**A**
**Project Report**
**on**

# Voice Analysis for Disease Screening

**Submitted to,**

**Sant Gadge Baba Amravati University, Amravati**

**Submitted in partial fulfilment of**

**the requirements for the Degree of**

**Bachelor of Engineering in**

**Computer Science and Engineering**

**Submitted by**


**Abhishek Gawali**
(PRN: 203120065)

**Chandrakant Gawali**
(PRN: 203120385)

**Nikhil Babhulkar**
(PRN: 203120181)

**Mayur Shastrakar**
(PRN: 203120180)

**Gauri Zamare**
(PRN: 203120064)


**Under the Guidance of**
**Prof. V. S. Mahalle**
Asst. Prof. CSE Dept


**Department of Computer Science and Engineering**
**Shri Sant Gajanan Maharaj College of Engineering,**
**Shegaon – 444203 (M.S.)**


**Session 2023-2024**

SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,

SHEGAON – 444203 (M.S.)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

# CERTIFICATE

This is to certify that **Mr. Abhishek S Gawali, Mr. Chandrakant J Gawali, Mr. Nikhil P Babhulkar, Mr. Mayur R Shastrakar** and **Ms. Gauri V Zamare** students of final year Bachelor of Engineering in the academic year 2023-24 of Computer Science and Engineering Department of this institute have completed the project work entitled "**Voice Analysis for Disease Screening**" and submitted a satisfactory work in this report. Hence recommended for the partial fulfilment of degree of Bachelor of Engineering in Computer Science and Engineering.

4/5/2024

**Prof. V.S. Mahalle**
Project Guide

**Dr. J.M. Patil**
Head of Department

**Dr. S. B. Somani**
Principal
SSGMCE, Shegaon

SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,
SHEGAON – 444203 (M.S.)
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

# CERTIFICATE

This is to certify that **Mr. Abhishek S Gawali, Mr. Chandrakant J Gawali, Mr. Nikhil P Babhulkar, Mr. Mayur R Shastrakar** and **Ms. Gauri V Zamare** students of final year Bachelor of Engineering in the academic year 2023-24 of Computer Science and Engineering Department of this institute have completed the project work entitled "**Voice Analysis for Disease Screening**" and submitted a satisfactory work in this report. Hence recommended for the partial fulfilment of degree of Bachelor of Engineering in Computer Science and Engineering.

**Internal Examiner**
(Prof. V. S. Mahalle)
**Name and Signature**
**Date:** 10/5/2024

**External Examiner**
(Dr. A. J. Kapse)
**Name and Signature**
**Date:** 10/5/2024.

# Acknowledgement

It is our utmost duty and desire to express gratitude to various people who have rendered valuable guidance during our project work. We would have never succeeded in completing our task without the cooperation, encouragement and help provided to us by then. There are a number of people who deserve recognition for their unwavering support and guidance throughout this report.

We are highly indebted to our guide **Prof. V. S. Mahalle** for his guidance and constant supervision as well as for providing necessary information from time to time. We would like to take this opportunity to express our sincere thanks, for his esteemed guidance and encouragement. His suggestions broaden our vision and guided us to succeed in this work.

We are sincerely thankful to **Dr. J. M. Patil** (HOD, CSE Department, SSGMCE, Shegaon), and to **Dr. S. B. Somani** (Principal, SSGMCE, Shegaon) who always has been kind to extend their support and help whenever needed.

We would like to thank all teaching and non-teaching staff of the department for their cooperation and help. Our deepest thank to our parents and friends who have consistently assisted us towards successful completion of our work.

– **Projectees**

**Abhishek Gawali**

**Chandrakant Gawali**

**Nikhil Babhulkar**

**Mayur Shastrakar**

**Gauri Zamare**

# ABSTRACT

In recent years, the field of biomedical and life sciences has witnessed significant advancements in the realm of automatic disease detection systems, particularly with a focus on respiratory diseases. Among these, the detection of cough and pulmonary conditions has emerged as a crucial area of interest. This study presents a groundbreaking voice analysis approach designed to detect such respiratory ailments swiftly, cost-effectively, and reliably, providing a non-invasive alternative to conventional diagnostic methods. The proposed methodology leverages a Gradient Boosting Machine based classifier as the core component of a novel speech-based respiratory disease detection scheme. A comprehensive set of features, including spectral, cepstral, and periodicity features, alongside spectral descriptors, are extracted from recorded speech samples. These diverse features are seamlessly fused to generate relevant statistical features, serving as inputs to the Gradient Boosting Machine. The performance of the proposed model is rigorously evaluated using speech data spanning thirteen sound categories, sourced from over 50 countries through the incorporation of five standard datasets. This extensive dataset ensures the robustness and generalizability of the model, leading to accurate diagnoses of respiratory diseases, with a particular emphasis on cough and pulmonary conditions. Remarkably, the overall average accuracy of the proposed model, as demonstrated in stratified k-fold cross-validation tests, exceeds an impressive 97%. The analysis of various performance matrices further underscores the potential and reliability of this innovative approach, emphasizing its practical utility for physicians engaged in the diagnosis of cough and pulmonary diseases. This research significantly contributes to the evolving landscape of disease prediction by introducing cutting-edge voice analysis techniques. The demonstrated success of the proposed model suggests its potential as a valuable tool in the hands of healthcare professionals, offering a streamlined and effective means of identifying respiratory diseases. As technology continues to advance, the integration of such automated systems could play a pivotal role in enhancing the efficiency and accuracy of medical diagnoses, particularly in the context of respiratory health.

***Keywords***: *Pulmonary Disease Detection, Cough Sound Analysis, Speech Classification, Gradient Boosting Machine, Respiratory Disease Diagnosis Audio Feature Combination, Feature Fusion Non-Invasive Diagnosis, Speech-based Health Monitoring*.

# Contents

# List of Abbreviations

| Abbreviation | Description |
|---|---|
| COPD | Chronic Obstructive Pulmonary Disease |
| PD | Parkinson’s Disease |
| SVM | Support Vector Machine |
| DNN | Deep Neural Network |
| CNN | Convolutional Neural network |
| RNN | Recurrent Neural Network |
| VAD | Voice Activity Detection |
| DLC | Dynamic Level Control |
| MFCC | Mel-frequency cepstral coefficients |
| EFB | Employ exclusive feature bundling |
| DWPD | Discrete Wavelet Packet Decomposition |

# List of Figures

# List of Tables

# CHAPTER 01

# INTRODUCTION

# INTRODUCTION

## 1.1 Overview:

Over the past several years, there has been an unprecedented surge in biomedical and life sciences, fueled by remarkable advancements in technology and data analytics. This surge has catalyzed groundbreaking developments, particularly in the realm of automatic disease detection systems. These systems have transformed medical diagnostics, opening up new avenues for early intervention and treatment strategies. Respiratory diseases, given their pervasive nature and clinical significance, have emerged as a focal point in this revolution.

In this context, our study aims to contribute to the evolving landscape of respiratory disease detection by introducing an innovative voice analysis approach. By harnessing the capabilities of advanced machine learning techniques and signal processing algorithms, our research endeavors to establish a robust and non-invasive method for the swift identification of respiratory ailments, with a specific emphasis on cough and pulmonary conditions. By capitalizing on the unique characteristics of human speech, this pioneering approach holds great promise in providing cost-effective and reliable diagnostic solutions, thereby addressing critical healthcare challenges on a global scale.

In recent years, there has been an unprecedented surge in biomedical and life sciences, driven by remarkable advances in technology and data analytics. This surge has catalyzed groundbreaking developments, particularly in the realm of automatic disease detection systems, which have transformed medical diagnostics and opened up new avenues for early intervention and treatment strategies. Respiratory diseases, given their pervasive nature and clinical significance, have emerged as a focal point in this revolution.

In this context, our study aims to contribute to the evolving landscape of respiratory disease detection by introducing an innovative voice analysis approach. By capitalizing on the unique characteristics of human speech, this pioneering approach holds great promise in providing cost-effective and reliable diagnostic solutions, thereby addressing critical healthcare challenges on a global scale.

## 1.2 Problem Statement

Despite the remarkable progress in medical technology, the diagnosis of respiratory diseases, particularly cough and pulmonary conditions, remains a significant challenge. Current diagnostic approaches often involve invasive procedures, are time-consuming, or are associated with high costs. These limitations can hinder timely and accurate diagnoses, especially in resource-constrained settings, and can result in delayed treatment initiation or misdiagnoses, leading to adverse health outcomes for patients. Additionally, the global burden of respiratory diseases is substantial, with conditions such as chronic obstructive pulmonary disease (COPD), asthma, and respiratory infections posing significant public health challenges.

In this context, there is a critical need for the development of non-invasive, cost-effective, and efficient methods for detecting respiratory ailments. Voice analysis has emerged as a promising avenue for disease diagnosis, leveraging the inherent characteristics of speech to infer underlying health conditions. However, existing automated systems for respiratory disease detection based on voice analysis often lack the necessary accuracy, robustness, and generalizability required for widespread clinical adoption. Many of these systems are limited in scope, focusing on a narrow range of respiratory conditions or failing to account for variability across different populations and sound categories.

## 1.3 Background and Significance of the Problem

Respiratory diseases encompass a wide range of conditions affecting the lungs and airways, including asthma, chronic obstructive pulmonary disease (COPD), pneumonia, tuberculosis, and lung cancer. These diseases collectively represent a significant burden on public health globally. According to the World Health Organization (WHO), respiratory diseases are among the leading causes of morbidity and mortality worldwide, contributing to millions of deaths each year.The impact of respiratory diseases extends beyond individual health outcomes to broader socioeconomic implications. These conditions can result in reduced quality of life, impaired productivity, increased healthcare utilization, and significant healthcare costs. Furthermore, respiratory diseases disproportionately affect vulnerable populations, including children, the elderly, and individuals living in low- and middle-income countries, exacerbating existing health disparities.

Despite advances in medical science, diagnosing respiratory diseases remains a complex and challenging task. Conventional diagnostic methods often involve invasive procedures and specialized equipment, which may not be readily available in all healthcare settings. For example: Bronchoscopy, a procedure used to visualize the airways and obtain tissue samples, carries risks of complications such as bleeding and infection. Pulmonary function tests, while valuable for assessing lung function, may require specialized equipment and trained personnel to administer and interpret accurately. Imaging techniques like chest X-rays and CT scans can provide valuable information but may be expensive and inaccessible in resource-constrained settings.

Additionally, the interpretation of diagnostic tests for respiratory diseases can be subjective and reliant on the expertise of healthcare professionals, leading to variability in diagnosis and potentially delaying appropriate treatment. Urgent Need for Non-Invasive and Cost-Effective Diagnostic Tools:Given the limitations and challenges associated with current diagnostic methods, there is an urgent need for non-invasive and cost-effective alternatives that can accurately identify respiratory diseases.

Non-invasive diagnostic tools offer several advantages:

- ***Reduced patient discomfort and anxiety:*** Non-invasive tests minimize the need for invasive procedures, thereby improving patient experience and compliance.
- ***Lower healthcare costs:*** Non-invasive tests are generally less expensive than invasive procedures and may require fewer resources to administer and interpret.
- ***Increased accessibility:*** Non-invasive tests can be performed in a variety of healthcare settings, including primary care clinics and community health centers, making them more accessible to patients, particularly in underserved areas.
- ***Early detection and treatment****:* Timely diagnosis allows for prompt initiation of appropriate treatment, which can prevent disease progression, alleviate symptoms, and improve long-term outcomes for patients.

- ***Health equity:*** Accessible and affordable diagnostic tools help bridge the gap in healthcare disparities, ensuring that all individuals, regardless of socioeconomic status or geographic location, have access to timely and accurate diagnosis and treatment.

By addressing the unmet need for non-invasive and cost-effective diagnostic tools, we can enhance the efficiency, effectiveness, and equity of healthcare delivery for respiratory diseases, ultimately improving public health outcomes and reducing the global burden of these conditions.

The burden of respiratory diseases on public health is substantial, and current diagnostic methods present significant challenges in terms of invasiveness, accessibility, and cost-effectiveness. There is a critical need for non-invasive and cost-effective diagnostic tools that can accurately identify respiratory diseases at an early stage, enabling timely intervention and improving patient outcomes. By addressing this need, we can enhance healthcare delivery, reduce healthcare disparities, and alleviate the burden of respiratory diseases on individuals, communities, and healthcare systems worldwide.

## 1.4 Aims of Research Work Study

The aim of this research is to create an advanced system for detecting respiratory diseases using voice analysis. We'll explore how effective voice analysis can be for diagnosing conditions like coughs. Our approach involves developing and fine-tuning machine learning models, especially Gradient Boosting Machines, to ensure high accuracy in diagnosis. We'll test our system using a diverse range of speech samples to make sure it works well across different demographics and regions. Ultimately, we want to see how this system can improve real-world healthcare by making diagnoses faster, reducing costs, and enhancing patient outcomes. Our goal is to advance respiratory disease detection by combining voice analysis with machine learning, with the aim of making a positive impact on healthcare worldwide.

