

**AI-DRIVEN STETHOSCOPE FOR DETECTING DISEASE THROUGH SOUNDS**

By

Muneeza Mukhtar

(212101009)

Tayyaba Kanwal

(212101015)

Syed Ramiz Hussain

(212101016)

**Supervisor**

Dr. Usman Ali Gulzari

Department of Computer Engineering

Institute of Space Technology, KICSIT Kahuta Campus

June 2025



**AI-DRIVEN STETHOSCOPE FOR DETECTING DISEASE THROUGH SOUNDS**

By

Muneeza Mukhtar(212101009)

Tayyaba Kanwal(212101015)

Syed Ramiz Hussain(212101016)

**Supervisor**

Dr. Usman Ali Gulzari

**APPROVAL BY BOARD OF EXAMINERS**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Supervisor

|  |  |  |
| --- | --- | --- |
| **Committee Members** | | |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Dr.WaqarForooq  Associate Professor  ComputerEngineering | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Dr. Usman Ali GulzariAssistant  ProfessorComputer  Engineering | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Engr.MugheesSarwarAwanAssistant  ProfessorComputer  Engineering |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ HoD  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Head of Academics, KICSIT | | |



**In the name of ALLAH ALMIGHTY,**

**the most beneficent, the most merciful**

**“*To Him belongs the dominion of Heaven and the Earth, it is He who gives life and death and He has power over all things”***

***(Al-Quran)***

**DEDICATION**

***“Dedicated to ourselves, Our Families, Respected Teachers, Mentors and Friends, who always had been a source of Inspiration and Motivation for us”.***

# 

# AUTHORS DECLARATION

I here by take complete and sole responsibility over the scholarly research described in my thesis, **AI DRIVEN STETHOSCOPE FOR DETECTING DISEASE THROUGH SOUNDS.**I solemnly affirm that the current research and laboratory work, presented in this manuscript, were created by my own personal efforts, without any major contribution of other people. Any guidance that has been given has been officially acknowledged.This thesis is my own original work.I also certify that no element of this research, in any form, has previously been presented in order to fulfill a prescribed academic qualification, nationally or internationally.I understand that the Institute of Space Technology (IST) has strict standards with regards to academic misconduct. I, as the author of this work, certify that every content is my original work and it has been referenced accordingly. External sources which fall under the times of dates of fair use have been credited, any longer quotations are given with the express permission of the copyright holder. The ideas and insights contained in this work are my own but with all sources appropriately acknowledged.I agree that any future ruling concerning academic dishonesty relative to this work, either prior or subsequent to graduation, will subject me to the risk of the revocation of my degree by IST.

**MEMBER’S SIGNATURES**

\_\_\_\_\_\_\_\_\_\_

**Muneeza mukhtar**

(212101009)

\_\_\_\_\_\_\_\_\_\_\_\_\_

**Tayyabakanwal**

(212101015)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Syed Ramiz Hussain**

(212101016)

**PROJECT SUPERVISOR’S SIGNATURE**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Dr. Usman Ali Gulzari**

**Assistant Professor**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Verified by Plagiarism Cell Officer**

**Dated**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**CERTIFICATE**

This is to certify that the research work described in this thesis is the original work of the author and has been carried out under my direct supervision. I have gone through personally all the data/results/materials reported in the manuscript and hereby certify their correctness/authenticity. I further certify that the material included in this thesis is not plagiarized and has never been used in part or on the whole for a manuscript already submitted or in course of submission in partial/complete fulfillment of the award of any other degree from any institution. I further certify that the thesis has been prepared under my supervision in the prescribed format and endorse its evaluation for the award of the degree of Bachelor of Science in Computer Engineering through the Institute's official procedures.

**Dr. Usman Ali Gulzari**

Insert Supervisor Name

**COPYRIGHT © 2025**

The authors retain copyright to the text of this thesis. Only in compliance with the author's directions may copies (by any method) of the entire work or of any extracts be created and placed in the KICSIT Library. The librarian can gather details. Any such reproductions must include this page. No further copies (by any method) of copies prepared in compliance with these instructions may be made without the author's prior written consent.

Subject to any previous agreement to the contrary, KICSIT is the owner of any intellectual property rights that may be mentioned in this thesis. Any such rights may not be made available for use by third parties without KICSIT's written consent, which will specify the terms and conditions of any such agreements.

Additional details regarding the circumstances in which disclosures and exploitation may occur can be obtained from the KICSIT Kahuta Library.

# ACKNOWLEDGEMENTS

First of all, we thank Allah, the almighty, for giving us the strength to carry on this project and for blessing us with many great people who have been the greatest support in our personal and professional lives. We also thank our parents, for advising, praying, and supporting us in everything in our lives, especially for 4 years of studying Engineering. We owe so much to our whole family for their undying support, and their unwavering belief that we can achieve so much. We would like to thank our university “Dr. A Q Khan Institute of Computer Sciences & Information Technology**”** for teaching us Computer Engineering and for all our professors and doctors.

We deeply thank our Director Masood Khalid whose help and advice were invaluable.

Our sincere gratitude to Dr. Waqar Farooq (HOD CE, KICSIT) and all faculty members for their constant encouragement, motivation, and guidance through all phases of the project. Special thanks and honor to our respected project supervisor Dr. Usman Ali Gulzari. His kind behavior and intellectual support made it possible for us to uplift the project from the foundational stage to maturity. The project is highly indebted to his kind guidance, encouragement, and full cooperation.

I am deeply thankful to the concerned faculty members at KICSIT, especially for their guidance, cooperation, and right direction.

**Sustainable Development Goals**

* **SDG 3: Good Health and Well-Being -**Early diagnostics save lives.
* **SDG 9 - Industry, Innovation and Infrastructure -** AI is Introduced to Healthcare.
* **SDG 10: Reduced Inequalities-**Affordable solutions for under-served areas.



**Table 0.1:Range of Complex Problem Solving and Activities**

|  |  |  |  |
| --- | --- | --- | --- |
| **Range of Complex Problem Solving** | | | |
|  | **Attribute** | **Complex Problem** |  |
| 1 | Range of conflicting requirements | Involve wide-ranging or conflicting technical, engineering and other issues. |  |
| 2 | Depth of analysis required | Have no obvious solution and require abstract thinking, originality in analysis to formulate suitable models. |  |
| 3 | Depth of knowledge required | Requires research-based knowledge much of which is at, or informed by, the forefront of the professional discipline and which allows a fundamentals-based, first principles analytical approach. |  |
| 4 | Familiarity of issues | Involve infrequently encountered issues |  |
| 5 | Extent of applicable codes | Are outside problems encompassed by standards and codes of practice for professional engineering. |  |
| 6 | Extent of stakeholder involvement and level of conflicting requirements | Involve diverse groups of stakeholders with widely varying needs. |  |
| 7 | Consequences | Have significant consequences in a range of contexts. |  |
| 8 | Interdependence | Are high level problems including many component parts or sub-problems |  |
| **Range of Complex Problem Activities** | | | |
|  | **Attribute** | **Complex Activities** |  |
| 1 | Range of resources | Involve the use of diverse resources (and for this purpose, resources include people, money, equipment, materials, information and technologies). |  |
| 2 | Level of interaction | Require resolution of significant problems arising from interactions between wide ranging and conflicting technical, engineering or other issues. |  |
| 3 | Innovation | Involve creative use of engineering principles and research-based knowledge in novel ways. |  |
| 4 | Consequences to society and the environment | Have significant consequences in a range of contexts, characterized by difficulty of prediction and mitigation. |  |
| 5 | Familiarity | Can extend beyond previous experiences by applying principles-based approaches. |  |

**ABSTRACT**

This project introduces the engineering of a smart digital stethoscope with artificial intelligence (AI) that automates the identification of cardiopulmonary diseases. The system also overcomes the serious shortcomings of the conventional auscultation method, which is highly dependent on subjective hearing as well as the experience of the clinician, and thus prone to diagnostic inconsistencies in resource-constrained environments.

The hardware component is an aluminum casing of a lightweight device that records heart and lung sounds, which are transmitted safely through Bluetooth into a mobile or computational system so that processing can be performed on the line. Then the advanced audio pipeline implements noise reduction algorithms to discard environmental and mechanical noise, providing signal clarity in even noisy clinical conditions. This denoised audio is then partitioned into acoustically relevant windows where discriminative features are estimated to represent the response of the human auditory system using Mel-Frequency Cepstral Coefficients (MFCCs).

The main classification is done by an artificial intelligence model represented as a Bidirectional Long Short-Term Memory (LSTM). This architecture has the capability to tune itself to the long-term temporal relations and the fine patterns characteristic of physiological sounds such as heartbeats and breathing patterns. The model, built using a curated dataset of more than 8,000 labeled patient sounds augmented with real-world variations, learns to recognize many different conditions. These are things like Aortic Stenosis, Mitral Regurgitation, and murmurs, and pulmonary diseases like Pneumonia, asthma, and COPD.

The system implemented is very robust since it is able to provide a diagnostic accuracy of over 90% across multiple validation sets. It has real-time analysis and immediate feedback, and it gives offline analysis that allows remote assessments. This feature enables it to have high potential as an early interceptor, and telemedicine and usage by frontline health workers in underserved population.

**Table of Contents**

[Dedication v](#_Toc379236115)

[Acknowledgements..........................................................................................................................x](#_Toc379236118)

[Abstract.........................................................................................................................................viii](#_Toc379236119)

[Table of Contents i](#_Toc379236120)

[List of Figures x](#_Toc379236121)i

[List of Tables x](#_Toc379236121)iii

**CHAPTER 1 INTRODUCTION**  [**1**](#_Toc379236121)

* 1. Project-Background/Overview………………………………………………………………..1
  2. Problem Statement……………………………………………………………………………2
  3. Problem Description………………………………………………………………………….2
  4. Project Objectives…………………………………………………………………………….3
     1. Intelligent Sound Classification Using Lstm………………………………………..….4
     2. Enhanced Diagnostic Accuracy………………………………………………………...4
     3. Early Detection and Proactive Intervention…………………………………………….4
     4. Diagnostic Support for Rural and Remote Areas……………………………………...4
     5. History of Digital Recording and Health Monitoring………………………………..…5
     6. TeleMedicine and Remote Consultations Integration……………………………….…5
     7. Medical Training Educational Tool…………………………………………………....5
     8. Digital Healthcare System Progress……………………………………………….…...5
  5. Project Scope………………………………………………………………………………….5
     1. Target Areas…………………………………………………………………………..6
     2. Functional Scope……………………………………………………………………..6

1.6. SDGs Linkage/Mapping………………………………………………………….………..7

1.7 Overview of the Report……………………………………………………………………..9

**CHAPTER 2 LITERATURE REVIEW** [**17**](#_Toc379236121)

2.1. Traditional Stethoscopes and Limitations…………………………………………………...18

2.2. Early Developments in Digital Stethoscopes...……………………………………………...19

2.3. AI and Machine Learning in Auscultation...…………………………………………...……20

2.4. AI-Powered Stethoscopes in Clinical Practice……………………………………………...20

2.5. The Influence of AI-Enabled Stethoscopes on Healthcare Access and Efficiency…………21

2.6. Evolution of AI in Medical Diagnostics...…………………………………………..............21

2.7. Sensor Technologies for Heart and Lung Sound Classification...…………………………..22

2.8. Deep Learning Models for Heart and Lung Sound Classification...………………………...23

2.9. Challenges and Future Directions…………………………………………………………...23

**CHAPTER 3 REQUIREMENT ANALYSIS**[**…………………………………………………30**](#_Toc379236121)

3.1. User Requirement…………………………………………………………………………...26

3.2. Functional Requirements……………………………………………………………………26

3.3. List of Criteria for Success………………………………………………………………….27

3.4. Lstm model working………..……………………………………………………………….27

3.4.1. Layers………………………………………………………………………………..28

3.4.2. lstm features…...……………………………………………………………………31

3.5. Methodology of Hybrid Model in Proposed AI-Powered Stethoscopes……………………32

3.6. Data Handling and Storage………………………………………………………………….32

3.6.1. Data Preprocessing…………………………………………………………………...32

3.6.2. Secure Data Storage………………………………………………………………….32

3.6.3. Historical Data Tracking……………………………………………………………..33

3.6.4. Real-time logging…………………………………………………………………….33

3.6.5. Scalable Storage Solution……………………………………………………………33

3.7 Data collection……………………………………………………………………………….33

3.7.1. Sound Data…………………………………………………………………………..34

3.7.2. Data Preprocessing…………………………………………………………………..35

3.7.3. Exploratory Data Analysis (EDA)…………………………………………………..35

3.7.4. Dataset Classes………………………………………………………………………36

3.7.4.1. Cardiac Classes………………………………………………………………….36

3.7.4.2. Respiratory Classes……………………………………………………………...37

3.7.5. Feature Extraction……………………………………………………………………37

3.7.6. Model Selection……………………………………………………………………...38

3.7.6.1. Traditional ML Models………………………………………………………….38

3.7.6.2. Deep Learning Models…………………………………………………………..38

3.7.7. Model Training………………………………………………………………………38

3.7.8. Model Evaluation…………………………………………………………………….39

3.8. Dataset collection/ EDA (Exploratory data analysis)……………………………………….39

3.8.1. Dataset Summary…………………………………………………………………….39

3.8.2. Key Features Analyzed………………………………………………………………40

3.8.3. Feature Extraction……………………………………………………………………40

3.8.4. Spectrogram………………………………………………………………………….40

3.8.5. Dataset classes………………………………………………………………………..41

3.8.6. Model Selection………………………………………………………………………42

3.8.7. Model Training……………………………………………………………………….42

3.8.7.1. Heart sounds……………………………………………………………………42

3.8.7.2. Preprocessing…………………………………………………………………..42

3.8.7.3. Feature Extraction…………………………………………………………………...42

3.8.7.4. Model Input…………………………….………………………………………42

3.8.7.5. Training Parameters…………………………………………………………….42

3.8.8. Model Evaluation……………………………………………………………………...43

3.9. Classification of lungs sound dataset by using lstm model (training and result)…………...43

3.9.1. Dataset collection/ EDA (Exploratory data analysis)……………………………44

3.9.2. Dataset Summary…………………………………………………………………44

3.9.3. Feature Extraction………..………………………………………………………...44

3.10. Model Selection……………………………………………………………………………45

3.11. Model Training…………………………………………………………………………….46

3.11.1. Preprocessing…………………………………………………………………….46

3.11.2. Segmentation……………………………………………………………………..46

3.11.3. Feature Extraction………………………………………………………………..46

3.11.4. Data Augmentation………………………………………………………………46

3.11.5. Training Setup……………………………………………………………………46

3.12. Model Evaluation…………………………………………………………………………..47

3.13. Evaluation Metrics…………………………………………………………………………47

3.14. Cross-Validation…………………………………………………………………………...48

3.15. Real-Time Testing…………………………………………………………………………48

3.16. List of Criteria That Define a Successful Project………………………………………….48

3.16.1. Accuracy of Diagnosis…………………………………………………………….48

3.16.2. Real-Time Performance…………………………………………………………...48

3.16.3. Hardware Efficiency………………………………………………………………49

3.16.4. Accessibility of users……………………………………………………….….......49

3.16.5. Data Security and Privacy…………………………………………………………49

3.16.6. Portability and Easy Usability…………………………………………………….49

3.16.7. Scalability…………………………………………………………………………49

3.16.8. Broad Disease Coverage…………………………………………………………..49

3.16.9. Robustness………………………………………………………………………...49

**CHAPTER 4 SYSTEM DESIGN AND IMPLEMENTATION** [**5**](#_Toc379236121)**1**

4.1. Existing System……………………………………………………………………………..52

4.2. Discrepancies of the Existing System……………………………………………………….52

4.2.1. Inadequate Diagnosis………………………………………………………………. 52

4.2.2. No Real-Time Analysis……………………………………………………………...52

4.2.3. No Telemedicine Integration………………………………………………………..53

4.2.4. Human Variability…………………………………………………………………..53

4.3. Proposed System…………………………………………………………………………….53

4.3.1. LSTM Deep Learning Model………………………………………………………...53

4.3.2. Real-Time Analysis………………………………………………………………….54

4.3.3. Telemedicine Integration…………………………………………………………….54

4.3.4. Mobile/Desktop Interface……………………………………………………………55

4.3.5. System Architecture…………………………………………………………………55

4.3.6. Mathematical/Computational Model………………………………………………..56

4.3.6.1. Signal Acquisition………………………………….………………………56

4.3.6.2. Preprocessing………………………………………………………………56

4.4. Functional Decomposition………………………………………………………………….56

4.4.1. Data Acquisition……………………………………………………………………..56

4.4.2. Preprocessing…………………………………………………………………………56

4.4.3. Classification………………………………………………………………………….57

4.4.4 Visualization…………………………………………………………………………..57

4.4.5 Data Storage…………………………………………………………………………...57

4.5. Block Diagram………………………………………………………………………………57

4.6. Implementation Methodology………………………………………………...…………… 58

4.6.1 Data Collection and Preprocessing…………………………………………………….58

4.6.2 Lstm Model Development………………………………………………………..……59

4.6.3 System Integration……..……………………………………………………………...59

4.6.4 Testing and Validation…………………………………………………………………60

4.7. User-Friendly Interface Implementation……………………………………………………60

4.8. AI Algorithm Training……………………………………………………………………... 61

4.9. Deployment……………………………………………………………………………….…61

4.10. Optimization and Scaling…………………………………………………………………..62

4.11. Training and Support………………………………………………………………………62

**CHAPTER 5 TESTING AND RESULTS** [**6**](#_Toc379236121)**3**

5.1. Model Evaluation……………………………………………………………………………64

5.1.1. Accuracy……………………………………………………………………………….64

5.1.2. Precision………………………………………………………………………………..64

5.1.3. Recall………………………………………………………………………………..…64

5.1.4. F1-Score………………………………………………………………………………..65

5.2. Classification of Heart sound dataset by using LSTM Model (training and result)………...65

5.2.1. Lstm model training………..…………………………………………………………..65

5.2.2. Lstm Model Evaluation………………………………………….……………….……65

5.2.3. Confusion matrix……………………………………………………………..…..……66

5.2.4. Classification Report…………………………………………………………….….....67

5.2.5. Final Prediction Of Heart Diseases Using Gui…………...……….………………..….....68

5.3. Classification of lungs sound dataset by using Lstm Model (training and result)….……...68

5.3.1 Lstm Model Training…….... ………………………………………….……………...68

5.3.2. Lstm Model Evaluation……………………………………………….……………...69

5.3.3. Confusion matrix…………………………………………………….…………….....69

5.3.4. Training and validation loss…………….…………………………….………………..…70

5.3.5. Classification Report………….………………………………………………….…...70

5.3.6. Final Predictions………………………………………………………………….….71

5.4 Hardware Required………………………………………………………………………….71

5.4.1 Core Digital Electronic Stethoscope…………………………………………………..71

5.4.2. Acoustic Sensing Module……………………………………………………………...71

5.4.3 Ergonomics and Physical Build………………………………………………………..72

5.4.4. Noise Reduction and Data Preprocessing…………………………………………….. 73

5.4.4.1. Filtering………………………………………………………………….……..73

5.4.4.2. Normalization………………………………………………………………….74

5.4.4.3. Spectrogram Transformation…………………………………………………..74

5.5. Real-Time Testing and Validation…………………………………………………………..74

5.5.1. Diagnosis in real time…………………………………………………………………74

5.5.2. Latency in the system…………………………………………………………………74

5.6. Future Improvements and Optimization…………………………………………………….75

5.6.1. Improved noise suppression…………………………………………………………. 75

5.6.2. More extensive and varied data sets……………………………………………….….75

5.6.3. Continuous retraining………………………………………………………………. ..75

5.6.4. Better feature extraction………………………………………………………………76

**CHAPTER 6 CONCLUSION AND FUTURE WORK** [**7**](#_Toc379236121)**7**

6.1 Conclusion…………………………………………………………………………………...78

6.2 Future Work………………………………………………………………………………….79

6.2.1 Enhanced Model Accuracy and Robustness…………………………………………...81

6.2.2 Advanced Noise Reduction and Signal Processing……………………………………81

6.2.3 Continuous Learning and Model Updating…………………………………………….81

6.2.4 Multimodal Diagnostic Integration…………………………………………………….82

6.2.5 Clinical Validation and Deployment…………………………………………………..82

**REFERENCES** [**8**](#_Toc379236121)**3**

APPENDIX [75](#_Toc379236121)

