WEB BASED AUTOMATIC TUBERCULOSIS COUGH DETECTION APPLICATION DEVELOPMENT

A MAJOR PROJECT REPORT

Submitted by

MEDA ANIL KUMAR

GUNDU BHARADWAJ

ATLA LOKESH

Under the Guidance of

DR D.SUBITHA

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

ELECTRONICS & COMMUNICATION ENGINEERING

Vel Tech
Rangarajan Dr. Sagunthala
R&D Institute of Science and Technology
(Deemed to be University Estd. u/s 3 of UGC Act, 1956)

MAY 2023

BONAFIDE CERTIFICATE

Certified that this major project report entitled **"WEB BASED AUTOMATIC TUBERCULOSIS COUGH DETECTION APPLICATION DEVELOPMENT"** is the bonafide work of **"MEDA ANIL KUMAR (19UEEC0327), GUNDU BHARADWAJ (19UEEC0175) and ATLA LOKESH (19UEEC0033)"** who carried out the project work under my supervision.

SUPERVISOR	**HEAD OF THE DEPARTMENT**
DR D.SUBITHA	**Dr.P.ESTHER RANI**
Associate Professor	Professor
Department of ECE	Department of ECE

Submitted for major project work viva-voce examination held on:_____________________

INTERNAL EXAMINER	**EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

We express our deepest gratitude and sincere thanks to our respected Founder President and Chancellor **Col. Prof. Dr. R. RANGARAJAN** , Foundress President **Dr. SAGUNTHALA RANGARAJAN** , Managing Trustee and Vice President.

We are very much grateful to our beloved Vice Chancellor **Prof. Dr. S. SALIVAHANAN**, for providing with an environment to complete our project successfully.

We are obligated to our beloved Registrar **Dr. E. KANNAN**, for providing immense support in all our endeavors.

We are thankful to our esteemed Dean of Academics **Dr. A.T.RAVICHANDRAN**, for providing a wonderful environment to complete work successfully.

We are extremely thankful and pay our gratitude to our Dean SOEC **Dr.R.S. VALARMATHI**, for her valuable guidance and support on completion of this major project.

It is a great pleasure for us to acknowledge the assistance and contribution of our Head of the Department **Dr. P. ESTHER RANI**, for her assistance and useful suggestions which helped us in completing the project, in time and we thank her for being instrumental in the completion of final year(7th sem) with her encouragement and unwavering support during the entire course of this project.

We are extremely thankful and pay gratitude to our project supervisor **DR. D.SUBITHA**, Associate Professor for her valuable guidance and support on completing this Major project.

MEDA ANIL KUMAR

GUNDU BHARADWAJ

ATLA LOKESH

TABLE OF CONTENTS

ABSTRACT

The diagnosis of tuberculosis and maintaining data on the disease now depend on quick, accurate non-invasive testing. Modern machine learning (ML) and deep learning (DL) techniques that employ characteristics of coughing noises for TB diagnosis are motivated by recent investigations. In this study, we present system concepts for TB cough detection, with the long-term goal of developing an application for user-friendliness. The sounds of the cough convey important information about the glottis' functioning under various respiratory pathological situations. As a result, respiratory illnesses like COVID-19 can be recognised by the peculiarities of cough sounds. Our suggested model is trained on a dataset of TB coughs and is based on convolution neural networks. Finally, we look into the application of several types of algorithms and evaluate their accuracy. Our findings show high performance in detection of virus.

Keywords: TB detection, machine learning, deep learning, respiratory diagnosis, healthcare, artificial intelligence

LIST OF FIGURES

LIST OF ABBREVIATIONS

CNN	-	Convolution Neural Network
ANN	-	Artificial Neural Network
RNN	-	Recurrent Neural Network
DL	-	Deep Learning

CHAPTER 1

INTRODUCTION

1.1 OBJECTIVE

Tuberculosis (TB) is one of the most common infectious diseases in the world and is one of the leading causes of morbidity and mortality. In addition to being the most frequent infectious agent-related cause of death, exceeding HIV/AIDS, tuberculosis (TB) is one of the top 10 global causes of mortality according to the World Health Organization (WHO). In 2020, it is anticipated that 10 million new cases of TB will be reported, and 1.4 million people will die from the disease.

Figure 1.1: Tuberculosis

The bacteria Mycobacterium tuberculosis[1.1], which mostly affects the lungs, is the cause of TB, an airborne illness. A chronic cough that lasts for weeks or even months is the most typical sign of TB. The prevention of the disease's spread and the reduction of morbidity and mortality depend on early TB testing and prompt treatment. Unfortunately, TB detection is frequently difficult, especially in areas with limited resources and access to diagnostic instruments.

The use of digital technologies to enhance TB detection and control has garnered more attention in recent years. Machine learning (ML) and artificial intelligence (AI) have in particular showed promise in TB detection from cough sounds. Compared to conventional diagnostic techniques, cough-based TB detection systems provide a number of benefits, including non-invasiveness, cheap cost, and usability.

As a result, the goal of this project is to create an automatic TB cough detection tool for the web that can quickly and effectively identify TB from cough sounds. The application will analyse cough noises and find suspected TB patients using ML and AI approaches. The application could enhance early identification of the disease and ultimately lower TB morbidity and mortality by offering a straightforward, affordable, and non-invasive method of TB detection. The application's web-based design will also make a wider range of users, such as medical professionals, researchers, and patients, able to use it.

A more thorough approach to TB control and prevention will be possible thanks to the TB cough detection application's web-based design. The programme can be used as a screening tool by medical practitioners for individuals who appear with a persistent cough, enabling early TB detection and prompt treatment. In order to enhance TB detection algorithms and create more precise and trustworthy diagnostic tools, researchers can utilise the programme to gather data on cough sounds from various populations.

The web-based tool also gives patients the ability to record and upload their cough sounds for study, encouraging them to take an active role in their own health. By encouraging users to get checked out if they think they may have TB, this feature can help spread awareness of the condition. The creation of a web-based tool for automatically detecting TB coughs has the power to transform global efforts in TB prevention and control. The application could contribute to lowering the expense of the disease and enhancing the health of those who are afflicted by it by offering a simple, non-invasive, and accessible method of TB detection. As a result of this project, it is envisaged that more research and development into digital health technologies would be stimulated for TB control and prevention.

1.2 DIFFERENT TYPES OF TUBERCULOSIS DETECTION

The greatest infectious cause of death prior to the COVID-19 pandemic was tuberculosis (TB), which resulted in approximately 10.0 million new infections and 1.4 million fatalities globally in 20191. Since COVID-192 was given primacy over COVID-19 in terms of resources and tools utilised to diagnose and treat TB, the COVID-19 pandemic and lockdowns have negatively impacted global TB programmes. To stop the spread of the disease and decrease the effect of COVID-19 on TB management, it is crucial to deploy cutting-edge technology and concepts to enhance TB prevention and care.

