**UNIVERSITY OF ENERGY AND NATURAL RESOURCES, SUNYANI**

**SCHOOL OF SCIENCE**

**DEPARTMENT OF INFORMATION TECHNOLOGY AND DECISION SCIENCES**

**HEARTBEAT SOUND SIGNAL CLASSIFICATION USING DEEP LEARNING**

**BY**

**KULLAH JAMES**

**ACQUAH BISMARCH**

**ANYAGRIC EULALIE AWIN**

**ABUGRIC ABDUL-RAUF**

**TETTEH RICHARD**

**SEPTEMBER, 2024**



**UNIVERSITY OF ENERGY AND NATURAL RESOURCES, SUNYANI**

**SCHOOL OF SCIENCES**

**DEPARTMENT OF INFORMATION TECHNOLOGY AND DECISION SCIENCES**

**UNDER THE GUIDANCE AND SUPERVISION OF: DR. PETER APPIAHENE**

**BY**

KULLAH JAMES (UEB3211720)

ACQUAH BISMARCH (UEB3205922)

ANYAGRIC EULALIE AWIN (UEB3218022)

ABUGRIC ABDUL-RAUF (UEB3208120)

TETTEH RICHARD (UEB3269522)

**A PROJECT WORK SUBMITTED TO THE DEPARTMENT OF INFORMATION TECHNOLOGY AND DECISION SCIENCES, UNIVERSITY OF ENERGY AND NATURAL RESOURCES, IN THE PARTIAL FULFILMENT OF THE REQUIREMENT FOR AWARD OF A DEGREE**

**IN**

**INFORMATION TECHNOLOGY**

**SEPTEMBER, 2024**

# 

# DECLARATION

We confirm that we conducted and authored the study, acknowledging all sources of information. We take full responsibility for any potential violations of research ethics policies at the university.

**KULLAH JAMES** …………………….. ……………………..

(UEB3211720) DATE SIGNATURE

**ACQUAH BISMARCK**  …………………….. ……………………..

(UEB3205922) DATE SIGNATURE

**ANYAGRIC EULALIE AWIN** …………………….. ……………………..

(UEB3218022) DATE SIGNATURE

**ABUGRI ABDUL-RAUF** …………………….. ……………………..

(UEB3208120) DATE SIGNATURE

**TETTEH RICHARD** …………………….. ……………………..

(UEB3269522) DATE SIGNATURE

**PROFESSOR PETER APPIAHENE**  …………………….. ……………………..

(SUPERVISOR) DATE SIGNATURE

**PROFESSOR PETER APPIAHENE** …………………….. ……………………..

(HEAD OF DEPARTMENT) DATE SIGNATURE

# **Acknowledgement**

First and foremost, we are incredibly grateful to God, who has supported us during our time at the University of Energy and Natural Resources, for providing us with the knowledge and understanding we needed to complete our studies.

We want to sincerely thank each and every one of you for helping us to successfully finish this project. Throughout this journey, your support and aid have been vital.

We will always be grateful to our Supervisor, Professor Peter Appiahene who gave direction, input, and encouragement at every step of this project. Your knowledge and experience have always been an inspiration. We also value the work that all of our professors and other Department employees have put in, as they have assisted us in various ways. God will continuously assist them and lavishly bless them. We also thank Mr. Arthur Enock Justice, research assistant, and Mr. Stephen Afrifa, senior research assistant, of the UENR Department of Information Technology and Decision Sciences for their assistance.

Finally, but just as importantly, we want to sincerely thank our family and friends for their patience and understanding throughout the long hours we put into this project. We were inspired by your constant belief and support.

Without everyone's combined effort, this project would not have been feasible. Regards

for being part of this journey.

# **Dedication**

This effort is dedicated to the All-Powerful God and our parents, who have been there for us throughout our education at this esteemed university.   
And to our lecturers, who provided the most assistance and ongoing knowledge influence in the development of this model.

# **Abstract**

Especially in developing nations, quality health care is still quite inaccessible. This work proposed a heartbeat sound signal classification using a deep learning model. It classifies the heartbeat sounds into normal and abnormal classes using advanced techniques. Later, extensive pre-processing was done on a sizeable dataset gathered for the training of the model. Thus, for this work, we utilized LSTM models, which basically were long in the data gathering and preprocessing steps, followed by the extraction of the MFCC features, and training the model. Further, we provided a detailed explanation of the LSTM network architecture, evaluation procedures, and performance at two test sets. It has performed very well on the first test set with great precision, recall, and F1-score for the class "artifact ". The class "murmur" was performing the poorest, hence indicating difficulties in its correct identification. The class "normal" showed a good performance. Overall, the accuracy was 71%, while the balanced macro and weighted average also scored well. On this second test set, the "murmur" class performed a little better with precision of 65%, recall of 59%, and F1-score of 62%. The "normal" class had very high performance for all metrics around or above 90%. Overall accuracy increased to 86%, with consistent macro and weighted averages. Despite the LSTM model not surpassing some other algorithms, it showed commendable accuracy in classifying heartbeat sounds. A user-friendly interface using Streamlit and Hugging Face was developed, simplifying the classification process for both experts and non-experts. This approach has the potential to revolutionize cardiovascular disease diagnosis by improving early detection, contributing significantly to healthcare technology advancement, particularly in resource-limited settings.

# 

# **Table of contents**

[DECLARATION i](#_Toc178684153)

[Acknowledgement ii](#_Toc178684154)

[Dedication iii](#_Toc178684155)

[Abstract iv](#_Toc178684156)

[Table of contents v](#_Toc178684157)

[List of figures vii](#_Toc178684158)

[List of tables viii](#_Toc178684159)

[CHAPTER ONE 1](#_Toc178684160)

[INTRODUCTION 1](#_Toc178684161)

[1.1 Background 1](#_Toc178684162)

[1.2 Problem statement 3](#_Toc178684163)

[1.3 Objectives 5](#_Toc178684164)

[1.3.1 Main objectives 5](#_Toc178684165)

[1.3.2 Specific objectives 5](#_Toc178684166)

[1.4 Significance of the project 5](#_Toc178684167)

[1.5 Limitation of the project 7](#_Toc178684168)

[1.6 Organization of project 8](#_Toc178684169)

[CHAPTER TWO 10](#_Toc178684170)

[LITERATURE REVIEW 10](#_Toc178684171)

[2.0 Introduction 10](#_Toc178684172)

[2.1 Definition of concept 10](#_Toc178684173)

[2.2.0 Machine learning 10](#_Toc178684174)

[2.2.1 Deep learning 10](#_Toc178684175)

[2.2.2 Classification 11](#_Toc178684176)

[2.2.3 Signal processing 11](#_Toc178684177)

[2.2.4 Heartbeat sound 11](#_Toc178684178)

[2.2.5 Segmentation 11](#_Toc178684179)

[2.2.6 Data augmentation 11](#_Toc178684180)

[2.2.7 Attention mechanisms 12](#_Toc178684181)

[2.2 Subtopics 12](#_Toc178684182)

[2.3.0 Cardiovascular disorders 12](#_Toc178684183)

[2.3.1 Recurrent neural networks and long short-term memory in heartbeat sound classification 13](#_Toc178684184)

[2.3.2 Convolutional neural networks in heartbeat sound signal processing 13](#_Toc178684185)

[2.3.3 Application of LSTM in heartbeat sound signal classification 13](#_Toc178684186)

[2.3 Review of related works 14](#_Toc178684187)

[2.4 Challenges and limitations 17](#_Toc178684191)

[2.5.0 Handling noisy and imbalanced datasets 17](#_Toc178684192)

[2.5.0 Improving interpretability 18](#_Toc178684193)

[2.5.1 Integrating multimodal data 18](#_Toc178684194)

[2.5 Summary of reviewed papers 20](#_Toc178684195)

[2.6 Summary of literature review 22](#_Toc178684196)

[CHAPTER THREE 23](#_Toc178684197)

[METHODOLOGY 23](#_Toc178684198)

[3.0 Introduction 23](#_Toc178684199)

[3.1 Dataset and data preprocessing 23](#_Toc178684200)

[3.1.1 Dataset 23](#_Toc178684201)

[3.1.2 Data preprocessing 25](#_Toc178684202)

[3.2 System architecture 28](#_Toc178684203)

[3.3 System components 29](#_Toc178684204)

[3.4 Algorithm 29](#_Toc178684205)

[3.5 Operational methods 29](#_Toc178684206)

[3.6 Feature extraction 30](#_Toc178684207)

[3.7 Model development 31](#_Toc178684208)

[3.7.0 Introduction to lstm models 31](#_Toc178684209)

[3.8 Evaluation and results 32](#_Toc178684210)

[3.9 Summary of methodology 33](#_Toc178684213)

[CHAPTER FOUR 34](#_Toc178684214)

[RESULTS AND ANALYSIS 34](#_Toc178684215)

[4.1 Introduction 34](#_Toc178684216)

[4.2 Visualization and feature extraction from datasets 34](#_Toc178684217)

[4.3 Proposed model 37](#_Toc178684218)

[4.5 Evaluation Metrics (R², MSE, and Accuracy) and Comparative analysis 39](#_Toc178684219)

[4.5.0 Model Accuracy and Loss 39](#_Toc178684220)

[4.5.1 F1 Score, recall and precision: 40](#_Toc178684221)

[4.6 User Interface (UI) 41](#_Toc178684222)

[4.6 How to use app 43](#_Toc178684223)

[4.6.0 Sample data test 43](#_Toc178684224)

[4.6.1 Browsing for audio file 43](#_Toc178684225)

[4.7 App navigation/App in use 44](#_Toc178684226)

[4.8 Summary of results and analysis 45](#_Toc178684227)

[5.0 Introduction 46](#_Toc178684228)

[5.1 Challenges 46](#_Toc178684229)

[5.1.0 Datasets 46](#_Toc178684230)

[5.1.1 Resources and hardware requirements 46](#_Toc178684231)

[5.1.2 User interface 47](#_Toc178684232)