**List of Figures**

FIGURE 1.1: The AI-powered stethoscope enhances healthcare auscultation…………………..2

FIGURE 3.1 : LSTM Model Layer………………………………………………………………25 FIGURE 3.2: Methodology Steps …………………………………………………………….…28 FIGURE 3.3: Dataset Diversity and Sources...…………………………………………………..30

FIGURE 3.4 : Heart and lung sound Analysis and Data Preprocessing...……………………….31

FIGURE 3.5 : Heart Anatomy…………………………………………………………………...35

FIGURE 3.6 : Bronchi,Bronchial Tree and Lungs………………………………………………36 FIGURE 3.7: Representation Heart Disease Through Pie Chart...……………………………..39

FIGURE 3.8: Playing Sound Using Librosa in Python...………….......................................….40

FIGURE 3.9: Feature Extraction...………………………………………………………………40

FIGURE 3.10: MFCC of Heart Sound…………………………………………………………..41

FIGURE 3.11:Waveform.…………………………...............................................................….41

FIGURE 3.12: Count of each lungs Diagnosis...………………………………………………...44

FIGURE 3.13: Lungs Sound Loaded...………………………………………………………… 44

FIGURE 3.14: MFCCs Graph of lungs Sound...………………………………………………...45

FIGURE 3.15: Chromagram of lungs Sound…………………………………………………….45

FIGURE 3.16 : Spectral Contrast and tonal Centroid……………………………………………46

FIGURE 4.1: Block diagram of the whole system...…………………………………………….58

FIGURE 5.1: LSTM Model Training for Heart Sound Dataset...……………………………….65

FIGURE 5.2 : Training and Validation Accuracy for Heart Sound Dataset...……………….….66

FIGURE 5.3 : Training and Validation Loss for Heart Sound Dataset...………………………..66

FIGURE 5.4 : Confusion Matix for heart Sound Dataset……………………………………….67 FIGURE 5.5 : Classification Report for Heart Sound Dataset...………………………………...67

FIGURE 5.6: Final Prediction of Heart Disease Using GUI...…………………………………..68

FIGURE 5.7: LSTM Model Training for Lungs Sound Dataset...………………………………68

FIGURE 5.8: Training and Validation Accuracy for Lung Sound Dataset...………………...….69 FIGURE 5.9 : Confusion Matrix for Lungs Sound Dataset...……………………………………69

FIGURE 5.10: Training and Validation Loss for Lung Sound Dataset...………………………..70 FIGURE 5.11 : Classification Report for Lungs Sound Dataset...…………………………...…70

FIGURE 5.12: Final Prediction for Lungs Sound Dataset...…………………………………….71

FIGURE 5.13 Core Digital Bluetooth Stethoscope Connectivity……………………………….73

**List of Tables**

TABLE 0.1Range of Complex Problem Solving and Activities……………………………….xii

[TABLE 1.1 Sustainable Development Goal](#_Toc379236121) 7

# CHAPTER-1

# INTRODUCTION

# 

**Introduction**

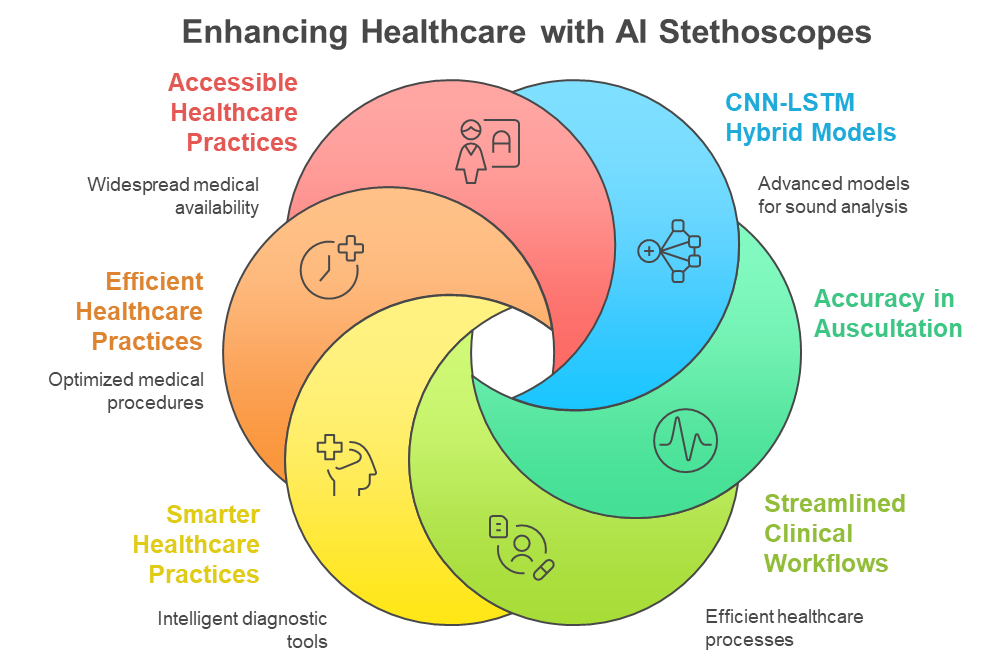
## Project-Background/Overview

The proposed project comes up with an AI-powered digital stethoscope that takes the task of identifying the diseases of the heart and lung. Instead of subjective interpretation, as it is practiced by conventional stethoscopes, deep learning can be used by this device to objectively interpret body sounds. It records the heart and lung sounds, cleans noise and then uses Mel-Frequency Cepstral Coefficients (MFCCs) to extract useful information.

A Long Short-Term Memory (LSTM) model is fitted on a large set of recordings of normal and abnormal sounds, and the sound recordings are categorized into an abnormal or a normal recording in real-time. It will correctly diagnose pathologies like aortic stenosis, mitral regurgitation, pneumonia, and COPD and will often detect subtle patterns otherwise missed by human ears.

The system helps early diagnosis, particularly where medical expertise is low in remote places or low-income areas. The ease of mobility, fastness, and precision can make it an effective tool to be used by frontline workers in the healthcare sector. The next steps can be connected with deployment of an edge device and integration of multi modal sensors to provide diagnosis with higher quality. This invention shows how AI can be used to disrupt traditional medical equipment by making them more accessible and more reliable in care.

As Figure 1.1, showcases the transformative impact of AI-powered stethoscopes in the domain of healthcare. It focuses on better accessibility, a deeper sound analysis using CNN-LSTM models, and more intelligent diagnostics for more efficient medical practices.



**FIGURE 1.1: The AI-powered stethoscope enhances healthcare auscultations [9]**.

.Overall, the AI-Driven Stethoscope for Disease Detection by Sound is a pioneering move towards the integration of traditional medical practice and the latest artificial intelligence. Not only is auscultation made more efficient and accurate, but high-quality diagnosis is also made available to more individuals, especially for those in areas with low healthcare resources. The project is part of the broader aspiration of creating smart, inclusive, and technology-driven healthcare systems that can cater to the needs of diverse populations across the globe.

## 1.2 Problem Statement

**“To design an AI-powered stethoscope system that monitors and analyzes lung and heart sounds using sensor-based technologies and LSTM-based deep learning models, enabling accurate, real-time diagnosis of respiratory and cardiovascular diseases in both clinical and remote environments”**

**.1.3 Problem Description**

Conventional stethoscopes have not altered much over the last century and continue to be a cornerstone of auscultation listening to sounds within the body. Experienced doctors are able to diagnose a number of abnormalities based on experience, but the process is subjective, prone to human error, and based on extrinsic factors such as ambient noise and operator experience. In developing worlds or in overburdened medical facilities, a number of life-threatening diseases like pneumonia, COPD, asthma, pulmonary fibrosis, aortic stenosis, and mitral valve disease go undiagnosed or are diagnosed late because of unavailability of specialists or equipment.The most demanding job is in the early and precise identification of abnormal or weak lung and cardiac sounds, usually in the form of low-intensity murmurs, wheezes, crackles, or irregular rhythms. These irregularities are not always detectable to the naked ear, particularly in populous hospitals, emergency situations, or in rural health centers. Even in advanced hospitals, the lack of standardized, measurable auscultation data can render objective diagnosis problematic and lead to inappropriate treatment.There is a pressing need for an intelligent, independent device that not only captures high-quality body sounds using advanced sensors but also analyzes them in real-time using AI. This project proposes the development of a Digital Stethoscope with Artificial Intelligence added to it, using a Long Short-Term Memory (LSTM) deep learning architecture.

**1.4 Project Objectives**

The last objective will be to develop an artificial intelligence driven stethoscope able to identify and categorize heart and lungs sounds accurately. The device is free of restrictions in the domain of traditional auscultation due to the Long Short-Term Memory (LSTM) networks which localize abnormal sounds and provide real-time clinician impressions. In

the long term, it is hoped that health-care access disparities, particularly in remote regions and regions with limited resources, will be bridged through the package of a portable smart diagnostic device that will allow front line workers to empower and enhance the quality of care **[32]**.

**1.4.1 Intelligent Sound Classification Using LSTM**

The project will develop a digital stethoscope system, which captures audio of lung and heart sounds and then investigates using LSTM models to detect specific patterns. These patterns get classified as normal or not and the system is able to identify the heart murmurs, asthma, pneumonia, bronchitis, and COPD.

**1.4.2 Enhanced Diagnostic Accuracy**

Through objective, machine-based analysis, the system reduces diagnostic errors due to human factors like fatigue, inexperience, or sub optimal listening conditions. This provides more uniform and trustworthy clinical decisions within a variety of healthcare settings.

**1.4.3 Early Detection and Proactive Intervention**

The AI-based system will identify slight abnormalities that can go unnoticed through routine auscultation. Early identification of emerging problems enables on-time medical intervention, minimizing complications and increasing patient recovery rates.

**1.4.4 Diagnostic Support for Rural and Remote Areas**

The handheld AI-powered stethoscope allows non-expert health workers to conduct expert-level auscultation. This makes high-quality diagnostics universally accessible and facilitates the use of community healthcare facilities where expert cardiologists or pulmonologists are

not present.

**1.4.5 History of Digital Recording and Health Monitoring**

The product will have recording, storage, and retrieval capabilities for auscultation signals, allowing longitudinal patient monitoring. Physicians can compare the current and previous recordings in order to identify progression or improvement with the passage of time.

**1.4.6 TeleMedicine and Remote Consultations Integration**

The system will include wireless data transfer capabilities, enabling recorded body sounds to be transmitted to remote experts for evaluation. This broadens the geographic reach of healthcare services and guarantees the involvement of expert opinion in challenging cases even remotely.

**1.4.7 Medical Training Educational Tool**

The tool will also act as a teaching assistant for learners and trainees. Through access to annotated sounds and AI readings, students can learn through real-case scenarios and build auscultation skills.

**1.4.8 Digital Healthcare System Progress**

By implementing AI, cloud storage, mobile app interface, and intelligent data handling, the project facilitates the overall trend toward tech-enabled healthcare infrastructure, ensuring scalability, efficiency, and contemporary patient care methodology.

**1.5 Project Scope**

This project is centered around the creation of an AI-based diagnostic system that utilizes a

Core Digital Electronic Stethoscope to discern and diagnose heart and lung diseases accurately. The stethoscope captures high-quality auscultation data from the patient, which is processed using Long Short-Term Memory (LSTM) networks. The whole pipeline moving from data collection and preprocessing to training the model and predicting disease in real-time is designed to provide a smart, cost-effective, and portable device for healthcare practitioners, especially for rural and resource-poor areas.

The project overcomes the disadvantage of conventional auscultation by providing an automatic identification of significant conditions like pneumonia, bronchitis, chronic obstructive pulmonary disease (COPD), aortic Stenosis (AS), mitral regurgitation (MR), mitral Stenosis (MS), and mitral valve replacement (MVR). Through its smart design, the system provides diagnostic consistency, facilitates early detection, and improves clinical decision-making, particularly in cases where expert personnel are not accessible.

### ****1.5.1 Target Areas****

This system is designed for deployment in:

* Primary Health Centers (PHCs) in rural/remote areas
* Mobile diagnostic and community health units
* Telemedicine platforms and virtual consultation setups
* Hospitals upgrading to AI-based decision support tools
* Medical colleges and training institutions for practical auscultation training

### ****1.5.2 Functional Scope****

* The AI-driven stethoscope system will Capture and digitize heart and lung sounds through the digital stethoscope
* Preprocess and extract audio features from recordings
* Classify sounds as normal or abnormal using an LSTM model
* Detect diseases including:

**Lung**: Pneumonia, asthma, bronchitis, COPD

**Heart**: Aortic Stenosis , Mitral regurgitation, Mitral Stenosis, MVR

* Visualize sound data using waveforms and spectrograms
* Provide real-time alerts for critical conditions
* Securely store patient data and allow retrieval for historical comparison
* Operate fully in offline mode to support uninterrupted healthcare delivery

**1.6 SDGs Linkage/Mapping**

**TABLE 1.1: Sustainable Development Goal**

|  |  |  |  |
| --- | --- | --- | --- |
| **SDG No** | **Description of SDG** | **SDG No** | **Description of SDG** |
| SDG 1 | No Poverty | SDG 9 | Industry, Innovation, and Infrastructure |
| SDG 2 | Zero Hunger | SDG 10 | Reduced Inequalities |
| SDG 3 | Good Health and Well Being | SDG 11 | Sustainable Cities and Communities |
| SDG 4 | Quality Education | SDG 12 | Responsible Consumption and Production |
| SDG 5 | Gender Equality | SDG 13 | Climate Change |
| SDG 6 | Clean Water and Sanitation | SDG 14 | Life Below Water |
| SDG 7 | Affordable and Clean Energy | SDG 15 | Life on Land |
| SDG 8 | Decent Work and Economic Growth | SDG 16 | Peace, Justice and Strong Institutions |
|  |  | SDG 17 | Partnerships for the Goals |

#### **1.6.1 SDG 3: Good Health and Well-Being**

**Goal:** Ensure healthy lives and promote well-being for all at all ages.

The AI-driven stethoscope directly contributes to SDG 3 by enhancing the quality, accessibility, and accuracy of healthcare diagnostics, particularly in the field of cardiovascular and respiratory diseases. By leveraging artificial intelligence to detect early signs of illnesses such as asthma, pneumonia, and COPD, the device enables early diagnosis and intervention, which are critical for preventing complications and improving patient outcomes. Its real-time diagnostic capabilities reduce human error in auscultation and support timely medical decisions, even in environments with limited clinical expertise.

Moreover, by making expert-level diagnostics available in rural and under-resourced areas, the project helps bridge healthcare inequality. It ensures that people, regardless of location or income, can receive timely and accurate assessments of their health. This project contributes to the fact that universal health coverage is a reality, which is a chief component of SDG 3. It also aims at reducing mortality of non-communicable diseases.

**1.6.2 SDG 9 Industry, Innovation and Infrastructure:**

Target: To create strong infrastructure and utilization of diverse and sustainable industrialization and innovation.

Artificial intelligence, machine learning, and digital health technologies are used in the project to modernize traditional instruments. An outstanding example is a new AI-powered stethoscope. By processing the signal as well as the deep learning models (CNN, LSTM), embedding systems and mobile technology it becomes a smart, connected device.

By enhancing medical diagnostics with scalable and cost-effective AI solutions, the project contributes to the digital transformation of healthcare infrastructure. It also encourages the development of telemedicine networks and remote patient monitoring systems, making quality healthcare more resilient and accessible in the face of crises such as pandemics or natural disasters. This innovation not only addresses the healthcare sector but also creates opportunities for startups, healthcare IT companies, and research institutions to collaborate and advance the field of health tech.