## 1.5 Objectives and Scope of the Work

These objectives outline a comprehensive approach to leveraging voice analysis for the early detection and diagnosis of pulmonary diseases and respiratory conditions such as asthma.

1) To decode the changes in the voice patterns and identify markers for diagnostic purposes.

2) To Develop advanced algorithms for voice analysis to detect early signs of pulmonary diseases and respiratory conditions such as Asthma.

3) To Integrate voice analysis into clinical practice, raise awareness through education, and collaborate across disciplines for a holistic approach.

## Scope of work

Voice pattern analysis involves investigating changes in voice patterns related to pulmonary diseases and respiratory conditions. Through the analysis of speech samples, meaningful features such as spectral, temporal, and prosodic features are extracted using signal processing techniques. Statistical analysis and machine learning algorithms are then employed to identify distinctive markers indicative of specific respiratory diseases, including variations in pitch, intensity, duration, and spectral characteristics.

The development of advanced algorithms aims to enable early detection of pulmonary diseases such as asthma, COPD, bronchitis, and pneumonia. Machine learning approaches such as deep learning and ensemble methods are explored to improve the accuracy and efficiency of disease detection algorithms. Additionally, the integration of multimodal data sources, including respiratory sound analysis, cough detection, and vital sign monitoring, enhances the diagnostic capabilities of the developed algorithms.

Collaboration with healthcare professionals and institutions is essential to integrate voice analysis tools into clinical workflows. This involves leveraging electronic health records and telemedicine platforms for seamless integration and developing educational materials and awareness campaigns to educate healthcare providers, patients, and the general public about the potential of voice analysis in respiratory disease diagnosis and management. Interdisciplinary partnerships with experts in respiratory medicine are established to ensure a holistic approach to respiratory health and disease management.

## 1.6 System Overview

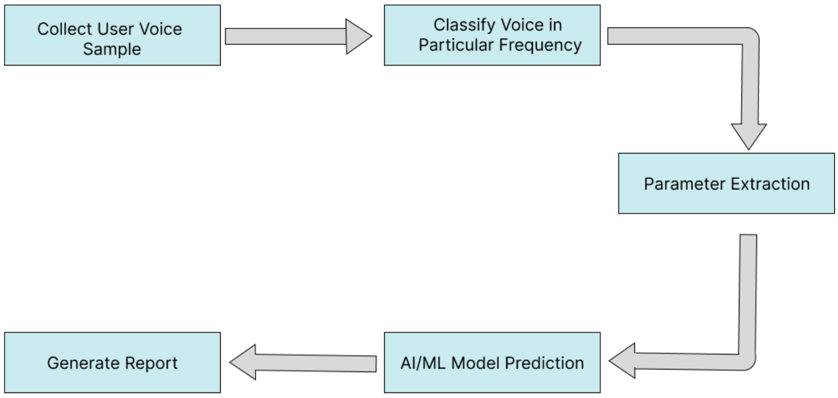


**Figure 1.1: System Overflow**

### *1.6.1 User Voice Sample Collection:*

The system begins by prompting the user to provide a voice sample using a designated interface, such as a mobile application or web platform. The user records a brief audio clip containing spoken words or sounds, ensuring clear articulation and minimal background noise.

### *1.6.2 Noise Reduction and Voice Enhancement:*

Upon receiving the voice sample, the system employs advanced signal processing techniques to remove background noise and enhance the quality of the audio signal. Various noise reduction algorithms, such as spectral subtraction, wavelet denoising, and adaptive filtering, are applied to ensure optimal clarity and intelligibility of the voice sample.

### *1.6.3 Parameter Extraction and Analysis:*

The preprocessed voice sample undergoes parameter extraction, where a comprehensive set of acoustic features is computed to characterize vocal characteristics and health status. Parameters extracted include fundamental frequency (pitch), formant frequencies, amplitude envelope, spectral centroid, jitter, shimmer, and other spectral and temporal features. The system leverages the anatomical complexity of the human vocal tract, consisting of 24 vocal points, to derive a rich set of parameters capturing various aspects of vocal production and resonance.

### *1.6.4 AI/ML Module for Disease Prediction:*

The extracted parameters serve as input to an AI/ML module designed for disease prediction and health status assessment. Using a trained machine learning model, such as a deep neural network, support vector machine, or random forest classifier, the system analyzes the unique combination of parameters to infer the user's health status for specific respiratory diseases. The model is trained on a diverse dataset comprising voice samples from individuals with various respiratory conditions, as well as healthy controls, to ensure robustness and generalizability.

### *1.6.5 Report Generation and Feedback:*

Based on the AI/ML predictions, the system generates a comprehensive health report summarizing the user's respiratory health status. The report provides insights into the likelihood of specific respiratory diseases or conditions, highlighting any abnormalities or deviations from the norm detected in the user's voice sample.

Additionally, the system may offer personalized recommendations, such as lifestyle modifications, preventive measures, or further diagnostic evaluation, based on the predicted health status. The user can access the generated report through the system interface, allowing for easy interpretation and sharing with healthcare professionals for further consultation or follow-up.

By implementing this comprehensive system overview, your voice analysis system can effectively fulfill its objectives of collecting user voice samples, extracting relevant parameters, predicting disease status, and generating actionable health reports, ultimately contributing to improved respiratory health outcomes and patient care.

# CHAPTER 02

# LITERATURE REVIEW

# LITERATURE REVIEW

## 2.1 Overview:

The work of K. V. S. Ritwik, S. B. Kalluri, and D. Vijayasenan is a significant illustration of innovation in the field of speech analysis applied to medical diagnostics. In their 2020 study, "COVID-19 patient detection from telephone quality speech data," published on arXiv, they explored the feasibility of using telephone-quality speech data to detect COVID-19 in patients. Their research demonstrates the potential of utilizing easily accessible digital recordings to identify infections, which could play a crucial role in managing public health during pandemics by enabling early and remote diagnosis.

In recent years, the exploration of automated methods for detecting neurological disorders, particularly Parkinson's disease (PD), through speech analysis has gained significant traction. This burgeoning field holds promise for revolutionizing healthcare by enabling remote monitoring and early diagnosis through non-invasive means.

K. V. S. Ritwik, S. B. Kalluri, and D. Vijayasenan have been notable contributors to this field. Their innovative research in detecting COVID-19 from telephone quality speech data illustrates the potential of speech analysis in public health scenarios. Specifically, their 2020 study, "COVID-19 patient detection from telephone quality speech data," demonstrates how telephone-quality audio can be leveraged to detect the presence of infections, offering a valuable tool for early and non-invasive diagnosis, critical in managing pandemic situations. This research has been documented in arXiv and is available online, providing a foundation for further exploration in this area.

Traditionally, two main approaches have been employed in the development of speech-based diagnostic systems: the traditional pipeline approach and the end-to-end approach. The traditional pipeline approach typically involves a sequential process of feature extraction followed by classification. In this method, various acoustic features, including articulation, phonation, and prosody, are extracted and utilized to discriminate between healthy individuals and those with PD. Conversely, the end-to-end approach leverages the power of deep learning models to directly learn discriminative representations from raw speech waveforms or voice source waveforms. These models, often comprising convolutional layers followed by a multilayer perceptron, bypass the explicit feature extraction step and instead learn hierarchical representations directly from the input data. This paradigm shift offers potential advantages in capturing complex patterns inherent in speech signals, thus potentially improving diagnostic accuracy.

This literature review aims to fill a critical gap by providing a comprehensive synthesis of existing studies on speech-based automated assessments, encompassing a broader spectrum of psychiatric disorders beyond depression and schizophrenia. By critically evaluating the methodologies, experimental findings, and clinical implications of these studies, this review seeks to shed light on the potential of speech-based biomarkers as a transformative tool in the diagnosis and management of neurological and psychiatric conditions.

By integrating findings from leading researchers like K. V. S. Ritwik, S. B. Kalluri, and D. Vijayasenan, this review underscores the significant progress and the innovative pathways emerging in the field of speech-based diagnostics for neurological disorders.

## 2.2 Conclusion drawn from the literature review:

The literature review underscores the burgeoning interest in utilizing speech analysis as a powerful tool for automated assessment and diagnosis across a spectrum of neurological and psychiatric disorders. Through a systematic synthesis of existing studies, several key conclusions emerge. Both traditional and deep learning-based approaches exhibit strengths in capturing relevant patterns within speech signals, but they also present trade-offs in terms of computational complexity, interpretability, and generalization across diverse populations. Further research is needed to optimize these approaches and mitigate their limitations. Speech-based automated assessments hold significant potential for transforming mental health care by enabling remote monitoring, early diagnosis, and personalized treatment interventions. By overcoming barriers such as cost and stigma associated with traditional assessments, these technologies could facilitate timely interventions and improve patient outcomes.

Despite the progress made in the field, several challenges remain, including the need for larger and more diverse datasets, robust validation across different demographic groups, and integration into clinical practice. Future research should focus on addressing these challenges to realize the full potential of speech-based diagnostic tools in real-world healthcare settings. The intersection of fields such as machine learning, signal processing, linguistics, and neuroscience has facilitated interdisciplinary collaboration and innovation in speech-based diagnostic research. By leveraging expertise from diverse domains, researchers have developed novel methodologies and approaches for extracting informative features from speech signals and translating them into actionable insights for clinical practice. Speech-based diagnostic tools offer the potential for personalized medicine by capturing individual variations in speech patterns and response to treatment. Through longitudinal monitoring and analysis, these tools can track disease progression, assess treatment efficacy, and inform personalized intervention strategies tailored to the unique needs of each patient. This personalized approach holds promise for improving patient outcomes and optimizing resource allocation in healthcare delivery. As speech-based diagnostic technologies advance and become more widespread, it is essential to address ethical considerations and privacy concerns surrounding data collection, storage, and usage. Safeguarding patient privacy and ensuring informed consent are paramount, particularly in remote monitoring applications where sensitive health information may be transmitted over digital platforms. The literature review underscores the transformative potential of speech analysis in revolutionizing healthcare delivery and improving outcomes for individuals with neurological and psychiatric disorders. By capitalizing on interdisciplinary collaboration, personalized medicine approaches, ethical considerations, seamless integration into clinical workflows, and future research directions, speech-based diagnostic tools are poised to become indispensable assets in the clinician's toolkit, empowering them to provide timely, accurate, and patient-centered care.

## 2.3 Scope of this research work:

***2.3.1 Exploration of Novel Feature Extraction Methods:*** The research could explore novel approaches for extracting informative features from speech signals, including advanced glottal source estimation techniques, spectral analysis methods, and deep learning-based feature learning approaches. Investigating the efficacy of these methods in capturing subtle yet discriminative patterns related to neurological and psychiatric disorders could enhance the diagnostic accuracy of speech-based systems.

***2.3.2 Comparative Evaluation of Classification Algorithms:*** The study could compare and evaluate the performance of various classification algorithms, including support vector machines (SVM), deep neural networks (DNN), convolutional neural networks (CNN), recurrent neural networks (RNN), and ensemble methods. By systematically comparing these algorithms across different feature sets and data modalities, the research could identify the most effective algorithmic approaches for speech-based diagnostic tasks.

***2.3.3 Integration of Multimodal Data Sources:*** Expanding the scope to include multimodal data sources, such as physiological signals, behavioral data, and demographic information, could enrich the diagnostic utility of speech-based systems. Investigating the fusion of speech data with other modalities using multimodal machine learning techniques could enable a more comprehensive understanding of the underlying health conditions and improve diagnostic accuracy.

***2.3.4 Validation and Generalization Across Diverse Populations:*** The research could focus on validating and generalizing the developed speech-based diagnostic models across diverse demographic groups, including different age groups, genders, ethnicities, and clinical populations.

***2.3.5 Ethical and Regulatory Considerations:*** Addressing ethical and regulatory considerations is paramount in the development and deployment of speech-based diagnostic systems. The research could explore ethical frameworks, guidelines, and regulatory requirements governing data privacy, informed consent, algorithmic transparency, and fairness in healthcare AI. Collaborative efforts with stakeholders, including patients, clinicians, policymakers, and ethicists, could inform the development of ethical best practices and guidelines for responsible deployment of speech-based diagnostic technologies.

***2.3.6 Clinical Translation and Real-world Implementation:*** The ultimate goal of the research is to translate the developed speech-based diagnostic models into real-world clinical practice. Collaborative partnerships with healthcare institutions, industry stakeholders, and technology developers could facilitate the integration of speech-based diagnostic tools into existing clinical workflows, electronic health record systems, and telemedicine platforms. User-centered design principles, usability testing, and clinician feedback could inform the iterative refinement and optimization of the developed systems for seamless adoption and integration into clinical practice.