TB is an infectious disease brought on by inhaling droplets containing the bacteria Mycobacterium tuberculosis. Latent tuberculosis, commonly referred to as asymptomatic tuberculosis, is not contagious, in contrast to active tuberculosis, which is contagious and symptomatic4. Between these two extremes of TB exist subclinical forms, in which people are assumed to be asymptomatic but yet have the potential to infect others.

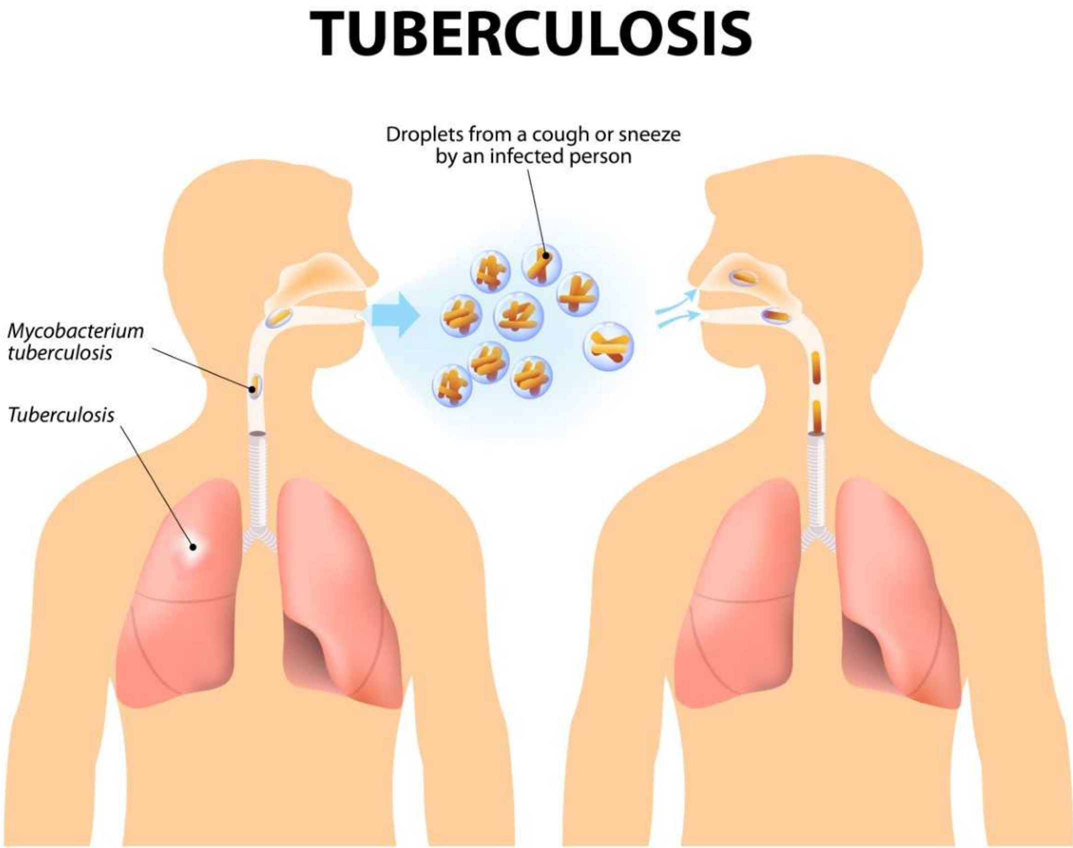

Figure 1.2: Tuberculosis

About 15-20% of active TB disease also affects other body organs, such as the lymph nodes, the abdomen, the meninges, the eyes, and the neurological system, to name a few5. Active TB disease most frequently affects the lungs (pulmonary TB), but it can also affect other regions of the body. Extrapulmonary TB (EPTB) is the name given to TB that exists outside of the lung. The most popular method for finding active pulmonary TB is microbiological testing on lung mucus (sputum) samples. The gold standard for TB testing is sputum culture. It costs money, takes time, and necessitates using centralised biosafety labs. Sputum smear microscopy is a common less expensive alternative in basic care settings in countries with limited resources, despite its low sensitivity. identifying drug resistance7.GeneXpert PCR equipment is an example of a smear-replacement technology that offers more sensitive TB detection and speedier turnaround times. 8,9 Reference standards like smear microscopy, culture, and GeneXpert are frequently used to assess the efficacy and precision of more contemporary diagnoses. Although active tuberculosis can be treated, the lengthy regimens (6 months for drug-susceptible TB) and side effects brought on by the antibiotics used to treat it make it more difficult and raise the possibility of drug resistance developing. Coughing is a typical TB symptom, therefore it can aid in TB diagnosis and assessment of treatment effectiveness. This Viewpoint examines how improvements in acoustic epidemiology and AI-based algorithms can enhance the evaluation of cough. utilised for tuberculosis detection and treatment (TB).

The fact that coughing is both a symptom of and a defence against respiratory infections makes it a complicated physiological phenomena. Cough is evaluated clinically as a defining symptom of pulmonary TB at several phases of TB care, including as a triage tool to begin TB testing or to track therapy response. The quantity of M. tuberculosis in the lungs affects cough patterns, and effective TB treatment typically results in coughs regressing....

The length of the cough and symptoms are often used in TB screening programmes to determine when TB testing is essential, although this method lacks sensitivity. The only strategy for identifying TB cases is symptom-based screening because triage technology like chest X-rays are not available in low-resource settings, peripheral health clinics, and communities. Anybody who exhibits TB-compatible symptoms, such as a protracted cough (often characterised as a cough lasting two weeks or longer)16, is advised to get tested. According to the WHO TB screening guidelines for 2021, the WHO community-based triage test target product profile (TPP) has a sensitivity of over 90% whereas the sensitivity of persistent cough alone is just 42

Both individuals and medical professionals have trouble describing the signs and symptoms of their coughs. Individuals may have trouble remembering how long their symptoms lasted, and the severity of their symptoms might change18,19. The routine translation of this information-rich symptom into subjective binary data prevents full clinical use and interpretation of cough data (e.g., cough vs no cough, chronic versus acute, getting better versus getting worse).It is possible to enhance patient management and clinical outcomes at various stages throughout the cascade of TB care by making cough an objective and quantifiable component of TB care, either by training people to recognise abnormal cough patterns or by using artificial intelligence (AI) technology (using computer systems to recognise and interpret the implications of a cough sound)20 to differentiate types of coughs.

Recent developments in this area have sparked the creation of screening techniques that use machine learning to accurately and reliably identify TB coughs[2]. These programmes potentially revolutionise both attempts to control the spread of the TB illness and the treatment of individual patients. A potential method for determining whether patients need to undergo pricey molecular downstream testing is automatic cough categorization for TB screening in a real-world setting.

