structure the user interface, as well as automatic reflow, features that allow the application to detect and adjust to changes in screen size. The applications developed in this environment can be published by directly packaging them into installation files from the Program Designer toolstrip, and deploying them as a standalone desktop application or online as a web application.

3.4.3.2 Training

Several classification models were experimented with in this thesis, however, this section will discuss only the final models selected to test the data. The data was trained using the MATLAB classification learner because it provided a single interface window to compare the performance of various classifiers as well as allowing for hyper-parameter tuning. All models and feature extraction were done utilizing the parallel computing toolbox in MATLAB. The parallel computing toolbox allows for tasks like parameter sweeps, optimizations, and Monte Carlo simulations, to run independent iterations in parallel on multicore CPUs, thus making it faster. Noteworthy is the fact that MATLAB file datastore was used to store the databases in order to save RAM for processing tasks and training algorithms. The datastore function creates a datastore, which is a repository for collections of data that are too large to fit in memory.

3.4.3.3 Binary Classification Using Ensemble

The binary classification using the Adaboost ensemble algorithm was also trained on 3000 audio signals. Preprocessing tasks such as zero padding and filtering were conducted on the audio signals before feature extraction. 27 features were extracted and used to train the model initially but neighborhood component analysis was utilized in feature selection to determine the most important features that contribute to the performance of the algorithm. Subsequently, the Adaboost algorithm was trained on the following 14 carefully selected predictors; sample kurtosis, signal entropy, dominant frequency, first 10 mel frequency cepstral coefficients, and the 13th mel frequency cepstral coefficient.

The dataset was imbalanced with the number of normal heart sounds exceeding the abnormal heart sounds in the ratio of 3:1. Also, in heart sound classification, it is more expensive to classify an abnormal heart sound as normal than to classify a normal heart sound as abnormal. In this regard, a misclassification cost of 10 for false negatives was introduced during the training of the algorithm.

4 CHAPTER FOUR - RESULTS AND DISCUSSION

The experiments that were undertaken to improve the performance of the models, as well as the outcomes of those trials, are discussed in this chapter. It's worth noting that this Chapter solely discusses the experiments that led to the final model and the findings of the study.

4.1 Complete Assembly of Electronic Stethoscope

The amplification unit and the power unit discussed in chapter three were assembled to give a complete working unit with an 8-0hm output speaker connected. The setup was able to amplify the sounds picked from the chest greatly as compared to using the earpiece from the acoustic stethoscope. Figure 4.3 shows the complete assembly of the electronic stethoscope used for testing. Figures 4.1 and 4.2 show the signal at the input and outputs of the amplifier respectively. As noticed, the signal is noisy, and as such the output amplified also contains a lot of noise. However, the graphs show that the input signal is amplified when passed through the LM386N-1 audio power amplifier.