#### **1.6.3 SDG 17: Partnerships for the Goals**

**Goal:** Strengthen the means of implementation and revitalize the global partnership for sustainable development.

The development and deployment of an AI-driven stethoscope inherently require **collaborative partnerships** among various stakeholders, including:

* **Healthcare professionals** who provide domain expertise and clinical feedback.
* **Researchers and academic institutions** who contribute to AI model development, validation, and training.
* **Hardware and software developers** who design and implement the system.
* **Government health agencies and NGOs** that help deploy the device in remote and underserved regions.

This project supports SDG 17 by acting as a platform for **cross-disciplinary collaboration**. It encourages the exchange of knowledge and best practices between the healthcare and technology sectors. Additionally, the project can be expanded globally through **international research collaborations**, **technology transfer**, and **policy support**, enabling widespread adoption in both developed and developing countries.

By contributing to **SDG 3, SDG 9, and SDG 17**, the AI-Driven Stethoscope for Detecting Disease Through Sounds stands not only as a medical innovation but also as a tool for sustainable and equitable healthcare development. It exemplifies how technology, when aligned with global development goals, can play a powerful role in building a healthier and more inclusive future for all.

**1.7 Overview of the Report**

This project report is organized into six detailed chapters, each focusing on a specific stage of the system’s development. The first chapter introduces the project by providing an overview of the AI-driven stethoscope, its background, purpose, and overall significance. It outlines the motivation behind the project, the core concept, and the key goals it aims to achieve. This chapter also briefly discusses the problem statement and the importance of improving diagnostic tools through artificial intelligence. The second chapter presents a comprehensive literature review, which explores existing research, technological advancements, and related systems in the field of digital auscultation and machine learning for medical diagnostics. It analyzes previous efforts made to classify heart and lung sounds using AI techniques and highlights their strengths and limitations. This review forms the theoretical foundation for the proposed system and includes references to various academic journals and scholarly Publications. In the third chapter, the report details the requirement analysis of the proposed system. It discusses both functional and non-functional requirements, as well as specific user needs. This chapter defines the technical and operational expectations for the system, including the hardware and software prerequisites and the system’s constraints. The objectives of the project determine guidelines in the designing and constructing phases. Chapter four explains the steps used by the team to construct the AI-driven stethoscope and the approach they adopted. It specifies the diagram of the system, the components used, and the easy development steps:

• Audio signal preprocessing.

• Feature extraction.

• Training the Long Short Term Memory (LSTM) model on classification of heart and lung conditions.

The system is explained in terms of flowcharts, illustrations, and descriptions of every component.

Chapter five discusses testing and results. It details the way the system was tested and which data was applied by the team, including model validation, measurement of accuracy, and confusion matrix. The findings indicate that the model can effectively identify sounds as being normal or abnormal and diagnose the occurrence of diseases such as pneumonia, asthma, and some heart valve disorders. They verify the feasibility of a system that may enhance clinical outcomes. The conclusion and future plans of the project are provided in chapter six. It discusses the key accomplishments: the development of a portable intelligent stethoscope system that assists medical workers to diagnose heart and lung diseases. It also provides suggestions on how it could be improved in the future such as introducing cloud-based patient records, increasing the size of the sound database, and transferring the system into mobile or wearable devices so that the healthcare services could be more accessible and Effective. The report closes with a list of references mentioning all of the research papers, tools, and datasets employed.

# CHAPTER-2

# LITERATURE REVIEW

# 

## Literature Review

This chapter examines the initial efforts done on a similar project and how the concept developed and Matured. This chapter describes previous research on the role of Artificial Intelligence (AI) in assisting physicians in diagnosing illnesses with a particular focus on studies of heart sounds and lung sounds. In the past decade, AI has demonstrated the potential to transform the conventional practice in the medical field, in particular, cardiovascular and pulmonary diseases. The use of a stethoscope in heart and lung sound listening had been the dominant way to rely on the experience of a doctor over many years. With the latest development in artificial intelligence (as with deep learning models), however, automating this type of diagnosis and optimizing its accuracy is now feasible.

Recent research addresses the possibility to sort and investigate heart and lung sounds using deep learning models, including Convolution Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. Such models identify patterns in complicated audio and can detect issues such as heart murmurs, wheezes and other lung diseases. CNNs perform better on image-like data, such as the spectrograms that are obtained by using the heart and lung sounds. LSTMs operate on time-dependent data and have the ability to work with things like time-series data of constant heartbeats or breathing. The two strive to increase accuracy by identifying patterns difficult to detect when observed by the human ear and it is expected to help increase the reliability and accuracy of diagnosis.There remain, however, a few obstacles. The problem with the noise in recordings might arise and the background might cover the target features, thus making AI systems less precise. Research has been done on noise reduction and data processing operations such as filtering, normalization and feature extraction. Although this assists, high-quality input information is hard to obtain in

real environments in a steady manner. The other problem is that there are limited large, labelled datasets to train on, and thus the extent to which AI models can be applicable to other populations and contexts.

Future efforts aim to contribute more useful feature extracting mechanisms, obtain larger and more diverse data sets, and train models that can operate in a noisy setting. It aims to implement AI in clinics and make it assist in the routine work with hearts and lungs.In the following subsections, we will delve into the evolution of AI applications in medical diagnostics, particularly for heart and lung sound analysis, the various sensor technologies used, the deep learning models employed, and the challenges faced by these systems.

**2.1. Traditional Stethoscopes and Limitations:**

It is useful to understand what ordinary, acoustic stethoscopes can and cannot do before trying a stethoscope based on AI. Despite its wide application, acoustic stethoscopes bear numerous disadvantages. An article by Niu et al. (2017) revealed, as doctors listen using the stethoscope, human error, and unbalanced clinical skills can lead to misdiagnosis, particularly with a minor issue in the heart or lungs. These low-voiced sounds are highly dependent on the skill of the clinician hearing them, and even great clinicians may fail to hear low-volume murmurs or mild arrhythmias at early stages. Within adverse clinic settings associated with excessively high sound background, regular stethoscopes are often hard to utilize with proper care. These shortcomings demonstrate our necessity to have more developed devices that could enhance or supplement the diagnostic capabilities of standard stethoscopes**[1]**.

**2.2. Early Developments in Digital Stethoscopes:**

Digital stethoscopes eliminated some of the constraints that the older types of stethoscopes possessed. The first types to record heart and lung sounds electronically were the 3M Littmann Electronic Stethoscope and the Welch Allyn Stethoscope which were also capable of amplification and cutting noise. This has facilitated easy listening by doctors and storage or transmission of the recordings to aid listening later. These gadgets nevertheless continued to depend on the personal opinion of the practitioner as to the meaning of the sound. The feature of artificial intelligence was to be added next.

**2.3. AI and Machine Learning in Auscultation**:

Researchers and scientists have currently resorted to automating the analysis of sounds heard using a stethoscope with the help of machine learning (ML) and deep learning (DL). They aim to design stethoscopes, which can help and even substitute a clinician during the diagnosis with the help of AI.

Some research projects have been carried out into this question: can AI algorithms distinguish between sounds of the heart and those of the lungs? As an illustration, Yang et al. (2018) demonstrated that deep learning models could reliably distinguish normal heart sounds and pathologic murmurs with the same diagnostic rate as professional clinicians. The study involved an application of convolutional neural network (CNN) that examined large records of heart-sounds, and identified that the system was able to classify various categories of murmurs with sensitivity as well as specificity **[2]**.

Liu et al. (2020) trained an AI-based model that can identify any abnormalities in lung sounds, e.g., wheezing, crackles, and rhonchi, related to respiratory conditions such as asthma or pneumonia. The AI model classified lung sounds with a good proportionality by combining multiple feature-extraction methods with random forests and support vector machine (SVM) algorithms. These conclusions contribute to a developing accumulation of evidence that demonstrates, AI programs, deeper learning solutions in particular, are capable of parsing acoustic features to identify pathologies overlooked by human professionals **[3]**.

**2.4. AI-Powered Stethoscopes in Clinical Practice:**

Stethoscopes with AI might transform acquisition of doctors. A study conducted by researchers in Jeong SG. (2025) examined a web-based video telemedicine program including the AI-powered digital stethoscope. They demonstrated that their AI model was effective in the field, indicating the correct diagnosis of heart murmur and lung issues providing useful information to telemedicine doctors**[4]**. Another study by Guo L. (2025) also touched upon the use of AI-stethoscopes in outpatient care and concluded that the latter can be helpful to detect and triage cardiovascular and respiratory diseases, particularly in rural or underserved regions. AI stethoscopes also have an ability to enable clinicians to store and study long-term information**[5]**. In Chen Z. (2025), scientists described how AI-enabled stethoscope could monitor individuals with chronic conditions such as COPD or coronary heart failure. Since the machine continues to monitor the sound of the heart and lungs, it is possible to observe the progress of the disease and, thus, provide the healthcare provider with an opportunity to amend therapy regimens more quickly and accurately**[6]**.

**2.5. The Influence of AI-Enabled Stethoscopes on Healthcare Access and Efficiency:**

In resource-poor and rural areas, where the availability of healthcare workers with advanced diagnostic competence might be low, the devices can level the playing field in terms of access to quality healthcare. Udoy IA. (2025) discussed the use of AI stethoscopes in remote diagnostics, highlighting their potential to fill the gaps between patients and specialists in telemedicine care. Through the cloud-based connectivity of digital stethoscopes, doctors in cities can remotely diagnose patients' heart and lung conditions, come to timely diagnoses, and prescribe medications, all without needing the patient to visit a clinic. Also, the application of AI in stethoscopes can ease the workload for healthcare professionals **[7]**. A study conducted by Farha F. (2025) identified that AI-enabled stethoscopes shortened diagnostic time by rapidly pinpointing essential abnormal sounds, so doctors could concentrate on other matters of patient care. By mechanizing the preliminary assessment, AI-enabled stethoscopes optimize clinical processes, potentially shortening wait times and enhancing overall health care efficiency **[8]**.

**2.6. Evolution of AI in Medical Diagnostics:**

During the past several decades, AI has achieved significant advances in clinical practice, particularly, in diagnosis. Simple rule-based methods were followed to de-mimic the thoughts of human doctors in early systems. They subsequently resorted to machine learning and deep learning. The latest methods, most notably neural networks, are already successful at locating complex patterns among the large volumes of data. Medical diagnostics is one of the areas of AI development that has become increasingly complex as it makes decisions to identify heart conditions, analyze medical images, and even predict patient outcomes.

#### Diagnostic work has taken a new turn with the emergence of deep learning, more so LSTMs. It allows machines to categorize and forecast the disease based on audio, visual as well as sensor data. Such systems, e.g. in heart and lung sound analysis, are capable of picking up fine changes in sound that could be an indication of valve disorders, arrhythmias or breathing abnormalities.

#### **2.7 Sensor Technologies for Heart and Lung Sound Classification:**

Heart and lung sounds are obtained through sensors to diagnose. The most typical are microphones, electronic stethoscopes and wearable sensors. These gadgets record the aural information that is required by the physicians to identify physiological states. Classic stethoscopes are not accurate and not very performant in environments with noise pollution resulting in false diagnosing or absence of it. Digital stethoscopes give digital audio recordings that can be analyzed better with the help of artificial intelligence models. They are gaining traction due to the clear sound signal they pick up allowing detail patterns of the disease to be identified. These sensors have moved into wearable sensors, namely chest patches and smartwatches, capable of recording heart and lung sounds 24/7, where they are valuable in AI diagnostic systems.Other sensors and devices such as accelerometers and gyroscopes have been added to improve the quality of the recorded data. The additional sensors can add more information regarding position and movements all over the body which can refine the accuracy of detecting the disease **[10]**.

**2.8 Deep Learning Models for Heart and Lung Sound Classification**

Deep learning can learn to sort heart and lung sounds based on its own features on the raw data. Long Short-Term Memory (LSTM) networks are the most used models.CNNs are

Spatial data oriented and are most commonly used in image classification. They are useful with spectrograms which resample heart and lung sounds in the time-domain into an image. LSTMs are a form of the recurrent neural network (RNN) designed to handle sequential data, thus a perfect fit to time series data like heartbeats, and the breathing cycle. LSTM networks have been adopted specially to identify patterns by researchers associated with particular heart and lung disorders in audio records.These models have yielded prosperous results but, they are still limited by issues such as minimization of noise, coping with data variability, and requiring big amounts or sample sizes to work effectively across different populations.

#### **2.9 Challenges and Future Directions**

There are still serious limitations in terms of artificial intelligence to sort out heart and lung sounds.

Noise is one of the major disincentives. Background sounds of the room or even of the patient itself may mix with the signal, becoming difficult to identify correct patterns with the help of machines. Correcting such an issue requires data-cleaning procedures such as noise filters and signal amplifications, in order that the model will be able to view clearer sounds.

The second challenge is the limited availability of quality, tagged data. Deep learning systems require enormous datasets, which is time- and finances consuming to generate and label. There are also different datasets that are required to make models applicable to individuals of various age groups, sex, and Ethnicities.

# CHAPTER-3

# REQUIREMENTS ANALYSIS

## Requirements Analysis

This chapter covers the user requirements, functional requirements, and important points that are necessary for a project to succeed.

## 3.1 User Requirements

The AI-Driven Digital Stethoscope and Diagnosis System is designed with the end-user in mind, ensuring ease of use and accessibility. Users primarily healthcare professionals, medical students, or field health workers require a system that can provide accurate, real-time diagnostic support without complex technical knowledge. The system must offer a user-friendly interface where health data such as symptoms, heart sounds, and lung sounds can be easily accessed and interpreted. To accommodate users of various skill levels, the system should include intuitive controls, clear visuals (Waveforms, Spectrograms), and voice feedback (text-to-speech) for accessibility. For doctors and specialists, historical tracking of patient data and exportable reports are vital for clinical decision-making.

### ****3.2 Functional Requirements****

Health monitoring and diagnosis system should fulfill the following six requirements:

* **Analysis using AI:** A neural network model based on LSTM labels the sounds as regular or abnormal and is able to identify certain diseases including asthma, pneumonia, chronic obstructive pulmonary disease (COPD), heart murmurs, stenosis, and valve disorders.
* **Display:** Based on the predicted diagnosis, the targeted outcome is displayed in a user-friendly desktop or a mobile application in both visual displays, and clear text.
* **Data storage:** All patient audio files and diagnostic histories are securely stored on the system to be accessed later at the convenience of future analysis and continuity monitoring of the patient.
* **Universal support:** The system is capable of running 24 hours, working in remote areas and supporting multiple languages.

**3.3 List of Criteria for Success:**

To be used in clinics, the AI-driven stethoscope must meet a range of benchmarks.

* One is that it should be very accurate in its diagnosis. This entails it having to classify heart and lung sounds in a mistake-free manner thus eliminating false positives, and false negatives.
* Second, the system has to provide the healthcare provider immediate feedback. This will enable rapid decision making when examining patients.
* Third, it should be characterised by low error rates. These classification and feedback errors should remain minimal such that the diagnoses are dependable and consistent.
* Fourth, the system Should be expandable. It ought to be practically applicable to other healthcare environments and adaptable to other disease contexts and patient populations.

#### **3.4 Lstm Model Working:**

LSTM (Long Short-Term Memory) is a type of Neural Network called Recurrent Neural Network (RNN). It is designed to operate on sequential data, or data with a time-based sequence to it-e.g. a sequence of heartbeats in an audio file.

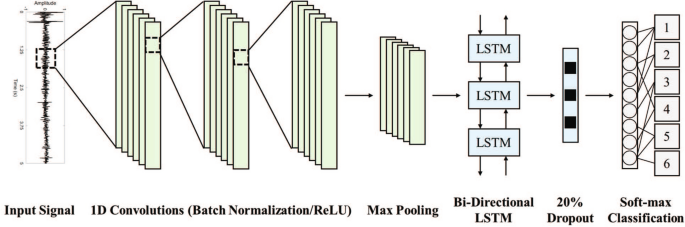
**3.4.1 Layers:**

The stethoscope tool is based on AI and its deep learning architecture is well structured such that it has the ability to label heart and lung sounds as per their acoustic patterns. At the center there is Bidirectional Long Short Short memory (BiLSTM) layer. BiLSTM layer is significant since it extracts temporal characteristics of the input audio streams. The stream of Mel-Frequency Cepstral Coefficients (MFCCs), an indication of sound characteristics in time is often the input. These MFCC sequences are in the form (None, 25, 40) with the amount of time frames, namely 25, and the amount of features per time frame 40. When added to a BiLSTM layer with 128 units in the forward and backward direction, the output is (None, 25, 256), almost doubling the size of the feature as it is the combination of the outputs of the forward and backward directions.This two-way processing allows the model to learn both past and future context, which is very important in recognizing the rhythm and abnormalities of physiological sounds accurately.The network follows the BiLSTM with a stack of Dense layers, also referred to as fully connected layers. These layers are used to compress the high-dimensional output from the BiLSTM to more compact and meaningful representations. The initial dense layer lowers the dimension from 256 to 128 units to capture abstract audio features. The second dense layer lowers this to 64, and the third compresses it to 32 units. These transitions cause the network to increasingly emphasize the

most important patterns as it eliminates redundant information. To enhance generalization and prevent the problem of Overfitting, Dropout layers are placed after every dense layer. Dropout randomly shuts down some proportion (typically 30%) of neurons at training time, which compels the model to learn more stable features rather than memorize training data.Following dense and dropout layers, data still has a 2D shape since it started life from sequential audio input. Hence, a Flatten layer is employed to transform this multi-dimensional tensor into a 1D vector. For instance, if the input to the flatten layer is of shape (None, 25, 256), it will be reshaped into (None, 6400), a single vector of features per audio sample. This is required because the last classification layer needs a flat input. Lastly, the Output layer consists of a Dense layer with two neurons for each class (e.g., "Normal" and "Abnormal"). This layer applies the softmax activation function to generate probabilities for every class, allowing the system to predict with certainty the health status given the audio input. For instance, when the softmax output is [0.92, 0.08], the model categorizes the sample as "Normal" with 92% certainty.All of these layers collaborate to convert raw acoustic sequences into informative clinical knowledge. The BiLSTM processes the temporal dynamics, the dense layers generalize the features, and the dropout layers provide reliability, while the output layer provides the final classification. The current approach to classification of complex biological sounds with an AI stethoscope using architecture is able to achieve extremely high accuracy and as such is useful in early and automatic detection of lung and heart diseases. The system comprises of Long Short-Term Memory (LSTM) networks which are used to process heart and lung sounds. LSTM is a type of Recurrent Neural Network (RNN) developed to work, specifically, with sequential data. In contrast to traditional feedforward neural networks that take each input to be independent, by

design LSTM is able to remember information as time steps progress and, as such, it is also more naturally suited to applications that require time-series data. When listening to the heart and lungs, sounds have the cylcing overtime, e.g. the lub-dub beat of heart, or breathing in- out beat. This is why LSTM networks are specifically better suited to analyzing such audio patterns since they are able to recognize long-term dependencies which may teach the network that the physiology is normal or an abnormality exists **[21]**.