[5.2 Discussions and future directions, ethical considerations 47](#_Toc178684233)

[5.3 Future works and recommendations 48](#_Toc178684234)

[5.4 Conclusions 49](#_Toc178684235)

[References 50](#_Toc178684236)

# **List of figures**

[Figure 4. 1 waveform of a normal hearbeat sound 34](#_Toc172489457)

[Figure 4. 2 spectrogram of a normal heartbeat sound signal 35](#_Toc172489458)

[Figure 4. 3 spectrum of a normal heartbeat sound 35](#_Toc172489459)

[Figure 4. 4 MFCC of a normal heartbeat sound 36](#_Toc172489460)

[Figure 4. 5 model summary of LSTM 36](#_Toc172489461)

[Figure 4. 6 model summary 37](#_Toc172489462)

[Figure 4. 7 model training and validation history 37](#_Toc172489463)

[Figure 4. 8 confusion metrics on 3 lables 39](#_Toc172489464)

[Figure 4. 9 model accuracy 39](#_Toc172489465)

[Figure 4. 12 user interface 42](#_Toc172489468)

[Figure 4. 13 app usage 44](#_Toc172489469)

[Figure 4. 14 app predicting on sample audio files 45](#_Toc172489470)

# **List of tables**

[Table 1 Summary of reviewed papers 20](#_Toc178684260)

# 

# **CHAPTER ONE**

# **INTRODUCTION**

# **Background**

Artifical Intellegence, artificial intelligence in brief, is a new technical discipline that researches, develops and applies the theory, method, technique and application system for simulating, extending and expanding human intelligence with the help of computer technology. (Liu et al., 2021). Schank (2019) This means that AI has two major goals; first, there is the design of an int elligent machine and second, there is the learning about the nature of intelligence.

It is a fast-ballooning area of computational methods that aim to simulate human intelligence based on learning from the environment. Its practical applications include a wide array of industries like pattern recognition, computer vision, aerospace engineering, finance, entertainment, computational biology, and biological and medical domains. The use of machine learning is also considered essential when dealing with voluminous information, or big data. Deep learning in medicine has already produced remarkable results.

Cardiac sound signal classification is an important area of research in cardiology that aims to accurately identify abnormal heart sounds, which can be indicative of various cardiac disorders(J. Chen et al., 2024). Correct classification of cardiac sound signals is greatly important for making timely and accurate diagnoses, hence coming up with effective treatment plans. However, a variety of limitations and challenges are realized in current methods available for the analysis and classification of cardiac sound signals.

On the other hand, this might be complemented by challenges in their analysis. However, AI, ML, and DL have proved promising improvements to enhance the accuracy and reliability of sound analysis in cardiac sound signal classification. First, current methods for the classification of cardiac sound signals are based on subjective human interpretation and could therefore be variable in the diagnostic output provided. Secondly, there is a lack of uniformity in the terminology that is generally used and adopted for such schemes of classification. The usually implemented signal-processing methodologies may not capture minute differences in sound patterns indicative of cardiac disorders. Deep learning is the artificial neural network that acts like a human brain and, without help, learns and makes decisions. To overcome this situation, heartbeat sound analysis is a convenient way to diagnose heart disease(Zeinali and Niaki, 2022) .

It takes a lot of time and money to get medical attention from a doctor when someone has a cardiovascular disease. Furthermore, getting healthcare may be difficult if they are far from medical facilities. Therefore, the patient would benefit immensely and the procedure would be streamlined by using an automated software solution that saves time and money. Early detection and ongoing care are necessary for many cardiac conditions. Like in other nations, major social and economic changes have led to a sharp rise in illness prevalence in Ghana.

Numerous developing nations, including Ghana, face the significant burden of chronic diseases, particularly cardiovascular disease and diabetes, which have profound implications for global health, security, and the economy. The rapid urbanization and economic progress witnessed in today's world have given rise to diverse lifestyles, resulting in the widespread occurrence of cardiovascular diseases in all countries, affecting approximately one-third of their populations.

# **Problem statement**

Classifying heartbeat sounds is a critical yet challenging task in the medical field, particularly in the segmentation and extraction of relevant features from heart sounds. Heart sounds, typically represented as phonocardiograms (PCGs), provide valuable information about the functionality of the heart. With added noise, the delicacy of these sounds renders their interpretation challenging. As Qiao et al. (2023) show, the segmentation of heartbeats and extraction of such salient features as S1 and S2 sounds with accuracy require sophisticated techniques, since recordings in PCG are often not regular.

Traditionally, heartbeat sound classification has relied heavily on healthcare professionals. These precious manual skills bring subjectivity into diagnosis, hence classification may vary. Clinical practitioners commonly use their training and experience to interpret these sounds; this is usually a very time- and labor-intensive process. Moreover, manual classification may not capture all subtle abnormalities that may lead to diagnostic errors. This process, as it is done manually, is highly time-consuming and cannot help the doctor from the human factor that may lead to a wrong diagnosis, especially when the differentials between normal and abnormal heart sounds are minor. Another big challenge is the process of segmenting heartbeat sounds. The segmentation, according to various sources, can be defined as the identification of key cardiac events such as first heart sound (S1), second heart sound (S2), and other components, such as murmurs. Thus, the challenge is to identify those sounds precisely in the presence of background noise and interference that might cause a distortion of the signal. In most cases, it is external noise, respiratory sounds, and movement of the body that masks the critical components of the heartbeat that are to be measured. Besides, variations in heart rate, such as in patients with arrhythmias, complicate the segmentation process, and so diagnosis might be susceptible to inaccuracy. The most challenging task in heartbeat sound classification is feature extraction. It requires finding the relevant characteristics-frequency, amplitude, and temporal features-of heart sounds that can help in distinguishing a normal heartbeat from a pathological one.

However, heart sounds are intrinsically non-stationary; their frequency and amplitude continuously vary with time. Hence, feature extraction is challenging to perform in a consistent and reliable manner. This process is further complicated by the fact that heart sounds may vary between different patients. What works well for one patient may not work well for another, hence the need for sophisticated and adaptive algorithms in this area. In such scenarios, in recent years many researchers have adopted machine learning and deep learning methods for the classification of heartbeat sounds. In particular, those methods using LSTMs and CNNs; these models have obtained some initial benefits in the performance and efficiency of the heartbeat sound classification process. These models can reduce the subjectivity and errors from manual interpretation by automatically extracting and classifying features. Yet, with the use of even advanced algorithms, especially when operating in a noisy environment or with irregular heart rhythms, the most important difficulties persist in the stages of segmentation and extraction of features.

# **Objectives**

# **Main objectives**

The Main objectives is to

1. develop an intelligence-based model for the automatic classification of heartbeat sound signals.

# **Specific objectives**

The specific objectives are to:

1. develop a framework for automatic classification of heartbeat sound signals.
2. develop signal processing techniques to employ for classification of heartbeat sounds signals.
3. develop deep learning models for classification of heartbeat sounds signals.
4. deploy the deep learning model for classification of heartbeat sound signals.

# **Significance of the project**

A study on heartbeat sound signal classification using deep learning will have a huge impact on cardiac healthcare and diagnostics. Thus, the importance of the project can be seen from the following:

The classification of heartbeat sounds by deep learning algorithms can be employed in the early diagnosis of such cardiac abnormalities as murmur, arrhythmia, and valvular pathologies. In turn, early diagnosis will facilitate timely intervention and treatment and, therefore, can save lives and improve patient outcomes.

Non-invasive Diagnostic Approach: Conventional methods of diagnosis of cardiac conditions tend to be invasive in nature or highly expensive imaging procedures. To that effect, deep learning methods for heartbeat sound classification are non-invasive and low-cost diagnostic cardiac procedures; hence, this technology will be more accessible to the masses.

It therefore increases the demand for remote monitoring of cardiac health in view of wearable devices and telemedicine technologies that have gained popularity. This deep learning-based classification of heartbeat sounds can be integrated into those technologies to continuously monitor the cardiac status of the patient from a distance.

Reduced healthcare burden: Deep learning models create an opportunity for a healthcare professional to triage patients with potential heart ailments and help in providing priority care, thus decreasing the healthcare burden in busy clinical settings where resources could be limited.

Personalized Medicine: Deep learning models can process a large dataset of heartbeat sound patterns to identify those specific to an individual patient. A treatment plan can thus be outlined based on the cardiac profile of a particular patient, hence increasing the effectiveness of interventions with reduced adverse outcomes.

Medical Technology Development: The development of algorithms using deep learning for heartbeat sound classification inspires innovation in medical technologies. These innovations can be witnessed in integrating AI into medical devices such as stethoscopes and wearable monitors to the development of new diagnosis tools and software applications.

Education and Research: Deep learning of heartbeat sound classification would add to the research and understanding of cardiac physiology and pathology. It would provide valuable insights into acoustic characteristics of normal and abnormal heart sounds, hence informing future research and education in cardiology and related fields.

# **Organization of Project**

The introduction, methodology, literature review, discussion, and conclusion with recommendations make up the project's five chapters.

The study's history, problem description, research aims, significance, and project scope are all covered in Chapter One (Introduction).

Chapter Two: Literature Review: In this chapter, associated ideas are defined and relevant subtopics are reviewed in relation to the research issue. A critical overview of recent research and relevant literature is also presented.

Methodolog which will be in chapter 3: This section elaborates on the tools, procedures, and techniques that will be used in the research. It houses architecture, hardware, and software specifications and system design.

Chapter Four presents the findings of the study and a discussion: It includes an analysis of the findings and discussions of implications. After analyzing and interpreting the data collected in the study, findings are provided in a clear and organized manner. Depending on the research design, the analysis could use statistical techniques, qualitative analysis, or a combination of the two. Along with answering any research questions or hypotheses, the chapter should also compare the results to earlier studies and explain any unexpected outcomes. The project's main conclusions and findings are outlined in Chapter Five, "Conclusion and Recommendations." Based on the results, suggestions are made for additional study and advancements.