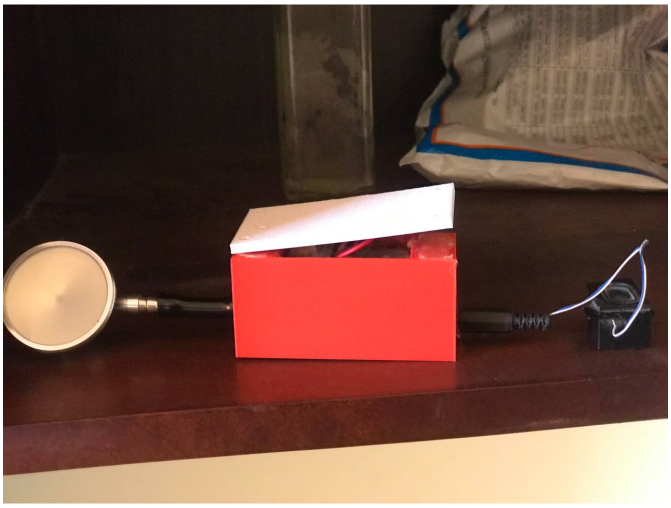


Figure 4.1: Complete assembly of electronic stethoscope

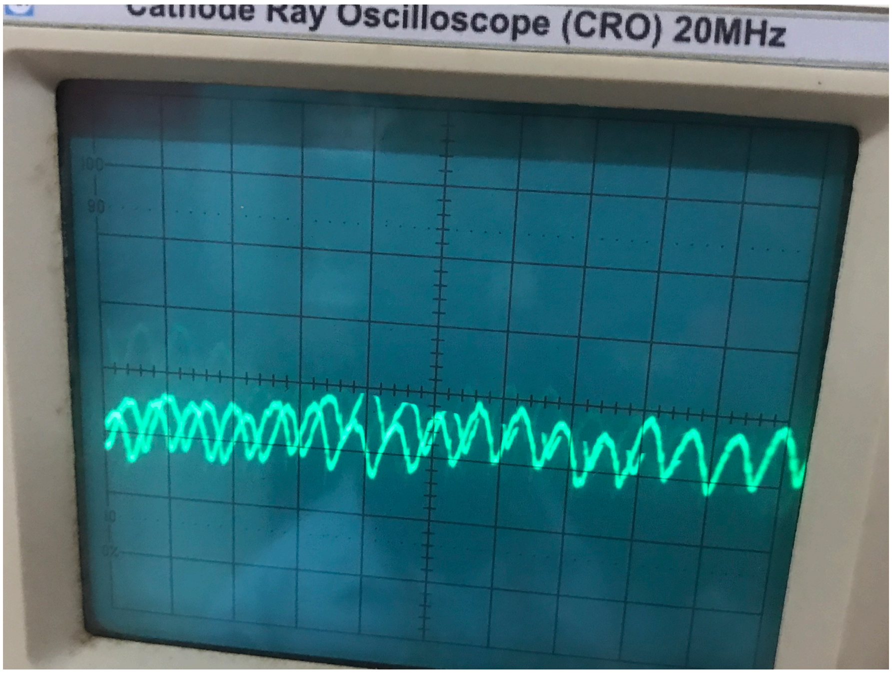
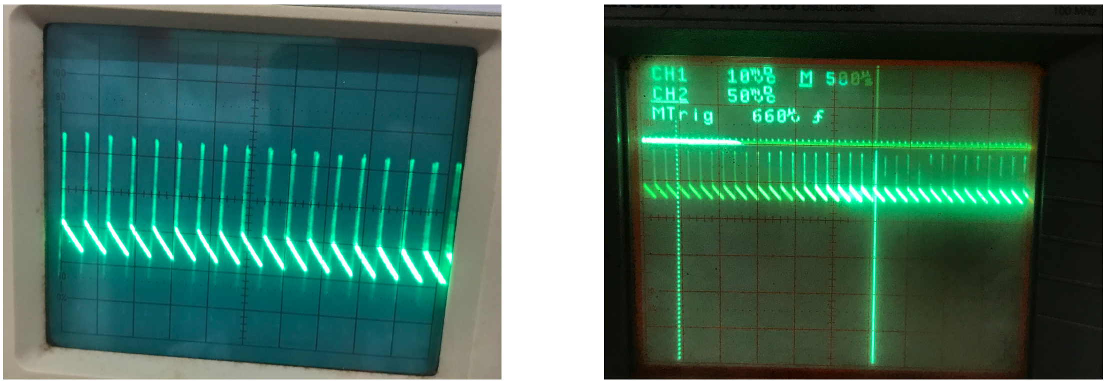


Figure 4.2: Input signal to Oscilloscope



(a) Output Signal on 20MHz Oscilloscope

(b) Output Signal on 10MHz Oscilloscope

Figure 4.3: Amplified signals output from LM386N-1

4.2 Software Results

The main concern during training was to have a learning model that would generalize on unseen data without compromising on accuracy. This meant a lot of tuning of the hyperparameters, feature transformations, and different estimators were utilized to achieve this outcome. The AdaBoost estimator was chosen based on its performance. Figure 4.4 shows an image of the application interface developed to allow easy interaction by medical professionals during the process of auscultation.