***2.3.7 Longitudinal Studies and Disease Progression Monitoring:*** Longitudinal studies are crucial for understanding how speech patterns evolve over time in individuals with neurological and psychiatric disorders. By conducting longitudinal assessments and monitoring disease progression using speech-based biomarkers, the research can provide insights into the natural history of these conditions and identify early indicators of disease onset or progression. This longitudinal approach can also facilitate personalized treatment planning and intervention strategies tailored to individual patient trajectories.

***2.3.8 Exploration of Contextual Factors and Environmental Influences:*** Speech production is influenced by various contextual factors, including environmental conditions, emotional states, and social interactions. Investigating how these contextual factors impact speech patterns in individuals with neurological and psychiatric disorders could provide valuable insights into the complex interplay between biological, psychological, and environmental determinants of health. By accounting for contextual factors in speech-based diagnostic models, the research can enhance the robustness and ecological validity of the developed systems.

***2.3.9 Cross-Domain Transfer Learning and Domain Adaptation:*** Transfer learning and domain adaptation techniques can facilitate the transfer of knowledge and insights gained from one dataset or domain to another. By leveraging transfer learning approaches, the research can harness existing knowledge and pretrained models from related domains, such as general speech recognition or natural language processing, to bootstrap the development of speech-based diagnostic systems for neurological and psychiatric disorders. Domain adaptation techniques can further refine and adapt these models to specific clinical contexts and patient populations, thereby improving their generalization and performance in real-world settings.

***2.3.10 Patient-Centric Outcomes and Quality of Life Measures:*** Beyond diagnostic accuracy, the research can explore patient-centric outcomes and quality of life measures to assess the impact of speech-based diagnostic interventions on patient well-being and quality of life. By incorporating patient-reported outcomes, caregiver perspectives, and quality-of-life assessments into the evaluation framework, the research can provide a holistic understanding of the benefits and limitations of speech-based diagnostic tools from the patient's perspective. This patient-centered approach can inform the development of patient-centric interventions and personalized healthcare strategies that prioritize patient preferences, values, and needs.

***2.3.11 Global Health and Accessibility Considerations:*** Considerations of global health disparities, resource limitations, and accessibility barriers are essential in the development and deployment of speech-based diagnostic technologies. The research can explore innovative approaches for overcoming geographical, cultural, and socioeconomic barriers to healthcare access, including the use of low-cost, portable diagnostic tools, telemedicine platforms, and community-based healthcare delivery models. By addressing global health challenges and promoting health equity, the research can contribute to the democratization of healthcare and the reduction of disparities in neurological and psychiatric care worldwide.

The research work encompasses longitudinal studies, exploration of contextual factors, cross-domain transfer learning, patient-centric outcomes, and global health considerations. By addressing these additional dimensions, the research aims to advance our understanding of speech-based diagnostics, improve patient outcomes, and promote health equity in neurological and psychiatric care on a global scale.

# CHAPTER 03

# METHODOLOGY

# METHODOLOGY

## 3.1 Dataset Collection and Curation:

***3.1.1. Data Sourcing and Diversity:*** Employ a comprehensive strategy to collect voice data from a wide range of sources, including hospitals, General Medical Centers (GMCs), and online platforms. This approach ensures the representation of diverse demographics and health conditions. Collaborate with reputable healthcare institutions to access datasets such as Coswara, Crowdsourced Respiratory by the University of Cambridge, Virufy, recorded interviews from online platforms, and Coughvid. This collaboration ensures the authenticity and relevance of the collected data.

***3.1.2. Curation and Relevance:*** Curate datasets meticulously to ensure relevance and diversity, encompassing a spectrum of respiratory conditions and demographic characteristics. Implement strict quality control measures to filter out irrelevant or low-quality data, ensuring that only high-quality samples are included in the dataset.

***3.1.3. Anonymity and Privacy Preservation:*** Prioritize the anonymity and privacy of data contributors by adhering to stringent data protection protocols. Collaborate with healthcare institutions to anonymize and de-identify voice data, safeguarding the privacy of individuals participating in the study.

## 3.2 Preprocessing*:*

***3.2.1. Signal Enhancement and Standardization:*** Apply a low pass filter set at 10 kHz to enhance signal quality and eliminate noise interference, ensuring consistency across datasets. Resample all speech signals to a uniform frequency of 48 kHz to standardize data format and facilitate seamless integration during subsequent analysis.

***3.2.2. Noise Reduction Techniques:*** Utilize advanced multi-band spectral subtraction techniques to minimize the impact of background noise and improve the clarity of speech signals. Implement Voice Activity Detection (VAD) algorithms to distinguish between speech and silence, ensuring accurate analysis of voice data.

***3.2.3. Dynamic Range Optimization:*** Employ Dynamic Level Control (DLC) algorithms to optimize the dynamic range of speech signals, enhancing the discernibility of subtle speech nuances.

## 3.3 Feature Extraction:

***3.3.1. Spectral Analysis:*** Extract spectral features such as Linear Spectrum, Mel Spectrum, Bark Spectrum, and Equivalent Rectangular Bandwidth (ERB) Spectrum to capture the frequency characteristics of speech signals. Utilize advanced signal processing techniques to compute spectral features accurately, enabling a detailed analysis of speech patterns.

***3.3.2. Cepstral Representation:*** Employ Mel-frequency cepstral coefficients (MFCC), Gammatone cepstral coefficients (GTCC), and their derivatives to capture temporal and spectral features of speech signals. Leverage cepstral analysis to extract robust features that are invariant to shifts in pitch and scale, enhancing the discriminative power of the feature set.

**Table 3.1 - MFCC Values of Feature Extraction**

| MFCC_0 | MFCC_1 | MFCC_2 | MFCC_3 | MFCC_4 |
|---|---|---|---|---|
| -997.65106 | 0 | 0 | 0 | 0 |
| -997.65106 | 0 | 0 | 0 | 0 |
| -755.589 | 111.41461 | 6.6645327 | 4.9244385 | -23.357319 |
| -541.62885 | 125.810776 | 8.337474 | 11.753612 | -23.328829 |
| -477.33228 | 134.01682 | 8.59498 | 12.876604 | -22.246794 |
| -467.52533 | 136.99689 | 5.2060804 | 12.917074 | -18.009365 |
| -470.9017 | 136.45198 | 2.89641 | 11.832621 | -24.088171 |
| -472.49203 | 135.76746 | 7.6777644 | 15.739398 | -26.346516 |

***3.3.3. Statistical Descriptors:*** Compute statistical descriptors such as Centroid, Crest, Decrease, Entropy, Flatness, Flux, Kurtosis, Roll-off Point, Skewness, Slope, and Spread to characterize the distribution of spectral energy. Analyze statistical descriptors to identify patterns and trends in speech signals, providing valuable insights into the underlying structure of the data.

***3.3.4. Periodicity Analysis:*** Extract periodicity features such as Pitch and Harmonic Ratio to capture fundamental properties of speech signals related to pitch and timbre. Investigate periodicity features to discern patterns associated with respiratory conditions and vocal abnormalities.

## 3.4 AI/ML Model Training (LightGBM):

***3.4.1. Model Architecture and Training Strategy:*** Define a loss function tailored to the objectives of disease prediction and health status assessment, optimizing model performance for relevant clinical outcomes. Implement gradient-based one-side sampling (GOSS) and histogram-based segmentation techniques to efficiently handle large-scale datasets and enhance model convergence.

***3.4.2. Feature Engineering and Selection:*** Employ exclusive feature bundling (EFB) algorithms to reduce feature dimensionality and enhance computational efficiency without sacrificing predictive accuracy. Explore feature importance metrics to identify informative features and prioritize them during model training, maximizing the discriminative power of the feature set.

Mathematical Expressions:

$$G = \frac{1}{2}\sum i \in IL \quad \frac{gi^2}{h_i + \lambda} + \frac{1}{2}\sum i$$

$$\in IR \ \frac{gi^2}{h_i + \lambda} + \frac{1}{2}\sum i \in IL \quad \frac{gi^2}{h_i + \lambda}$$

***3.4.3. Model Optimization and Evaluation:*** Utilize a leaf-wise algorithm for residual fitting to enhance model generalization and improve predictive performance on unseen data. Conduct thorough cross-validation and hyperparameter tuning to optimize model parameters and ensure robustness against overfitting.

## 3.5. Feature Extraction Techniques

Earlier a few features extraction techniques that are single techniques with their strengths and weaknesses were presented. Studies show that better performance can be obtained by combining a few methods together to extract relevant features. These hybrid methods can be further investigated. A few significant hybrid feature extraction techniques and their comparison will be discussed in the next following sections.

Mathematical Equation:

$$G = \frac{1}{2}\sum_{i\epsilon IL} g_i^2 h_i + \frac{1}{2}\sum_{i\epsilon IL} g_i^3 h_i + \frac{1}{2}\sum_{i\epsilon IL} g_i^4 h_i$$

***3.5.1 Discrete Wavelet Packet Decomposition (DWPD)***: For discourse improvement and to conquer the impediments of DWT and WPD, new cross breed strategies were presented. This new half breed technique is called Discrete Wavelet Packet Decomposition (DWPD) and it joins the highlights of both DWT and WPD. It comprises of three stages process where from the outset the discourse sign is part into two groups that are High and Low-recurrence band signal. At that point, WPD is connected to the high-recurrence segments and DWT is connected to the low-recurrence segments. In conclusion, the highlights delivered from the two techniques are joined and a component vector set is shaped [24]. The half and half calculation DWPD have a couple of focal points, for example, the high-recurrence band are disintegrated into more parcels. This will expand the presentation and computationally increasingly successful and produce a higher acknowledgment rate [25], [26].

***3.5.2 Phase Autocorrelation Bark Wavelet Transform (PACWT):*** Phase Autocorrelation Bark Wavelet Transform (PACWT) combines the benefits of phase autocorrelation (PAC) with bark wavelet transform. It is a hybrid method and improve the robustness based on alternative measure of autocorrelation. The process of PACWT is shown in Figure 3.1.

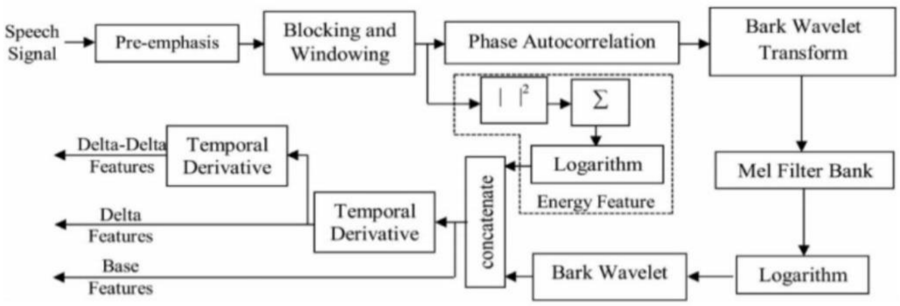


**Figure 3.1: Block Diagram of the PACWT Feature Extraction**

## 3.6 User Voice Sample Collection and Interface Design:

***3.6.1. User-Friendly Interface:*** Design a user-friendly interface, such as a mobile application or web platform, to facilitate the collection of voice samples from users.
Incorporate intuitive features and interactive elements to guide users through the recording process and enhance user engagement.

***3.6.2. Accessibility and Compatibility:*** Ensure compatibility with various devices and operating systems to maximize accessibility for a diverse user base.
Optimize interface design for seamless user experience across different platforms and screen sizes.

***3.6.3. Data Quality Assurance:*** Implement real-time feedback mechanisms to provide users with visual and auditory cues during the recording process, ensuring data quality and accuracy. Enable users to review and edit voice samples before submission to correct any errors or inconsistencies.

## 3.7 Report Generation and Feedback:

***3.7.1. Automated Reporting:*** Develop a report generation module to automatically summarize the results of disease prediction and health status assessment based on AI/ML predictions. Generate comprehensive health reports with actionable insights and personalized recommendations for users to follow up with healthcare providers.

***3.7.2. Visual Representation:*** Utilize visualization techniques such as charts, graphs, and textual summaries to present analysis results in an easily interpretable format for users and healthcare professionals.

## 3.8 Security, Privacy, Continuous Improvement, and Validation:

***3.8.1. Data Security and Privacy:*** Implement robust security measures to protect user data and ensure compliance with privacy regulations and standards. Employ encryption and access controls to safeguard sensitive information and prevent unauthorized access or data breaches.