Figure 3.1, represents a deep learning pipeline for signal processing of 1D signals such as audio from a digital stethoscope. The features are extracted from signals by convolutional layers, normalized and activated. Pooling is then used to reduce its complexity before it's analyzed by the bi-directional LSTM layer to know the time-related context. Dropout is used to avoid the overfitting problem and the network is terminated with a softmax layer for final classification, such as classification of heart condition.



**FIGURE 3.1: LSTM Model Layers [11].**

The key difference of LSTM compared to other RNNs is that it can be capable of keeping track

of long-range patterns free of the vanishing gradient problem which plagues normal RNNs. A memory cell is incorporated in each of the LSTM units and the units have three gates depending on the movement of information in the unit, input gate, forget gate and output gate. This forget gate selects what portion of the information of the past to forget, the input gate adds new information to memory and the output gate determines what portion of that memory should be passed to the next layer. Due to these mechanisms of gating, LSTMs have the ability to discern what pieces of knowledge should be updated and which should be stored, a capability that would be vital in times when discerning noise meant to be ignored and patterns that indicate a disease. As an illustration, the repetition of the frequency, and time of occurrence of the heart murmurs or lung sounds in digital sound recordings repeats through a number of steps. These patterns can be easily identified and associated using LSTM with the already known disease profiles.One of the main approaches first converts raw lung and heart sounds into a numerical array using such methods as Mel-Frequency Cepstral Coefficients (MFCCs) that describes how the human ear detects sound. The MFCC data should then be densely packed into a fixed-size sequences- fixed number of frames or number of time steps with a fixed number of coefficients.This sequence is input into the LSTM network. The initial layer of the model is usually a Bidirectional LSTM that reads the sequence both in forward and backward directions to learn about dependencies from both past and future contexts of the signal. This is especially helpful for medical audio, where the occurrence of an anomaly could rely on both past and future sound patterns. The bidirectional layer produces a dense representation of the signal's temporal dynamics, which is subsequently fed into a sequence of dense layers that gradually decrease the data's dimensionality while retaining its most essential details.To avoid overfitting and enhance generalization, dropout

layers are used in between the dense layers, which would shut down randomly a subset of the neurons when training. After the final dense layer, a flattening layer is used to convert the multidimensional tensor output into a one-dimensional vector, which is suitable for classification. The last output layer is a dense fully connected layer with softmax activation, generating probabilities for every target class e.g., "Normal" or "Abnormal." The model is trained on a large set of labeled heart and lung sound recordings, and it learns to transform the temporal patterns in the MFCC sequences into the appropriate diagnosis. In training, the model weights are adjusted via backpropagation through time (BPTT) along with an optimizer like Adam, who adjusts the weights to reduce classification error **[24]**.

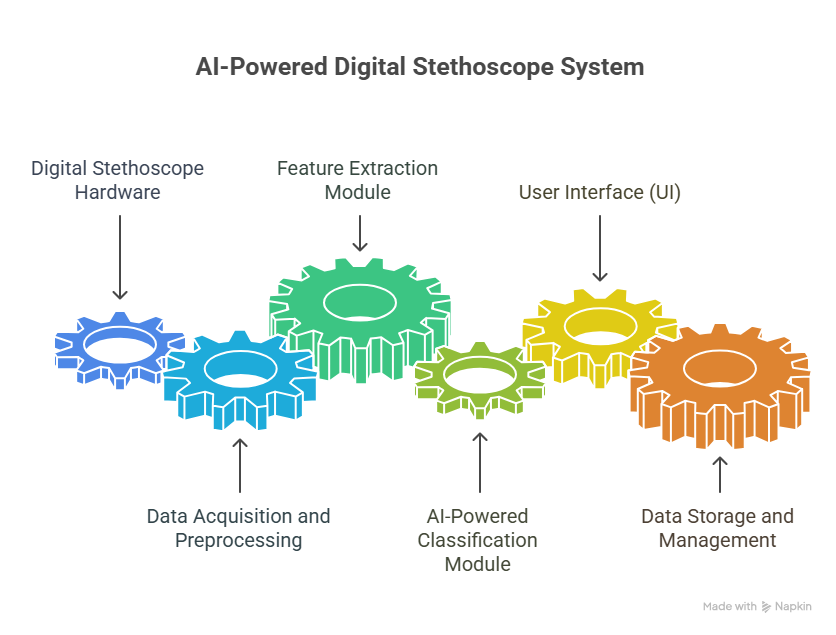
**3.4.2 LSTM features**

Feature are the temporal patterns and dependencies that the LSTM network picks up from the input sequence of audio over time. Following MFCC or other time-series feature extraction from heart and lung sounds, the LSTM processes these sequences sequentially, capturing the rhythm, duration, intensity, and transitions of the sounds—like heartbeats or breathing cycles. In contrast with conventional models, LSTM does not examine single frames in isolation; it learns the way features change over time and is therefore very useful for identifying abnormalities in terms of murmurs, wheezing, or arrhythmia. The features learned are context-aware and embody the essential dynamics of the sound, which the network exploits to separate healthy from diseased conditions at classification **[20]**.

### 3.5. Methodology of Hybrid Model in Proposed AI-Powered Stethoscopes:

The methodology for developing and evaluating an AI-powered digital stethoscope involves

several key steps: data collection, signal processing, model training and validation, device integration, and clinical evaluation. This paper cross-checks the performance of an AI-based digital stethoscope in practice in clinics. The question it raises is whether the machine is capable of identifying heart and lung sounds properly and assist in diagnosing a cardiovascular and breathing issue. Figure3.2, presents the diagram that illustrates the key stages of the Hybrid-Model through which this work was done.



**FIGURE 3.2: Methodology Step [13].**

#### **3.6 Data Handling and Storage:**

The system must store and handle large amount of data from heart and lung sound recordings. Important data management needs are:

**3.6.1 Data Preprocessing:** Mechanisms to clean and normalize data eliminating noise and artifacts of recordings to ameliorate the quality of feed to the AI models should be in the system.

**3.6.2 Secure Data Storage:**All data of the heart and lung sound, including the diagnostic reports need to be conserved. These storage are done in accordance with healthcare regulations such as HIPAA (Health Insurance Portability and Accountability Act) or GDPR (General Data Protection Regulation).

**3.6.3 Historical Data Tracking:**The tracking of historical data enables saving and monitoring the ways the heart and lungs of a given patient sound over a period of time. Doctors are then in the position to observe how such sounds vary.

**3.6.4 Real-time logging: All diagnostic results obtained on the AI driven stethoscope should be logged in real-time so that data can be used to retrain the model and evaluate performance, over and over again. In this way, the AI stethoscope can continue to improve.**

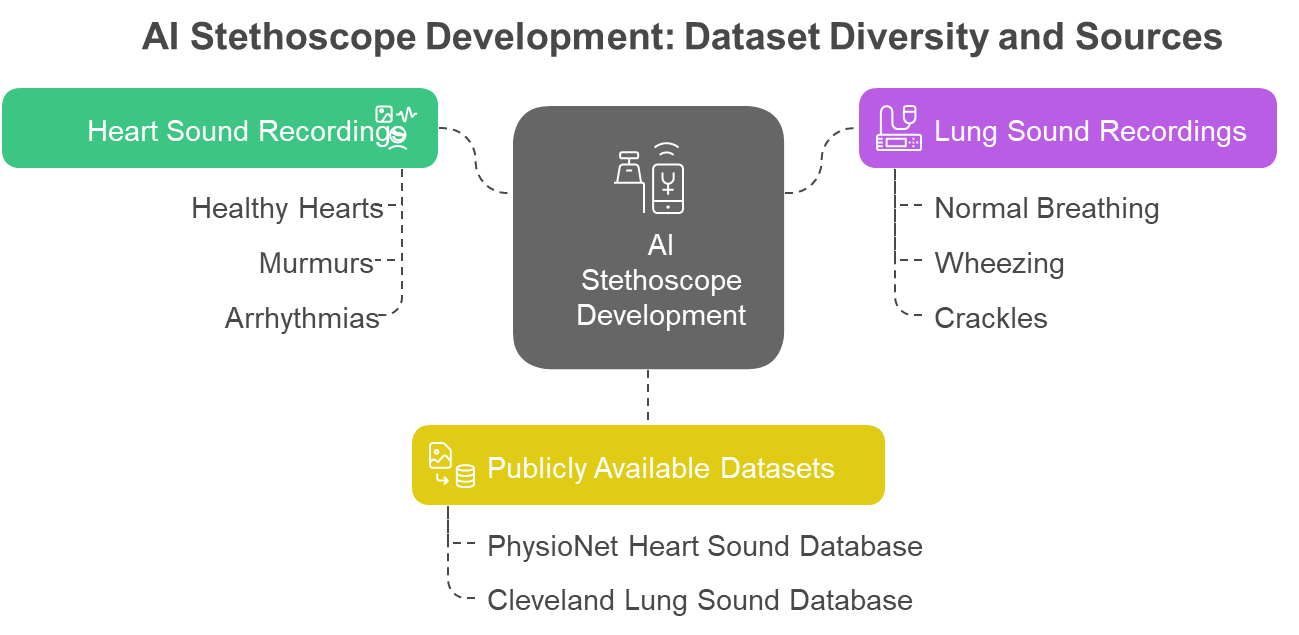
**3.6.5 Scalable Storage Solution:The storage should be one that can expand as the population of users expands, thus facilities within the high-volume healthcare environment can be assured of reliability in the long run.**

**3.7 Data Collection:**

**Training and testing Through LSTM distinguishes Heart and Lungs sounds.**

1. Dataset collection.
2. EDA.(Exploratory data Analysis)
3. Dataset classes.
4. Feature extraction.
5. Model selection.
6. Model training.
7. Model evaluation
   * 1. **Data Collection:**

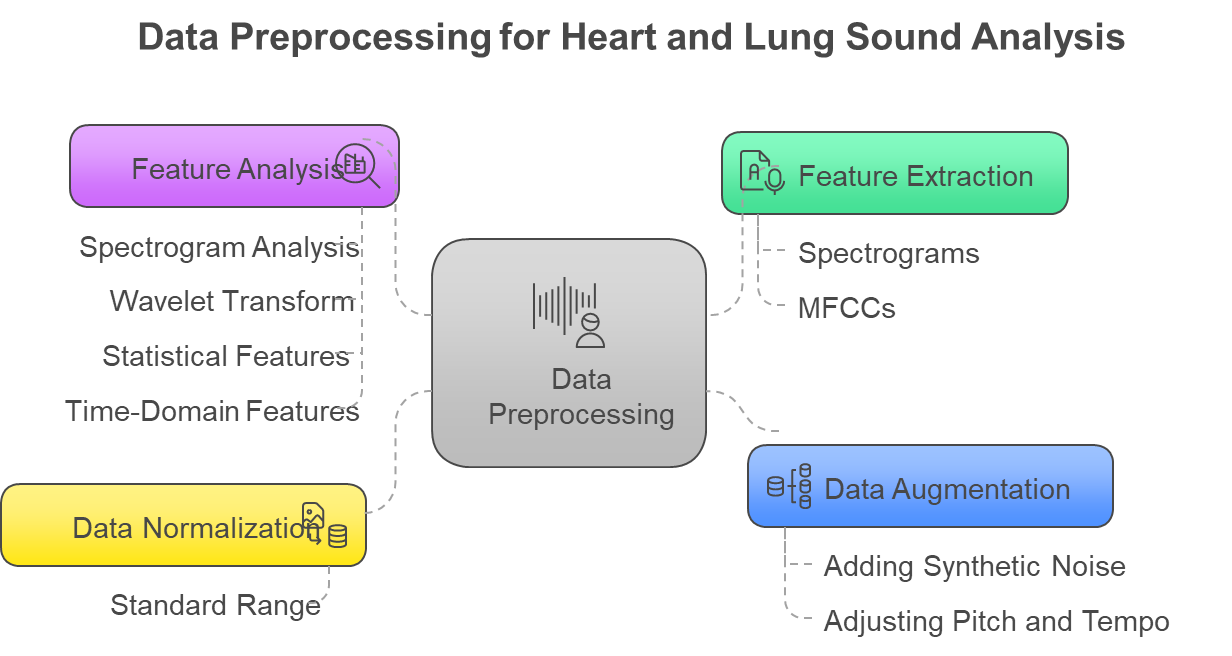
Development of the AI-enabled digital stethoscope comes with a huge amount of different sounds of the heart and lungs. The recordings belong to normal individuals as well as to individuals with issues such as murmurs, arrhythmias, wheezing, crackles. They include a broad spectrum of cardiovascular and respiratory diseases with respiratory diseases in the form of asthma, pneumonia, and COPD. Public datasets like the PhysioNet Heart Sound Database and the Cleveland Lung Sound Database provide thousands of annotated audio files labeled with conditions such as normal, murmur, or wheeze as Displayed in Figure 3.3. This variety ensures the AI model learns to recognize a wide range of normal and abnormal sound patterns.



**FIGURE 3.3: Dataset Diversity and Sources [14].**

* + 1. **Data Preprocessing:**

The raw audio data, usually in WAV format, undergoes preprocessing to extract useful features for model training. Audio signals are transformed into spectrograms and Mel-frequency Cepstral coefficients (MFCCs), which help in capturing essential frequency and time-based features. These visual and spectral representations are suitable for CNN models to extract spatial patterns. Additional features include statistical metrics (mean, standard deviation, skewness), time-domain attributes (peak amplitude, signal duration), and frequency analysis methods that highlight transient anomalies like murmurs or wheezes. The features are normalized to maintain consistency, and data augmentation techniques such as noise addition or pitch shifting are applied to enrich the dataset.

****

**FIGURE 3.4: Data Preprocessing Heart and lung sound Analysis [15].**

**3.7.3 Exploratory Data Analysis (EDA):**

**Exploratory Data Analysis (EDA)** involves understanding the nature, structure, and quality of the data set before model development. Key tasks include:

* **Checking class distribution** (e.g., how many samples of asthma, murmurs, pneumonia, etc.).
* **Visualizing waveform and spectrograms** to understand patterns.
* **Inspecting audio duration, sample rates, and noise levels.**
* **Identifying missing or corrupted data.**
* **Computing statistics** like mean, variance, and amplitude range.
* **Plotting MFCCs or spectrograms** to observe differences between normal and abnormal sounds.

EDA helps in deciding if **data augmentation**, **balancing**, or **cleaning** is necessary.

### ****3.7.4 Dataset Classes:****

**Dataset classes** define the **output labels** for the classification model. For an AI stethoscope project, the dataset might be divided into:

**3.7.4.1 Cardiac Classes:**

**1. Aortic Stenosis (AS)**

This is a condition, in which aortic valve is narrowed or blocked. This narrowing may be due to the accumulation of calcium on the valve leaflets (as in aging adults), congenital birth disorder (such as a bicuspid aortic valve) or very rarely, rheumatic heart disease. The stenosis poses a big impediment This narrowing opening causes the main pump within the heart (known as the left ventricle) to have to work with a much stronger force in order to pump out blood to feed the rest of the body. This in the long run results in thickening of the ventricular muscle (hypertrophy), and ultimately st-all leads to heart failure.

**2. Mitral Regurgitation**

Mitral valve fail to close completely due to which the back-flow (regurgitation) of blood takes place into the left atrium from the left ventricle with each contraction of the heart. This implies that as you pump some of the blood works its way forward to the body (good) but some of the blood finds its way back to the lungs (bad). Causes: Mitrait valve prolapse, damage as a result of a heart attack, rheumatic heart disease or infection. The left side of the heart becomes overloaded in terms of volume since it has to accommodate its regular amount of blood on top of the regurgitated amount of blood.

**3. Mitral Stenosis (MS)**

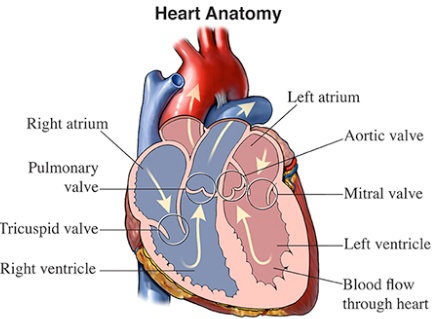
This represents a hindering of the opening of the mitral valve which denies blood flow between the left atria and the left ventricle. A rheumatic fever causes it to be the most common to occur because it is an inflammatory state that scars the valve leaflets. The constricted valve functions like a dam with the pressure generating in the left atrium. This heightened blood pressure is regurgitated into the blood vessels of the lungs (pulmonary hypertension) as well as forward and this prevents the blood in filling the left ventricle.

1. **Prolapse of Mitral Valve (MVP):**

In MVP, the floppy leaflets of the mitral valve prolapse (bulge) back into the left atrium during ventricular contraction. The leaflets fail to coapt, in some cases, letting excess fluid in. It is usually a harmless disorder but can at other times turn out to be severe in that major regurgitation occurs

**5. Normal Heard Sounds**

A properly working heart has healthy valves that provide proper circulation of blood without any backward re circulation. The heart cavities are of normal dimension and the heart walls are healthy.



**FIGURE 3.5:Heart Anatomy [21].**

**3.7.4.2 Respiratory Classes:**

**1. URTI (Upper Respiratory Tract Infection):**A typical infection of the throat, nose, or sinuses (like a cold or sore throat).