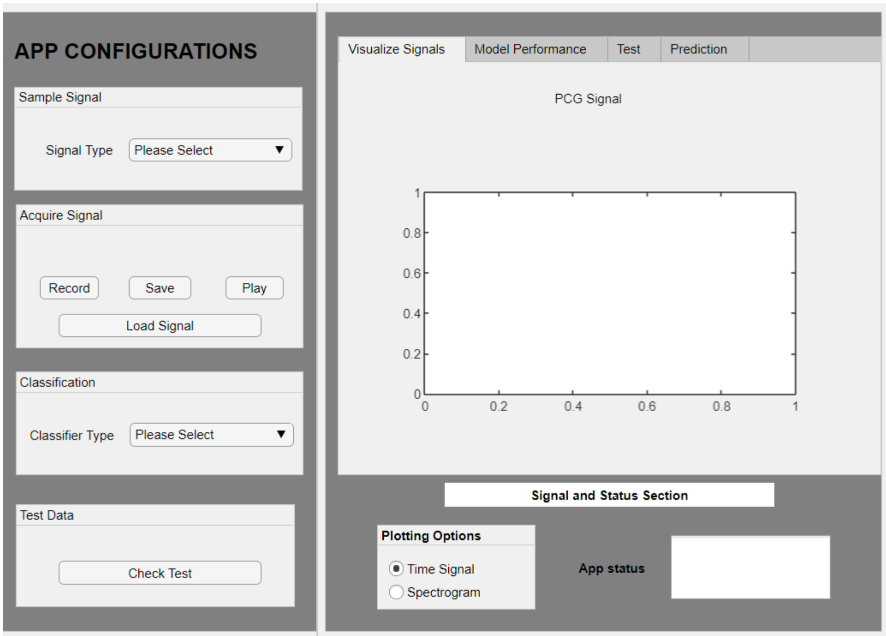


Figure 4.4: Interactive GUI Desktop application for stethoscope

4.2.1 Binary Classification using AdaBoost

The AdaBoost Ensemble after several training iterations on a 10-fold cross-validation gave a final training accuracy of 93.95% on 14 features. The model when trained on all 27 features gave a performance of 95.873% but the time used for this training was significantly higher than with 14 features. The ensemble algorithm training was optimized using the Bayesian optimization method with an Expected improvement plus acquisition function and 20 iterations. The base classifier was a decision tree and after the Bayesian optimization over a range of other ensemble boosting classifiers, the AdaBoost algorithm was selected to have the best performance with a learning rate of 0.99906 and 485 decision learners as the other optimally chosen hyperparameters. As shown in table 4.1, the Adaboost classifier has a high specificity score of 92.06% which means only 7.94% of the classifications return a false positive and a perfect sensitivity of 100%. Figures 4.5 and 4.6 show the Adaboost classifier hyperparameters used in training the dataset.