In an effort to transform subjective cough reporting into objective data, the intensity of coughs with various aetiologies have been gathered and examined using questionnaire-based procedures and scales. These tools include, among others, cough diaries21, the visual analogue scale (VAS), and the cough symptom score (CSS). Based on how a patient feels about their cough, the VAS and CSS both make an effort to gauge how intense their coughing is. There are several different kinds of cough diaries, but they all ask the patient to record the frequency and intensity of their coughs over time. By questioning participants about their quality of life in connection to their health, several surveys gauge cough. For instance, a self-report assessment called the Leicester Cough Questionnaire (LCQ)that has been validated to evaluate the quality of life of those with a chronic cough, can be used to make this assessment. Moreover, it has been applied to evaluate TB patient cohorts receiving anti-TB treatment. Despite being straightforward to use and implement in clinical settings, these tools are nevertheless subject to bias based on the user's perception of their own health and attention to symptoms, which ultimately restricts the therapeutic applicability of such tools.

1.3 MACHINE LEARNING APPLICATION IN TB DETECTION

Machine learning methods are used in the web-based automatic TB cough detection application[1.2] to find TB coughs. The application will be developed using the Python programming language and the Flask web framework. The application's user-friendly layout will allow users to upload audio samples of cough noises for analysis. Pre-processing will be applied to the audio files to remove extraneous sounds and background noise. The pre-processed audio samples will be analysed using a machine learning system that has been trained on a dataset of TB cough sounds.The computer will categorise the cough sounds as TB or non-TB based on their acoustic properties. The results of the analysis will be displayed on the user interface, indicating whether the cough sound is suggestive of TB or not. The programme will also provide information on the sensitivity and specificity of the classification method. Also, the programme will contain a feedback option that would allow users to report any classification errors or research inconsistencies.

The use of recording equipment and computer-assisted acoustic interpretation techniques enhances the objectivity of cough analysis. Loudon and Spohn began recording and counting the nighttime coughs of TB patients in the 1960s25 using tape recorders. Early ambulatory cough metres contained electromyogram (EMG) electrodes26, which recorded the user's coughing as well as chest spasms, and audio recorders for audio recording. A self-contained cough monitor was created and tested by Paul et al. in 2006 utilising a CompactFlash memory card and an accelerometer (to detect vibrations associated with coughing). This device was placed in the patient's suprasternal notch, or jugular notch, and it showed good agreement with the coughing captured on video. Systems for 24-hour recording that are increasingly sophisticated have been created throughout time. These devices generally include a microphone that communicates the patient's cough sounds to a digital sound recorder, usually fastened to the patient's hip28. The microphone can be a free-field necklace

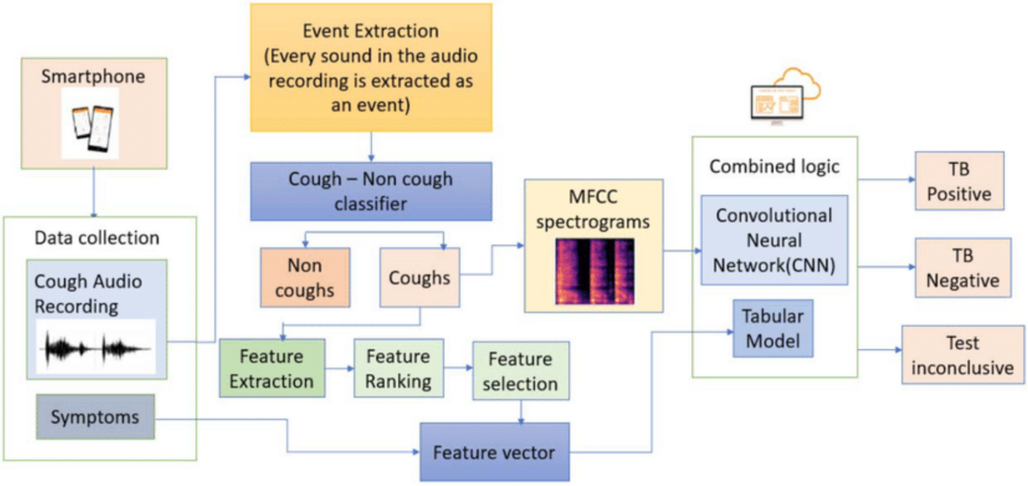

Figure 1.3: Tuberculosis detection

microphone or one that fastens to the patient's lapel. These recording devices include the VitaloJak, the Leicester Cough Monitor (LCM), and the Cayetano Cough Monitor (CayeCoM).

Cough counts and patterns were the initial objective indicators used to evaluate cough intensity and temporal change. For measuring cough frequency, the LCM, CayeCoM, and VitaloJak have all demonstrated validity29,30,31. The LCM and VitaloJak are now the most widely used cough monitoring devices, with stated cough detection sensitivities of 91 and ¿99

While continuous cough recording has been made possible by ambulatory recording devices, many of the current models are large and intrusive. Cough is a noticeable and stigmatising symptom, and its prevalence among TB patients has significantly increased this stigma35. Effective monitoring of people who cough requires covert recording techniques to prevent contributing to the stigmatisation of respiratory illnesses. Monitoring TB coughs can be done covertly using smartphones that include cough recognition and recording apps. There are already a number of cough recording programmes available, such as Hyfe Research, AI4COVID-19, and ResAppDx..
Especially in resource-constrained regions with limited access to healthcare institutions, the creation of this web-based automatic tuberculosis cough detection programme offers enormous potential for enhancing early detection and treatment of TB. Healthcare professionals, patients, and carers can all use this programme to conduct non-invasive TB screenings. It can also be used as a tool for patients with active TB to track their progress through treatment and identify relapses.

Data collection, data preprocessing, feature extraction, model training, model evaluation, and deployment are some of the processes that go into the development of this programme [5]. The process of gathering data requires gathering a sizable dataset of cough sounds from both healthy people and patients with proven diagnoses of tuberculosis. Data preparation involves removing background noise and other unimportant noises from the raw audio data by cleaning and filtering it. In feature extraction, pertinent acoustic features from the preprocessed audio data are taken out and applied to categorization.

Creating a machine learning algorithm that can distinguish between TB coughs and non-TB coughs based on their acoustic characteristics is known as "model training." Testing the effectiveness of the created algorithm using a separate dataset of cough sounds is known as model evaluation. The proposed algorithm is deployed via being incorporated into a publicly accessible web-based platform.

Numerous cough monitoring algorithms have been developed as a result of advances in machine learning, a field of artificial intelligence that enables machines to use algorithms on data to automatically "learn" and make judgements on their own 20. These algorithms can be used with digital recording devices, including smartphones. With the use of this cutting-edge technology, cough sounds may be studied in terms of frequency and kind. For instance, some systems first transform audio recordings into spectrograms, which are pictures that show how loud a noise is at various frequencies and times. They then use an algorithm to the spectrogram to visually evaluate the cough's characteristics.