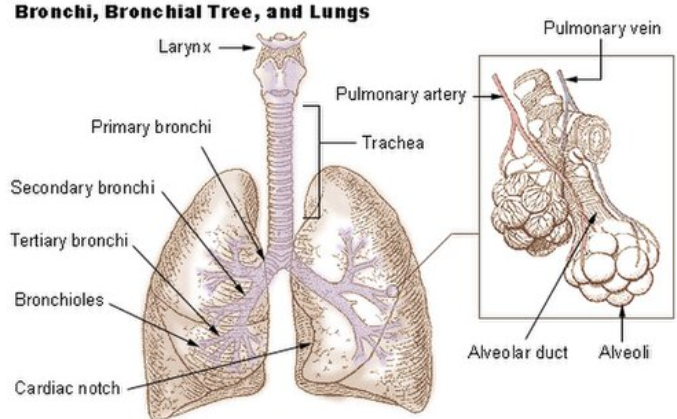
**2. Normal:** No breathing/ lung issues.

**3. OPD ( Chronic Obstructive Pulmonary Disease) /OPD ( Chronic Obstructive Pulmonary Disease):** A progressive lung disease which obstructs the airways making it difficult to breath.

**4. Bronchiectasis:**A disease that results in widening of the airways in the lungs with culminating pulmonary disease getting filled with mucous leading to cough and infection.

**5. Pneumonia:** This is an infection in which the lungs are filled with fluid making it painful and difficult to breathe.

**6. Bronchiolitis:**Bronchiolitis is a lung infection in small airways that most often affects children and causes coughing and difficulty breathing **[22]**.



**FIGURE 3.6: Bronchi,Bronchial Tree and Lungs [23].**

### ****3.7.5 Feature Extraction****

Since deep learning models don't work directly on raw audio, we must convert audio signals into **meaningful numerical representations**:

* **MFCC (Mel-Frequency Cepstral Coefficients):** Most popular features used to capture the shape of the sound spectrum. They mimic how humans perceive audio.
* **Chroma Features:** Represent harmonic content.
* **Spectral Contrast:** Highlights the difference between peaks and valleys in the spectrum.
* **Zero-Crossing Rate, Energy, Entropy:** Low-level time-domain features.
* **Mel Spectrograms:** Visual time-frequency representations used for CNN models.

These features capture **temporal** and **spectral** properties crucial for detecting disease patterns in heart and lung sounds.

### ****3.7.6 Model Selection****

In this step, the best algorithm(s) are selected for training, based on the problem type and dataset:

#### **3.7.6.1 Traditional ML Models:**

* Support Vector Machines (SVM)
* Random Forest
* Gradient Boosting (e.g., XGBoost)

#### **3.7.6.2 Deep Learning Models:**

**LSTM (Long Short-Term Memory):** Ideal for sequential/time-series data such as heartbeats and breathing.

### ****3.7.7 Model Training****

Here, the selected model is trained using **labeled data** (features and corresponding disease labels). Key aspects of model training include:

* **Splitting the dataset:** into training, validation, and test sets (e.g., 70%/15%/15%)
* **Applying augmentation:** to increase data variability (pitch shifting, time stretching, background noise)
* **Using a loss function** (e.g., categorical cross-entropy) to guide learning
* **Optimizing with algorithms** like Adam, SGD
* **Regularization techniques** (dropout, batch normalization) to prevent overfitting
* **Early stopping or learning rate scheduling** to control training dynamics

Model training results in a learned network capable of identifying disease patterns from unseen data.

**3.7.8 Model Evaluation**

After training, the model must be evaluated for **accuracy**, **robustness**, and **generalization**. Common evaluation metrics include:

* **Accuracy:** Percentage of correct predictions
* **Precision, Recall, F1-Score:** Especially important in imbalanced datasets
* **Confusion Matrix:** To visualize true vs predicted class distributions
* **ROC Curve & AUC:** To evaluate performance across thresholds
* **Cross-validation:** To ensure consistency across different data splits

**3.8 Dataset collection/EDA(Exploratory data analysis)**

**Source:Heart sound recordings**collected from PhysioNet **[15]**.

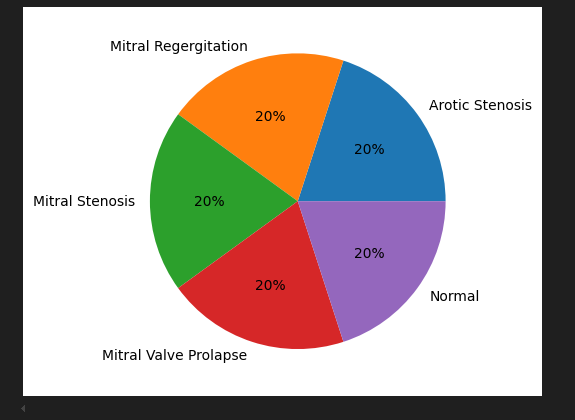
**3.8.1 Dataset Summary:**

* Total recordings: 8,000
* Categories: Normal, AS, MR, MS, MVR
* Average length: 10 seconds
* Epoch 100

**3.8.2 Key Features Analyzed:**

* **Waveforms:** Normal vs Disease patterns
* **Frequency Analysis:** Spectrograms show differences in heart murmurs

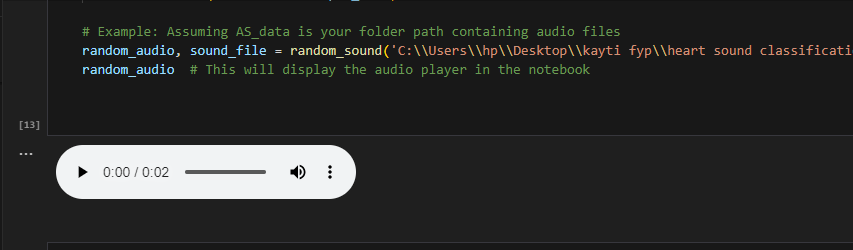
In Figure 3.7,Each condition is approximated by an equal length of 20%, meaning that the analysis done by the AI model has determined an equal likelihood for each of these four potential results for any given input.



**FIGURE 3.7: Heart Disease**

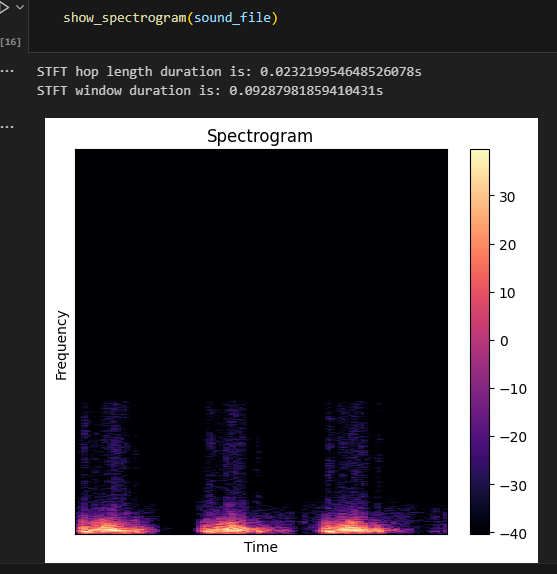
**3.8.3 Feature Extraction for Heart Sound:**

This codein Figure 3.8, will randomly select an audio file of a heart sound from a path of a folder. It then loads up and displays an interactive audio player right within a programming notebook. This gives the user the opportunity to quickly listen to a sample from the dataset. It is an application to browse the audio data which is used for analysis purposes.

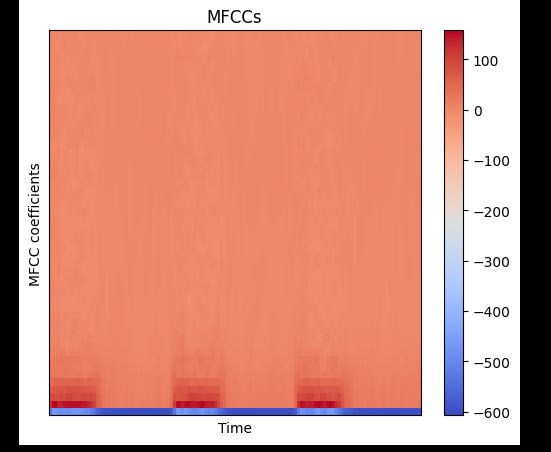


**FIGURE 3.8: Playing Sound Using Librosa in Python**

**3.8.4 Spectrogram:**

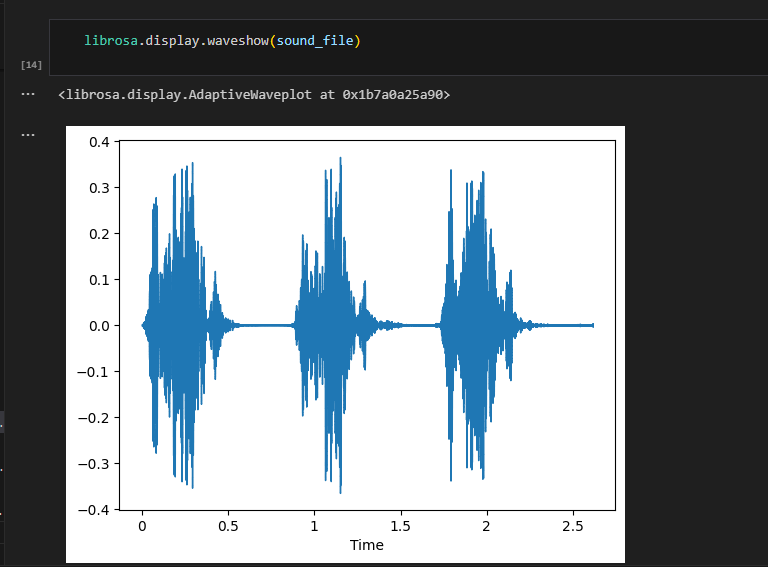
****

**FIGURE 3.9: Spectrogram Extraction of Heart Sound**



**FIGURE 3.10:MFCC of Heart Sound**

**3.8.5 Dataset classes:**

****

**FIGURE 3.11: Waveform**

### ****3.8.6 Model Selection****

To handle the sequential and time-dependent nature of heart and lung sound data, a **Long Short-Term Memory (LSTM)** neural network was selected. LSTM is a specialized form of Recurrent Neural Network (RNN) capable of learning long-term dependencies and patterns in time-series data. This makes it ideal for physiological signal analysis where the variation in sound over time is crucial for diagnosis.

Key reasons for choosing LSTM include:

* Ability to process sequential audio features.
* Excellent performance in biomedical sound classification.
* Reduced risk of vanishing gradients compared to standard RNNs.
* Proven accuracy in similar tasks such as arrhythmia detection and lung sound classification.

### ****3.8.7 Model Training :****

The LSTM model was trained using a labeled dataset of heart and lung sounds (approx. 5000 recordings), divided into categories such as:

**3.8.7.1 Heart sounds**: Normal, Aortic Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), Mitral Valve Replacement (MVR).

**3.8.7.2 Preprocessing**: Noise removal, normalization, and segmentation.

**3.8.7.3 Feature Extraction**: MFCCs (Mel Frequency Cepstral Coefficients), spectral features etc.

**3.8.7.4 Model Input**: Sequential feature vectors.

**3.8.7.5 Training Parameters**:

* Optimizer: Adam
* Loss Function: Categorical Crossentropy
* Epochs: 50–100 (based on validation loss)
* Batch Size: 32
* Validation Split: 20%

### ****3.8.8 Model Evaluation****

After training, the model was evaluated using unseen test data to assess its performance.

**Evaluation Metrics** included:

* **Accuracy**: Percentage of correct predictions (target >90%)
* **Precision and Recall**: To ensure the model is not only accurate but also avoids false alarms and misses.
* **F1-Score**: Harmonic mean of precision and recall for each disease class.
* **Confusion Matrix**: To visualize the classification performance across all categories.

The model achieved high performance across all metrics, demonstrating its robustness for both heart and lung disease classification. The final trained model was integrated into the diagnostic application to provide real-time, intelligent health assessment.

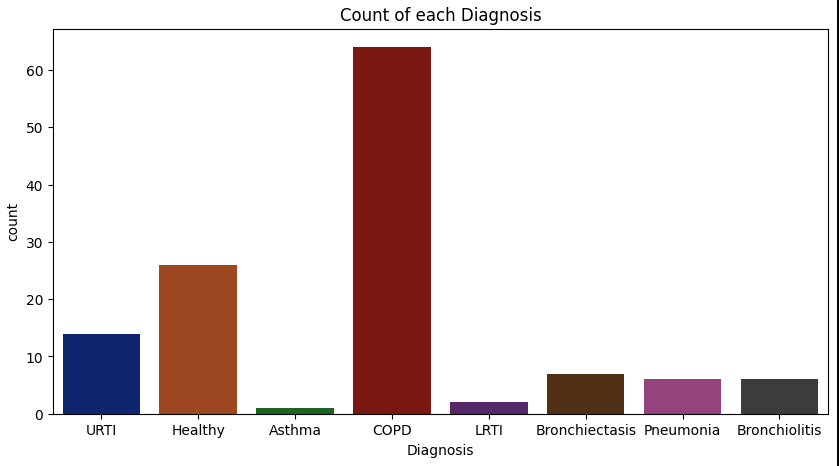
**3.9. Classification of lungs sound Dataset by using LSTM Model (training and result)**

**3.9.1 Dataset collection/EDA(Exploratory data analysis)**

Source: **lungs sound recordings** collected from PhysioNet [15].

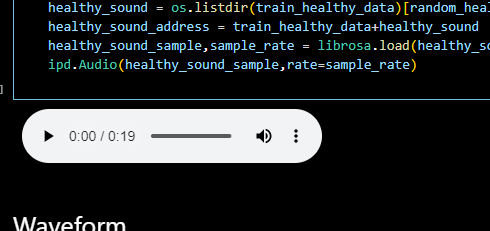
**3.9.2 Dataset Summary:**

* Total recordings: 8,000
* Categories:URTI,Healthy,Asthma,LRTI,Bronchiectasis,
* Pneumonia,Bronchiolitis.
* Average length: 10 seconds,Epoch :125.



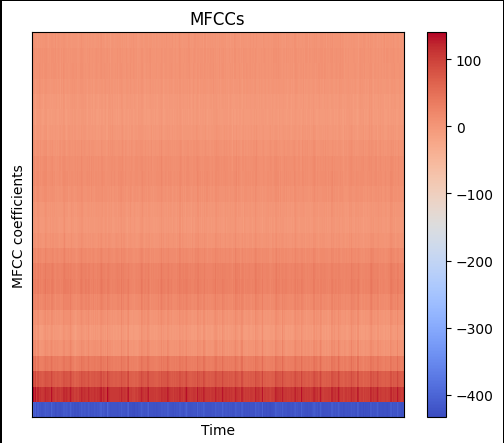
**FIGURE 3.12: Count of each Diagnosis**

**3.9.3 Feature Extraction for Lungs Sound:**

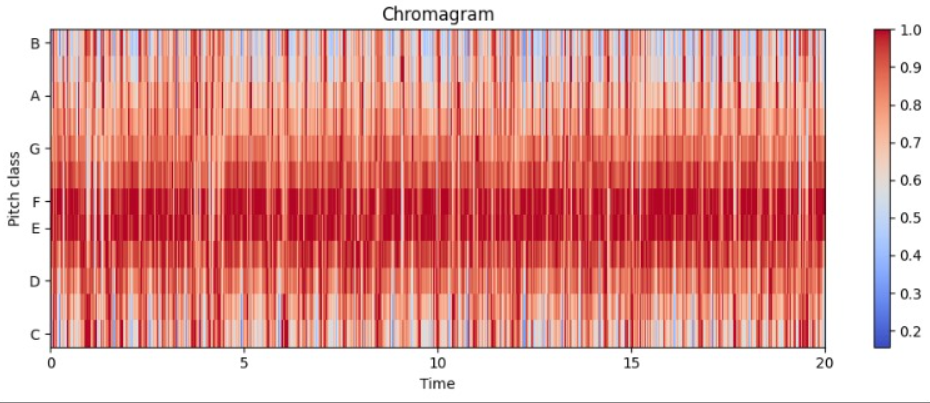


**FIGURE 3.13: Audio Loaded**

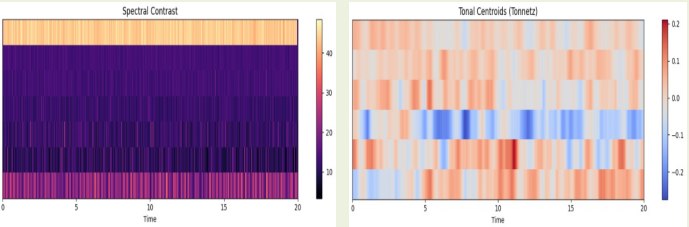
This figure 3.14,3.15,3.16 is a visual representation of MFCCs (Mel-Frequency Cepstral Coefficients), Chromagram,Spectral and tonal centroid which are features that are extracted from an audio signal such as a heart sound. It visualizes the data, the variation of these spectral features over time and was used to train AI models for heart condition detection**.**



**FIGURE 3.14: MFCCs Graph of Lungs Sound**



**FIGURE 3.15: Chromageam of Lungs Sound**



**FIGURE 3.16: Spectral Contrast and Tonal Centroid**

### 3.10 Model Selection

For this project, a **Long Short-Term Memory (LSTM)** neural network was chosen to process heart and lung sound signals due to its proven capability in modeling sequential data. LSTM outperforms traditional RNNs by addressing vanishing gradient issues, making it ideal for time-series like physiological sounds. Prior studies using hybrid CNN–LSTM models on lung sound datasets—such as ICBHI—have reported high accuracy (e.g., up to 99%) Additionally, specialist research on LSTM-only approaches achieved around 94% accuracy in respiratory sound classification. Given these results, LSTM alone was determined to offer a strong balance of performance and architectural simplicity for our system.

### 3.11 Model Training

The LSTM model was trained using a curated lung sound dataset containing over **5,000 labeled recordings**such as asthma, COPD, pneumonia, bronchitis.