Table 4.1: Classification results of Adaboost classifier on training data

Classifier	Sensitivity	Specificity	F1 Score
Adaboost	100%	92.06%	79.77%

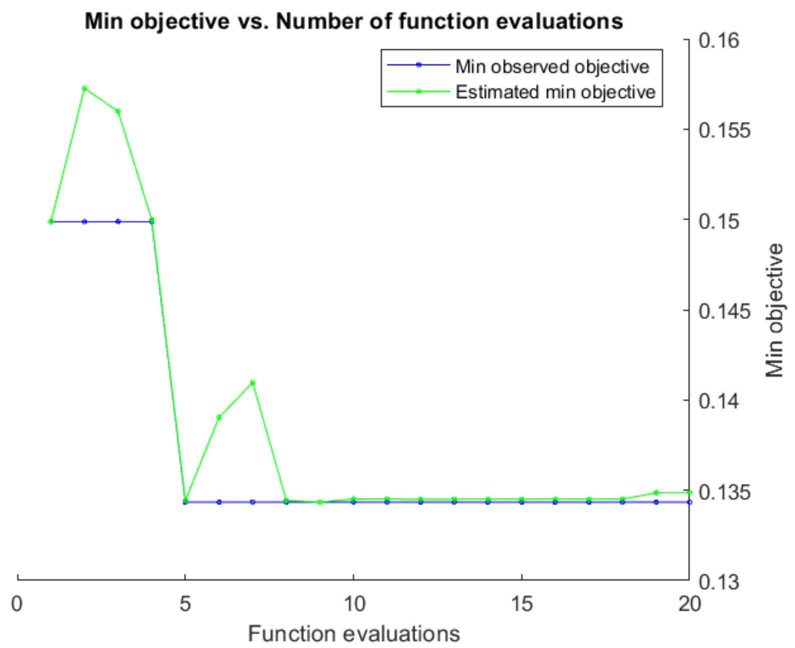


Figure 4.5: Misclassification error plot

4.3 Testing Results

30% of the heart audio signals from the physioNet sound database were used in training the performance of the classifier. As expected, the classifier performed well on this test data as the training was conducted on similar data. Figure 4.4 shows the confusion ma-

```
Optimization completed.
MaxObjectiveEvaluations of 20 reached.
Total function evaluations: 20
Total elapsed time: 1398.8458 seconds
Total objective function evaluation time: 1374.2472

Best observed feasible point:
      Method       NumLearningCycles    LearnRate
    __________    _________________    _________

    AdaBoostM1           485            0.99906

Observed objective function value = 0.13435
Estimated objective function value = 0.13487
Function evaluation time = 180.9593

Best estimated feasible point (according to models):
      Method       NumLearningCycles    LearnRate
    __________    _________________    _________

    AdaBoostM1           485            0.99906

Estimated objective function value = 0.13487
Estimated function evaluation time = 179.0153
trained_model_featsel =
  ClassificationEnsemble
                       PredictorNames: {'sampleKurtosis'  'signalEntropy
                         ResponseName: 'Y'
                CategoricalPredictors: []
                           ClassNames: {'Abnormal'  'Normal'}
                       ScoreTransform: 'none'
                      NumObservations: 9014
    HyperparameterOptimizationResults: [1×1 BayesianOptimization]
                           NumTrained: 485
                               Method: 'AdaBoostM1'
                         LearnerNames: {'Tree'}
                 ReasonForTermination: 'Terminated normally after compl
                              FitInfo: [485×1 double]
```

Figure 4.6: Hyperparameters of Optimized AdaBoost ensemble

trix which shows that the model does well on unseen test data. The AdaBoost ensemble classifier gave a sensitivity score of 89.6% and a specificity score of 85.5% which implies the model is able to detect and classify false negatives rather than false positives as expected due to the misclassification cost assigned to it. The test data achieved an accuracy of 86.46%.

4.4 Discussion

The results from the hardware implementation demonstrate that the audio sounds picked from the heart using the chest-piece coupled microphone system are able to be magnified to a variable gain of 20-200Hz. However, the results also show that due to the sensitivity of the microphone, several ambient sounds are picked up as well which corrupts the