These algorithms are being trained to discriminate coughs from patients with different clinical states or at different stages of disease as well as to recognise human coughs from ambient noises, however the latter use case has not yet been validated39,40,41,42,43 (cough detection). Many early cough categorization methods have been developed using COVID-19 and TB. A classification algorithm was reported to detect COVID-19 infections among people with a cough with 98 sensitivity and 94 specificity based on a sample of 5320 people, of whom 50

When compared to a composite reference standard of sputum smear microscopy, GeneXpert, and chest X-ray, TimBre, a TB screening tool, uses machine learning to identify TB coughs from a sample of 5 bacteriologically positive and 469 bacteriologically negative individuals44 with a sensitivity of 80 and specificity of 92. Another study by Botha et al. developed a cough-based screening system that could distinguish cough sounds produced by 16 people with TB from those produced by 35 people with other lung diseases with 93 sensitivity and 95 specificity against a bacteriological (laboratory method not specified) reference standard, meeting the WHO's TPP requirements of 90 sensitivity and 70 specificity for a community-based TB triage test3,45.Also, an AI system for classifying TB cough was developed using a sample of 17 TB patients and 21 healthy individuals. When compared to a sputum culture reference standard46, this algorithm has a sensitivity of 95, a specificity of 72, and an accuracy of 78. These early studies suggest that digital cough monitoring, including the identification and classification of cough events, may aid in TB screening (see Supplementary Table 1). Nonetheless, further analysis and testing are required to develop the field..

The qualities of the training dataset affect how accurate these AI systems are. The sample sizes utilised to evaluate these algorithms have been quite small47, and there has only been a limited amount of external validation of different AI algorithms conducted. Early diagnostic investigations of novel tests, such as AI algorithms, also have a tendency to overstate the diagnosis accuracy, mostly as a result of the preferred exclusion of more complex cases48. The therapeutic use of these AI algorithms will remain constrained until more replication studies utilising sizable, diverse cough datasets, representative of various demographics, have been done..

Health systems and TB initiatives must efficiently use their limited resources to detect people who have TB or who display TB symptoms. As part of a syndromic monitoring strategy, people who are at risk of contracting TB or who have already had TB should passively and prospectively monitor their cough (i.e., detection and aggregation of individual and population health indicators, such as symptoms, prior to establishing a definitive diagnosis). Hence, case-finding efforts and high-risk locations may be targeted using aggregates of cough episodes over time and place. Spatiotemporal variations in cough frequency can be used as a surrogate measure of COVID-19, TB, or other respiratory illness prevalence at the community level36.It should be investigated to see if devoting public health resources just to investigating such cough clusters might speed up the identification of further prevalent cases and improve sickness case notifications. This cough surveillance analytic approach might identify cough hotspots where the risk of TB transmission has previously been, and likely still is, much greater by concentrating on those with a history of active pulmonary TB.

Early data shows that classification algorithms for coughs that adhere to WHO TPPs for a community-based TB screening test may be developed. It is necessary to conduct additional validation using cohorts with large sample sizes and diverse demographics before any definitive conclusions about their sensitivities and specificities can be made. AI-based cough screening should supplement chest X-rays and other community-based screening techniques to boost the proportion of people with suspected TB who are promptly and correctly sent to facilities for confirmation testing. In fact, using cough to predict chest X-ray abnormalities may lead to radiology testing, for which various automated interpretation algorithms have already been fully validated.If AI-based cough screening were deployed on mobile devices, it might enable low-cost remote active case detection and self-screening, with referral to a healthcare facility for confirmation TB testing and connection to care following. The vignette explains how a cough monitoring device can make it simpler to send individuals who have a cough to a doctor.

Cough detection and longitudinal monitoring[1.3] could objectively document a rise in TB-compatible symptoms in people who are more likely to acquire active pulmonary TB, such as household contacts, encouraging early care-seeking and preventing transmission. Subclinical pulmonary TB51 might potentially be addressed using this strategy. Subclinical51 refers to people who experience little symptoms but do not consider them to be serious. In these situations, digital cough monitoring may

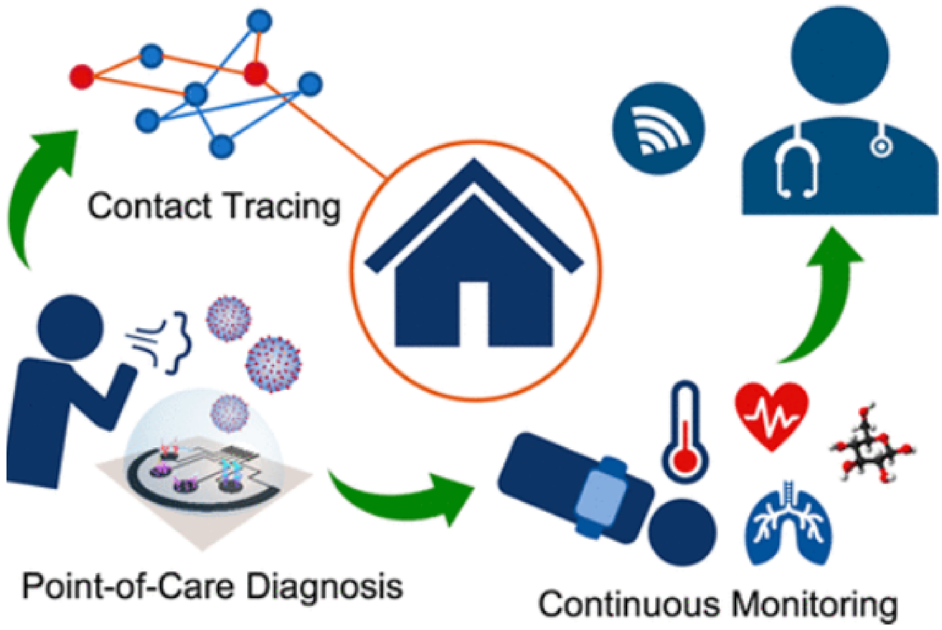

Figure 1.4: Tuberculosis Diagnosis using app

be utilised to detect the presence or importance of coughing that would have gone unnoticed or unreported otherwise. Contrarily, digital cough monitoring would not be relevant to really asymptomatic TB patients, limiting its use as an active case-finding method in this population. According to a research of 24 TB patients15, coughing frequency may not be connected to M. tuberculosis production found on face masks. This indicates that some participants, even those who did not cough regularly, nonetheless coughed up a significant amount of M. tuberculosis (and vice versa). Although while additional study is needed, this illustrates the risks of relying on cough monitoring for determining active case-finding and controlling TB transmission. preclinical TB patients, which limits its usage as an active case-finding method in this population..