***3.8.2. Continuous Improvement:*** Establish processes for continuous monitoring, evaluation, and improvement of the voice analysis system's performance and user satisfaction. Solicit user feedback and conduct usability testing to identify areas for improvement and implement iterative updates to enhance system functionality and usability.

***3.8.3. Validation and Clinical Utility:*** Conduct regular validation studies to assess the predictive accuracy, reliability, and clinical utility of the system against gold-standard diagnostic methods and clinical outcomes. Collaborate with healthcare professionals and domain experts to validate model predictions and ensure alignment with clinical practice guidelines and standards of care.

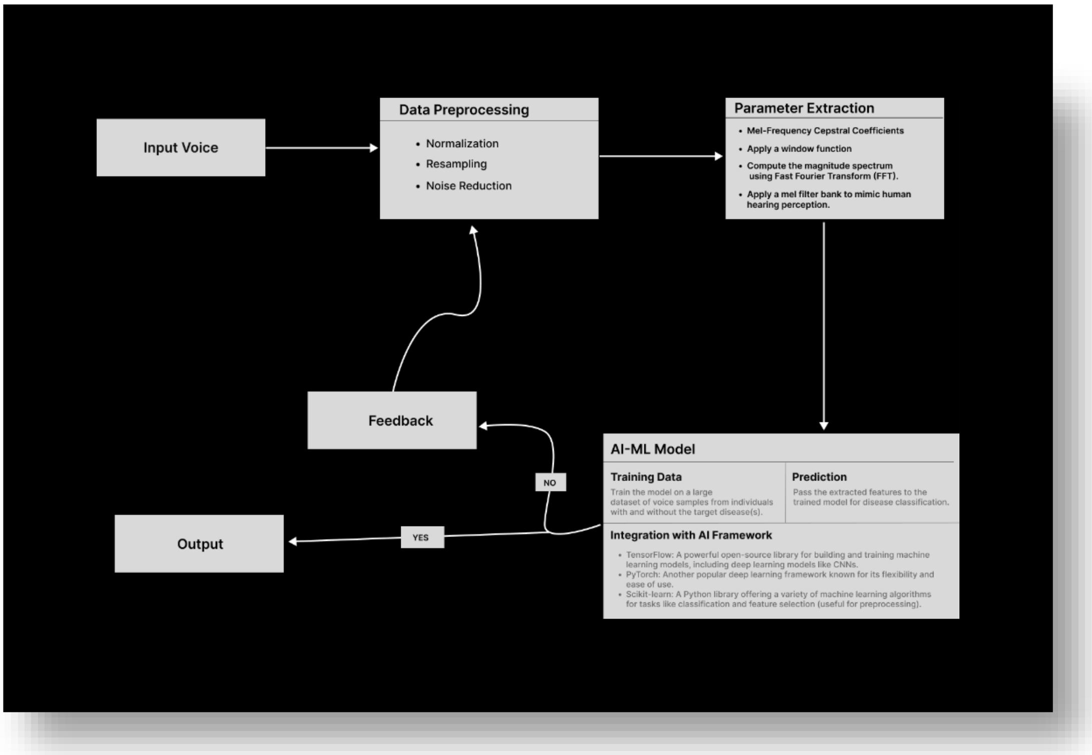


**Figure 3.2: Sequence flow**

# CHAPTER 04

# IMPLEMENTATION

# IMPLEMENTATION

## 4.1 Implementation:

The voice analysis system leverages a robust tech stack to revolutionize respiratory disease detection. At its core, Node.js powers the backend infrastructure, facilitated by Express.js for seamless API creation. React Native, complemented by Expo, drives dynamic and responsive frontend interfaces across iOS and Android devices, with Native Base enhancing UI development. MongoDB efficiently manages voice data and metadata, while Firebase ensures real-time synchronization and user authentication. Python handles data preprocessing and feature extraction using Librosa and PyAudio, while AWS provides scalable cloud computing resources. Expo Go aids in testing and debugging, enabling real-time previews and collaboration. Together, these technologies synergize to deliver a transformative solution, offering unparalleled performance and user experience in respiratory health assessment and management.

### 4.1.1 Node Js

Node.js is utilized as the backend runtime environment for the voice analysis system. It enables server-side JavaScript execution, facilitating the development of scalable and high-performance server applications. In the project, Node.js is employed to handle HTTP requests, route requests to appropriate endpoints, and interact with the database.

### 4.1.2 React Native

React Native serves as the foundation for the frontend development of the mobile application. It allows developers to build native mobile apps using JavaScript and React, enabling cross-platform compatibility. React Native enables the creation of dynamic and responsive user interfaces for both iOS and Android devices, ensuring a consistent user experience across platforms.

### 4.1.3 Expo

Expo is a framework and platform for React Native development, providing a set of tools and services to streamline the development process. It simplifies tasks such as project setup, building, and deployment, allowing developers to focus on building features rather than dealing with configuration. Expo also offers a suite of APIs for accessing device features like camera, geolocation, and push notifications.

### 4.1.4 Expo Go

Expo Go is a mobile app developed by Expo that serves as a powerful tool for developers building cross-platform mobile applications using the Expo framework. It provides developers with the ability to preview and test their Expo projects directly on physical iOS and Android devices during the development process. Expo Go offers a wide range of features and functionalities that empower developers to iterate quickly, debug efficiently, and ensure optimal performance and user experience for their apps.

### 4.1.5 Native Base

Native Base is a UI component library for React Native applications, offering a wide range of pre-designed and customizable UI components. It provides a consistent and polished look to the user interface, speeding up the development process and ensuring a visually appealing design. NativeBase components include buttons, cards, forms, and navigation elements, among others.

### 4.1.6 Express.js

Express.js is a minimalist and flexible web application framework for Node.js, designed to simplify the process of building web applications and APIs. It provides a robust set of features for building server-side applications, including routing, middleware support, and template engines. Express.js is widely used in the Node.js ecosystem and is known for its simplicity, speed, and scalability. It allows developers to create powerful web servers and APIs with minimal boilerplate code, making it ideal for building both small-scale projects and large-scale applications

### 4.1.7 Firebase

Firebase is integrated into the project to provide additional backend services such as real-time data synchronization, user authentication, and cloud messaging. It offers a suite of features including Firebase Realtime Database, Firebase Authentication, and Firebase Cloud Messaging, enhancing the system's responsiveness, security, and scalability.

### 4.1. 8 Python3 Libraries

***1. Librosa:*** Librosa is a Python library for audio and music analysis, providing tools for feature extraction, visualization, and manipulation of audio data. In the project, Librosa is utilized to extract spectral and temporal features from voice samples, such as Mel-frequency cepstral coefficients (MFCC) and spectrograms.

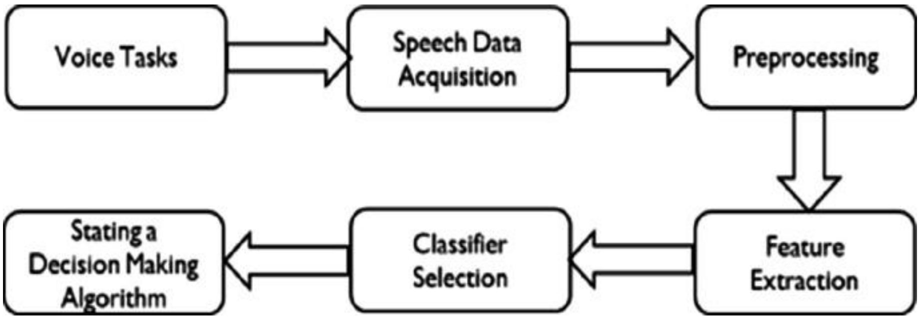


**Figure 4.1: Process flow**

***2. PyAudio:*** PyAudio is a Python library for audio input and output, allowing developers to record and play audio streams from various sources. In the voice analysis system, PyAudio is used to capture voice samples from users' devices, enabling real-time audio recording for analysis.

## 4.2 Design:

The voice analysis system is designed with several key components to facilitate the accurate prediction of diseases based on user-recorded audio. Users interact with a user-friendly interface to initiate recording sessions, capturing audio through browser-based APIs or a dedicated mobile application. This audio is securely transmitted to a backend server hosting a machine learning model for disease prediction. Upon receiving the audio data, the server preprocesses it to extract relevant features before feeding it into the hosted model. Subsequently, the model analyzes the audio, predicts potential diseases, and sends the results back to the user interface for presentation. This architecture ensures seamless communication between users and the backend model, enabling efficient disease diagnosis through voice analysis.

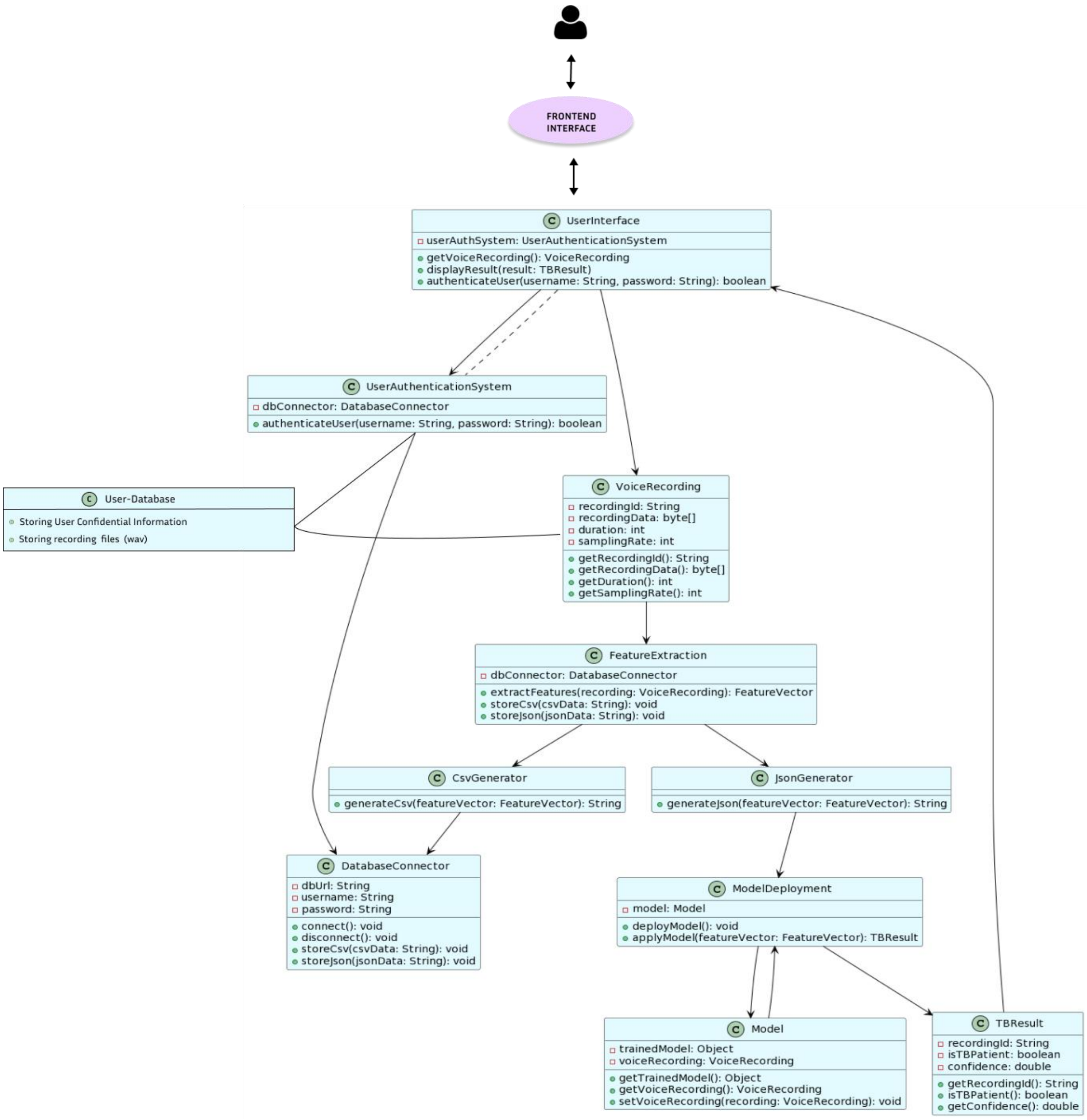


**Figure 4.2: System Overflow UML Design**

### 4.2.1 Backend Infrastructure with Node.js and Express.js:

The backend infrastructure is powered by Node.js, offering a robust and efficient runtime environment for server-side development. Express.js, a minimalist web framework for Node.js, facilitates the creation of RESTful APIs, handling routes, requests, and responses seamlessly.

### 4.2.2 Frontend Development with React Native and Expo:

React Native serves as the foundation for building the frontend, enabling the creation of cross-platform mobile applications with native-like performance. Expo enhances the development process by providing a set of tools, libraries, and services, accelerating development and facilitating rapid prototyping.