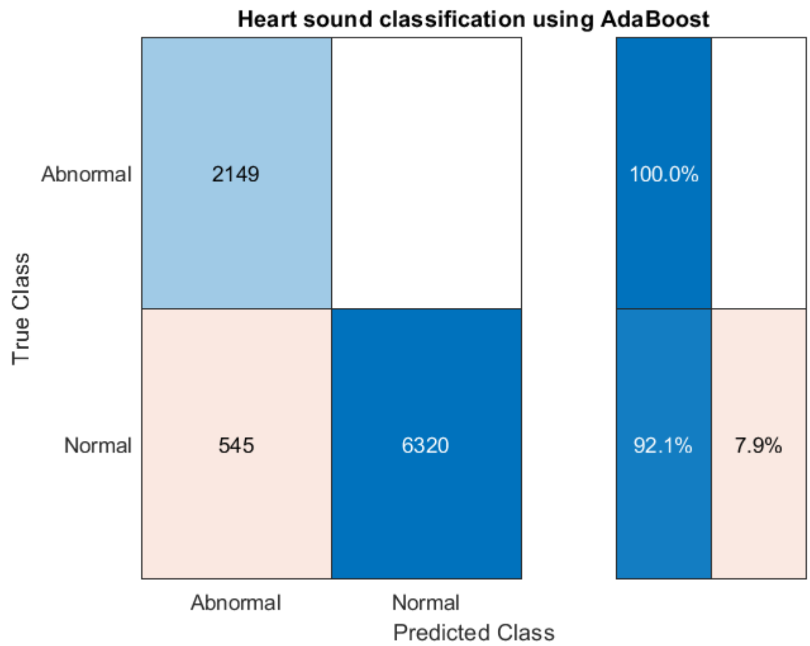


Figure 4.7: Confusion matrix of training observations using the AdaBoost ensemble

signal as noise. Thus a noisy amplified signal is obtained which is undesirable behaviour in the system.

In the software implementation, the specificity and sensitivity results obtained during training depict that the system does a good job in classifying false positives and has a very good reduction of false negatives. The false negatives are especially important since classifying an abnormal heart which is the positive class as normal is very dangerous as this means heart conditions could go undetected. The results also show that the classifier is able to perform well given that the dataset was imbalanced with the majority of the classes being normal.

Similarly, the test results buttress the training results with high accuracy, high specificity, and sensitivity scores. The key insight here is that the system does not overfit during training and is able to generalize over unseen data.

Finally, the application developed to allow users to obtain sounds, preprocess, visualize,

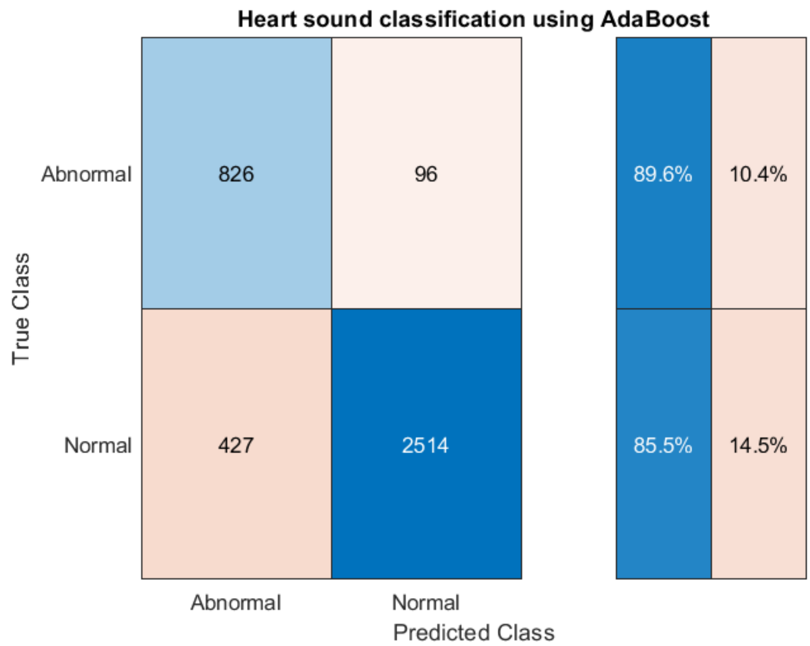


Figure 4.8: Confusion matrix of test observations using the AdaBoost ensemble

and obtain classification results demonstrates the capability of this system to be deployed to consulting rooms, and medical lecture rooms to be used by physicians and students to aid in the process of auscultation as an end-to-end system.

5 CHAPTER FIVE- CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion

This research study explored the design of an electronic stethoscope that is able to amplify faint heart sounds and the application of the AdaBoost classifier in detecting normal and abnormal heart sounds. As a result of this research, preprocessing techniques such as introducing misclassification costs, applying neighborhood component analysis to select the most relevant features, and audio filtering, Ensemble models gave an accuracy of 86.46% on classifying heart sounds as normal or abnormal respectively.