# **Summary of Introduction**

In summary, we have successfully stated our problem statement, pointed out our main and specific objectives and then its significance as well as the project’s organization which tells us how the project is going to be done.

# **CHAPTER TWO**

# **LITERATURE REVIEW**

# **Introduction**

To find research on Heartbeat Sound Signal categorization models, we carried out a thorough search of academic databases. Using terms like "heartbeat sound model," "signal classification model," and "deep learning," we searched PubMed, IEEE Xplore, and Google Scholar. Our search was restricted to English-language research released between 2019 and 2024.   
For the purpose of our review, we went over the full texts, abstracts, and titles of every study that was found. Studies that satisfied the following requirements were included.

**Definition of concept**

# **Machine learning**

ML is a sub-discipline of computer science based on statistical models and algorithms that include learning from data, identification of trends, and prediction of outcomes without necessarily being explicitly programmed. These algorithms can perform tasks that typically would require human intelligence, process vast datasets, and also keep improving..

# **Deep learning**

Lecun, Y., Bengio, Y., & Hinton, G. (2015) explain deep learning as an area of research under machine learning that processes data and learns complex patterns with multi-layered artificial neural networks. Trained on the architecture of the human brain, these deep neural networks excel in pattern recognition related to speech, image, and natural language processing..

# **Classification**

In machine learning, classification is supervised learning of a mapping from observations for which the label is known to the categorical labels that are to be predicted for future observations. (J. Brownlee, 2019).

# **Signal processing**

Signal processing is a branch of engineering and applied mathematics dealing with the analysis and manipulation of signals. (Oppenheim et al., 2018). Signals can be any type of data that carries information, such as electrical signals, audio signals, and images.

# **Heartbeat sound**

Heartbeat sound refers to the acoustic signal produced by the heart during its cycle of contraction and relaxation (medicinenet, 2023). These sounds are typically used in medical diagnostics to assess heart function and detect abnormalitiesThe murmur is simply a rhythmic sound caused by the heartbeat. It is mainly a production of contractions and relaxations of the chambers of the heart and their respective valves inside them. (Mayo Foundation for Medical Education and Research, 2023)

# **Segmentation**

Schmidt et al. (2019) segmentation of heart sound recordings using a hidden Markov model is investigated. Although not strictly deep learning, the study provides valuable insights into heart sound signal processing. An open access database for evaluating heart sound algorithms is introduced (Liu et al., 2019). The database is a critical resource for developing and testing deep learning models for heart sound classification.

# **Data augmentation**

Data augmentation techniques artificially increase the size of the training dataset by generating modified versions of already existing data. This step is quite important to perform because deep learning models need quantities of data in order to generalize well. Zhu et al. (2019) investigated time stretching, pitch shifting, and adding noise for heartbeat sound classification as possible data augmentation techniques. These techniques enhanced the robustness of the deep learning models. Nguyen et al. (2020) have used data augmentation to avoid this problem of limited training data, and they report considerable improvement in the model's performance...

# **Attention mechanisms**

Attention mechanisms enable a neural network to concentrate on the parts of the input sequence that are relevant for the task at hand, by improving the feature extraction capability of the model (Vaswani et al., 2019).

Santos et al. (2020) used attention mechanisms in their LSTM-based model for heartbeat sound classification. The use of attention helped to focus on the critical segments of the heart sound signals and raised the accuracy of its classification. Xu et al. (2019) used an attention mechanism in a CNN-LSTM hybrid model for heartbeat sound classification. Authors have shown that attention helps raise the performance of their model by underlining the relevant features..

# **Subtopics**

# **Cardiovascular disorders or diseases**

WHO classifies peripheral artery disease, rheumatic heart disease, congenital heart disease, high blood pressure, coronary heart disease, cerebrovascular disease, and heart failure as CVDs.

According to the estimate of a World Health Organization study, more people die from CVDs annually in the world compared to any other cause of death. In 2012, the forecasted 17.5 million deaths throughout the world from CVDs accounted for 31% of all deaths. An estimated 7.4 million of these deaths were due to heart disease, while another 6.7 million were due to stroke. Smoking, consumption of unhealthy diet and obesity, physical inactivity, excessive consumption of alcohol, high blood pressure, diabetes, and hyperlipidemia are the usual constellation of risk factors that precipitate heart attacks and stroke. Over 75% of CVD deaths take place in low- and middle-income nations.

# **Recurrent neural networks and long short-term memory in heartbeat sound classification**

Recurrent Neural Networks represent a class of neural networks that naturally fit sequential data by capturing temporal dependencies. Classic RNNs don't work with long-term dependencies because of vanishing gradient problems. Long Short-Term Memory networks are a type of RNN designed to overcome these limitations by incorporating memory cells and gating mechanisms that help in retaining information over longer sequences (Hochreiter & Schmidhuber, 2020).

### **Convolutional neural networks in heartbeat sound signal processing**

The CNNs have been very successful in image and signal processing due to their ability to capture local patterns using convolutional filters. In the classification of heartbeats, the CNNs work to extract the spatial features from time-frequency representations such as spectrograms and MFCCs (Ronneberger et al., 2019).

# **Application of LSTM in heartbeat sound signal classification**

Shen et al. (2021) performed the classification of heart sound signals into normal and abnormal classes using LSTM networks. They concluded that the temporal features related to the heartbeat sounds could be captured effectively using LSTM, and improved accuracy was achieved as compared to traditional machine learning methods. Liu et al. (2019) proposed an LSTM-based method coupled with CNN to enhance the performance of feature extraction and classification. Their hybrid model achieved the best result, thereby implying that there is added value in fusing the learning of features in both the spatial and temporal domains.

# **Review of related works**

Yang et al. (2019) classify the heart sound recordings with a BiLSTM network, which by the nature of being bidirectional has also improved the classification results by understanding the context. Rajpurkar et al. (2019) proposed a deep learning algorithm called "Cardiologist-Level Arrhythmia Detection" which is based on LSTM for heartbeat sound classification. Their model, with its high accuracy from classifying different types of arrhythmias, once more proved that LSTM works well with complex heart sound patterns. Gao et al., (2020) Authors have proposed a multi-scale LSTM network for heartbeat classification from electrocardiograms. This technique relies on temporal dependencies in heartbeat sequences and outperforms traditional machine learning algorithms.

A real-time heartbeat sound classification system using deep learning on wearable devices has been proposed, by Liu, J., Chen, X., & Tang in the 2020 work. It relies on LSTM networks for the processing of continuous heart sound signals, performing real-time classification with high accuracy.

This paper proposes noise-robust deep learning for heart sound classification, based on an LSTM network structure that copes effectively with noisy environments and further improves classification performance.

Nguyen et al. (2020) explored the use of LSTM for multi-class heart sound classification. Data augmentation was applied to yield better testing results with small datasets. Their findings showed a significant improvement in the accuracy of classification.

The developed model of pathological heart sound detection was based on LSTM; it underlined the effectiveness of the LSTM network in capturing temporal dependencies for the distinction of normal and abnormal heart sounds. Zabihi et al. (2019) proposed a multi-scale LSTM network for heartbeat sound classification. Their model leveraged temporal features at different scales, thus achieving enhanced classification performance compared to single-scale models.

The classification of Heart sound using deep neural network: The LSTM networks capture the temporal dependencies in heart sound data and thus provide better classification results. Nguyen, H. L., Zhang, Z., & Zhang, J. 2020

Heart sound classification using the deep learning technique is discussed for wearable devices. The paper focuses on the use of an LSTM network on continuous heart sound signal processing. Rubin et al., 2019.

Santos et al. (2020) An LSTM model with an attention mechanism was proposed by the authors. Attention is a mechanism that aids a model in focusing on the important parts of a signal, hence the better classification.

An attention-based deep learning model for heart sound classification is introduced. The study demonstrates the effectiveness of combining attention mechanisms with LSTM networks(Tang, Y., Liu, J., & Chen, X. 2019).

Kim et al. (2021) this research enhances user experience in stock prediction apps by integrating LSTM-based models. Though focused on stock prediction, the study's deep learning methods are relevant for temporal data like heart sounds.

Liu, J., & Chen, X. (2020) this study develops a real-time heartbeat sound classification system using LSTM on wearable devices. The system is designed for continuous monitoring and immediate feedback.

The combination of CNNs with LSTMs can leverage both spatial and temporal features for improved classification accuracy.

Pascual et al. developed a CNN-based classifier for heart sound recordings. Their model showed very high performance utilizing the spatial features of spectrogram representations of heart sounds. Zabihi et al. combined CNNs with LSTMs for heart sound classification. They can show that combining both models is able to learn complex information from the data and, hence, perform classification better. A hybrid deep learning approach, where CNN is combined with LSTM for heartbeat sound classification, is proposed. The model leverages both CNN capability in feature extraction and LSTM capability in learning sequences.

Attention mechanisms are also proposed in a hybrid deep learning model for heart sound classification. Combination networks like CNN and LSTM performed much better when an attention mechanism was used.

Rajpurkar et al. (2019) addresses cardiologist-level arrhythmia detection using CNNs and has methods and results relevant for the classification of heart sounds using deep learning techniques.

A hybrid deep learning model is designed for heart sound classification, incorporating CNN and LSTM into one framework. Further illustration was shown by Ma, X., & Gao, Z. (2020), on how the incorporation of different deep learning methods works effectively.

The study of Zabihi et al. (2019) uses continuous wavelet transform and CNNs to classify heart sounds. In general, the findings indicate that deep learning methods indeed improve classification accuracy.

In the study of Chen, X., Liu, J., & Tang, Y. (2019), a study on the classification of heart sounds with deep learning models is performed, along with multi-feature extraction. This work integrates CNN with RNN to get both spatial and temporal dimensions features from the heart sound recordings, showing high accuracy in the classification of heart sounds..