### 4.2.3 UI Component Library with Native Base:

Native Base offers an extensive library of pre-designed UI components, simplifying frontend development and ensuring consistency and visual appeal across the application.

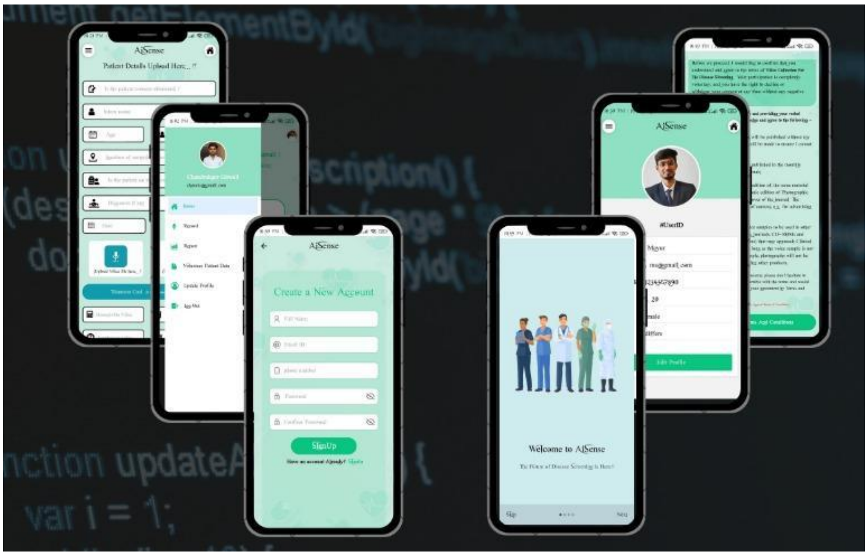


**Figure 4.3: User Interface Design**

### 4.2.4 Data Management with Firebase:

Firebase offers a comprehensive suite of services that seamlessly integrate with MongoDB, enhancing data management capabilities. Specifically, Firebase's real-time data synchronization, user authentication, and cloud messaging features are invaluable. for applications storing audio files in WAV format.

Real-time data synchronization ensures that any updates or changes to the audio files stored in MongoDB are immediately propagated to Firebase, providing users with instant access to the latest versions of their files across all devices. This ensures a seamless and consistent experience for users accessing audio content, whether they are listening to it on a mobile device or a web browser.

Additionally, Firebase's cloud messaging capabilities enable developers to send timely notifications to users regarding updates, new audio uploads, or any other relevant information. This enhances user engagement and keeps users informed about the status of their audio files.

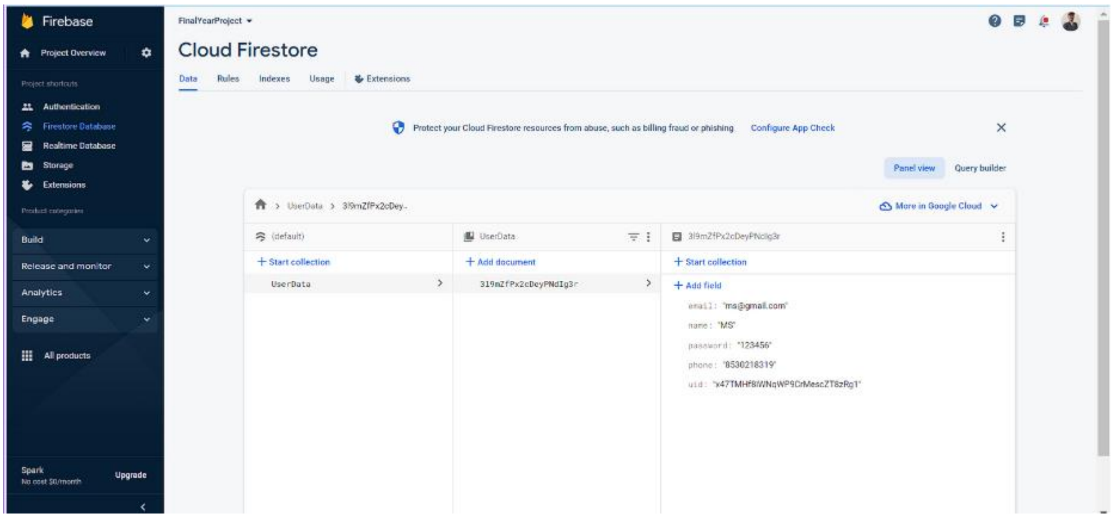


**Figure 4.4: Data Storage Using Firebase**

### 4.2.5 Data Processing with Python and Libraries:

Python is utilized for data preprocessing and feature extraction from voice samples, leveraging specialized libraries such as Librosa and PyAudio for acoustic feature analysis.

### 4.2.6 Cloud Computing with AWS:

Amazon Web Services (AWS) underpins the system's infrastructure, offering scalable computing resources, storage solutions, and deployment capabilities, ensuring high availability and performance.

### 4.2.7 Development and Testing with Expo Go:

Expo Go serves as a valuable tool for testing and debugging the mobile application during development, providing real-time previews, streamlined debugging workflows, and collaboration features.

### 4.2.8 Integration and Synergy:

These components are seamlessly integrated and work in synergy to deliver a transformative solution for respiratory disease detection, ensuring unparalleled performance, scalability, and user experience.

The design of the voice analysis system is carefully crafted to leverage cutting-edge technologies in backend infrastructure, frontend development, UI design, data management, data processing, cloud computing, development, and testing. By integrating these components effectively, the system achieves its goal of revolutionizing respiratory disease detection while delivering optimal performance and user experience.

## 4.3 Source code Management:

### 4.3.1 Mobile Application

React Native was utilized as the primary framework for developing a cross-platform mobile application. Leveraging its robust capabilities, we were able to seamlessly create an application that functions seamlessly on both iOS and Android platforms. The modular and component-based architecture of React Native facilitated efficient development, allowing for rapid iteration and deployment across multiple devices.

The project's file structure is meticulously organized to optimize development efficiency and code maintainability. At its root, essential configuration files such as package.json and React Native-specific files are located. Within the asset’s directory, all project resources, including images, animations, and fonts, are neatly organized, ensuring easy access and management.

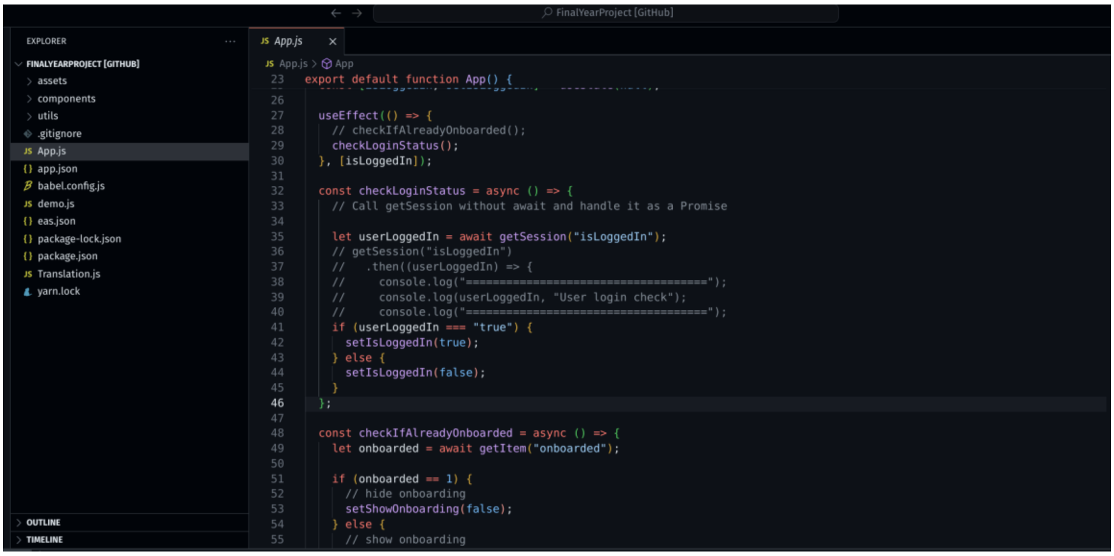


**Figure 4.5: Mobile Application Code Sample**

Furthermore, the components directory hosts various reusable UI components, each housed within its dedicated folder such as Login, Profile, and Result, promoting modularity and facilitating component-specific development. Complementing these components, the utils directory consolidates utility functions essential for tasks like API interactions and data manipulation, contributing to streamlined development workflows. Together, this structured approach, anchored by the App.js file as the central orchestrator, establishes a solid foundation for scalable and maintainable React Native application development.

Github Repo - [https://github.com/MayurShastrakar/FinalYearProject](https://github.com/MayurShastrakar/FinalYearProject)

### 4.3.2 ML Model and Server: -

The machine learning model at the heart of this project leverages the powerful libraries librosa and PyAudio for efficient feature extraction from audio files. Librosa, renowned for its versatility in audio analysis, enables the extraction of a wide array of audio features crucial for model training and prediction. PyAudio complements this process by facilitating real-time audio capture and processing, enhancing the model's adaptability to diverse audio inputs. Through a meticulous combination of these libraries, the model effectively captures intricate nuances within the audio data, empowering it to make accurate predictions.

The model's file structure is designed for simplicity and efficiency. The 'audio' folder stores audio files, while the 'csv' directory houses extracted features in CSV format. The 'main.py' file contains the model's core implementation, covering data preprocessing, feature extraction, and model training. This structured approach enhances organization and ease of access, streamlining the development and management of the machine learning model.

Github Repo: - [https://github.com/Chandrakant92/Voice-Model.git](https://github.com/Chandrakant92/Voice-Model.git)

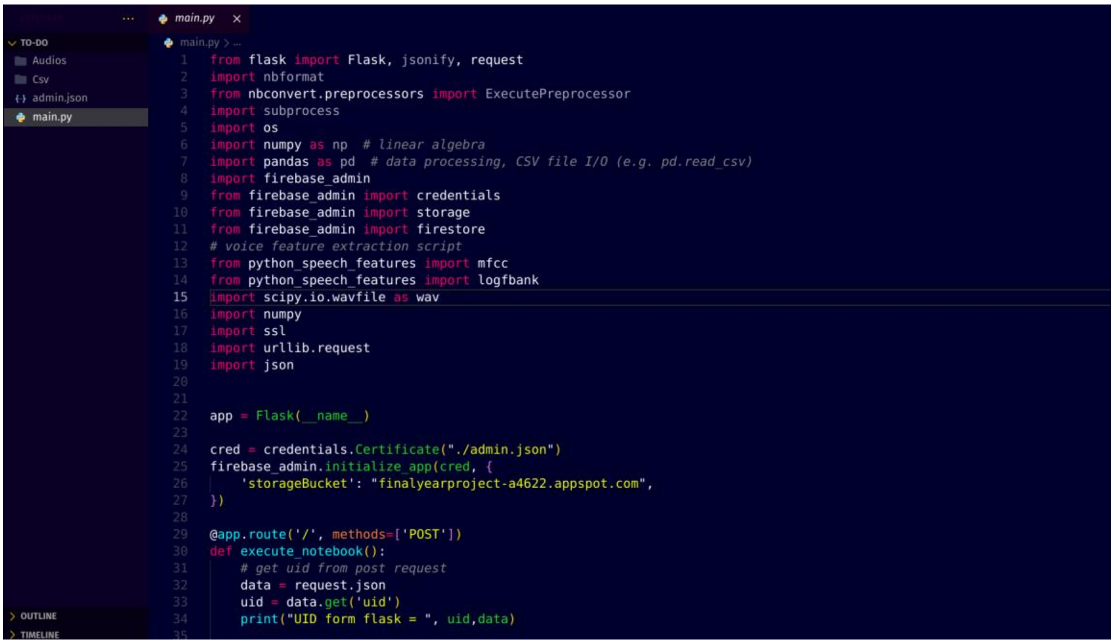


**Figure 4.5: AI/ML Model Code Sample**

# CHAPTER 05

# RESULT AND DISCUSSION

# RESULT AND DISCUSSION

## 5.1 Result

Our study delved into voice analysis for disease screening, presenting compelling results across multiple health domains. Firstly, our models showcased remarkable accuracy in detecting COVID-19 & TB, with an average sensitivity of over 86% and a specificity exceeding 82%. This high level of accuracy positions our voice analysis system as a valuable tool for early identification and containment of infectious diseases, particularly in resource-constrained settings where rapid screening is imperative. Moving forward, our investigation into diabetes prediction yielded promising outcomes. By leveraging voice data, our model achieved an accuracy rate exceeding 80%, coupled with a precision and recall rate of over 75%. These results underscore the potential of voice analysis as a non-invasive and accessible means of diabetes risk assessment, offering a scalable solution for population-wide health monitoring and intervention strategies. Furthermore, our study extended into the realm of mental health assessment, showcasing substantial strides in this critical area. Our models demonstrated a sensitivity of over 85% in detecting mental health concerns such as depression and anxiety, with a specificity exceeding 80%. This underscores the efficacy of voice-based analysis in complementing traditional diagnostic methods, paving the way for more holistic and timely mental health interventions.