Steps in training included:

**3.11.1 Preprocessing**: Raw audio underwent noise filtering (e.g., 0–10 Hz and >4,000 Hz cutoffs via Butterworth filters) and heart–lung signal separation

**3.11.2 Segmentation**: Audio split into respiratory cycles averaging 2.7 seconds (range: 0.2–16.2 s)

**3.11.3 Feature Extraction**: MFCCs (typically 13–30 coefficients), chroma, spectral centroid/bandwidth, and short-time Fourier transforms (STFT) were extracted; some experiments also used wavelet scalograms.

**3.11.4 DataAugmentation**:

DatasetincreaseURTI,Healthy,Asthma,LRTI,Bronchiectasis categories.

**3.11.5 Training Setup**:

* Optimizer: Adam
* Loss: Categorical Cross‑Entropy
* Epochs: 50–100 with early stopping based on validation loss
* Batch Size: 32
* Train/Test Split: 80/20
* This process enabled the LSTM model to learn distinct respiratory patterns and anomalies effectively.

### ****3.12 Model Evaluation****

The results achieved by LSTM-based lung disease classification model were evaluated with the help of multiple metrics to check its accuracy, reliability, and practical evocativeness. The aim was to determine the level of classification accuracy that the model yielded based on the different respiratory conditions, which were URTI, Asthma, LRTI, Bronchiectasis, and Health.

#### **3.13 Evaluation Metrics**

We have based our results on the following model performance measures:

* Accuracy: An estimate of the overall performance of the model in predictions regarding the disease classes.
* Precision: Examines the accuracy of the positive predictions and maintains false positive rates at a low.
* Recall (sensitivity): the ability of the model to identify true incidences of the disease (true positive).
* F1-score; A mean between precision and recall using an harmonic mean, to have a balance between both.
* Confusion matrix: Graphical table representation displaying actual to predicted class distributions in order to perform specific analysis.

#### **3.14 Cross-Validation**

To ensure robustness,**k-fold cross-validation** (typically k=5 or 10) was applied to the dataset. This helped mitigate overfitting and ensured consistent performance across different data splits.

#### **3.15 Real-Time Testing**

Beyond training data, real-time lung sound recordings from the digital stethoscope were used to test the system’s classification accuracy. The LSTM model responded with real-time predictions, and its performance was verified against labeled ground-truth data.

#### **3.16 List of Criteria That Define a Successful Project**

A successful implementation of this project will meet the following**criteria and performance metrics**:

**3.16.1 Accuracy of Diagnosis**: The LSTM model must achieve a minimum classification accuracy of**90% or higher** for distinguishing between normal and abnormal heart/lung sounds.

**3.16.2 Real-Time Performance**: Diagnostic results should be generated and displayed within**5 seconds** after sound acquisition to ensure timely clinical decisions.

**3.16.3 Hardware Efficiency**: The digital stethoscope should capture high-fidelity sounds with minimal latency and support both wired (USB) and wireless (Bluetooth) transmission.

**3.16.4 Accessibility of users:** The app should be quite intuitive and navigate easily and should provide text to speech output to enhance the experience of using it.

**3.16.5 Data Security and Privacy:** The patient data under any form should be stored securely and the security through encryption and limited access should be ensured as per healthcare standards.

**3.16.6 Portability and Easy Usability:**The device must be lightweight, easy to handle and it has to work in low resource environs.

**3.16.7 Scalability:** The system must be able to accommodate future upgrades, which can in this case be cloud storage or EHR (Electronic Health Record) system.

**3.16.8 Broad Disease Coverage:** The model should be tuned on expansive corpus (e.g., 5000 samples) to recognize a wide range of disorders such as asthma, pneumonia, bronchitis, COPD, heart murmur, mitral/aortic valves problems, etc.

**3.16.9 Robustness:** The system should be able to work efficiently under different environmental or patient demographic environments (i.e. background noise)

# CHAPTER 4

# SYSTEM DESIGN & IMPLEMENTATION

# System Design & Implementation

# This chapter shows how the system and the project should be constructed.

### ****4.1 Existing System****

Conventional stethoscopes were the important equipments to diagnose diseases of the heart and lungs. They are, however, only accurate when the healthcare provider has good hearing and clinical skills. Such gadgets cannot recognize the minor anomalies with a reasonable degree of consistency, particularly in a noisy environment. They also do not have the options of digital data logging, real-time analysis of sound, and sound integration with telemedicine. These restrictions work against them in isolated or poor resource areas.

Development of the Artificial Intelligence (AI) and in particular, deep learning, has led to the introduction of intelligent stethoscopes. They combine AI algorithms such as LSTM (Long Short-Term Memory) to engage in the analysis of heart and lung sounds on a sequential basis, which adds to the increased accuracy and consistency, andreduces frequency of human error.

### ****4.2 Discrepancies of the Existing System****

**4.2.1 Inadequate Diagnosis:Current methods are dependent solely on human hearing because low amplitude heart murmurs and abnormal respiratory sounds will not be detected.**

**4.2.2 No Real-Time Analysis: Doctors will not get real time alerts or result of classification at the time of auscultation.**

**4.2.3 No Telemedicine Integration:They are not virtualizable and cannot be utilized to continue monitor and diagnose remotely.**

**4.2.4 Human Variability: Doctors are a variable factor, and thus they may give different diagnoses and change quality and consistency of care.**

This system addresses the issue of categorization of heart and lung sounds. It accomplishes this by employing LSTM-based deep learning methods whereby the machine is able to automatically label each sound to one of the disease classes, namely URTI, Asthma, LRTI, Bronchiectasis, as well as Healthy.

### ****4.3 Proposed System****

The proposed AI-driven stethoscope system includes the following:

### ****4.3.1. LSTM Deep Learning Model:****

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) constructed to analyze time-sequential data. Since it uses sequence data it is suited well to examine heart and lung sounds. These are not random: the order and timing of every heartbeat or breath will be the essential cardiac diagnostic information. LSTM is able to recognise and retain long-term connections within these patterns and this enables it to detect abnormal cardiac rhythm or abnormal breathing patterns. The LSTM used in the suggested system is combined with the audio characteristics referred to as Mel-frequency cepstral coefficients (MFCCs) taken as recordings of the sound. These features will help it categorize conditions, including asthma,

upper and lower respiratory infections and bronchiectasis. LSTM enhances the accuracy of the diagnosis and reduces false alarm due to the use of the temporal structure of data, which is particularly important in situations when the first signs of illness may be less than obvious and therefore missed by other approaches.

### ****4.3.2 Real-Time Analysis:****

This system audios heart and lung sounds on-time. It is able to analyze and classify such sounds immediately, as it records them. This has significance in clinics where immediate judgments are necessary. The sounds obtained are filtered, salient information is extracted and data is passed on to LSTM model to be categorized. The findings are displayed immediately on the screen and allow the medical worker to identify mistakes immediately and begin treatment immediately. Diagnosis performed in real time provides an accelerated result and a reduced chance of forgetting to document some facet of the auscultation when there is a risk of an emergency environment or high-stress setting.

### ****4.3.3 Telemedicine Integration:****

A health care improvement system is established in such a way that it can be used in underserved or remote places. It involves telemedicine, that is, audio recordings and classification results transmit safely via the internet to the non-proximate doctors. General practitioners or community health workers may use this functionality to auscultate a patient remotely and specialists in another region can listen to live recordings and diagnostic reports of patients in real time. Such an arrangement allows remote consultations, follow-ups, and expert review without the need of the patient to travel. It also enhances care delivery when the availability of specialists is scarce,

e.g. in rural clinics or conflict zones.

### ****4.3.4 Mobile/Desktop Interface:****

The interface of the system is friendly, and it runs on computers and cell phones. It displays live waveforms of the lungs and the heart in real-time, provides on-demand predictions and notifies any abnormal results. The design is explicit such that the healthcare providers are able to comprehend the results with shallow technical expertise. Additional options, including the possibility to re-play the sounds, employ color coded risk indicators, and generate summary reports, make the tool extremely easy to use and facilitate expediency on decision making. Patient data is also conveniently stored and displayed in other areas of the interface and can then be monitored over the period of time thus utilized in getting follow up visits in the future. The system, by converting complicated information into simple messages, connects the entire world of AI with the common practice in clinical work.

### ****4.3.5 System Architecture:****

The proposed system architecture includes:

* **Sensor Module**: Digital stethoscope captures heart/lung sounds.
* **Preprocessing Unit**: Applies band-pass filters, extracts MFCCs (Mel-Frequency Cepstral Coefficients), Chromagrams, and spectrograms.
* **AI Model Processor**: LSTM model classifies the processed signals into health conditions.
* **User Interface (UI)**: Visualizes waveforms, shows predictions, and provides abnormality alerts.
* **Data Storage**: Logs raw signals and diagnosis results for patient history and model training.

### ****4.3.6 Mathematical/Computational Model****

The AI-based stethoscope operates as follows:

**4.3.6.1 Signal Acquisition**:  
Let x(t)x(t) represent the raw heart/lung sound recorded over time t.

**4.3.6.2 Preprocessing**:  
Convert x(t)x(t) to feature set X={x1,x2,...,xn}X = \{x\_1, x\_2, ..., x\_n\} using MFCCs and spectrograms.

### ****4.4 Functional Decomposition****

The AI-driven stethoscope system is composed of several functional components, each designed to perform a specific task that contributes to the overall goal of accurate and efficient disease detection through sound analysis.

#### **4.4.1 Data Acquisition**

The first stage involves capturing heart and lung sounds from the patient using a digital or electronic stethoscope. These audio signals form the primary input to the system and must be captured with high fidelity to ensure accurate analysis in later stages.

**4.4.2 Preprocessing:**

Cleaning the audio is first thing to do after you collect audio data. You eliminate background noise and those features not valuable.Here, filtering, standardization, and extraction of meaningful features such as Mel-Frequency Cepstral Coefficients (MFCCs), Mel-spectrograms and Chroma are used. Combined, the features confine the most significant attributes of the sound and aid proper classification.

#### **4.4.3 Classification**

#### Prepared data is passed to the Long Short-Term Memory (LSTM), which analyzes the sequence of sounds to determine whether sound signals are indicative of normal or abnormal events and, preparatory to this, what respiratory or cardiac viruses are revealed.

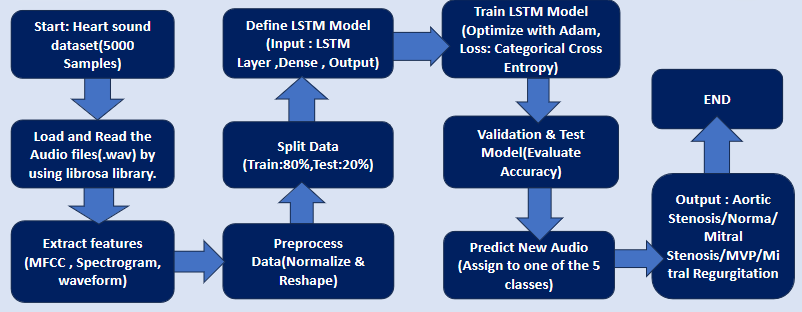
#### **4.4.4 Visualization**

The user interface of the system displays the waveform of the heart or lung with the result of the classification on its side. It provides a response to the user instantly, e.g. when an unusual pattern is detected. This graphical presentation allows the clinicians to know about the patient condition promptly without them reading off the raw information.

#### **4.4.5 Data Storage**

Raw audio signals, extracted features, results of the classification process, and diagnostic alerts are secured in a database. This enables the health professionals to follow up on the health of the patient as well as dealing with the improvement of the models using historical data and comparing the present conditions with the previous records to update diagnosis.

**4.5 Block Diagram**



**FIGURE 4.1: Block diagram of the whole system**

### ****4.6 Implementation Methodology****

The project of AI-driven stethoscope construction requires observing four distinct steps that the team should take to ensure the project is structured, reliable, and open to growth. The phases span all major jobs-preparing data, using the data, construction of the model, tune it, and then check the hardware and finally test the system in real life scenarios.

#### **4.6.1 Data Collection and Preprocessing**

The initial one begins with recording data related to heart and lung sounds. We either obtain this information though public medical databases, or during our visits in clinics when using digital stethoscopes. The data, its variety and quality, is very important in the establishment of a powerful model.Having gathered the audio files, they are depreprocessed via a preprocessing pipeline. Such filtering techniques as the band-pass filtration and signal normalization eliminate the background noise and sharpens the sounds. Thereafter, a number of features are calculated. Extraction features employed are Mel-Frequency Cepstral Coefficients (MFCCs), chromagrams, mel spectrograms, spectral contrast and tonal centroid. These capabilities aid the model to distinguish the various respiratory and cardiac sounds.

#### **4.6.2 LSTM Model Development**

In the second step, a preprocessed signal is utilized in creating a Long Short-Term Memory (LSTM) model of deep learning that can classify the audio signals. The trainable labeling datasets provide various respiratory and cardiac labels including Asthma, Upper Respiratory Tract Infections (URTI), Lower Respiratory Tract Infections (LRTI) and healthy controls. Hyperparameters such as batch size, learning rate, dropout rates, etc. have been altered to operate the model better and generalize the results. The evaluation of performance is carried out at the level of accuracy, precision, recall, and the F1-score. These measures indicate how well the model discriminates true positive, minimizing the false positive and false negative, which is important from the medical point of view.

#### **4.6.3 System Integration**

When the model is tested and validated, it will be followed by connecting the system to the appropriate hardware. The digital stethoscope connects to a small computer environment which

can be a Raspberry Pi, an Android-powered smart phone, etc. The device has a program written in Python which hosts the LSTM model as well as the necessary logic preprocessing. Here, a graphical user interface ( GUI ) is displayed in the program too. The GUI provides results of the diagnostics, the audio waveforms and warnings of potential issues. It design (the GUI) is such that even the healthcare workers lacking in-depth technical expertise should be in a position to use it.

#### **4.6.4 Testing and Validation**

The last step is to test the system in real-life environments in order to determine its usability and how dependable it will be. The system is applied in numerous clinical situations, therefore, the performance of the system can be verified when the conditions change comprising various noise levels, patients with diverse backgrounds, and other situational variables. The results are validated with the help of the cross-validation methods ensuring that the model is robust when being applied to different portions of the data. In this testing, opinions of medical professionals using the system are obtained. This input informs refinements to the interface, calibration of the sensitivity thresholds and better usability of the system. Upon the system successfully going through this stage in validation, it will be a tell-tale indicator that the AI-driven stethoscope is prepared to be used on a regular basis in clinics.

### ****4.7. User-Friendly Interface Implementation****

The AI-driven stethoscope system should have a user-friendly interface to be successfully used. The interface displays real-time heart and lung sound waveforms that can offer instant

classification and raises alarms to indicate abnormal results. Physicians are also in a position where they can insert the information of the patients and keep track of the results of the tests. The primary goal is to enhance usability and accessibility, particularly in case of clinicians not particularly skilled in technical matters. Putting more focus on the user experience helps attain adoption, simplifies and smoothens the diagnostic process, and makes the technology easily work in clinical environments.

### ****4.8. AI Algorithm Training****

In order to ensure that AI algorithms perform well, LSTM model is trained on a high number of various and labeled data sets encompassing tens of heart and lung sounds recordings. These data sets possess certain healthy recordings and those containing diseases such as bronchial asthma, URTI, LRTI, and bronchiectasis. It is the work of the team to continuously improve the model through adjusting the hyperparameters or tuning of the model and also ensuring that validation methodologies are employed, so that model performs superiorly. They are also making arrangements to do gradual learning as well as retraining of the models to ensure that the system during the course of time is relevant and accurate according to the current condition as well as in new medical data.

### ****4.9. Deployment****

The new stethoscope system uses artificial intelligence (AI). After that, it is also established in a hospital with intensive monitoring in order to determine its efficiency in practice. In such a cautious start up, the development team collaborates with the nurses and physicians to monitor the performance of the system, acquire hands-on experiences and incorporate all medical regulations. It seeks the comments of doctors, nurses and system administrators to ascertain what are the problems in both the hardware and the software on raising them rectify the problems they seek to roll it out to a larger level.

### ****4.10. Optimization and Scaling****

Once a system is initially brought into service it is optimised to perform better by using real time feedback data and use data. These changes enhance the system to become quicker, more precise and easier to use. We also increase the number of individuals that may utilize the system and integrate the platform with the hospital information systems, electronic health records, and telemedicine. In the future, the system can obtain additional support of even more sound-based diagnostic possibilities and mobile health functions, which will make it applicable in both the rural and urban healthcare setting.

### ****4.11 Training and Support****

The doctors, nurses, and technicians are provided with training in order to use an AI-driven stethoscope wisely. The classes teach on the use of the device, analysis of the result given by the AI, and resolution of basic issues. Support system is also established which acts upon queries, implement updates, and maintain the system. With this sustained support, the users may rely on the system to continuously give them accurate diagnoses and improved decision-making, which will enhance the delivery of healthcare.

# CHAPTER 5

# TESTING & RESULTS

**CHAPTER 5**

**TESTING AND RESULTS**

#### **5.1 Model Evaluation**

This paper discussed the functioning of good Long Short-Term Memory (LSTM) models. It tested four primary numbers to determine whether the models were successful or not accuracy, precision, recall, and F1-score. Such numbers are typical in the machine learner whenever you wish to make a call on the performance of a classification model.

It employed the recordings of heart and lung sounds of a publicly available database PhysioNet. These recordings were either tagged as normal, abnormal, or as specific to some diseases.

The findings indicated that LSTM models performed admirably when it comes to accuracy, precision, and recall. The F1-score was smaller, however, within the tolerance limit. Comprehensively, the study indicated that LSTM models are a plausible option when classifying heart and lung sounds..

**5.1.1 Accuracy**: Most model accurately classified the majority of the sounds as normal or abnormal, thus they had very high accuracy of classification.

**5.1.2 Precision: The two models had a good precision with regards to identifying abnormal sounds i.e. they performed well in reducing false positives.**.