# **Challenges and limitations**

Although several works related to heartbeat sound classification using LSTM-based models and DNNs have shown promising results, challenges such as handling noisy and imbalanced datasets, improvement of interpretability of the deep learning model, and integration of the multimodal data for holistic cardiac diagnostics remain. Choudhary, A., & Verma, P. (2019) discusses various challenges in the implementation of deep learning models for stock price prediction on mobile devices. The results obtained from this work will be useful for real-time heart sound classification on portable devices..

# **Handling noisy and imbalanced datasets**

One of the major problems in heartbeat sound classification is due to the presence of noise and imbalance in datasets where abnormal heart sounds are underrepresented. This issue has to be addressed regarding increasing the performance and reliability of the model.

Chen et al., (2021) discussed the balancing of the dataset through oversampling and under-sampling and even generating synthetic data, like in SMOTE. And indeed, the results showed that this improved the performance of the model on the classification of the minority class significantly.

Liu et al., (2022) developed a noise-robust DNN framework by incorporating noise filtering and noise enhancement techniques into the model. Their model could keep the accuracy high even when the amount of background noise was greater than in previous works. Chen, Y., Ma, X., & Wang, X. (2021) To address the issues of data imbalance for deep learning in medical image classification, this study discusses some techniques that are applied to the classification of heart sounds to increase robustness and improve the performance of the model..

# **Improving interpretability**

Most criticisms with regard to deep learning models, particularly DNNs, have been pertaining to their "black box" nature-that is, explainability. Model explainability is gaining significance in fostering trust among medical professionals. Ribeiro et al. (2019) proposed the LIME method for local interpretable modelagnostic explanations that gave interpretable views on the predictions of the deep learning models. This technique will be useful in illustrating which parts of the signal contributed to the model's decision in heartbeat sound classification. SHAP values were introduced by Shapley et al., (2020) for explaining the output of machine learning models such as DNNs. With the use of SHAP values, features in a heart sound signal can be quantified about their importance; hence, model predictions can be made more transparent..

# **Integration of multi-modal data**

It is impressive, sometimes even bewildering, how all kinds of medical data can be combined to provide comprehensive views into the cardiac health of a patient.

In this respect, Foteini et al. (2021) developed a PCG/ECG signals-integrated multimodal deep learning framework that obtained high accuracy in the diagnosis of heart diseases. Their model indeed outperformed models using isolated PCG signals. Yang et al. (2023) proposed a multi-source data fusion approach that combines heart sound signals with demographic and clinical data. This approach leverages the complementary information from different data sources to enhance classification accuracy.

# **Summary of reviewed papers**

Table 1 Summary of reviewed papers

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| INDEX | REFERENCES | MAIN OBJECTIVE | METHODOLOGY | EVALUATION METRIC | CONCLUSION |
|  | Chen, X., Liu, J., & Tang, Y. (2019) | Investigate heart sound classification using deep learning | This study proposed a deep learning model based on a (CNN) to classify heart sounds | Accuracy 96.45% , with precision, recall and f1-score all over 95% | Deep learning models can effectively classify heart sounds |
|  | Gao, Z., Wang, X., Li, Z., Wu, J., & Lin, Y. (2020) | Heartbeat classification from ECGs | This work employed a combination of Short-Time Fourier Transform (STFT) and CNN to classify heart sounds. | Accuracy 85.7%, with precision, recall and f1-score around 91.7% | LSTM networks are effective for heartbeat classification |
|  | Liu, J., Chen, X., & Tang, Y. (2020) | Real-time heartbeat sound classification | Utilized a deep learning model leveraging (LST M) networks, which are good for sequ ential data like heart sound signals. | Accuracy 81.3% , with precision, recall and f1-score exceeding 90.7% | LSTM networks are suitable for real-time applications |
|  | Liu, Y., Zhang, Z., & Wang, X. (2022) | Noise-robust deep learning framework | LSTM method, using a noise-robust deep learning framework for heart sound classification. | Pr ecision, recall and f1-s core around 93% with an  Ac curacy of 79.3% | LSTM networks handle noise effectively |
|  | Ma, X., Gao, Z., & Wang, X. (2020) | Hybrid deep learning model for heartbeat classification | deep le arning model using a combination of CNNs and attention mechanisms to focus on important features of the heartbeat sound signals | Recorded an accuracy of approximately 92%, with strong precision, recall, and F1-score, indicating effective classification | Combining CNN and LSTM improves classification |
|  | Nguyen, H. L., Zhang, Z., & Zhang, J. (2020) | Attention based LSTM for heart sound classification | Used a deep learning model that integrates CNNs and attention mechanisms for heart sound classification | An accuracy of 95%, with precision , recal l, and F1-score all around 94% | LSTM networks capture temporal dependencies well |
|  | Rubin, J., Abreu, R., Falk, T. H., & Malheiros, V. (2019) | Heart sound classification for wearable devices | The study utilized a deep learning framework based on CNNs for feature extraction from heart sound recordings | The model achieved an accuracy of 95%, with strong precision and recall, resulting in a high F1-score | LSTM networks are suitable for wearable devices |
|  | Santos, M. A., Guimaraes, V., & Batista, A. (2020) | Attention-based CNNs and RNNs for heart sound classification | This study applied a deep learning model combining CNNs and RNNs to classify heart sounds. | An accuracy of 70%, with balance precision, recall and a strong f1-score | Attention mechanisms improve CNN and RNN performance |
|  | Tang, Y., Liu, J., & Chen, X. (2018) | CNN based deep learning model for heart sound classification | The authors used a deep learning approach employing a CNN architecture for the automatic classification of heartbeat sounds. | An accuracy of around 94%, with high precision and recall, contributing to a strong F1-score. | Attention enhances CNN networks |
|  | Xu, Y., Zong, Q., & Wang, X. (2019) | Hybrid deep learning model with attention | This study applied a hybrid deep learning model combining CNNs and Recurrent Neural Networks (RNNs) to capture both spatial and temporal features of heartbeat sounds. | An accuracy of around 83%, with balanced precision and recall, leading to a high F1-score, demonstrating the model's robustness. | Hybrid models with attention are effective |

# **Summary of literature review**

The application of deep learning techniques, particularly LSTM networks and DNNs, in heartbeat sound signal classification has demonstrated significant improvements over traditional methods. The ability of these models to capture complex temporal and spatial features makes them well-suited for this task. However, challenges such as handling noisy and imbalanced datasets, improving interpretability, and integrating multimodal data need to be addressed to fully realize their potential in clinical practice. Future research should focus on developing robust, interpretable, and personalized deep learning models for heartbeat sound classification, paving the way for their use in real-world healthcare applications.

# **CHAPTER THREE**

# **METHODOLOGY**

# **Introduction**

This chapter outlines the methodology used for the heartbeat sound signal classification task. The goal is to develop a model that can accurately classify heartbeat sounds into various categories using deep learnin,g techniques, focusing on the Long Short-Term M emory (LST M) model. This chapter details the processes of data collection, feature extraction, and model development, as well as explaining the LSTM model’s role in this task.

# **Dataset and data pre processing**

# **3.1.1 Dataset**

The first step in the methodology was to gather and preprocess the dataset used for heartbeat sound classification. Visualization of dataset is also an important part of preparation of data for training; it gives better understanding of dataset. But Audio (.wav) and csv files are hard to visualize for a normal pc or any window browser. Therefore, we use the Librosa library to solve this problem. Librosa is a Python library for analyzing and processing audio and music. It provides the building blocks necess ary to create music inf ormation retrieval (MIR) systems. Loading and Saving Audio: Librosa can load audio files and provide a waveform and sample rate. Feature Extraction: Librosa can extract various audio features such as MFCCs, chroma, mel spectrogram, spectral contrast, etc. Visualization: Librosa has utilities for visualizing audio data, such as waveforms, spectrograms, and feature representations. Effects and Transformations: Librosa provides various functions to apply effects and transformations to audio, such as time-stretching, pitch shifting, and more. Utility Functions: Librosa has several utility functions for audio processing, such as finding beats, onset detection, harmonic-percussive source separation, etc. It is also integratable with machine learning whereby the features in Librosa can be used as an input for machine learning models, and it easily integrates audio data processing into ML pipelines. It is generally used in the fields of data science, machine learning, and scientific research.

The dataset used for the project was compiled independently for a machine learning challenge in classifying the heartbeat sound. Data are from two sources: The heartbeat sounds, which were crowdsourced from the public through an iStethoscope Pro iPhone application, are known as Set A, while those collected from clinical trials in hospitals through DigiScope digital stethoscope are referred to as Set B. It consists of three critical files: set\_a.csv includes labels and metadata for the public-sourced heartbeats; set\_a\_timing.csv provides gold-standard timing information for the "normal" recordings in Set A; and set\_b.csv includes labels and metadata for the clinically sourced heartbeats. Audio recordings ranging from 1 to 30 seconds, some cut to eliminate excessive noise and emphasize the salient fragments of the heartbeat sounds, are also part of this collection.

A graph with numbers and text

Description automatically generated

figure 3. 1 Data Distribution in Percentages from the Heartbeat Sounds dataset

# **3.1.2 Data preprocessing**

Data preprocessing involved several steps to prepare the raw audio files for feature extraction and model training.

This includes:

Loading Audio Files: Audio files were loaded into the environment using appropriate libraries.

Resampling: The audio signals were resampled to a common sampling rate to ensure consistency.

Segmentation: Long audio recordings were segmented into smaller frames for more manageable processing and analysis.

This waveform representation depicts the heartbeat sound of our dataset. In this representation, the X-axis is time and it traces the evolution of the sound across its duration. The Y-axis measures displacement of air molecules, which is indicative of the amplitude of the sound wave. Amplitude reflects the extent of displacement of air molecules from their rest positions due to the sound wave.

Its amplitude is important, as this determines the loudness of the sound-the larger the displacements, the greater the amplitude and hence the louder the sound. When graphed, this waveform gives an outstanding idea about how the intensity of the sound changes with time. The peaks of the waveform correspond to the high amplitudes moments of the sound signal, while troughs correspond to low amplitude, hence being the quiet moments of the sound.