It is not always guaranteed that patients with suspected TB will have access to the appropriate confirmatory tests when they arrive at the medical facility. This is caused in part by the lack of knowledge and training among medical professionals on the crucial TB symptoms. This problem has been demonstrated in studies using standardised patients (SPs), healthy people who have been told to attend medical facilities feigning TB symptoms without the healthcare professionals being aware that these symptoms are false52. According to a comprehensive study of SPs in India, only 50

In order to supplement less accurate symptom-based triage procedures and increase the number of persons with presumptive TB who get confirmatory testing, healthcare practitioners may deploy AI-based cough categorization software. Active case-finding and community-based screening are comparable to this. Given how non-specific symptom screening is, tools for identifying coughs may also help reduce the proportion of people without TB who get needless TB testing. Smartphones may be used as recording devices and are commonly available. They are already used for video directly observed treatment (vDOT) TB treatment adherence monitoring, which frees TB patients from needing to travel to a clinic to take their anti-TB medicine as needed by conventional DOT methods54.Given that cough symptoms subside with good treatment13, cough detection apps might be used as a low-cost, person-centered technique for physicians to remotely monitor TB patients' clinical response to therapy. People could even self-monitor their cough as treatment progresses. Unfavorable patterns of cough evolution that have been objectively seen might cause patients and medical professionals to doubt the efficacy of the existing treatment strategy, enabling for the early identification of medication resistance or poor adherence.

Within the first year after finishing anti-TB treatment, a sizable portion of those who are effectively cured of TB are at risk of relapse55. To spot early indications of TB recurrence during this time of increased risk, prospective cough surveillance that was employed throughout treatment could be resumed. An area of TB care that is frequently disregarded in TB management pathways56 is post-TB lung damage, which is more likely to occur even if persons do not experience TB recurrence or relapse. Hence, cough monitoring could also be helpful as a starting point in identifying people with post-TB lung illness and associated lung function reduction if it is validated.

The development of AI-based cough detection technology might also be useful for TB research. Clinical drug development trials might include digital cough monitoring as a secondary outcome. Sputum culture test results changing from positive to negative during the first 8 weeks of medication have been used as a proxy for anti-TB treatment effectiveness in drug development trials up to this point57. Such culture techniques require a lot of time and resources, and they prevent the tracking of intermediate results like patient symptoms. Regulatory bodies may also ask for information on patient-reported improvements in cough, albeit this information is similarly subjective and subject to varying degrees of accuracy58,59. Self-assessment of cough in the context of experimental therapy-efficacy testing is likely to be less accurate than symptom-based screening. objective observation Cough may operate as a supplement to traditional culture-based endpoints and patient-reported outcomes, allowing for more subtle monitoring of intermediate endpoints.

The audio signal is used to extract characteristics including zero crossing rate (ZCR), kurtosis, and mel-frequency cepstral coefficients (MFCCs). Successful use of MFCCs as features in audio analysis, particularly in automatic speech recognition, have been reported [44, 45]. They have been used successfully to distinguish between dry and wet coughs as well as to recognise coughs connected to COVID-19 and tuberculosis. Using first-order difference and higher resolution MFCCs in addition to the standard MFCC extraction method) and acceleration (second-order difference), as their addition has previously demonstrated improved classifier performance [49]. The ZCR measures the amount of signal variability by counting the times a signal changes sign within a frame. The kurtosis, in addition, shows that samples of an audio signal tend to have greater amplitudes. The hyperparameters listed in Table 2 were applied to all cough and non-cough audio events to extract these features.

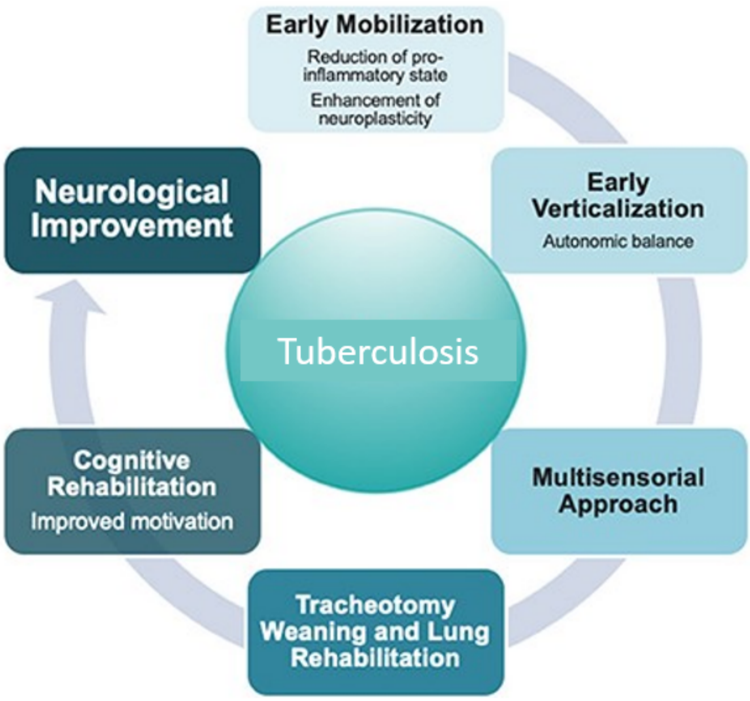

Figure 1.5: Rehabilitation of Tuberculosis affected patients

Mild TB patients'[1.4] recovery typically takes one to three weeks. In that regard, a number of elements, such as the enormous demand for highly responsive, more effective, and smarter detection systems, blatantly show that addressing the first phase is unquestionably the most important for controlling the TB pandemic. The main issue is that the solutions that are now on the market take a lot of time, are prone to late detection, and may contain inaccuracies. On the other hand, coronavirus epidemics are discovered using medical technologies like medical detection kits. Its accuracy in identifying TB, however, is currently being evaluated for its promise as well as its limitations. These instruments are expensive and need to be installed for diagnosis. The following contributions are made by this study in this context: The data gathered from individual users is combined with a machine learning approach to create a new framework that offers a preliminary diagnosis tool for TB. The methodology is put into practise by creating a mobile application where data are gathered via a self-reported questionnaire. A number of smartphone embedded sensors, including a thermometer, a microphone, etc., are also utilised to collect data. Based on the gathered data, a tailored convolutional neural network (CNN) is suggested for TB prediction. To emphasise the value of selecting essential features, experiments are conducted with various sets of features. Outcomes are talked upon with regard to cross-validation, computation time, precision, recall, and F1 score.The last five portions of the essay are as follows: The research that use the most recent AI-based methods and applications to predict TB are described in Section 2 along with other related studies. In Section 3, the characteristics and architecture design of the suggested system for TB preliminary diagnosis are thoroughly presented. It also goes into detail on the database that was used for model training and validation. The models were trained using the settings described in the section on hyperparameter settings for training.