The results followed the trajectory of similar works by several researchers but more importantly, it utilized a lesser number of features to make the classification. This is important in deploying the system to resource-constrained areas as a standalone system. Since, the project is localized to Ghana, and possibly extensible to the sub-region, it should aim at utilizing fewer resources of any form (power, computational complexity, cost) and this project achieved that. As expected, the model performed well on the validation data set which consisted of audio sets from several batches of the same audio dataset.

In performing this research, and to the best of my knowledge, this is the first appearance of using the neighborhood component as the feature selection algorithm in heart sound classification. This research provides a good foundation for investigating the classification of heart sounds into abnormal and normal and holds the tremendous potential of being extended to actually improve the auscultation process and to classify various cardiovascular complications with greater detail.

5.2 Limitations

This section describes the constraints that potentially affected the performance and results of the classification algorithms and the conversion of the acoustic stethoscope to an electronic stethoscope The following were constraints on the sound clarity and the audio performance.

1. The audio signals were from a single microphone source. The audio signals from the physioNet challenge were recorded with a single microphone and as such the model might not be able to generalize properly when tested with sounds recorded with microphones with different properties.

2. The amplifier and pre-amplifier unit introduced noise. The pre-built amplifier and pre-amplifier modules reduced the amount of customization that could be done to the circuit which could some noise through the wire connection and the positioning of the electronics during the assembly. The assembly of the electronic circuitry on the bed of the assembly case exposed and made the signal susceptible to noise.

5.3 Recommendations

This section offers suggestions to extend this research. Apart from addressing the limitations of this research, the findings of this study can be widely applicable if the following recommendations are considered.

1. The use of a double or single-plated copper-clad Printed Circuit Board for assembly of the electronic circuit. This addition would ensure the wires are close to each other and are the signal is protected from noise during transmission.

2. An improvement to the sound acquisition system by utilizing digital microphones such as the Micro Electro-mechanical Systems microphones. The current assembly places the electret microphone at a distance from the chest piece which still makes the sound susceptible to noise and results in low clarity. The MEMS microphone technology allows for the miniaturized microphones that could be fitted in the chest piece directly and also eliminates the effects of using an analog microphone.

3. Application of adaptive noise cancellation to improve sound clarity. The model developed in this thesis uses Butterworth filters to eliminate noise but a better option to improve audio clarity greatly would be to use the adaptive noise cancellation.

4. Utilising heart sounds from different databases for training algorithms. One of the challenges was generalizing the models for audio signals with recorded with different microphones. Due to the different characteristics obtained from these audios, it was difficult to standardize and apply the classification algorithms which resulted in poor performances during testing. Thus, this feature will make it possible to create a modular application that could work for any stethoscope setup, and by so doing create a generalized application system.

5. An improvement of the binary classification model to multi-class classification model for cardiovascular diseases. This current model is only able to distinguish between normal and abnormal heart sounds, and doctors have to investigate further in the case of abnormal heart sounds. Future works can extend this model to be able to offer a distinctive classification of various cardiovascular compli-

cations such as pneumonia, Chronic Obstructive Pulmonary Disease, and lower respiratory tract infections.

REFERENCES

[1] W. H. Organization. “The-top-10-causes-of-death.” (2020).

[2] S. K. Majumdar, “History of the stethoscope an overview,” vol. XXXII, pp. 137–150, 2002, issn: 0304-9558.

[3] C. Smith, “Transducer for sensing body sounds,” 2002, issn: USOO6498854B1. [Online]. Available: https://patents.google.com/patent/US6498854B1/en.

[4] W. H. Organization., *Telemedicine : opportunities and developments in member states : report on the second Global survey on eHealth*. World Health Organization, 2010, p. 93, isbn: 9789241564144.

[5] E. Mbunge, B. Muchemwa, and J. Batani, “Are we there yet? unbundling the potential adoption and integration of telemedicine to improve virtual healthcare services in african health systems,” *Sensors International*, vol. 3, p. 100 152, 2022, issn: 2666-3511. doi: https://doi.org/10.1016/j.sintl.2021.100152. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666351121000735.