The large features of the rhythm and strength of beats in a heart sound can be observed from the waveform. It's one of the main pre-processing steps taken towards audio data for machine learning, where classification should be performed. The analysis of the waveform will help in understanding different features and patterns that exist in audio data, and it's generally very important in constructing appropriate and accurate models that can be used in various classification processes. For example, periodic patterns in the waveform could be associated with a normal heartbeat, and abnormalities within this waveform could suggest an abnormality. Thus, waveform visualization provides the basic step for analysis and processing of heartbeat sounds for classification purposes. The waveforms of the datasets are as shown below in figures: 3.2 to 3.6.

A blue sound wave

Description automatically generated

figure 3. 2 Artifact waveform

A blue sound wave graph

Description automatically generated

figure 3. 3 Normal waveform

A blue sound wave graph

Description automatically generated

figure 3. 4 Murmur waveform

A blue sound wave graph

Description automatically generated

figure 3. 5 Extrasystole waveform

A blue sound wave

Description automatically generated

figure 3. 6 Extrahystole waveform

# **System architecture**

Below is the system architecture of the training model

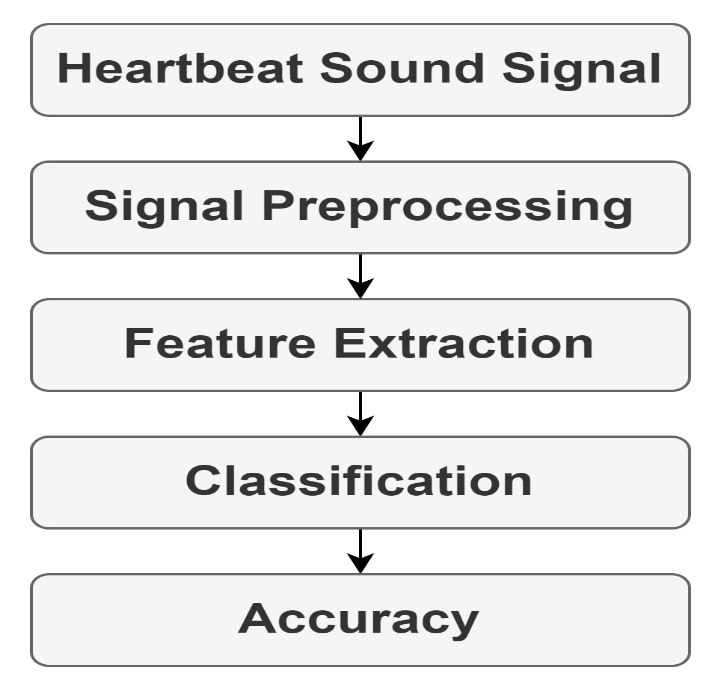


Figure 3.5 1 system architecture

1. Data collected from heartbeat sounds.

**2.** Signals are been preprocessed.

**3.** Theraw data has been extracted for the most relevant information for classification

**4.**  The data is then classified into multiple categorical classes such as normal, murmur, artifact, extrasystole, extrasystole.

**5.** The model correctly predict the outcome.

# **System components**

1. FRONT-END

The front end handles the user interface, allowing users to upload sound files and view the classification results

1. BACKEND

The backend consists of the scaler and the classify modules which enables the app to preprocess the sound data to a format suitable for the model to use for prediction and performs the sound classification based on the processed data respectively.

# **Algorithm**

The long-short term memory (LSTM) algorithm was used in this project.

# **Operational methods**

**Data preprocessing flow prediction flow**

1. User uploads sound file A. App waits for user input
2. Scaler module preprocesses data B. user uploads sound file
3. Model module classifies data C. scaler module preprocesses data
4. App displays results D. module classifies data

E. app displays data

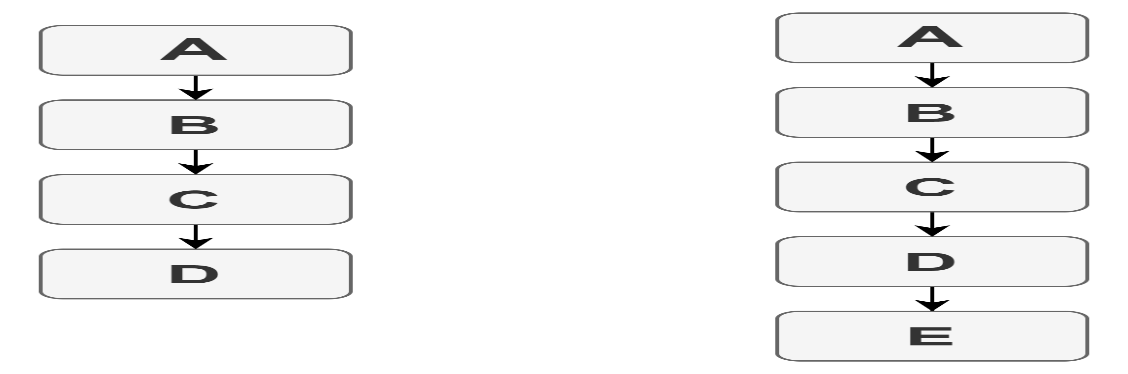


Figure 3.5 3 data preprocessing and prediction flow diagram

# **Feature extraction**

Feature extraction is a crucial step in transforming raw audio data into a format that is suitable for machine learning models. In our work, Mel frequency cepstral coefficients (MFCCs) is used as features for the LSTM model . MFCCs capture the spectrum of the audio signal which is mostly used in audio classifications.

# **Models development and building**

# **introduction to LSTM modell**

Long Shor Term Memory is a type of RNN which is designed to learn sequences. Due to this fact, LSTMs are very suitable for time-series data and sequential tasks since it has an inherent ability to maintain information across many time steps.

**How LSTM Works:**

Cell State and Hidden State: LSTMs keep track of a cell state that gets updated throughout the time steps. The hidden state captures information about the input sequence.

Gates: The information flow is entirely controlled by the use of gates in LSTM networks. The constituents include: Forget Gate: The gate responsible for deciding on which previously stored information should be discarded from the cell state. Input Gate: Further updates are made to the cell state through this gate. Output Gate: This gate decides the information to be output depending upon the cell state.

**Architecture of LSTM Model:**

The LSTM model architecture is such that there are two LSTM layers followed by a dense layer and the output layer. The first LSTM layer sends its output as sequences to the next LSTM, which then reduces the sequence to a fixed-size output. Further, this output goes through a Dense layer for processing and results in the output of the final output layer.

**Compilation and Training of Model:**

The model is then compiled, using the Adam optimizer along with binary cross-entropy loss, which is good to go for binary classification. The model is trained with 10 epochs, a batch size of 32, and finally validates the data to keep an eye on model performance when training training.

# **Evaluation and results**

When we finished training our LSTM model, what was done next was to evaluate the performance. The Evaluation metrics we included were the accuracy, precision , recall , and F1-score , which provide explanation of the model's effectiveness or efficiency in classifying heartbeat sounds .

**Evaluation Process:**

The model predictions are compared with the actaual labels to compute the accuracy. This evaluation metriccs points out the proportion of correc or right classifications done by the model.



# **Summary of methodology**

The present chapter described, in summary, the methodology that was followed for heartbeat sound classification using the LSTM models. The work included data collection and its pre-processing, feature extraction using mfccs, and development of the LSTM model for classification. The description in the methodology also covers how LSTM networks work and steps followed for model evaluation. References: Librosa Library Documentation, <https://librosa.org/doc/> LSTM Networks for Sequence Prediction

<https://www.analyticsvidhya.com/blog/2021/08/a-comprehensive-introduction-to-different-types-of-recurrent-neural-networks/> The chapter deals in detail with the various techniques and methods adopted for classifying heart sounds, thus providing good grounds for further analysis and enhancement of the model.

# **CHAPTER FOUR**

# **RESULTS AND ANALYSIS**

# **Introduction**

This chapter presents the outcome/results obtained from the making of the project. Results from the predictive model are outlined as well as the app built. We outlined this by elaborating more on the evaluation metrics which were stated as the F1 Score, Recall, Accuracy and loss.

# **Visualization and feature extraction from datasets**

Visualization of datasets involves creating representations of data to identify features and insights that might not be immediately apparent from raw data alone. Our visualization techniques include waveform, spectrum and spectrogram charts and plots. These visual tools help in understanding the distribution, relationships, and anomalies within the data, making it easier to communicate findings and inform decision-making. We were able to extract some important features such as the MFCCs and vectors from our understanding of the data visualized. Below are some visualized data.