Despite the long-standing availability of a successful treatment, tuberculosis (TB) is one of the most lethal infectious illnesses in the world. It is concerning to see the gradual rise of TB strains that are highly drug-resistant (XDR) and multi-drug resistant (MDR). Worldwide, MDR TB causes about 3.6

Our best LR system uses 23 features chosen from a set of 78 high-resolution mel-frequency cepstral coefficients to obtain an area under the ROC curve (AUC) of 0.94 when combined with feature selection using sequential forward selection (SFS) (MFCCs). This approach surpasses the 90

These algorithms are being trained to discriminate coughs from patients with different clinical states or at different stages of disease as well as to recognise human coughs from ambient noises, however the latter use case has not yet been validated39,40,41,42,43 (cough detection). Many early cough categorization methods have been developed using COVID-19 and TB. A classification algorithm was reported to detect COVID-19 infections among people with a cough with 98 sensitivity and 94 specificity based on a sample of 5320 people, of whom 50

A more thorough approach to TB control and prevention will be possible thanks to the TB cough detection application's web-based design. The programme can be used as a screening tool by medical practitioners for individuals who appear with a persistent cough, enabling early TB detection and prompt treatment. In order to enhance TB detection algorithms and create more precise and trustworthy diagnostic tools, researchers can utilise the programme to gather data on cough sounds from various populations.

The audio signal is used to extract characteristics including zero crossing rate (ZCR), kurtosis, and mel-frequency cepstral coefficients (MFCCs). Successful use of MFCCs as features in audio analysis, particularly in automatic speech recognition, have been reported [44, 45]. They have been used successfully to distinguish between dry and wet coughs as well as to recognise coughs connected to COVID-19 and tuberculosis. Using first-order difference and higher resolution MFCCs in addition to the standard MFCC extraction method) and acceleration (second-order difference), as their addition has previously demonstrated improved classifier performance [49]. The ZCR measures the amount of signal variability by counting the times a signal changes sign within a frame. The kurtosis, in addition, shows that samples of an audio signal tend to have greater amplitudes. The hyperparameters listed in Table 2 were applied to all cough and non-cough audio events to extract these features.

While continuous cough recording has been made possible by ambulatory recording devices, many of the current models are large and intrusive. Cough is a noticeable and stigmatising symptom, and its prevalence among TB patients has significantly increased this stigma35. Effective monitoring of people who cough requires covert recording techniques to prevent contributing to the stigmatisation of respiratory illnesses. Monitoring TB coughs can be done covertly using smartphones that include cough recognition and recording apps. There are already a number of cough recording programmes available, such as Hyfe Research, AI4COVID-19, and ResAppDx..

Especially in resource-constrained regions with limited access to healthcare institutions, the creation of this web-based automatic tuberculosis cough detection programme offers enormous potential for enhancing early detection and treatment of TB. Healthcare professionals, patients, and carers can all use this programme to conduct non-invasive TB screenings. It can also be used as a tool for patients with active TB to track their progress through treatment and identify relapses.

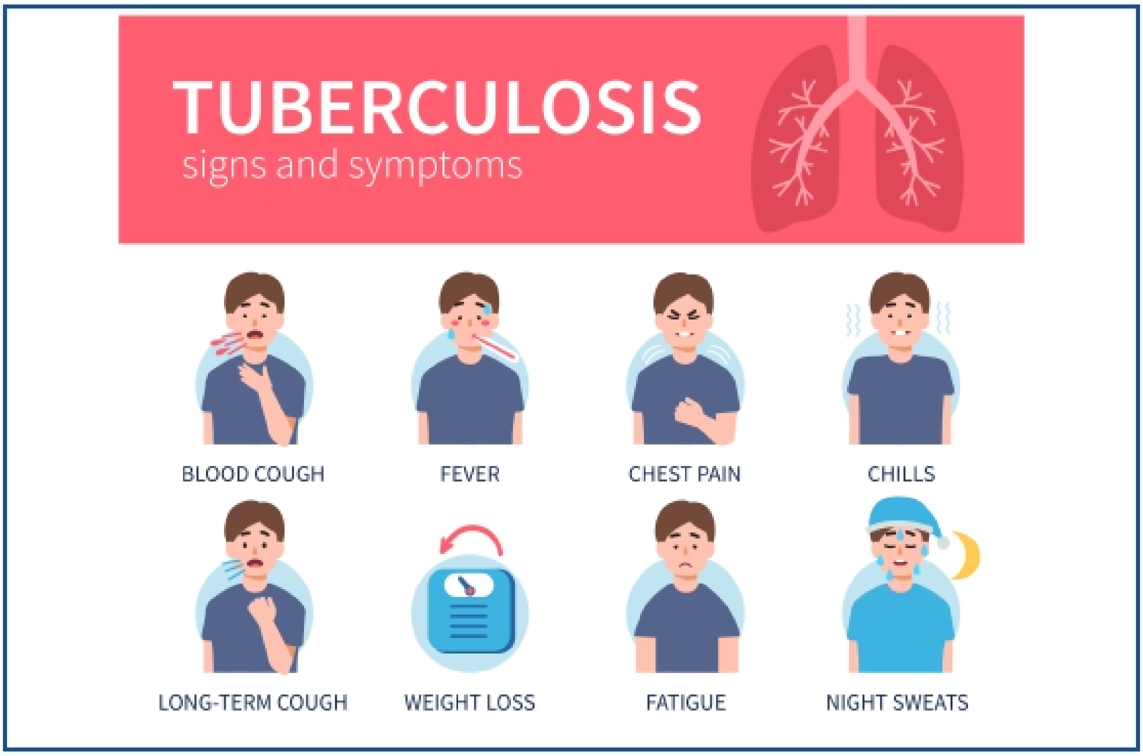

Figure 1.6: Tuberculosis Symptoms

The following contributions are made by this study in this context: The data gathered from individual users is combined with a machine learning approach to create a new framework that offers a preliminary diagnosis tool for TB. The methodology is put into practise by creating a mobile application where data are gathered via a self-reported questionnaire. A number of smartphone embedded sensors, including a thermometer, a microphone, etc., are also utilised to collect data. Based on the gathered data, a tailored convolutional neural network (CNN) is suggested for TB prediction.

To emphasise the value of selecting essential features, experiments are conducted with various sets of features. Outcomes are talked upon with regard to cross-validation, computation time, precision, recall, and F1 score.The last five portions of the essay are as follows: The research that use the most recent AI-based methods and applications to predict TB are described in Section 2 along with other related studies. In Section 3, the characteristics and architecture design of the suggested system for TB preliminary diagnosis are thoroughly presented. It also goes into detail on the database that was used for model training and validation. The models were trained using the settings described in the section on hyperparameter settings for training.