[6] J. F. Manuel Abella and D. G. Penney, “Comparison of the acoustic properties of six popular stethoscopes,” *Journal of Acoustical Society of America*, p. 91, 1992. doi: 10.1121/1.403655.

[7] H. Hafke-Dys, A. Bręborowicz, P. Kleka, J. Kociński, and A. Biniakowski, “The accuracy of lung auscultation in the practice of physicians and medical students,” *PLOS ONE*, vol. 14, no. 8, pp. 1–16, Aug. 2019. doi: 10.1371/journal.pone.0220606. [Online]. Available: https://doi.org/10.1371/journal.pone.0220606.

[8] S. McGee, “Evidence-based physical diagnoses,” in 4th Edition. 1600 John F. Kennedy Blvd.Ste 1800 Philadelphia, PA 19103-2899: Elsevier, 2018, ch. 30, pp. 261–274, isbn: 978-0-323392761.

[9] D. K. Degbedzui, M. Tetteh, E. E. Kaufmann, and G. A. Mills, “Bluetooth-based wireless digital stethoscope with mobile integration,” *Biomedical Engineering - Applications, Basis and Communications*, vol. 30, 3 Jun. 2018, issn: 10162372. doi: 10.4015/S1016237218500102.

[10] M. E. Chowdhury, A. Khandakar, K. Alzoubi, *et al.*, "Real-time smart-digital stethoscope system for heart diseases monitoring," *Sensors (Switzerland)*, vol. 19, 12 Jun. 2019, issn: 14248220. doi: 10.3390/s19122781.

[11] S. A. Fattah, N. M. Rahman, A. Maksud, *et al.*, "Stetho-phone: Low-cost digital stethoscope for remote personalized healthcare."

[12] W. Wang, Q. Xu, G. Zhang, *et al.*, "A bat-shape piezoresistor electronic stethoscope based on mems technology," *Measurement: Journal of the International Measurement Confederation*, vol. 147, Dec. 2019, issn: 02632241. doi: 10.1016/j.measurement.2019.106850.

[13] C. Chun-Tang, M. Nopadon, and W. Chi-Jo, *2012 35th International Conference on Telecommunications and Signal Processing, TSP : proceedings : July 3-4, 2012, Prague, Czech Republic.* IEEE, 2012, isbn: 9781467311182.

[14] R. Jorge, F. José, B. Juan, L. Vivian, R. Luis, and L. Carlos, "Electronic stethoscope with wireless communication to a smart-phone, including a signal filtering and segmentation algorithm of digital phonocardiography signals," vol. 60, Springer Verlag, 2017, pp. 553–556, isbn: 9789811040856. doi: 10.1007/978-981-10-4086-3_139.

[15] A. Lakhe, I. Sodhi, J. Warrier, and V. Sinha, "Development of digital stethoscope for telemedicine," *Journal of Medical Engineering and Technology*, vol. 40, pp. 20–24, 1 Jan. 2016, issn: 1464522X. doi: 10.3109/03091902.2015.1116633.

[16] M. Elhilali and J. West, "A smart stethoscope puts ai in medics' ears," *IEEE Spectrum*, 2019. [Online]. Available: https://spectrum.ieee.org/a-smart-stethoscope-puts-ai-in-medics-ears.

[17] S. Y. Jung, C. H. Liao, Y. S. Wu, S. M. Yuan, and C. T. Sun, "Efficiently classifying lung sounds through depthwise separable cnn models with fused stft and mfcc features," *Diagnostics*, vol. 11, 4 2021, issn: 20754418. doi: 10.3390/DIAGNOSTICS11040732.

[18] D. Preuveneers, I. Tsingenopoulos, and W. Joosen, "Resource usage and performance trade-offs for machine learning models in smart environments," *Sensors (Switzerland)*, vol. 20, 4 Feb. 2020, issn: 14248220. doi: 10.3390/S20041176.

[19] T. Grzywalski, R. Belluzzo, M. Piecuch, *et al.*, "Fully interactive lungs auscultation with ai enabled digital stethoscope," vol. 11526 LNAI, Springer Verlag, 2019, pp. 31–35, isbn: 9783030216412. doi: 10.1007/978-3-030-21642-9_5.