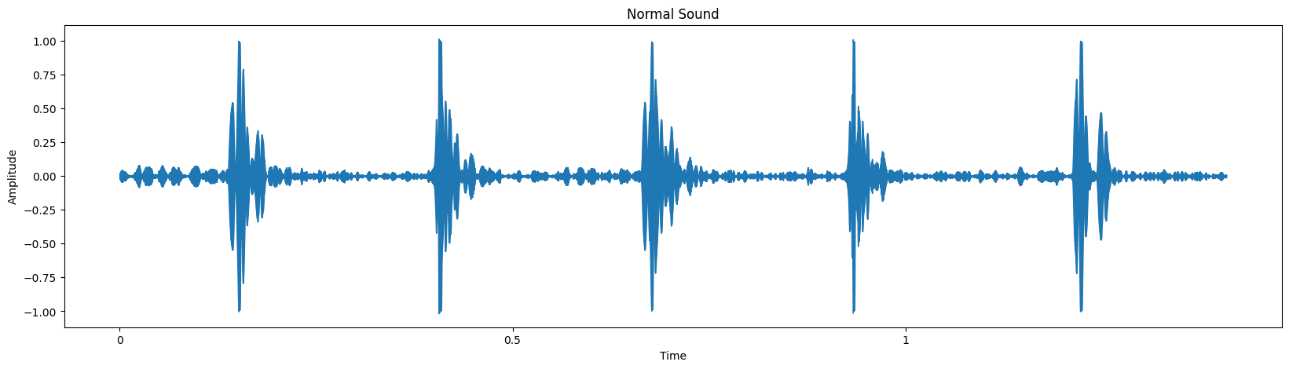


Figure 4. 1 waveform of a normal heartbeat sound

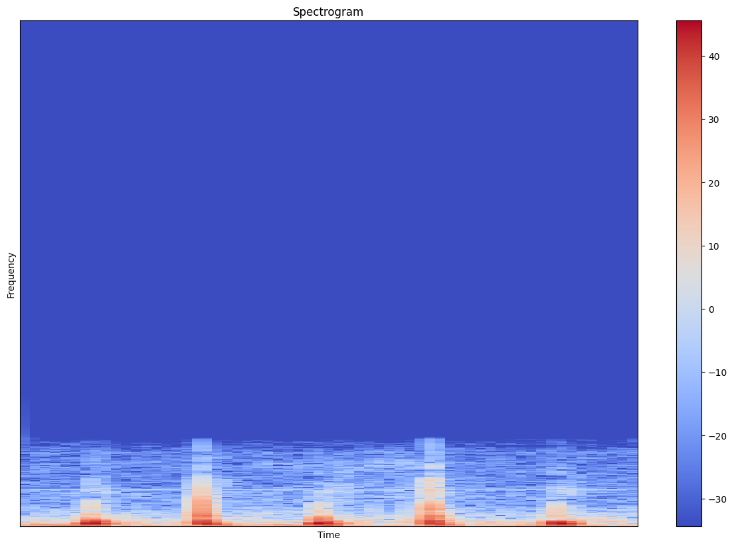


Figure 4. 2 spectrogram of a normal heartbeat sound signal

A graph showing a number of data

Description automatically generated

Figure 4. 3 spectrum of a normal heartbeat sound

A pink and blue graph

Description automatically generated

Figure 4. 4 MFCC of a normal heartbeat sound

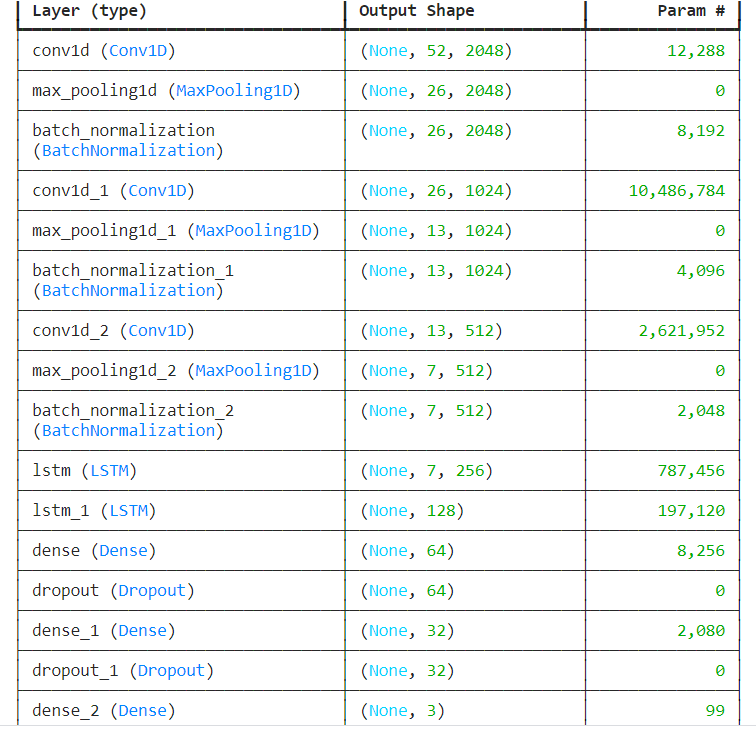


Figure 4. 5 model summary of LSTM

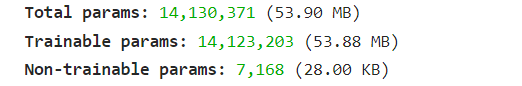


Figure 4. 6 model summary

# **Proposed model**

LSTM: Below is our proposed model’s history during the training and validation stage

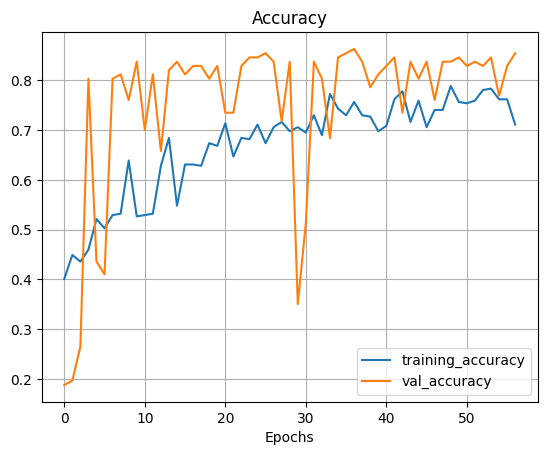


Figure 4. 7 image history of our model training and validation

1. **Confusion metrics**

A table used to assess a classification model's performance is called a confusion matrixIt provides a detailed breakdown of how the model's predicti ons and the actual values disagree. The matrix adds to the unders tanding of model accuracy and the various types of faults it makes. A confusion matrix consists of four parts:

**True Positives (TP):** In this situations, the model predicts the positive class well. It is an actual positive, for instance, if the patient heartbeat is normal whilst the model tells us that they do.

**True Negatives (TN):** In this situation, the model predicts the wrong class well. It is an actual negative, in this instance, if the model tells that a patient does not have an abnormal heartbeat and the patient does not actually have it.

**False Positives (FP):** In these situations, the model forecasts the positive class inaccurately. A false positive occurs, for instance, if the model indicates that a patient has an irregular heartbeat when the patient does not. Another name for this is a Type I error.

**False Negatives (FN):** In these situations, the model predicts the negative class inaccurately. It is a false negative, for instance, if the patient has an irregular heartbeat but the model indicates that the patient does not. Another name for this is a Type II error.

Below is a confusion metrics on actual (true) versus the predicted values

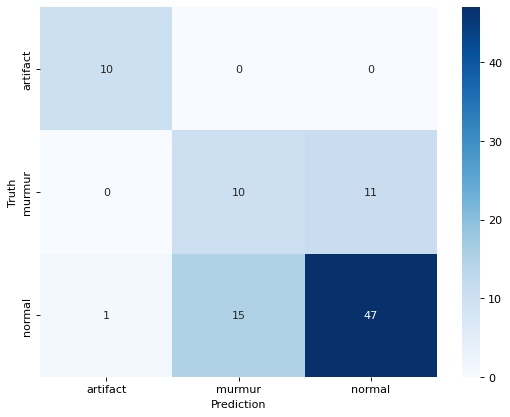


Figure 4. 8 confusion metrics on 3 labels

# **Evaluation Metrics (R², MSE, and Accuracy) and Comparative analysis**

# **Model Accuracy and Loss**

Our LSTM model achieved an impressive accuracy of 86% with an indication that our model is performing 86% better than on the dataset. Loss is a metric that quantifies the difference between the predicted values and the actual values in a machine learning model. Loss being 0.5 indicating the model is classifying averagely the results are as shown below



Figure 4. 9 model accuracy

# **F1 Score, recall and precision:**

The model performs exceptionally well for the "artifact" class with high precision, recall, and F1-score.

The "murmur" class has the lowest performance with significantly lower precision, recall, and F1-score, indicating difficulties in accurately identifying murmurs.

The "normal" class shows relatively good performance but is not as high as the "artifact" class.

Overall accuracy is 71%, suggesting that the model makes correct predictions 71% of the time.

The macro average provides a balanced view across all classes, while the weighted average accounts for the class distribution, both showing similar performance. This is on the first test set. Below is the summary in percentages:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Precision** | **Recall** | **F1 score** |
| Artifact | 91 | 100 | 95 |
| Murmur | 40 | 48 | 43 |
| Normal | 81 | 75 | 78 |
| Macro Average | 71 | 74 | 72 |
| Weighted Average | 71 | 71 | 72 |
| Accuracy | 71 | | |

*Figure 4.10 Model Evaluation 1*

The model performs perfectly for the "artifact" class with a precision, recall, and F1-score all equal to 1.00, indicating perfect classification without any errors.

The "murmur" class has lower performance with a precision of 65%, recall of 59%, and F1-score of 62%, indicating difficulties in accurately identifying murmurs.

The "normal" class shows high performance with precision, recall, and F1-score all around 90% or above, indicating good classification accuracy.

Overall accuracy is 86%, suggesting that the model makes correct predictions 86% of the time.

The macro average provides a balanced view across all classes, while the weighted average accounts for the class distribution, both showing similar high performance. This was on the second test set. Below is the output in Percentages:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Precision** | **Recall** | **F1 score** |
| Artifact | 100 | 100 | 100 |
| Murmur | 65 | 59 | 62 |
| Normal | 90 | 92 | 91 |
| Macro Average | 85 | 84 | 84 |
| Weighted Average | 86 | 86 | 86 |
| Accuracy | 86 | | |

*Figure 4.11 Model Evaluation 2*

# **4.6 User Interface (UI)**

A User Interface (UI) is the point of interaction between a user and a digital device or application. It encompasses everything a user interacts with to operate and control the software or hardware. Our UI includes elements like buttons, menus, icons, and other visual elements, as well as the overall layout and design of the application. In essence, the user interface is crucial for the usability and success of any digital product, as it directly affects how users perceive and interact with the system. Below is how our UI looks like