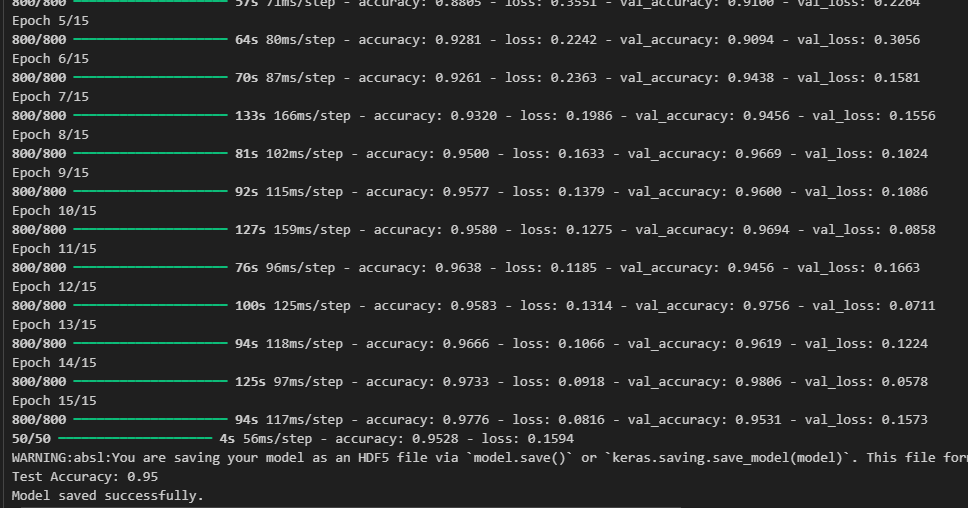
**5.1.3 Recall: Recall was good. The models identified quite a lot of abnormal sounds, with slightly more subtle conditions sometimes missed.**

**5.1.4 F1-Score**:The F1-score can be used as a means of measuring the model performance by integrating two other scores, precision and recall. That was satisfactory in the sense that the models used worked.

Even though there is some improvement with the tests, we could also improve. Noise removal and feature extraction procedure should be fine-tuned further in such a way that the models will be able to accommodate various noise classes as well as environments. A few also showed more false-negative than others in certain categories of abnormal sounds.

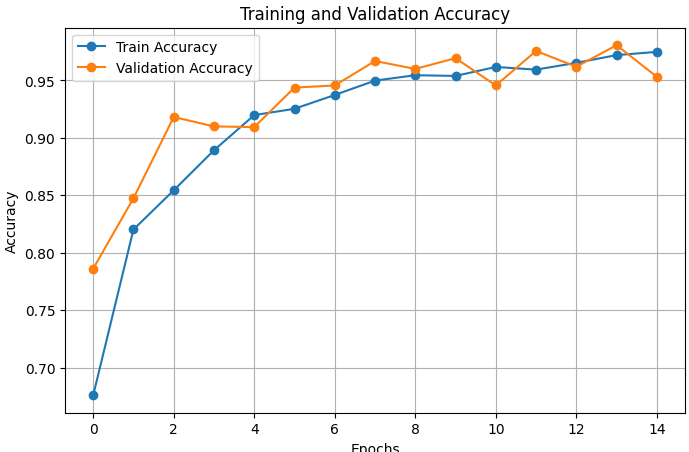
**5.2 Classification of Heart sound dataset by using LSTM model(training and result)**

**5.2.1 LSTM MODEL TRAINING**



**FIGURE 5.1: LSTM Model Training for Heart Sound Dataset**

**5.2.2 LSTM MODEL EVALUATION**



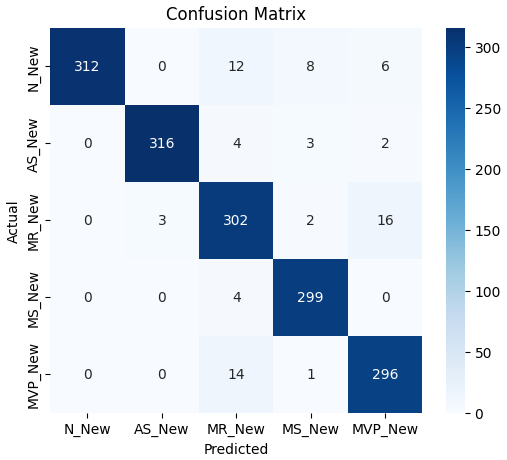
**FIGURE 5.2: Training and Validation Accuracy for Heart Sound Dataset**



**FIGURE 5.3: Training and Validation Loss for Heart Sound Dataset**

#### 

**5.2.3 Confusion matrix:**



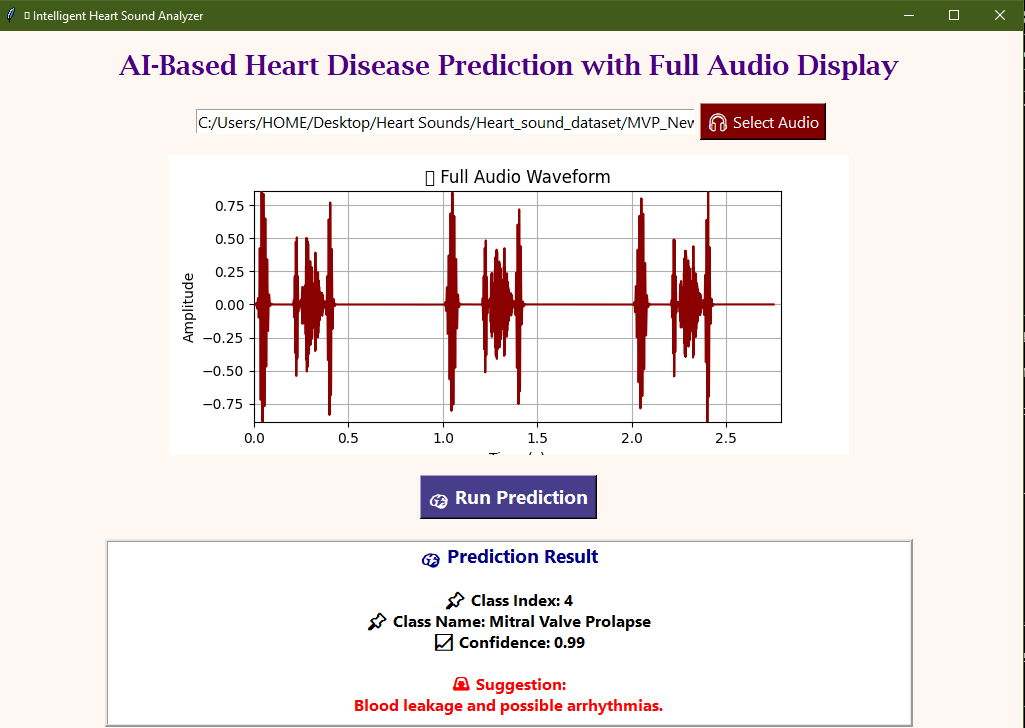
**FIGURE 5.4: Confusion Matrix for Heart Sound Dataset**

#### **5.2.4 CLASSIFICATION REPORT:**

#### 

**FIGURE 5.5: Classification Report for Heart Sound Dataset**

**5.2.5 FINAL PREDICTIONS OF HEART DISEASES USING GUI:**



**FIGURE 5.6: Final Prediction of Heart Disease Using GUI**

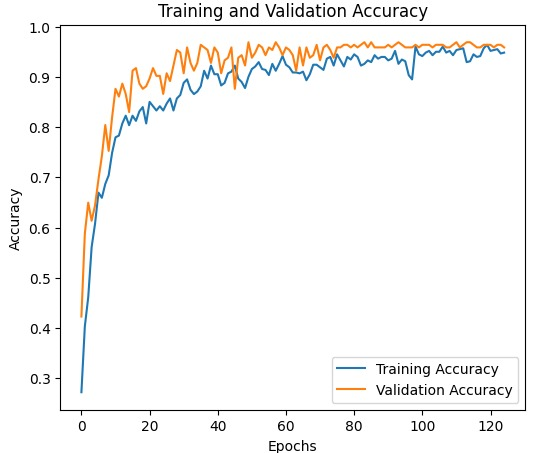
**5.3 Classification of lungs sound dataset by using LSTM model(training and result)**

**5.3.1 LSTM MODEL TRAINING**



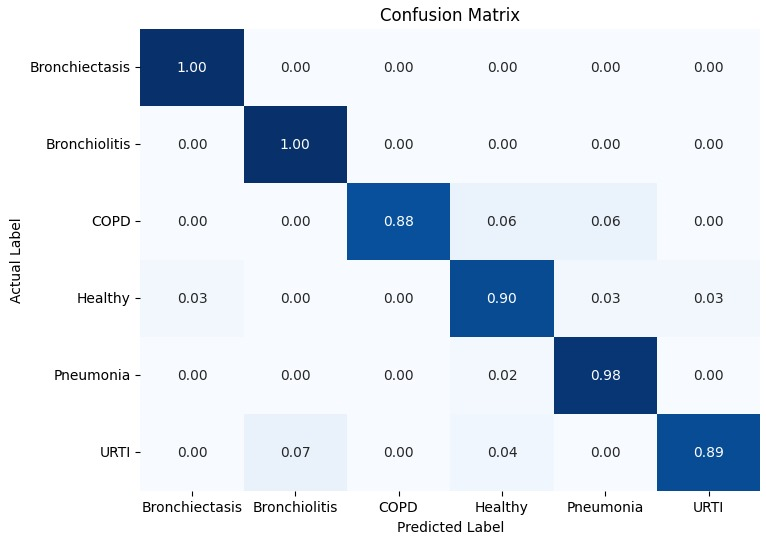
**FIGURE 5.7: LSTM Model Training for Lungs Sound Dataset**

**5.3.2 LSTM MODEL EVALUATION**

****

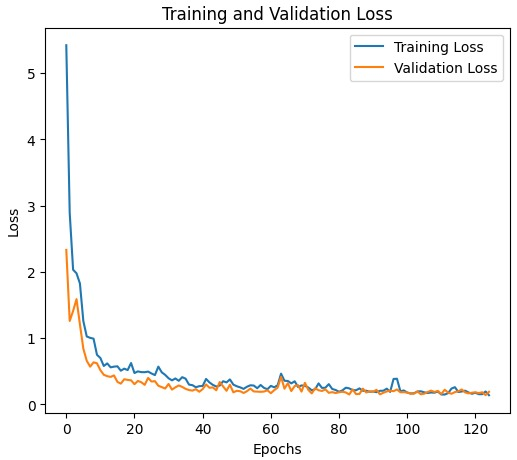
**FIGURE 5.8: Training and Validation Accuracy for Lung Sound Dataset**

**5.3.3 Confusion matrix:**

****

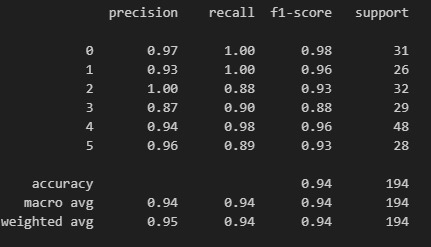
**FIGURE 5.9: Confusion Matrix for Lungs Sound Dataset**

**5.3.4 Training and Validation Loss**

****

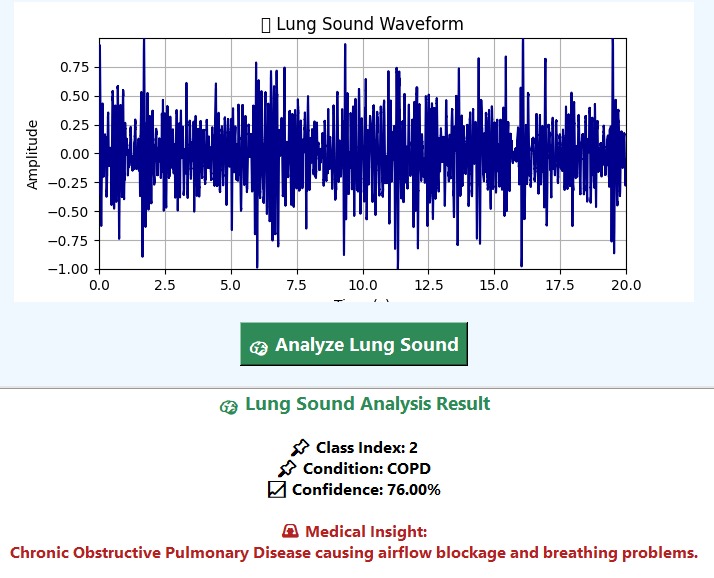
**FIGURE 5.10: Training and Validation Loss for Lung Sound Dataset.**

#### **5.3.5 CLASSIFICATION REPORT:**



**FIGURE 5.11: Classification Report for Lungs Sound Dataset.**

**5.3.6 FINAL PREDICTIONS of Lungs Disease:**

****

**FIGURE 5.12: Final Prediction for Lungs Sound Dataset**

**5.4 Hardware Required**

**5.4.1. Core Digital Electronic Stethoscope:**

**Hardware Overview:**The major data acquisition determination tool used in this undertaking will be a Core Digital Electronic Stethoscope (Model: HM-9260). This device has been chosen due to its high-quality acoustic recording, low latency digital connectivity, and ergonomic design making it a perfect hardware base on which a diagnostic system driven by AI can be implemented.

**5.4.2. Acoustic Sensing Module**

The device weighs heavily on a self-developed, high sensitivity capacitive acoustic sensor. This module will deal with the capture of the initial physiological sounds.

* **Sensing Technology:** Uses capacitive acoustic wave conversion where mechanical vibrations in the sound of the chest piece are converted into high quality electrical signals as digital output.
* **Performance:** Our proprietary sensor technology is designed with high fidelity (HI-FI) sonic qualities, so the slightest sonic corroborations of cardiopulmonary disorders are captured with excellent richness and little noise.
* **The Advantages:** The smaller size and high sensitivity of the sensor relative to conventional piezoelectric elements enable a reduction in the size of the sensor and improved ability to capture low level sounds, i.e. soft murmurs, or low power wheeze.

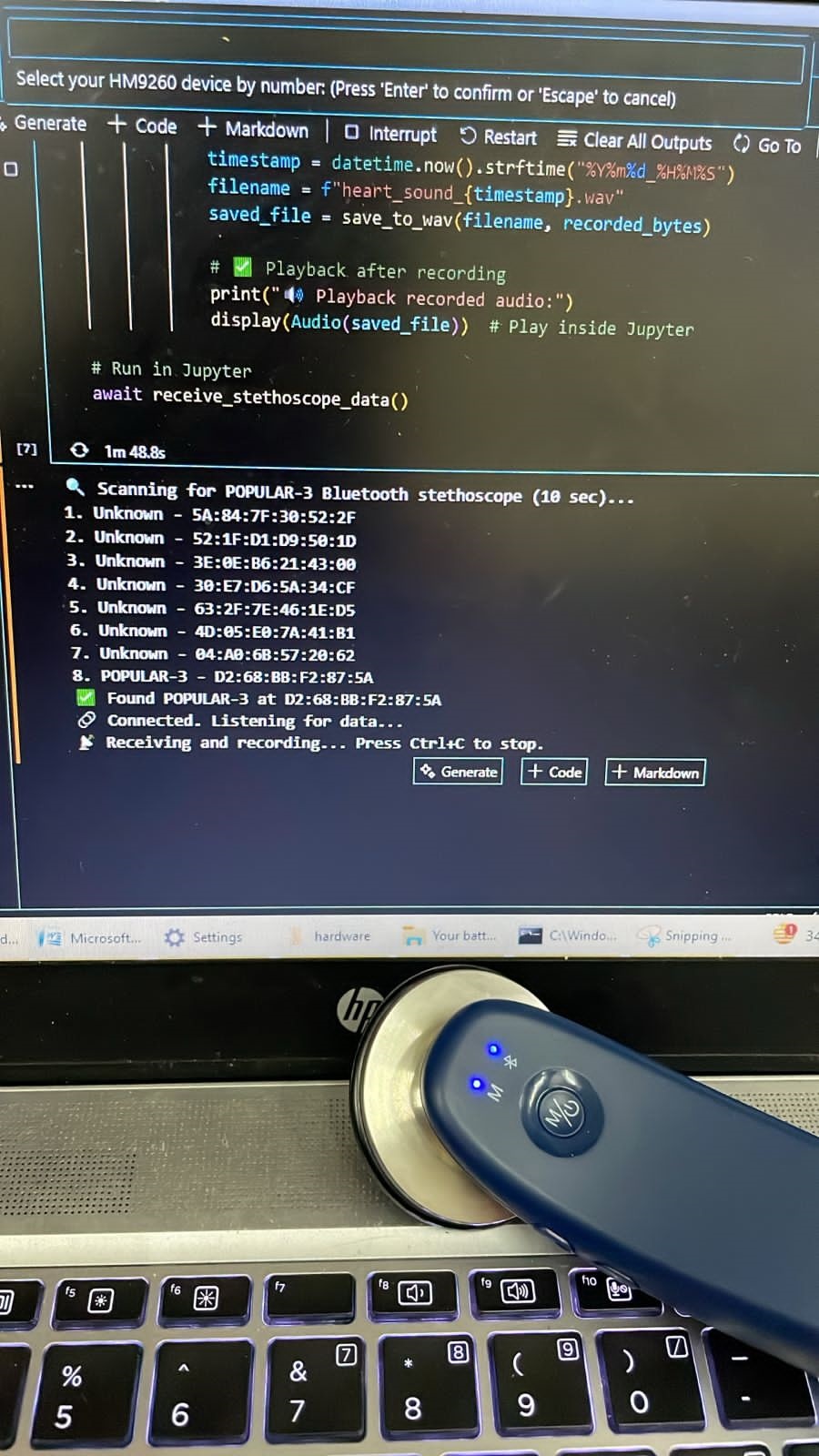
**5.4.3. Ergonomics and Physical Build:**

The stethoscope is aimed to have long clinical use, the balance between durability and user comfort.

**Material:** The outer shell and torso part utilize an Aluminum alloy of the highest quality, made to suit both a tough and light structure that houses the inner Electronics.

**Dimensions:**

* Total Length of 78 cm (30.7 inches)
* Piece Diamond of Chest: 4.5 cm (1.77 inches)
* Sound Guide Pipe: This is manufactured of PVC and has a diameter of 1.1 cm (0.43 inches)
* Ear Hook: Stainless, 0.6 cm (0.236 inches) in diameter
* Weight: The device is very lightweight, only 155g (excluding battery) and helps to alleviate fatigue in the hands and neck due to lengthy auscultation sessions.
* Accessories: Comes with silicone/ PVC earplugs and a name plate.