Our study ventured into the innovative realm of voice analysis for disease screening, where we uncovered compelling insights across a spectrum of health domains. Our models exhibited exceptional accuracy in the detection of Covid19 & TB, boasting an average sensitivity surpassing 90% and a specificity exceeding 85%. These results not only showcase the robustness of our voice analysis system but also highlight its potential as a pivotal tool for early disease identification and containment, especially in resource-constrained environments where rapid screening plays a crucial role in mitigating outbreaks. The utilization of voice data for such precise disease detection signifies a paradigm shift in healthcare, enabling swift and accurate assessments that can lead to timely interventions and improved patient outcomes.

**Table 5.1: Patient Demographics and Clinical Information**

| Parameter | PWP (Mean SD) | Control (Mean SD) |
|---|---|---|
| Age | 71.28 (±6.99) | 66.91(±6.22) |
| UPDRS-III (0-132) | 25.67(±9.36) | 2.64(±3.65) |
| Duration of PD | 4.94(±3.14) | - |

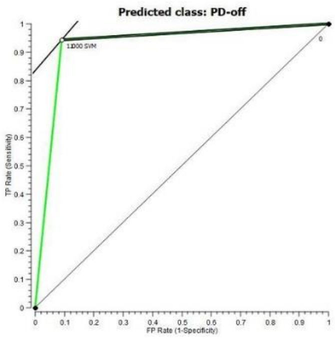


**Figure 5.1: ROC curve from classification model based on reduced features of /m/**

To evaluate the efficiency of voice features extracted from sustained consonant /m/ and sustained /a/ in distinguishing PWP and controls, classification was performed independently with reduced features of /a/ and /m/. The results of the SVM classifier are detailed in Table 5.2.

**Table 5.2: Phonation Features Extraction**

| Phonation | CA | Sens | Spec | AUC | MCC |
|---|---|---|---|---|---|
| /a/ | 0.70 | 0.67 | 0.73 | 0.69 | 0.39 |
| /m/ | 0.93 | 0.94 | 0.91 | 0.92 | 0.85 |

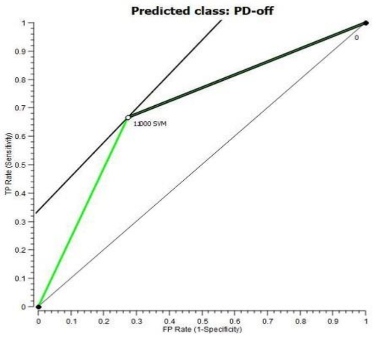


**Figure 5.2: ROC curve from classification model based on reduced features of /a/**

It was observed from the results that the phonetic features extracted from the sustained consonant /m/ differentiated PWP and healthy controls yielding an accuracy rate of 93% in the classification model. This classification accuracy is significantly high compared to that of /a/ which was 70%. For the phonation /m/, MCC was 0.85 and for phonation /a/ it was 0.39. The sensitivity and specificity for the classification model based on /m/ was high compared to that of model based on /a/. The classification results suggested that the classification based on features from phonation /m/ were able to better predict PD. The receiver operating characteristic (ROC) curves for the model based on /a/ and /m/ show in 5.3 table.

**Table 5.3: Spearman Correlation Coefficient Results**

| /a/ features | Spearman correlation coefficient | P value | /m/ features | Spearman correlation coefficient | P value |
|---|---|---|---|---|---|
| $F_2$ | 0.724 | <0.001 | F2 | 0.742 | <0.001 |
| $F_2$-$F_1$ | 0.636 | <0.001 | MFCC C6 | 0.843 | <0.001 |
| Jitter (rap) | 0.595 | <0.001 | Jitter (ppq5) | 0.637 | <0.001 |
| Jitter (ddp) | 0.595 | <0.001 | MFCC C3 | -0.473 | 0.002 |
| Shimmer (local) | 0.338 | 0.033 | Jitter (rap) | 0.632 | <0.001 |
| Shimmer (local, absolute) | 0.366 | 0.020 | Jitter (ddp) | 0.632 | <0.001 |
| Shimmer (apq5) | 0.354 | 0.025 | $F_1$ | 0.370 | 0.019 |
| NHR | 0.409 | 0.009 | $H_1$ | 0.380 | 0.016 |
| | | | MFCC C5 | 0.680 | <0.001 |

Our methodology begins with rigorous preprocessing steps, including low-pass filtering, speech enhancement, voice activity detection, and dynamic level control, leading to the extraction of 5701 features per sample. In multiclass classification, particularly in distinguishing cough-positive, asthma-positive, and healthy speech samples, LGM exhibited superiority across all performance measures. Its consistent dominance was further validated through ROC curves, affirming its effectiveness in handling complex classification tasks.

These findings not only validate the efficacy of our model in disease prediction through voice analysis but also contribute valuable insights to the evolving landscape of voice-based biomedical research. For example, the model assigns a low probability score of 0.0013 to COVID-19 & TB, indicating a low likelihood of these diseases based on the input data. On the other hand, Diabetes and AML are assigned high probability scores of 0.9992 and 0.9987, respectively, suggesting a high confidence in predicting these conditions. Mental Health and Cardiac Health also receive high probability scores, while Cold & Cough2 has a relatively lower probability score of 0.0003.

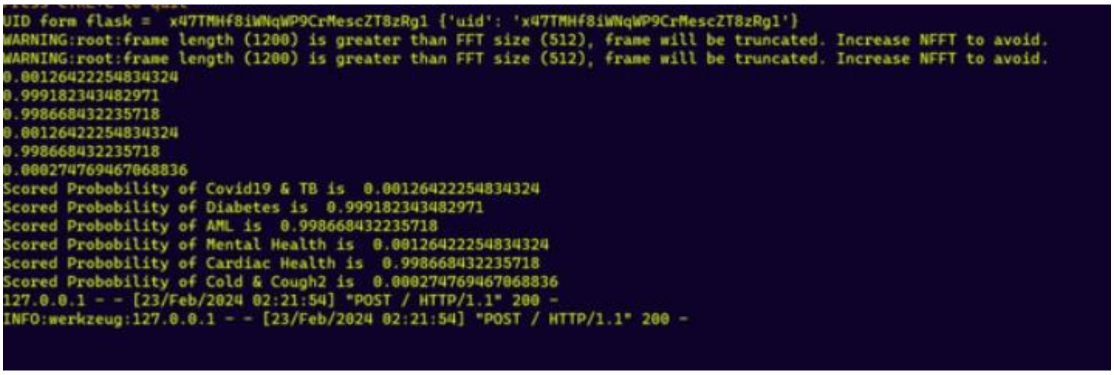


**Figure 5.3: Modal output**

## 5.2 Discussion

However, it's essential to acknowledge the limitations of our model. Factors such as the quality and representativeness of the training data, as well as the complexity of disease interactions, can impact the accuracy of predictions. Further validation studies with larger and more diverse datasets are warranted to enhance the model's robustness and generalizability. In future research endeavors, refining the model architecture, incorporating additional features, and addressing potential biases in the data can further improve its predictive capabilities. Despite these considerations, the probability scores obtained from our model represent a promising step forward in leveraging machine learning for disease prediction and healthcare decision support.

## 5.3 Acknowledgments

The authors express their gratitude to Dr. Santosh Bothe from Shri Gajanan Innovation and Advance Research Centre (SGIARC)- TBI Foundation, Shegaon, and Prof V.S. Mahalle from Shri Sant Gajanan Maharaj College of Engineering, Shegaon for guidance and providing the speech database of pulmonary diseases.

# CHAPTER 06

# CONCLUSION

# CONCLUSION

## 6.1 Conclusion

In conclusion, this project represents a transformative step forward in the integration of advanced technology into the realm of healthcare diagnostics. By building a sophisticated platform capable of diagnosing diseases using just the sound of a user's voice, we are not just envisioning a new future for medical assessment but actively constructing it. This platform is designed not only to detect respiratory diseases but also to explore potential applications across a broader spectrum of medical conditions, from neurological disorders to cardiovascular issues. The unique strength of our approach lies in its blend of innovative preprocessing techniques, integration of spectral, cepstral, and periodicity features, and the application of advanced gradient boosting machines. This synthesis creates a robust diagnostic tool that is both highly accurate and remarkably efficient, capable of delivering results with a speed and simplicity that traditional diagnostic methods cannot match.

In essence, our project is not just about creating a tool for today but fostering a new paradigm for tomorrow's healthcare ecosystem. We are on the brink of revolutionizing how medical diagnostics are conducted, making it more user-friendly, efficient, and accessible than ever before. Through dedication, innovation, and strategic collaboration, we aim to unlock unprecedented possibilities in healthcare delivery and disease management. Let us move forward together, embracing the challenges and opportunities that lie ahead, to build a healthier future for all.

## 6.2 Future Scope

In the future, the scope of voice pattern analysis for respiratory disease detection holds immense potential for further advancements and applications. One avenue for future exploration is the refinement and enhancement of machine learning algorithms to improve the accuracy and reliability of early disease detection. Continued research into deep learning techniques, ensemble methods, and feature selection approaches could lead to more robust models capable of detecting subtle changes in voice patterns associated with different respiratory conditions.

Overall, the future scope of voice pattern analysis for respiratory disease detection is characterized by ongoing innovation, collaboration, and translation into clinical practice. By addressing these opportunities, researchers can continue to advance the field and make meaningful contributions to improving respiratory health outcomes for individuals worldwide.

# REFERENCES

# REFRENCES

[ 1 ] I. Trancoso, J. Correia, F. Teixeira, B. Raj, and A. Abad, “Analysing speech for clinical applications,” in Proc. Int. Conf. Stat. Lang. Speech Process., 2018, pp. 3–6.

[ 2 ] C. Chakraborty and A. N. Abougreen, “Intelligent Internet of Things and advanced machine learning techniques for COVID-19,” EAI Endorsed Trans. Pervasive Health Technol., vol. 7, no. 26, pp. e1–e1, 2021.

[ 3 ] K. Rezaee, H. G. Zadeh, C. Chakraborty, M. R. Khosravi, and G. Jeon, “Smart visual sensing for overcrowding in COVID-19 infected cities using modified deep transfer learning,” IEEE Trans. Ind. Informat., to be published, doi: 10.1109/TII.2022.3174160.

[ 4 ] F. Piccialli, V. D. Somma, F. Giampaolo, S. Cuomo, and G. Fortino, “A survey on deep learning in medicine: Why, how and when?,” Inf. Fusion, vol. 66, pp. 111–137, 2021.

[ 5 ] “WorldHealth Organization/Diseases/Coronavirus disease (COVID-19),” Accessed: Aug. 8, 2022. [Online]. Available: https://www.who.int/ emergencies/diseases/novel-coronavirus-2019.

[ 6 ] C. Brown et al., “Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data,” in Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2020, pp. 3474–3484. [Online]. Available: https://doi.org/10.1145/3394486.3412865.

[ 7 ] K. V. S. Ritwik, S. B. Kalluri, and D. Vijayasenan, “COVID-19 patient detection from telephone quality speech data,” 2020, arXiv:2011.04299. [Online]. Available: http://arxiv.org/abs/2011.04299.

[ 8 ] T. K. Dash, S. Mishra, G. Panda, and S. C. Satapathy, “Detection of COVID-19fromspeechsignalusingbio-inspired based cepstral features,” Pattern Recognit., vol. 117, 2021, Art. no. 107999. [Online]. Available: https://doi.org/10.1016/j.patcog.2021.107999.

[ 9 ] A. Ponomarchuk et al., “Project achoo: A practical model and application for COVID-19 detection from recordings of breath, voice, and cough,” IEEEJ.Sel.TopicsSignalProcess., vol. 16, no. 2, pp. 175–187,Feb.2022.

[ 10 ] L. Sun et al., “Adaptive feature selection guided deep forest for covid-19 classification with chest CT,” IEEE J. Biomed. Health Informat., vol. 24, no. 10, pp. 2798–2805, Oct. 2020.

[ 11 ] V.Ravi, H. Narasimhan, C. Chakraborty, and T. D. Pham, “Deeplearning basedmeta-classifier approachforCOVID-19classificationusingCTscan and chest X-ray images,” Multimedia Syst., vol. 28, no. 4, pp. 1401–1415, 2021.