[20] M. Zabihi, A. B. Rad, S. Kiranyaz, M. Gabbouj, and A. K. Katsaggelos, "Heart sound anomaly and quality detection using ensemble of neural networks without segmentation."

[21] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds."

[22] G. D. Clifford, C. Liu, B. Moody, *et al.*, "Classification of normal/abnormal heart sound recordings: The physionet/computing in cardiology challenge 2016," 2016.

[23] M. N. Homsi, N. Medina, M. Hernandez, *et al.*, "Automatic heart sound recording classification using a nested set of ensemble algorithms."

[24] M. Márton´, M. Goda, and P. Hajas, "Morphological determination of pathological pcg signals by time and frequency domain analysis."

[25] "A brief guide to microphones," Audio-Technica Inc, 2021, pp. 5–13. [Online]. Available: https://www.audio-technica.com/en-us/support/a-brief-guide-to-microphones-introduction/.

[26] "Microphone specifications explained," 2016, pp. 1–11. [Online]. Available: https://invensense.tdk.com/.

[27] J. Lewis, "Analog dialogue 46-05 back burner, april (2012)," 2012, p. 1. [Online]. Available: www.analog.com/analogdialogue.

[28] D. Steven and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Aug. 1980. doi: 0096-3518/80/0800-0357.

[29] S. A. Alim and N. K. A. Rashid, “Some commonly used speech feature extraction algorithms,” *From Natural to Artificial Intelligence - Algorithms and Applications*, Dec. 2018. doi: 10.5772/intechopen.80419.

[30] S. S. Stevens, J. E. Volkmann, and E. B. Newman, “A scale for the measurement of the psychological magnitude pitch,” *Journal of the Acoustical Society of America*, vol. 8, pp. 185–190, 1937.

[31] K. S. Rao and A. K. Vuppala, *Speech Processing in Mobile Environments*. Springer International Publishing, 2014, isbn: 978-3-319-03115-6. doi: 10.1007/978-3-319-03116-3. [Online]. Available: http://link.springer.com/10.1007/978-3-319-03116-3.

[32] T. F. Zheng, G. Zhang, and Z. Song, “Comparison of different implementations of mfcc,” *Journal of Computer Science and Technology*, vol. 16, pp. 582–589, 2008.

[33] J. Picone, “Signal modeling techniques in speech recognition,” *Proceedings of the IEEE*, vol. 81, no. 9, pp. 1215–1247, Sep. 1993, issn: 1558-2256. doi: 10.1109/5.237532.

[34] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *J. Mach. Learn. Res.*, vol. 3, no. null, pp. 1157–1182, Mar. 2003, issn: 1532-4435.

[35] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, “Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis,” *Informatics*, vol. 8, 4 Dec. 2021, issn: 22279709. doi: 10.3390/informatics8040079.

[36] “Understanding ffts and windowing overview,”

[37] “Understanding fft windows application note an014,” 2003.

[38] *What is matlab?* [Online]. Available: https://www.mathworks.com/discovery/what-is-matlab.html.

[39] G. D. Clifford, C. Liu, B. Moody, *et al.*, “Classification of normal/abnormal heart sound recordings: The physionet/computing in cardiology challenge 2016,” 2016. doi: 10.22489/CinC.2016.179-154.

[40] M. Fraiwan, L. Fraiwan, B. Khassawneh, and A. Ibnian, "A dataset of lung sounds recorded from the chest wall using an electronic stethoscope," *Data in Brief*, vol. 35, Apr. 2021, issn: 23523409. doi: 10.1016/j.dib.2021.106913.

[41] D. B. Springer, L. Tarassenko, and G. D. Clifford, "Logistic regression-hsmm-based heart sound segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 63, pp. 822–832, 4 Apr. 2016, issn: 15582531. doi: 10.1109/TBME.2015.2475278.

APPENDIX

Project Code : https://github.com/Awontiirim/heart-sound-classification