The use of recording equipment and computer-assisted acoustic interpretation techniques enhances the objectivity of cough analysis. Loudon and Spohn began recording and counting the nighttime coughs of TB patients in the 1960s25 using tape recorders. Early ambulatory cough metres contained electromyogram (EMG) electrodes26, which recorded the user's coughing as well as chest spasms, and audio recorders for audio recording. A self-contained cough monitor was created and tested by Paul et al. in 2006 utilising a CompactFlash memory card and an accelerometer (to detect vibrations associated with coughing). This device was placed in the patient's suprasternal notch, or jugular notch, and it showed good agreement with the coughing captured on video. Throughout time, increasingly sophisticated round-the-clock recording systems have been created.A digital sound recorder, often attached to the patient's hip28, is equipped with a microphone that transmits the patient's cough noises. The microphone may be attached to the patient's lapel or worn freely as a necklace. The VitaloJak, the Leicester Cough Monitor (LCM), and the Cayetano Cough Monitor are some of these recording devices (CayeCoM).

The use of digital technologies to enhance TB detection and control has garnered more attention in recent years. Machine learning (ML) and artificial intelligence (AI) have in particular showed promise in TB detection from cough sounds. Compared to conventional diagnostic techniques, cough-based TB detection systems provide a number of benefits, including non-invasiveness, cheap cost, and usability.

As a result, the goal of this project is to create an automatic TB cough detection tool for the web that can quickly and effectively identify TB from cough sounds. The application will analyse cough noises and find suspected TB patients using ML and AI approaches. The application could enhance early identification of the disease and ultimately lower TB morbidity and mortality by offering a straightforward, affordable, and non-invasive method of TB detection. The application's web-based design will also make a wider range of users, such as medical professionals, researchers, and patients, able to use it.

A more thorough approach to TB control and prevention will be possible thanks to the TB cough detection application's web-based design. The programme can be used as a screening tool by medical practitioners for individuals who appear with a persistent cough, enabling early TB detection and prompt treatment. In order to enhance TB detection algorithms and create more precise and trustworthy diagnostic tools, researchers can utilise the programme to gather data on cough sounds from various populations.

The web-based tool also gives patients the ability to record and upload their cough sounds for study, encouraging them to take an active role in their own health. By encouraging users to get checked out if they think they may have TB, this feature can help spread awareness of the condition. The creation of a web-based tool for automatically detecting TB coughs has the power to transform global efforts in TB prevention and control. The application could contribute to lowering the expense of the disease and enhancing the health of those who are afflicted by it by offering a simple, non-invasive, and accessible method of TB detection. As a result of this project, it is envisaged that more research and development into digital health technologies would be stimulated for TB control and prevention.

To emphasise the value of selecting essential features, experiments are conducted with various sets of features. Outcomes are talked upon with regard to cross-validation, computation time, precision, recall, and F1 score.The last five portions of the essay are as follows: The research that use the most recent AI-based methods and applications to predict TB are described in Section 2 along with other related studies. In Section 3, the characteristics and architecture design of the suggested system for TB preliminary diagnosis are thoroughly presented. It also goes into detail on the database that was used for model training and validation. The models were trained using the settings described in the section on hyperparameter settings for training.

Our earlier research demonstrated that it is feasible to differentiate between the coughs of TB patients and healthy individuals, and our work is an obvious extension of that research.

controls using logistic regression (Botha et al. 2018). Yet, studies that only involve people with a problem and healthy controls have well-known disadvantages (Rutjes et al. 2005). We now show that it is also possible to distinguish between TB patients and those who have other lung problems but TB was ruled out as the cause of their symptoms. Contrary to the controlled environment in which our prior recordings were made, the audio data we use in this work were captured at a TB clinic and contain substantial background noise.

As the majority of patients at these clinics are ill, it's crucial to ascertain whether TB illness is likely before further sending them. This referral is frequently made by collecting an infectious and difficult-to-handle sample (sputum), which is then evaluated using a costly (in the context of the developing world) test that necessitates laboratory expertise and specialised equipment. Similar to our prior studies, we focus on automatically categorising coughs and speculatively identifying the start and end of coughing. We realise that by skipping this identification stage, difficult practical issues like managing cough spasms have been delayed for more research..

It was necessary to construct a new corpus of coughing noises from persons who have TB symptoms but may not truly have the disease in order to conduct our testing. This demographic is prioritised in the World Health Organization's (WHO) target product profile for the TB triage test. Similar in size to the corpus used in our prior research, our new corpus was built in a more realistic environment that more closely matches the basic healthcare environment in developing nations where TB screening is most likely to be conducted. This gives us the first concrete evidence that automated sound analysis is a viable and promising tool for TB screening in high-incidence areas.

In Figure 1, you can see how transportable recording equipment was set up in an open, cross-ventilated sputum collection booth in a packed primary healthcare facility in Cape Town, South Africa. This configuration mimics the usual clinic environment encountered in developing countries, which is where low-cost, quickly deployable automated TB screening is most critical. The recordings were made between the hours of 10 am and 4 pm by two medical practitioners without technological backgrounds who were trained to utilise the recording equipment. Without further acoustic shielding, the recording room was susceptible to noise from the clinic's usually high patient and staff population, as well as from the streets' busy traffic, pedestrians, animals, and cars.

Because of this, our dataset has a lot of ambient noise and accurately depicts a scenario where a TB screening test would likely be performed. No attempts were made to de-noise the data because our main focus was on the performance that may be predicted in a real-world situation.

Noise reduction performed before model training has also been shown to be commonly ineffective and may reduce resistance to a range of input circumstances in the related field of automated voice recognition (Caballero et al. 2018). This study was approved by the City of Cape Town and the Faculty of Health Sciences Research Ethics Committee at Stellenbosch University (N14/10/136) (10483). All those who took part in the study gave their informed consent.

The audio was captured while being shielded by a normal N95 mask that was switched out after each patient, using a ZOOM F8N field recorder. According to Hsu et al. (1998) and Todorovi'c et al. (2015), the microphone used was an RDE M3 condenser microphone. Informal listening tests showed that there was little to no effect of the mask on the quality of the audio signal. The medical personnel ensured that the patient and the microphone were separated by 10 to 15 centimetres (Figure 2). Each patient was told to cough while counting from one to 10, causing at least two cough bursts. They were told to cough again after taking a few long breaths.

The sampling rate for all audio recordings was 44.1 kHz. As seen in Figure 3, the coughing-containing sections of the generated audio recordings were manually annotated using the ELAN multimedia programme (Wittenburg et al., 2006). We see that numerous cough onsets frequently occur during these manually annotated stretches of purposeful coughing. Table 2's total of 3124 of these onsets for the 1358 cough episodes indicates an average of 2.3 onsets per cough event. There are 973 and 2151 cough onsets among the 402 TB cough events and the 956 non-TB cough events, respectively, meaning that there are 2.420.83 onsets per TB cough event and 2.250.91 onsets per non-TB cough event.