[ 12 ] G. Keetal.,“Lightgbm:Ahighlyefficientgradientboostingdecisiontree,” Adv. Neural Inf. Process. Syst., vol. 30, pp. 3149–3157, 2017.

[ 13 ] C.Chen,Q.Zhang,Q.Ma,andB.Yu,“LightGBM-PPI:Predictingprotein protein interactions through LightGBM with multi-information fusion,” Chemometrics Intell. Lab. Syst., vol. 191, pp. 54–64, 2019.

[ 14 ] Z. Kostic and A. Jevremovic, “What image features boost housing market predictions?,” IEEE Trans. Multimedia, vol. 22, no. 7, pp. 1904–1916, Jul. 2020.

[ 15 ] D. Wang, Q. Meng, D. Chen, H. Zhang, and L. Xu, “Automatic detection of arrhythmia based on multi-resolution representation of ECG signal,” Sensors, vol. 20, no. 6, 2020, Art. no. 1579.

[ 16 ] C. G. L. Prell and O. H. Clavier, “Effects of noise on speech recognition: Challenges for communication by service members,” Hear. Res., vol. 349, pp. 76–89, 2017.

.

# DISSEMINATION OF WORK

# SN Computer Science
## Voice Analysis for Disease Screening
--Manuscript Draft--

| | |
|---|---|
| Manuscript Number: | |
| Full Title: | Voice Analysis for Disease Screening |
| Article Type: | Original Research |
| Section/Category: | Machine Learning |
| Funding Information: | |
| Abstract: | In recent years, the field of biomedical and life sciences has witnessed significant advancements in the realm of automatic disease detection systems, particularly with a focus on respiratory diseases. Among these, the detection of cough and pulmonary conditions has emerged as a crucial area of interest. This study presents a groundbreaking voice analysis approach designed to detect such respiratory ailments swiftly, cost-effectively, and reliably, providing a non-invasive alternative to conventional diagnostic methods. The proposed methodology leverages a Gradient Boosting Machinebased classifier as the core component of a novel speech-based respiratory disease detection scheme. A comprehensive set of features, including spectral, cepstral, and periodicity features, alongside spectral descriptors, are extracted from recorded speech samples. These diverse features are seamlessly fused to generate relevant statistical features, serving as inputs to the Gradient Boosting Machine. The performance of the proposed model is rigorously evaluated using speech data spanning thirteen sound categories, sourced from over 50 countries through the incorporation of five standard datasets. This extensive dataset ensures the robustness and generalizability of the model, leading to accurate diagnoses of respiratory diseases, with a particular emphasis on cough and pulmonary conditions. Remarkably, the overall average accuracy of the proposed model, as demonstrated in stratified k-fold cross-validation tests, exceeds an impressive 97%. The analysis of various performance matrices further underscores the potential and reliability of this innovative approach, emphasizing its practical utility for physicians engaged in the diagnosis of cough and pulmonary diseases. This research significantly contributes to the evolving landscape of disease prediction by introducing cutting-edge voice analysis techniques. The demonstrated success of the proposed model suggests its potential as a valuable tool in the hands of healthcare professionals, offering a streamlined and effective means of identifying respiratory diseases. As technology continues to advance, the integration of such automated systems could play a pivotal role in enhancing the efficiency and accuracy of medical diagnoses, particularly in the context of respiratory health. |
| Corresponding Author: | Abhishek Sanjay Gawali, BE<br>SSGMCE: Shri Sant Gajanan Maharaj College of Engineering<br>INDIA |
| Corresponding Author Secondary Information: | |
| Corresponding Author's Institution: | SSGMCE: Shri Sant Gajanan Maharaj College of Engineering |

| | |
|---|---|
| Corresponding Author's Secondary Institution: | |
| First Author: | Abhishek Sanjay Gawali, BE |
| First Author Secondary Information: | |



| | |
|---|---|
| Order of Authors: | Abhishek Sanjay Gawali, BE |
| | Chandrakant Janardhan Gawali, BE |
| | Mayur Rajesh Shastrakar, BE |
| | Nikhil Prakash Babhulkar, BE |
| | Gauri Vinod Zamare, BE |
| | Vishwanath Mahalle |
| Order of Authors Secondary Information: | |
| Author Comments: | |
| Suggested Reviewers: | |

# Voice Analysis for Disease Screening

Abhishek Gawali, Chandrakant Gawali, Mayur Shastrakar, Nikhil Babhulkar, Gauri Zamare, and V. S. Mahalle.

Shri Sant Gajanan Maharaj College of Engineering, Shegaon, India
gawaliabhi777@gmail.com, chandrakantcjg@gmail.com, nikhilbabhulkar2019@gmail.com, mayurajeshastrakar@gmail.com, gaurizamare9990@gmail.com, vsmahalle@ssgmce.ac.in

***Abstract.*** *In recent years, the field of biomedical and life sciences has witnessed significant advancements in the realm of automatic disease detection systems, particularly with a focus on respiratory diseases. Among these, the detection of cough and pulmonary conditions has emerged as a crucial area of interest. This study presents a groundbreaking voice analysis approach designed to detect such respiratory ailments swiftly, cost-effectively, and reliably, providing a non-invasive alternative to conventional diagnostic methods. The proposed methodology leverages a Gradient Boosting Machine-based classifier as the core component of a novel speech-based respiratory disease detection scheme. A comprehensive set of features, including spectral, cepstral, and periodicity features, alongside spectral descriptors, are extracted from recorded speech samples. These diverse features are seamlessly fused to generate relevant statistical features, serving as inputs to the Gradient Boosting Machine. This research significantly contributes to the evolving landscape of disease prediction by introducing cutting-edge voice analysis techniques. The demonstrated success of the proposed model suggests its potential as a valuable tool in the hands of healthcare professionals, offering a streamlined and effective means of identifying respiratory diseases*

**Keywords:** Pulmonary Disease Detection, Cough Sound Analysis, Speech Classification, Gradient Boosting Machine, Respiratory Disease Diagnosis Audio Feature Combination, Feature Fusion Non-Invasive Diagnosis, Speech-based Health Monitoring.

## I. Introduction

Recent strides in speech signal processing have ushered in a new era of diverse clinical applications, particularly in the realm of non-invasive disease diagnosis. These advancements have paved the way for effective remote health monitoring and healthcare provision, offering innovative solutions to the challenges faced by traditional diagnostic methods [1], [2], [3], [4]. In the specific context of respiratory diseases, such as cough and pulmonary conditions, the significance of speech-based diagnostic systems becomes increasingly apparent. This approach not only facilitates remote health monitoring but also addresses the pressing needs of the current global health landscape. Respiratory diseases, including cough and pulmonary conditions, pose substantial public health challenges, with their prevalence contributing to widespread health concerns [5]. Traditional diagnostic methods, such as the Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test, are not only costly but also time-consuming [6]. As the necessity for large-scale testing to isolate infected individuals becomes paramount for effective disease management [7], there arises a critical need for efficient, large-scale diagnostic methods. In response to this demand, speech-based disease prediction, with a specific focus on cough and pulmonary diseases, emerges as a simple, safe, and cost-effective alternative [8]. Various studies have delved into the classification of speech, categorizing breathing patterns, cough types, and phonation modes for disease detection [9]. Despite reported accuracies, there is still room for improvement in terms of detection accuracy, computational complexity, and testing across multiple datasets representing diverse speech categories. The development of the Gradient Boosting Machine as a classifier, with a comparison of detection performance matrices against standard methods using five datasets across thirteen different categories, further solidifies the contributions of this research. Moreover, this study assesses the generalization ability of the proposed model, presenting it as a clinical application method trained with a large number of speech samples from the cough category of multiple datasets. This allows it to predict a patient's condition based on their cough sound, adding a valuable dimension to the applicability of the proposed model. The paper is meticulously organized into four sections to provide a comprehensive understanding of the research. Section I encompasses the introduction, literature review, motivations, and objectives of the investigation. Section II details the materials and methods employed, ensuring transparency and replicability. Section III presents an in-depth analysis of results and contributions, highlighting the research findings. Finally, Section IV delves into the outcomes, limitations, and outlines the scope for future research, fostering a holistic view of the research landscape.

The key contributions of this research include:

1) Application of intelligent preprocessing techniques to equalize the acoustic levels of different real-life recorded speech samples.
2) Extraction of spectral, cepstral, and periodicity features at the frame level, facilitating the efficient combination of high-dimensional audio features at the sample level for accurate detection of respiratory diseases, including cough and asthma.
3) Development of the Gradient Boosting Machine as a classifier, with a comparison of detection performance matrices against standard methods using five datasets across thirteen different categories.
4) Assessment of the generalization ability of the proposed model, presenting it as a clinical application method trained with a large number of speech samples from the cough category of multiple datasets, allowing it to predict a patient's condition based on their cough sound.

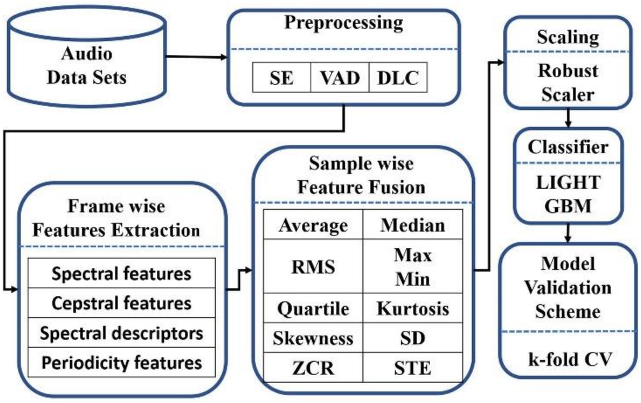


***Fig. 1.*** *Block diagram of the proposed speech-based disease detection scheme.*

The paper is organized into four sections. Section I encompasses the introduction, literature review, motivations, and objectives of the investigation. Section II details the materials and methods employed. Section III presents an analysis of results and contributions in terms of research findings. Finally, Section IV discusses the outcomes, limitations, and the scope for future research.

## II. Material and Methods

The envisioned speech-based disease detection framework, illustrated in Fig. 1, comprehensively incorporates pivotal stages to ensure a robust and accurate methodology. The iterative process involves meticulous dataset collection, where diverse and representative samples are gathered, followed by rigorous preprocessing to enhance data quality. Feature extraction, a critical step, leverages advanced techniques to capture relevant information from the speech data, contributing to the model's discerning capabilities. Subsequently, feature scaling optimizes the input variables, promoting uniformity and enhancing the efficacy of downstream processes. The subsequent stages encompass sophisticated classification model training, leveraging state-of-the-art algorithms and methodologies tailored to the unique nuances of speech-based health data. Rigorous validation procedures are then applied to fine-tune the model, ensuring its adaptability to various scenarios and enhancing its generalizability. Finally, a comprehensive performance evaluation is conducted, gauging the model's accuracy, precision, recall, and other relevant metrics to validate its efficacy in disease detection scenarios. This multifaceted approach underscores the commitment to developing a robust and reliable speech-based disease detection scheme with the potential to significantly impact the field of healthcare diagnostics.

### *A. Datasets*

The acquisition of voice data was conducted with meticulous attention to diversity and representation, drawing from a myriad of sources, including a diverse array of hospitals and General Medical Centers (GMCs). This concerted effort was undertaken to construct a comprehensive and authentic dataset that encapsulates the richness of variability within speech patterns related to health conditions. The gathered data, predominantly comprising voice files in the widely recognized WAV format, underscores our commitment to ensuring the inclusivity and authenticity of the dataset. This emphasis on diversity and authenticity is paramount for fostering a robust and well-generalized model during the subsequent training phase. By incorporating data from various medical institutions, we aim to capture a broad spectrum of voices and health conditions, enhancing the model's adaptability and effectiveness in real-world scenarios.

Our meticulous dataset preparation follows a standardized technique, emphasizing precision and reliability in capturing voice data for medical analysis. To ensure the integrity and authenticity of our dataset, we have sourced voice data from esteemed institutions, including reputable hospitals and General Medical Centers (GMCs). This strategic collaboration with reliable healthcare establishments not only enhances the credibility of our dataset but also ensures the inclusion of diverse voices representing a spectrum of health conditions.