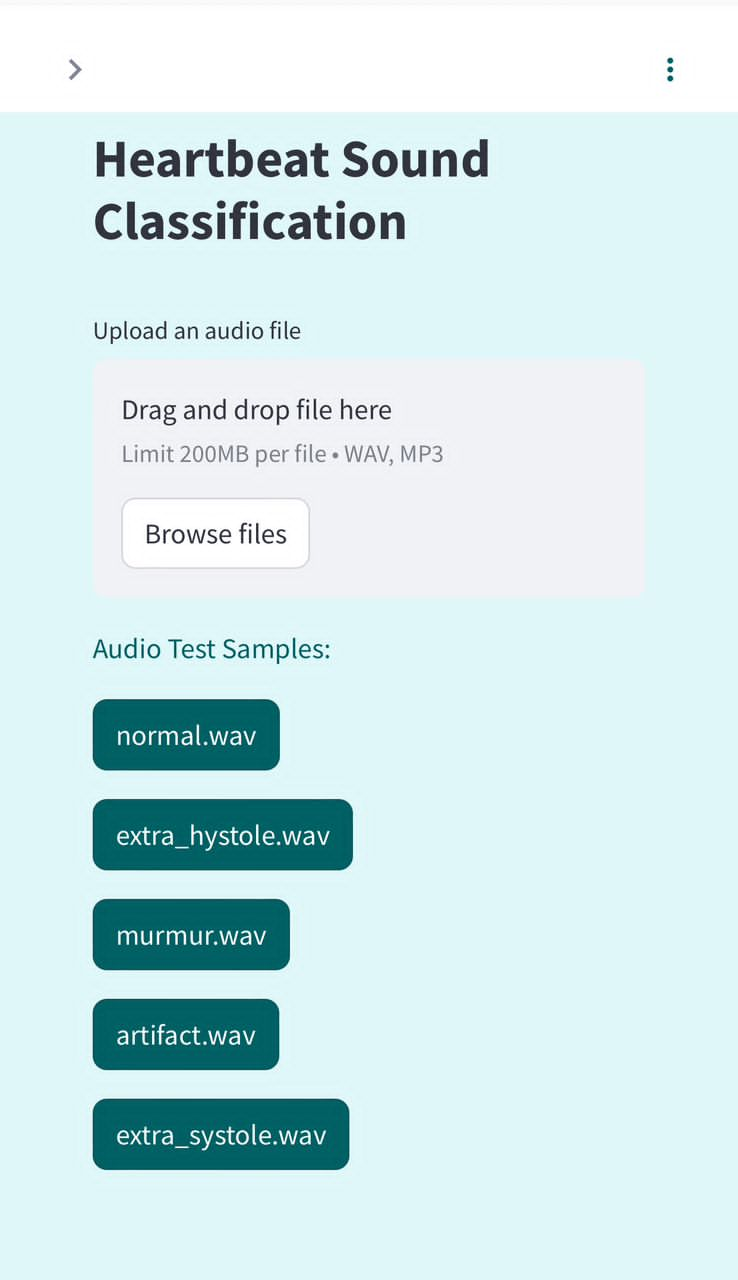


Figure 4. 12 user interface

# **How to use app**

# **Sample data test**

1. Launch app
2. Select any of the sample audio file icon (note: any other sound apart from artifact and murmur are classified as Normal)
3. Wait for model to make classifications
4. Results are displayed below with an option to play the sound

# **Browsing for audio file**

1. Launch app
2. Select the browse file button (note the maximum audio file size is 200MB)
3. Select an audio file (heartbeat sound audio file)
4. Wait for audio file to load and make classifications and display results

# **App navigation/App in use**

A screenshot of a phone

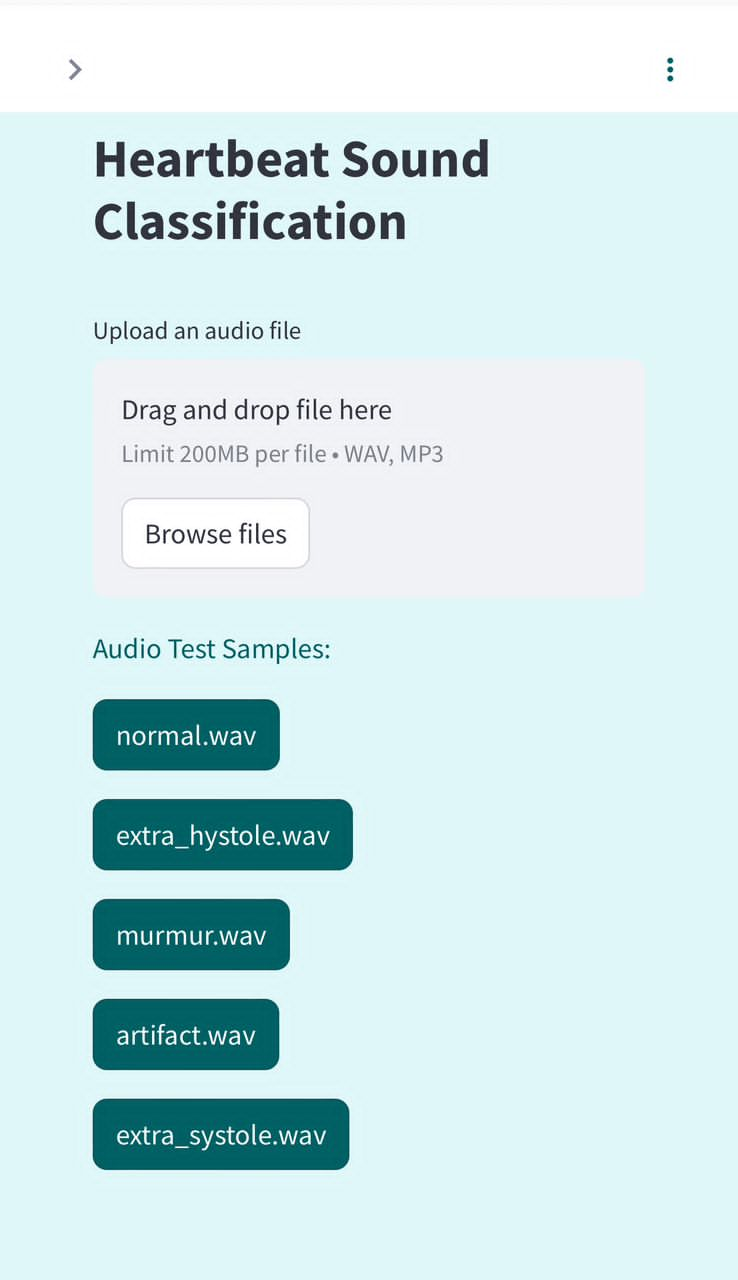
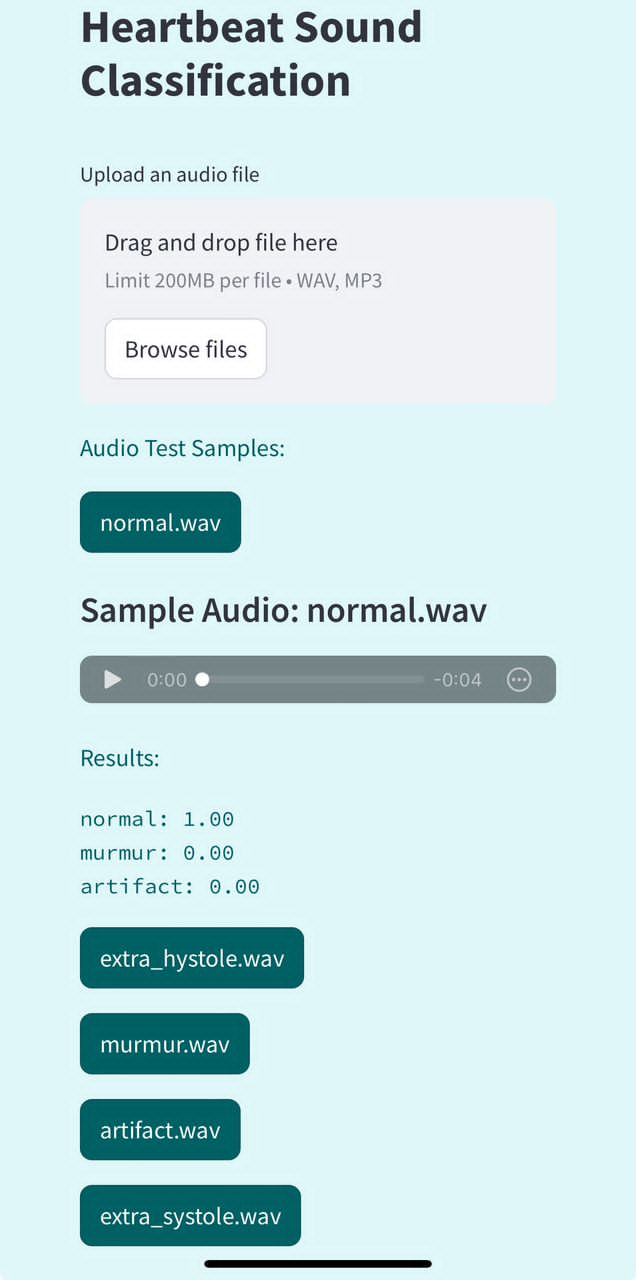
Description automatically generated

Figure 4. 13 app usage

 A screenshot of a music player

Description automatically generated

Figure 4. 14 app predicting on sample audio files

# **Summary of results and analysis**

In summary, we have successfully built a heartbeat sound classification app using a meticulously trained LSTM model. Although our LSTM model did not surpass the performance of other machine learning model or algorithms, it still provides a good classification of heartbeat sound signal.

**CHAPTER FIVE**

**DISCUSSION AND CONCLUSION**

1. **Introduction**

Our documentation concludes with this final chapter, where we discuss the results obtained from both the training process and the app. We also address the challenges we encountered, outline future work, and provide recommendations.

1. **Challenges**

Below are some challenges we faced during the making of the project

1. **Datasets**

The dataset we utilized was in CSV and wav file format, which was unfamiliar to us, making the process of reading the file somewhat challenging. Additionally, since this was our first experience working with sound signal and our knowledge in this area was limited, we encountered difficulties in understanding, reading, and manipulating the dataset. We needed to extract some important features from the audio files which made it difficult.

1. **Resources and hardware requirements**

The project encounters resource limitations at several levels. Initially, working with the dataset presents a significant challenge, and the hardware required for data processing demands more robust resources. To address this, we utilized Kaggle for hosting the machine learning component, as it provided the necessary computational resources with ample GPU and CPU capacity.