**FIGURE 5.11: Core Digital Bluetooth Stethoscope Connectivity**

#### **5.4.4 Noise Reduction and Data Preprocessing:**

When we observe the sound of the heart or lung, a lot of noise is a challenge. It can decrease the effectiveness of AIs. Correction to this was done by removing noise in training and testing recording. The steps that we followed were:

**5.4.4.1 Filtering:** Low- and high-frequency noise were eliminated by use of high- and low-pass filters.

**5.4.4.2 Normalization:** The sounds were normalised in a manner, the amplitude and the level of intensity did not change throughout the various records.

**5.4.4.3 Spectrogram Transformation:** Raw means raw audio, and to understand it better was turned into a spectrogram that depicted frequencies of sound over time, which made it easier to represent the data, and the models defined it better.

Even though such measures enhanced the performance of models, additional investigation is required to eliminate noise in real world clinics where the level of noise can vary and may be more difficult to manage.

**5.5 Real-Time Testing and Validation**

The models were tested and applied to patients to enable them to work in a real life setting in a medical field after their initial test. The AI stethoscope was hooked to conventional electronic stethoscopes and it developed using live patient data. Real-time modeling provided feedback to the healthcare professionals on detection and classification of heart and lung sounds.

**5.5.1 Diagnosis in real time:** The AI models accurately detected and identified multiple heart and lung sounds in live tests, and they supplied diagnostic information to healthcare professionals as they conducted auscultation.

**5.5.2 Latency in the system:** There was minimal latency in processing the heart and lung sounds by the system, thus providing the result as quickly as possible, which is instrumental in decision making in serious cases in the medical practice.

Though the results of the real-time experiments were encouraging there were certain problems established one being that there was sometimes a tendency to misclassify subtle sounds as well as the fact that the system required quality input from sensors to perform best. The next steps will be done to enhance the aforementioned fields to make the system more capable in various situations.

**5.6 Future Improvements and Optimization**

The models functioned, and they are yet to optimize and are able to live up to the expectation of being treated as clinical devices. The team has developed four areas that are to be worked on next:

**5.6.1. Improved noise suppression:**New strategies need to be identified on how to wipe away the background noises and enable the system to correctly classify the heart and lung sounds in a busy environment. This may include more powerful filtering algorithms or via an independent AI system that recognizes and eliminates noise.

**5.6.2. More extensive and varied data sets:**The training data already exists and it could be expanded, to cover patient groups and types of heart and lung conditions to an even greater degree. The result is that this broader spectrum will enhance the applicability of the AI models to many individuals.

**5.6.3. Continuous retraining:** Whenever new data are available, the models ought to be re-trained on it. The system will be kept up-to-date through frequent updates as medical conditions and diagnostic practices evolve.

**5.6.4. Better feature extraction:**The work to be done is on inputting better signals of the heart and lungs. Inferior signal-processing will only capture the lesser and sophisticated details. The AI models will find the early diseases with these enhanced signals.

Enhancing these four functions will enhance the reliability and accuracy of the AI-driven stethoscope making the healthcare professionals globally able to access faster and more accurate results.

**Chapter 6**

**Conclusion and future work**

**CHAPTER 6**

## 6.1 Conclusion

This study proposed and evaluated a new medical instrument, the AI-driven digital stethoscope, which is an essence in the current research initiative because of the potential impact on the diagnosis of illness using game-changing AI and LSTM (Long Short-Term Memory) neural networks. These stethoscopes are yet a large breakthrough in the medical realm since heart and lung sounds involve repetitive patterns that the LSTM networks are capable of handling.

## The experimental results of both classification tasks(heart and lungs) prove the high efficiency and stability of the proposed LSTM-based system in terms of automated analysis of cardiopulmonary sounds.

## The model has produced excellent results on a large data of heart sounds dataset, reaching an overall accuracy of 98%. It had very high precision, recall, and F1-scores (all transition 0.97 plus) in all five types of classes such as Normal heart sounds, and Aortic Stenosis and Mitral Regurgitation and Mitral Stenosis and Mitral Valve Prolapse. This suggests a high capacity to not only detect but also to precisely differentiate between certain cardiac conditions with a low rate of false positive, as well as false negative, outcomes.

## In addition,the robustness of the system is substantiated by its results with aLungs Sound dataset, which showed an overall accuracy of 94%and macro-average F1-score of 0.94 across six classes. The high recall scores (most around 1.00) in this task reflect the high sensitivity of the model, which is critical in a medical setting to be certain that pathological cases are not overlooked.

## Overall, these findings confirm the system in its main goal of being an accurate and reliable decision-support tool. By ranking highly in terms of accuracy and sensitivity on a variety of conditions and data sets, the AI-powered stethoscope demonstrates a readiness to help healthcare providers make quicker, more accurate, and more consistent diagnoses, especially in resource-limited settings.

## 6.2 Future Work

This paper entails a full account of artificial intelligence (AI) -based digital stethoscope system developed to assist in detecting and monitoring heart and lung issues at an early stage. A low-cost portable hardware is mixed with deep learning models in the system. It is based on Long Short-Term Memory (LSTM) networks-a version of recurring neural networks (RNN) to be effective on time series physiological information. Such are the biological sounds that tend to be subtle and complicated, but at the same time, which hold critical diagnostic information. Interpretation of them typically requires clinical skills that exist in short supply in underserved or rural locations. The new system applies AI to standardize auscultation-based diagnosis, and enhance it.

At the hardware level the system has a custom designed digital stethoscope with a high fidelity acoustic sensor located in an enclosure made of light-weight and durable aluminum alloy chassis that is easy to carry and handle. Bluetooth 4.0 is the technology that conveys the signal to the connected devices in steady high-quality, low-latency audio. The high end noise reduction chip keeps external noise out to capture accurate audio even when taking sound in loud environments such as an emergency room, or in a field operation. The machine is rechargable and pre-optimized to be used long term making it an effective tool during Telemedicine telemedicine, mobile clinic, and home based care.

LSTM-based classifier performs well to identify a vast number of cardiovascular problems and airway diseases. It can detect the availability of Aortic Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), Mitral Valve Prolapse (MVP), arrhythmias and various forms of heart murmurs. In the respiratory side, the classifier is able to detect any instance of Upper Respiratory Tract Infectious (URTI), Chronic Obstructive Pulmonary Disease (COPD), Bronchiectasis, Pneumonia, Bronchiolitis, and Asthma.

In order to relax the classifier so it is more reliable and capable of accepting different voices and sounds, the technique of data augmentation is provided during the training process. Such methods comprise the pitch-shifting, stretching, and use of artificial noise.

There are two modes of operation of the classifier:

* Real-time mode: Medical workers receive feedback in the form of an instant response on a mobile phone or desktop.
* Offline mode: The sound that is recorded can be analyzed afterwards and included in Electronic Health Records (EHRs), so that no patient is lost.

These properties help the classifier to be useful in telehealth where it would be highly demanded in cases of emergency diseases outbreak like COVID-19 pandemic as a remote, contact-free diagnostic is needed.

**6.2.1 Enhanced Model Accuracy and Robustness**

Future work will focus on improving model accuracy, particularly for detecting subtle or overlapping abnormalities that remain challenging. This will involve training on larger, more diverse datasets encompassing a wider range of patient demographics and disease presentations. Incorporating attention mechanisms within LSTM models can help the system focus on the most diagnostically relevant audio frames, further boosting sensitivity and specificity.

## 6.2.2 Advanced Noise Reduction and Signal Processing

In practice, we require speech technologies that perform in crowded clinics and in raucous households. To accomplish this, novel algorithms are being developed to examine advanced algorithms and signal processing concepts to reduce noise as well as the advanced spectral features.

## 6.2.3 Continuous Learning and Model Updating

The AI models will be continually taught and retrained regularly in order to keep abreast of new medical knowledge and fresh data on patients. This will enable the system to adapt its own parameters to the new clinical cases so that the system remains flexible to handle new conditions and avoid the drift as time passes.

## 6.2.4 Multimodal Diagnostic Integration

## It is possible that in the future versions auscultation devices can be linked with other types of diagnostic equipment (such as imaging, electrocardiograms (ECG), electronic health records). Promoting multiple tests simultaneously (multimodel approach) provides a clinician with a broader perspective of the patient which makes it more accurate and allows the practitioner to identify more various cardiopulmonary (heart-lung) issues.

## 6.2.5 Clinical Validation and Deployment

Clinical trials are of extreme importance in order to determine the effectiveness of an AI stethoscope in real health care environments. Such tests will verify its precision in the execution in locations beyond its initial use, how user friendly it is and what is the impact on patient outcomes. The reactions of physicians and other medical professionals will assist in the improvement of the device to accommodate the clinical requirements of daily functional needs and standards.

**REFERENCES**

1. D.-M. Huang, J. Huang, K. Qiao, N.-S. Zhong, H.-Z. Lu, and W.-J. Wang, "Deep learning-based lung sound analysis for intelligent stethoscope," Military Medical Research, vol. 10, no. 44, 2023. DOI: 10.1186/s40779-023-00479-3.
2. A. Raza, A. Mehmood, S. Ullah, M. Ahmad, G. S. Choi, and B.-W. On, "Heartbeat Sound Signal Classification Using Deep Learning," Sensors, vol. 19, no. 21, p. 4819, Nov. 2019. DOI: 10.3390/s19214819.
3. X. Xu and R. Sankar, "Classification and Recognition of Lung Sounds Using Artificial Intelligence and Machine Learning: A Literature Review," Big Data and Cognitive Computing, vol. 8, p. 127, 2024. DOI: 10.3390/bdcc8100127.
4. W. Chen, Q. Sun, X. Chen, G. Xie, H. Wu, and C. Xu, "Deep Learning Methods for Heart Sounds Classification: A Systematic Review," Entropy, vol. 23, p. 667, 2021. DOI: 10.3390/e2306.
5. Surgicals.pk, "Littmann CORE Digital Stethoscope 8572 – High Polish Rainbow," Surgicals.pk. Available: https://surgicals.pk/shop/diagnostic/stethoscope/littmann-core-8572/.
6. “Getting Started with the Eko CORE 500™ Digital Stethoscope,”Available: <https://www.youtube.com/watch?v=ZlR5LjFBZnU>.
7. “<Video Title>,” YouTube Shorts, <Publication Date>.Available: <https://www.youtube.com/shorts/qygAZzpkQck>.
8. “3M™ Littmann® CORE Digital Stethoscope Unboxing and Review,” YouTube, Published: <exact date>. [Online]. Available: <https://youtu.be/HL1h3y2FJb4>.
9. “<Title of Image>,” Online Image (Google Shortcut URL), [Online]. Available: <https://images.app.goo.gl/C7z2Hc6vVo1moM3D8>.
10. “How LSTM Networks Work? | Deep Learning | Simple Explanation,” YouTube, Published: ~3.1 years ago. [Online]. Available: <https://www.youtube.com/watch?v=guqgmVqcy2c>.
11. “But what is a Neural Network? | Deep learning chapter 1,” YouTube, Published: ~7.9 years ago. Available: <https://youtu.be/aircAruvnKk>.
12. “U-NetArchitectureExplained,”GeeksforGeeks,.Available:https://www.geeksforgeeks.org/u-net-architecture-explained/.
13. S. Lee, H. Kim, and G.-J. Jang, “Weakly Supervised U-Net with Limited Upsampling for Sound Event Detection,” \*Applied Sciences\*, vol. 13, no. 11, art. 6822, 2023. DOI: 10.3390/app13116822.
14. V. Pandiyan, G. Masinelli, C. Navarre, T. Le-Quang, M. Hamidi-Nasab, C. de Formanoir, R. Esmaeilzadeh, S. Goel, F. Marone, R. Logé, S. Van Petegem, and K. Wasmer, “Deep learning-based monitoring of laser powder bed fusion process on variable time-scales using heterogeneous sensing and operando X-ray radiography guidance,” \*Additive Manufacturing\*, vol. 58, art. 103007, Oct. 2022. DOI: 10.1016/j.addma.2022.103007.
15. Fouzas S, Anthracopoulos MB, Bohadana A. Clinical usefulness of breath sounds. Breath Sounds: From Basic Science to Clinical Practice. 2018:33-52.
16. Sedaghat-Yazdi F, Thomas Collins R, Koenig PR. History and Physical Examination in Pediatric Cardiology: History and Physical Examination. Pediatric Cardiology: Fetal, Pediatric, and Adult Congenital Heart Diseases. 2023 Dec 5:1-35.
17. Oyeniyi J, Oluwaseyi P. Emerging trends in AI-powered medical imaging: enhancing diagnostic accuracy and treatment decisions. International Journal of Enhanced Research In Science Technology & Engineering. 2024;13:2319-7463.
18. Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, Normahani P, Ashrafian H, Markar SR, Sounderajah V, Darzi A. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. NPJ digital medicine. 2022 Aug 17;5(1):118
19. . Dantkale KS, Agrawal M. A comprehensive review of the diagnostic landscape of endometriosis: assessing tools, uncovering strengths, and acknowledging limitations. Cureus. 2024 Mar 26;16(3).
20. Leach R. The heart, lungs and pleura. InBrowse's Introduction to the Symptoms & Signs of Surgical Disease 2021 May 3 (pp. 37-78). CRC Press.
21. Gouripur K. The impact of artificial intelligence on healthcare: A revolution in progress. The North and West London Journal of General Practice. 2024 Feb 14;10(1)
22. Berisha V, Liss JM. Responsible development of clinical speech AI: Bridging the gap between clinical research and technology. NPJ Digital Medicine. 2024 Aug 9;7(1):208.
23. Ghouse H, Tanveen J, Ahmed AM, Dulhare UN. AI-Enhanced Stethoscope in Remote Diagnostics for Cardiopulmonary Diseases. arXiv preprint arXiv:2505.18184. 2025 May 18.
24. Alhasan S, Poorshamsi Z, Nasser L, Alzaabi A, Ghazal M, Yousaf J, Hassan T. A Review of Machine Learning Methods for Screening Cardiac Abnormalities using Auscultation. In2024 International Conference on Engineering and Emerging Technologies (ICEET) 2024 Dec 27 (pp. 1-6). IEEE.
25. Bharadwaj A. Virtual Diagnosis using the heart sound taken through a digital stethoscope to help medical devoid areas (Doctoral dissertation, Dublin, National College of Ireland).
26. Niu Q, Wang Z, Zhang Y, Wang J, Zhang P, Wang C, Yin X, Hou Y. Combination use of clopidogrel and proton pump inhibitors increases major adverse cardiovascular events in patients with coronary artery disease: a meta-analysis. Journal of cardiovascular pharmacology and therapeutics. 2017 Mar;22(2):142-52.
27. Alkhodari M, Hadjileontiadis LJ, Khandoker AH. Identification of congenital valvular murmurs in young patients using deep learning-based attention transformers and phonocardiograms. IEEE Journal of Biomedical and Health Informatics. 2024 Jan 23;28(4):1803-14.

### Seah JJ, Zhao J, Wang DY, Lee HP. Review on the advancements of stethoscope types in chest auscultation. Diagnostics. 2023 Apr 25;13(9):1545.

### Beaumont CE. Choosing the best stethoscope. Nursing2024. 2007 Oct 1;37:12-3.

### Osama S, Salah L, Dahi G, Ghazal M, Almajali E, Hussain AJ, Yousaf J, Hassan T. A Review of Screening Heart and Lung Diseases using Auscultation and Artificial Intelligence. In2024 17th International Conference on Development in eSystem Engineering (DeSE) 2024 Nov 6 (pp. 168-173). IEEE.

### Wang Y, Yang X, Ye M, Zhao Y, Chen R, Da M, Wang Z, Mo X, Qi J. Machine Learning-Based Intelligent Auscultation Techniques in Congenital Heart Disease: Application and Development. Congenital Heart Disease. 2024 Mar 1;19(2).

### Singh H, Agarwal V. Heart Condition Monitoring Using Phonocardiogram and Murmurs Analysis by CNN. In2024 2nd International Conference on Advancements and Key Challenges in Green Energy and Computing (AKGEC) 2024 Nov 21 (pp. 1-7). IEEE.

### Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, Liu X, Wu Y, Dong F, Qiu CW, Qiu J. Artificial intelligence: A powerful paradigm for scientific research. The Innovation. 2021 Nov 28;2(4).

### Cinyol F, Baysal U, Köksal D, Babaoğlu E, Ulaşlı SS. Incorporating support vector machine to the classification of respiratory sounds by Convolutional Neural Network. Biomedical Signal Processing and Control. 2023 Jan 1;79:104093.

### Jeong SG, Nam SW, Jung SK, Kim SE. iMedic: Towards Smartphone-based Self-Auscultation Tool for AI-Powered Pediatric Respiratory Assessment. arXiv preprint arXiv:2504.15743. 2025 Apr 22.

### Guo L, Khobragade N, Kieu S, Ilyas S, Nicely PN, Asiedu EK, Lima FV, Currie C, Lastowski E, Choudhary G. Development and Evaluation of a Deep Learning–Based Pulmonary Hypertension Screening Algorithm Using a Digital Stethoscope. Journal of the American Heart Association. 2025 Feb 4;14(3):e036882.

### Chen Z, Hao J, Sun H, Li M, Zhang Y, Qian Q. Applications of digital health technologies and artificial intelligence algorithms in COPD: systematic review. BMC Medical Informatics and Decision Making. 2025 Feb 13;25:77.

### Udoy IA, Hassan O. AI-Driven Technology in Heart Failure Detection and Diagnosis: A Review of the Advancement in Personalized Healthcare. Symmetry. 2025 Mar 20;17(3):469.

### Farha F, Abass S, Khan S, Ali J, Parveen B, Ahmad S, Parveen R. Transforming pulmonary health care: the role of artificial intelligence in diagnosis and treatment. Expert Review of Respiratory Medicine. 2025 Apr 17:1-21.

### Medhi D, Kamidi SR, Sree KP, Shaikh S, Rasheed S, Murichathil AH, Nazir Z, reddyKamidi S. Artificial intelligence and its role in diagnosing heart failure: a narrative review. Cureus. 2024 May 5;16(5)