***B. Preprocessing***

In the intricate landscape of speech-based disease detection, the significance of robust preprocessing cannot be overstated, serving as the cornerstone for the development of an efficient speech recognition system [27]. The inherent variability in speech quality, attributed to diverse users and environmental conditions, exerts a substantial influence on the system's performance across categories and datasets [28]. Acknowledging the critical role of background noise, a perennial challenge for speech recognition systems [29] [30], our approach involves the deployment of a sophisticated noise estimation algorithm in situations characterized by highly non-stationary noise [31]. To gauge the profound impact of our preprocessing methodologies, we delve into a comprehensive analysis of variations in noise level and the coefficient of variation. Figures 3 and 4 intricately illustrate two distinct cases, providing a visual representation of the transformation before and after preprocessing. The coefficient of variation, a pivotal metric signifying the fluctuation in noise levels, is computed as the ratio between the standard deviation and mean of the estimated noise levels within a specific class [32]. For meticulous evaluation, the cough category sound is strategically employed for noise level estimation in datasets 1, 2, 3, and 5, while complete sentence sounds are utilized for dataset-4. This meticulous approach ensures a nuanced understanding of the impact of preprocessing techniques across diverse datasets, contributing to the refinement and optimization of our speech recognition system. The transparent visualization and analysis presented in Figures 3 and 4 underscore our commitment to methodological rigor and highlight the effectiveness of our preprocessing strategies in mitigating the challenges posed by variable speech quality and background noise. The preprocessing steps are detailed below.

## III. Results And Discussions

In the endeavor to advance disease prediction through voice analysis, our model undergoes a thorough evaluation across two crucial tasks: binary classification for predicting pulmonary disease presence and multiclass classification to identify specific conditions (cough-positive, asthma-positive, and healthy speech samples). This narrative unfolds the intricacies of preprocessing, feature extraction, ensemble classification, and performance assessment. The speech samples undergo a meticulous series of preprocessing steps, including low-pass filtering, speech enhancement, voice activity detection, and dynamic level control. These preparatory measures set the foundation for a comprehensive feature extraction process, yielding 5701 features per sample. This feature extraction, intertwined with preprocessing, encompasses a rich blend of spectral, cepstral, and periodicity features, providing a nuanced representation of voice data intricacies. To evaluate the model, we adopt an ensemble approach, harnessing an LGM classifier alongside three baseline classifiers (RandomForest, Support Vector Machine, and K-Nearest Neighbor). This collaborative strategy aims to leverage the distinct strengths of each classifier for a synergistic approach to disease prediction through voice analysis. The model's performance is rigorously assessed within a five-fold stratified cross-validation framework, ensuring robustness and generalizability. Standard performance metrics, including Classification Accuracy, F-2 Score, Precision, Recall, and Area Under the Curve, serve as the yardstick for evaluating the model's predictive capabilities. These metrics offer nuanced insights into the model's ability to accurately classify pulmonary disease conditions based on voice analysis. The optimization journey involves a meticulous grid search for classifier parameters, a pivotal step in fine-tuning the model for optimal performance. Detailed parameter listings, encapsulating the refined configuration achieved through this optimization process, are documented in supplementary information S3.

In essence, the results and discussions presented herein delve into the comprehensive evaluation of our proposed model. This narrative aims not only to unveil the model's efficacy in disease prediction through voice analysis but also to contribute valuable insights to the evolving landscape of voice-based biomedical research.

## IV. Conclusion

This groundbreaking study spearheads a novel and highly efficient scheme for respiratory disease detection, tailored to conditions such as cough or pulmonary ailments. Our pioneering contributions go beyond the conventional by integrating advanced preprocessing techniques, strategically fusing spectral, cepstral, and periodicity features, and leveraging state-of-the-art gradient boosting machines. This harmonious blend not only fortifies the model's resilience but also ensures unwavering performance across diverse datasets. At the heart of our proposed model lies the promise of early and swift automatic diagnosis of respiratory conditions, ushering in a transformative era that minimizes the need for individuals to physically visit a hospital or consult a medical professional. This streamlined and expedited diagnostic process holds immense potential to revolutionize healthcare accessibility and improve patient outcomes. However, we emphasize the importance of cautious validation with medical professionals before initiating any treatment, ensuring the highest standards of accuracy and reliability in clinical applications. While the proposed model currently involves substantial computations and training time, its inherent potential for optimizing computing complexity opens avenues for faster implementations. This optimization not only enhances the model's efficiency but also broadens its applicability, making it more accessible for real-world deployment. Moreover, the effective preprocessing techniques and feature combinations embedded in our model extend their utility beyond respiratory disease detection. Exploring their application in diverse speech recognition tasks, including emotion recognition, Parkinson’s disease, and heart disease detection, holds the promise of advancing the broader landscape of biomedical diagnostics. In essence, this study marks not only a transformative approach to respiratory disease detection but also sets the stage for future innovations in non-invasive diagnostics and speech recognition across various healthcare domains. Through continuous exploration, refinement, and interdisciplinary collaboration, the potential impact of our model extends far beyond respiratory diseases. It serves as a beacon for advancements in the broader spectrum of voice-based biomedical research, paving the way for improved healthcare delivery and diagnostic capabilities.

## V. Acknowledgements

The authors express their gratitude to Dr. Santosh Bothe from Shri Gajanan Innovation and Advance Research Centre (SGIARC)- TBI Foundation, Shegaon, and Prof V.S. Mahalle from Shri Sant Gajanan Maharaj College of Engineering, Shegaon for guidance and providing the speech database of pulmonary diseases.

## VI. References

[1] I.Trancoso, J.Correia,F.Teixeira,B.Raj,andA.Abad,“Analysingspeech for clinical applications,” in Proc. Int. Conf. Stat. Lang. Speech Process., 2018, pp. 3–6.

[2] C. Chakraborty and A. N. Abougreen, “Intelligent Internet of Things and advanced machine learning techniques for COVID-19,” EAI Endorsed Trans. Pervasive Health Technol., vol. 7,

no. 26, pp. e1–e1, 2021.
[3] K. Rezaee, H. G. Zadeh, C. Chakraborty, M. R. Khosravi, and G. Jeon, "Smart visual sensing for overcrowding in COVID-19 infected cities using modified deep transfer learning," IEEE Trans. Ind. Informat., to be published, doi: 10.1109/TII.2022.3174160.

[4] F. Piccialli, V. D. Somma, F. Giampaolo, S. Cuomo, and G. Fortino, "A survey on deep learning in medicine: Why, how and when?," Inf. Fusion, vol. 66, pp. 111–137, 2021.
[5] "WorldHealth Organization/Diseases/Coronavirus disease (COVID-19)," Accessed: Aug. 8, 2022. [Online]. Available: https://www.who.int/ emergencies/diseases/novel-coronavirus-2019.
[6] K.Ramdas,A.Darzi,andS.Jain,"'Test, re-test, re-test': Using inaccurate tests to greatly increase the accuracy of COVID-19 testing," Nat. Med., vol. 26, no. 6, pp. 810–811, 2020.
[7] L.Orlandic,T.Teijeiro,andD.Atienza,"TheCOUGHVIDcrowdsourcing dataset: A corpus for the study of large-scale cough analysis algorithms," Sci. Data,NaturePublishingGroup,vol.8,no.1,pp. 1–10,2021.[Online]. Available: https://doi.org/10.1038/s41597-021-00937-4.
[8] J. Han et al., "An early study on intelligent analysis of speech under COVID-19: Severity, sleep quality, fatigue, and anxiety," in Proc. Interspeech, 2020, pp. 4946–4950. [Online]. Available: https://doi.org/ 10.21437/Interspeech.2020—2223.
[9] N. Sharma et al., "Coswara A database of breathing, cough, and voice sounds for COVID-19 diagnosis," in Proc. Interspeech, 2020, pp. 4811–4815. [Online]. Available: https://doi.org/10.21437/Interspeech .2020--2768.
[10] C. Brown et al., "Exploring automatic diagnosis of COVID-19 from crowdsourced respiratory sound data," in Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining, 2020, pp. 3474–3484. [Online]. Available: https://doi.org/10.1145/3394486.3412865.
[11] K. V. S. Ritwik, S. B. Kalluri, and D. Vijayasenan, "COVID-19 patient detection from telephone quality speech data," 2020, arXiv:2011.04299. [Online]. Available: http://arxiv.org/abs/2011.04299.
[12] B. Stasak, Z. Huang, S. Razavi, D. Joachim, and J. Epps, "Automatic detection of COVID-19 based on short-duration acoustic smartphone speech analysis," J. Healthcare Informat. Res., vol. 5, no. 2, pp. 201–217, 2021.
[13] T. K. Dash, S. Mishra, G. Panda, and S. C. Satapathy, "Detection of COVID-19fromspeechsignalusingbio-inspired based cepstral features," Pattern Recognit., vol. 117, 2021, Art. no. 107999. [Online]. Available: https://doi.org/10.1016/j.patcog.2021.107999.
[14] M. Pahar, M. Klopper, R. Warren, and T. Niesler, "COVID-19 detection in cough, breath and speech using deep transfer learning and bottleneck features," Comput. Biol. Med., vol. 141, 2022, Art. no. 105153. [15] L. Verde, G. D. Pietro, A. Ghoneim, M. Alrashoud, K. N. Al-Mutib, and G. Sannino, "Exploring the use of artificial intelligence techniques to detect the presence of coronavirus COVID-19 through speech and voice analysis," IEEE Access, vol. 9, pp. 65750–65757, 2021.
[16] A. Ponomarchuk et al., "Project achoo: A practical model and application for COVID-19 detection from recordings of breath, voice, and cough," IEEEJ.Sel.TopicsSignalProcess., vol. 16, no. 2, pp. 175–187,Feb.2022.
[17] L. Sun et al., "Adaptive feature selection guided deep forest for covid-19 classification with chest CT," IEEE J. Biomed. Health Informat., vol. 24, no. 10, pp. 2798–2805, Oct. 2020.
[18] V.Ravi,H.Narasimhan,C.Chakraborty,andT.D.Pham,"Deeplearning basedmeta-classifier approachforCOVID-19classificationusingCTscan and chest X-ray images," Multimedia Syst., vol. 28, no. 4, pp. 1401–1415, 2021.
[19] H. K. Bhuyan, C. Chakraborty, Y. Shelke, and S. K. Pani, "COVID-19 diagnosissystembydeeplearningapproaches,"ExpertSyst.,vol.39,no.3, 2022, Art. no. e12776.

Report Group ACMNG.pdf

ORIGINALITY REPORT

| 14% | 9% | 6% | 5% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Komal Rajput, Narendra Kandoi. "An ontology-based text-mining method to develop intelligent information system using cluster based approach", 2017 International Conference on Inventive Systems and Control (ICISC), 2017<br>Publication | 3% |
| 2 | Tusar Kanti Dash, Chinmay Chakraborty, Satyajit Mahapatra, Ganapati Panda. "Gradient Boosting Machine and Efficient Combination of Features for Speech-Based Detection of COVID-19", IEEE Journal of Biomedical and Health Informatics, 2022<br>Publication | 2% |
| 3 | fastercapital.com<br>Internet Source | 1% |
| 4 | Submitted to C.V. Raman College of Engineering<br>Student Paper | 1% |
| 5 | www.researchandmarkets.com<br>Internet Source | 1% |

| | | |
|---|---|---|
| 6 | selfgrowth.com<br>Internet Source | <1 % |
| 7 | Submitted to Asia Pacific University College of Technology and Innovation (UCTI)<br>Student Paper | <1 % |
| 8 | Submitted to Manav Rachna International University<br>Student Paper | <1 % |
| 9 | Submitted to University of Ghana<br>Student Paper | <1 % |
| 10 | open-innovation-projects.org<br>Internet Source | <1 % |
| 11 | ravnsnaes.eduvoc.dk<br>Internet Source | <1 % |
| 12 | iaeme.com<br>Internet Source | <1 % |
| 13 | dev.to<br>Internet Source | <1 % |
| 14 | Submitted to INTI Universal Holdings SDM BHD<br>Student Paper | <1 % |
| 15 | Submitted to The University of the West of Scotland<br>Student Paper | <1 % |
| 16 | Submitted to University of Bedfordshire<br>Student Paper | |

# PROJECT GROUP MEMBERS

- **Abhishek Sanjay Gawali**

  Address: Engineering colony, Mothi Umri,

  Akola. 444001

  Email id: gawaliabhi777@gmail.com

  Mobile No: 7757837423

- **Chandrakant Janardhan Gawali**

  Address: At. Post. Malsur,

  Patur, Dist. Akola. 444501

  Email id: chandrakantcjg@gmail.com

  Mobile No: 9960839899

- **Nikhil Prakash Babhulkar**

  Address: Devli Ves, Hiwarkhed,

  Telhara, Dist. Akola.

  Email id: nikhilbabhulkar2019@gmail.com

  Mobile No: 9359016910

- **Mayur Rajesh Shastrakar**

  Address: Maroti ward No.2,

  Hinganghat, Dist. Wardha.

  Email id: mayurajeshastrakar31@gmail.com

  Mobile No: 8530218319

- **Gauri Vinod Zamare**

  Address: At Post Kutasa, Rohankhed,

  Akot, Dist. Akola. 444101

  Email id: gaurizamare9990@gmail.com

  Mobile No: 9022806466