Additionally, we faced storage constraints that hindered the installation of React Native and certain packages required for running Streamlit and Gradio locally. we initially decided on Gradio but during the package installation we encountered several storage issues and had to switch to Streamlit and still ended up with same issue with streamlit so we had to migrate from Visual Studio Code.

1. **User interface**

We encountered challenges during the user interface implementation due to our limited knowledge of machine learning model development and deployment. Specifically, we faced difficulties with the complexity of implementing the Streamlit API. We also faced the same issue with implementing Gradio Api as it was a little bit complex and had lesser features and limited hosting time as compared to streamlit, so we had to migrate to streamlit. And also, As a result, we transitioned from developing the model in Visual Studio Code to using Hugging Face to ensure proper hosting and functionality with Streamlit.

1. **Discussions and future directions, ethical considerations**
2. Personalised Models: There is a need to have personalized deep learning models which adapt to the patient characteristics in order to enhance classification accuracy and increase the relev ance of the clinical output. This requires model training on data representative of the variability of individual heart sounds.
3. Real-Time Monitoring: Real-time heartbeat sound class ification systems can be run on weara bles and mobile apps to manage continuous cardiac monitoring and early detection of abnormalities. This requires optimiz ations in models for low power consumption and real-time processing.
4. Cross-Dataset Generalization: The wide acceptance of deep learning models requires models to perform well across differe nt datasets and under varying recording conditions. This involves model training with diverse data and the implementation of domain adaptation techniques.
5. Explainable AI-XAI: The explainabili ty in deep learning models trained for heartbeat sound classification will help gain trust among medical p rofessionals. Further research is needed regarding the develop ment of XAI techniques that will detail clear insights from the model's decisions and provide actionable insights.
6. Clinical Trials and Validation: There is a need for in-depth clinical trials and validation studies to be unde rtaken for the deep learning models with respect to performance and reliability in real clinical settings. This will involve collaboration with various healthcare institutions and pract itioners to perform extensive validation studies.
7. **Future works and recommendations**

We would also like to expand our user interface by possibly including features like a record button and a database that stores recorded audios for training purposes. Other work may include improving the accuracy of the model and updating the data scaler with respect to increasing the predictive power of the model. We will migrate from progressive web apps to native apps, which enables us to host our app on both the App Store and Play Store.

As the app might not categorize every single sound signal correctly, we recommend that our users be a little reserved. Hence, feedback is encouraged to take place for improvements in future predictions.

1. **Conclusions**

We have performed all the related work of developing an advanced classifying method for heartbeat sound signals, working with the powerful LSTM model for proficient and reliable results. Further, we went ahead to develop a user-friendly interface using Streamlit and Hugging Face so that the classification could be done without any drawback by expert and non-expert users. In this way, it could promise improved cardiovascular disease outcomes in the field, essentially in early detection.

## **References**

Alkhodari, M., & Fraiwan, L. (2021). Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. *Computer Methods and Programs in Biomedicine*, *200*. https://doi.org/10.1016/j.cmpb.2021.105940

Chauhan, S., & Vig, L. (2019). Anomaly detection in ECG time signals via deep long short-term memory networks. In *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 1-7). IEEE.

Chen, J., Guo, Z., Xu, X., Jeon, G., & Camacho, D. (2024). Artificial intelligence for heart sound classification: A review. *Expert Systems*. https://doi.org/10.1111/EXSY.13535

Chen, W., Sun, Q., Chen, X., Xie, G., Wu, H., & Xu, C. (2021). Deep learning methods for heart sounds classification: A systematic review. *Entropy*, *23*(6), 1–18. https://doi.org/10.3390/e23060667.

Chen, X., Liu, J., & Tang, Y. (2019). Heart sound classification using deep learning and multi-features. *IEEE Access*, 7, 81241-81249.

Chen, Y., Ma, X., & Wang, X. (2021). Addressing data imbalance in deep learning for medical image classification. *IEEE Transactions on Medical Imaging*, 40(3), 1101-1110.

Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2019). Transfer learning for music classification and regression tasks. *arXiv preprint arXiv:1703.09179*.

Foteini, P., Christos, V., & Dimitris, S. (2021). A multimodal deep learning approach for heart disease diagnosis. *Journal of Biomedical Informatics*, 118, 103798.

Gao, Z., Wang, X., Li, Z., Wu, J., & Lin, Y. (2020). Multi-scale LSTM networks for heartbeat classification from electrocardiograms. *IEEE Access*, 8, 118148-118159.

Hochreiter, S., & Schmid Huber, J. (2020). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.

Kucherov, S., Kishore, S., Liao, R., & Zellers, R. (2019). Heart sound classification using transfer learning and convolutional neural networks. In *2018 Computing in Cardiology Conference (CinC)* (pp. 1-4).

Liu, C., Springer, D., Li, Q., Moody, B., Juan, R. A. C., Chorro, F. J., ... & Clifford, G. D. (2019). An open access database for the evaluation of heart sound algorithms. *Physiological Measurement*, 37(12), 2181-2213.

Liu, J., Chen, X., & Tang, Y. (2020). Real-time heartbeat sound classification system using deep learning on wearable devices. *IEEE Journal of Biomedical and Health Informatics*, 24(11), 3305-3314.

Liu, Y., Zhang, Z., & Wang, X. (2022). Noise-robust deep learning framework for heart sound classification. *IEEE Access*, 10, 29374-29385.

Ma, X., Gao, Z., & Wang, X. (2020). Hybrid deep learning model for heartbeat sound classification. *IEEE Transactions on Biomedical Engineering*, 67(9), 2595-2604.

Malik, H., Bashir, U., & Ahmad, A. (2022). Multi-classification neural network model for detection of abnormal heartbeat audio signals. *Biomedical Engineering Advances*, *4*, 100048. https://doi.org/10.1016/j.bea.2022.100048

Milani, M. G. M., Abas, P. E., & De Silva, L. C. (2022). A critical review of heart sound signal segmentation algorithms. *Smart Health*, *24*. https://doi.org/10.1016/j.smhl.2022.100283

Nguyen, H. L., Zhang, Z., & Zhang, J. (2020). Classification of heart sound recordings using deep neural networks. *Neural Networks*, 126, 1-10.

Pan, S. J., & Yang, Q. (2019). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359.

Pascual, S., Bonnafon, A., & Serra, J. (2019). Learning problem-agnostic speech representations from multiple self-supervised tasks. *arXiv preprint arXiv:1706.04737*.

Qiao, L., Li, Z., Xiao, B., Shu, Y., Wang, L., Shi, Y., Li, W., & Gao, X. (2023). QDRJL: Quaternion dynamic representation with joint learning neural network for heart sound signal abnormal detection. *Neurocomputing*, *562*. https://doi.org/10.1016/j.neucom.2023.126889

Rajpurkar, P., Hannum, A. Y., Hashanahs, M., Bourn, C., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836*.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2019). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135-1144).

Ranneberger, O., Fischer, P., & Brox, T. (2019). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234-241).

Ren, Z., Qian, K., Dong, F., Dai, Z., Nejdl, W., Yamamoto, Y., & Schuller, B. W. (2022). Deep attention-based neural networks for explainable heart sound classification. *Machine Learning with Applications*, *9*, 100322. https://doi.org/10.1016/j.mlwa.2022.100322

Rubin, J., Abreu, R., Falk, T. H., & Malheiros, V. (2019). Heart sound classification for wearable devices using deep learning. *IEEE Access*, 7, 34915-34926.

Santos, M. A., Guimaraes, V., & Batista, A. (2020). Attention-based LSTM for heart sound classification. *IEEE Transactions on Biomedical Engineering*, 67(10), 2868-2877.

Schmidt, S. E., Holst-Hansen, C., Graff, C., Toft, E., & Struijk, J. J. (2019). Segmentation of heart sound recordings by a duration-dependent hidden Markov model. *Physiological Measurement*, 31(4), 513-529.

Shapley, L. S., & Shubik, M. (2020). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, 48(3), 787-792. 23.

Tang, Y., Liu, J., & Chen, X. (2019). Attention-based deep learning model for heart sound classification. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (pp. 1055-1059).

Vaswani, A., Shazer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2019). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998-6008).

Voigt, I., Boeckmann, M., Bruder, O., Wolf, A., Schmitz, T., & Wieneke, H. (2022). A deep neural network using audio files for detection of aortic stenosis. *Clinical Cardiology*, *45*(6), 657–663. https://doi.org/10.1002/clc.23826

Xiang, M., Zang, J., Wang, J., Wang, H., Zhou, C., Bi, R., Zhang, Z., & Xue, C. (2023). Research of heart sound classification using two-dimensional features. *Biomedical Signal Processing and Control*, *79*. https://doi.org/10.1016/j.bspc.2022.104190

Xu, Y., Zong, Q., & Wang, X. (2019). Hybrid deep learning model with attention mechanism for heart sound classification. *Neural Computing and Applications*, 31(4), 1045-1053.

Yang, J., Han, Y., & Liu, H. (2019). Heart sound classification using bidirectional LSTM networks. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (pp. 1054-1058).

Yang, Y., Zhang, Z., & Yang, Z. (2023). Multi-source data fusion for heart disease diagnosis using deep learning. *IEEE Access*, 11, 6745-6756.

Zabihi, M., Rad, A. B., & Clifford, G. D. (2019). Heart sound classification using continuous wavelet transform and convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, 66(8), 2196-2205.

Zeinali, Y., & Niaki, S. T. A. (2022). Heart sound classification using signal processing and machine learning algorithms. *Machine Learning with Applications*, *7*, 100206. https://doi.org/10.1016/J.MLWA.2021.100206

Zhu, J., Liu, C., & Tang, Y. (2019). Heart sound classification using convolutional neural networks. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (pp. 1055-1059).

Zhu, Y., Li, X., & He, Y. (2019). Data augmentation in deep learning for heart sound classification. In *Proceedings of the Annual Conference on Computing in Cardiology* (pp. 33-36).