The number of cough bursts or onsets per cough event does not appear to be a significant factor in the TB cough categorization task, according to these averages and standard deviations. We don't try to automatically determine the limits of such onset subdivisions in this study, but this is still something we're working on. In the remaining sections of this study, we will refer to these sections of audio that feature coughing as cough events, such as those highlighted in Figure.

All of the participants had a self-reported involuntary cough that was possibly indicative of underlying lung disease and were all TB suspects. For the purposes of the study, clinical practise was restricted to bacteriological TB diagnosis. The subjects were not examined by a doctor, but rather, the healthcare staff took audio samples from them. The collection of differential diagnoses for illnesses other than TB was impractical since, even with additional testing, the diagnosis was frequently made based on how symptoms associated to the treatment changed. For the sake of the study, patients were therefore only subjected to standardised TB testing, and no other diagnoses other than TB were made. Table 1 contains the inclusion and exclusion standards for the participants.

AUC is more discriminatory than certain other established performance metrics, such accuracy for unbalanced datasets (Rakotomamonjy 2004, Huang Ling 2005, Fawcett 2006). We choose it as the performance statistic for this reason. Coughs total 1045 seconds (17.42 minutes) in our dataset. TB coughs last, on average, 0.74 seconds with a 0.31 standard deviation, compared to non-TB coughs, which last 0.78 seconds on average with a 0.39 standard deviation. We notice that the length of coughs produced by TB patients and those brought on by other lung disorders are comparable as a result. Contrary to what we had previously noticed, TB patients cough more frequently and for a longer period of time than healthy people do.

We consistently employed nested 5-fold cross-validation across all experiments due to the modest size of our dataset. Figure 7 illustrates how an outer loop partitions the dataset into training (80The training part is once more split into two independent inner loops within this outer loop, one of which performs 4-fold cross-validation and the other 2-fold. While the latter is used to calculate the equal error rate that is used to inform the classifier choice, the former is used to optimise the hyperparameter stated in Table 4. Inside the inner loops, there was no patient overlap, and the gender distribution was equal.

Unfortunately, it appears that our dataset is too tiny and possibly too noisy for neural networks' higher flexibility to be useful. We take notice of the fact that other studies have also claimed that LR is superior in several clinical prediction tasks (Christodoulou et al. 2019). We also want to point out that the same nested k-fold crossvalidation approach was used to evaluate each of our classifiers. So, even though more training data may one day be beneficial for more complex neural network architectures like the CNN, our comparison of the classifiers is justified, and the achieved performance is an accurate representation of what is currently achievable.

The added benefit of mean normalisation offered by MFCCs over the more straightforward log-filterbank energies is that it offers a quick and efficient technique to account for convolutional channel variability. The dataset we utilise in this study is less well-controlled and noisier than the one we used in our earlier research. As the microphone does not stay in the collecting booth overnight, for instance, its position is typically a bit altered between recordings. So, for the dataset we are considering here, the benefits of channel normalisation might be more significant.

Finally, we have demonstrated that performance can be improved via greedy feature removal. The greatest AUC is reached by using 23 out of the 78 potential features, while near-optimal performance may be achieved with as little as four features.

This is essential for doing audio TB screening on portable computing devices, such cellphones, as less computational labour is needed when the dimensionality of the feature vector is decreased. If the algorithm were implemented on a consumer mobile device, it would be portable, inexpensive, and easy to use, which would make it desirable in situations with limited resources.

For the first time, we have shown that it is feasible to automatically distinguish between patients with various lung problems and tuberculosis (TB) patients' forceful coughing noises. This backs up prior studies we conducted that claimed it could discriminate between healthy controls and TB patients' coughs. The sounds of TB sufferers' coughs do not appear to have any outwardly recognisable features, in contrast to conditions like croup coughs, which may be more clearly detected (Sharan et al. 2017, Sharan et al. 2018).

Our findings are built on a freshly compiled dataset that was taken in a busy primary healthcare clinic. Because of this, we also show that categorising TB cough is possible in the type of real-world environment that may be predicted in a screening centre in a developing country. Using layered cross-validation, the effectiveness of five machine learning classifiers was evaluated. By selecting the best 23 of 78 high-resolution MFCC characteristics using logistic regression (LR) and sequential forward selection (SFS), an area under the ROC curve (AUC) of 0.94 was obtained, proving that MFCCs without velocity or acceleration may give performance that is almost perfect..

The suggested screening method, which uses automatic analysis of coughing noises, is non-intrusive, can be used without the need for specialised medical knowledge or laboratory resources, produces results quickly, and can be implemented on readily available and reasonably priced consumer hardware, such as a smartphone. Therefore it may be a useful tool in the battle against TB, especially in poor countries with a high TB burden, like Cape Town, South Africa, where we are located (Blaser et al. 2016, Mulongeni et al. 2019). According to recent studies, South Africa now has 600–700 cases of TB per 100,000 people.

Second, at this moment, cough sounds are the sole recordings we are utilising for categorization. Although this study is still under progress, incorporating the speech audio that was also gathered as part of our data collection may be able to improve classifier accuracy. Last but not least, there are occasionally many bursts of cough onsets in the manually annotated cough occurrences. We are now looking on ways to automatically identify these bursts during a coughing incident. Fourthly, we've only tested our classifier on one dataset. To more clearly illustrate the classifier's ability to effectively analyse totally unseen input, an additional validation-only dataset is required. It is now planned to conduct such a data collection project, in which recordings will be produced in many languages.

CHAPTER 2

LITERTAURE SURVEY

Zimmer[1] "This research paper, discusses that the use of machine learning algorithms to detect, record, and analyze cough sounds for TB control and individual patient management. The paper proposes a largely automated algorithm for detecting cough sounds, which requires operator input for calibrating the device. The algorithm can be used for TB screening and diagnosis, as well as monitoring treatment progress. The paper highlights the importance of accurate and timely diagnosis of TB, especially in developing countries where access to healthcare facilities is limited. The proposed algorithm has shown promising results in detecting TB coughs with high sensitivity and specificity.One approach is to use artificial intelligence (AI) to detect, record, and analyze coughs. This technology can be used in both TB control and individual patient management[1]. By analyzing coughing patterns, healthcare providers can monitor patients' progress and adjust treatment plans as needed. AI-powered cough detection systems can also help identify patients who may be at risk of developing drug-resistant TB.In addition to AI-powered cough detection systems, other technologies are being developed to improve TB care. For example, researchers are exploring the use of mobile health (mHealth) tools such as text messaging and smartphone apps to support patient self-management and improve treatment adherence. These tools can help patients stay on track with their medication regimens and provide them with information about their condition. Overall, the paper suggests that using machine learning algorithms to analyze cough sounds can be a valuable tool for improving TB care and management.

Botha[2] In a research paper titled "The Detection of Tuberculosis by Automatic Cough Sound Analysis," a straightforward and readily implemented approach for TB screening based on the automatic analysis of coughing sounds is proposed. The research emphasises the significance of early TB diagnosis, particularly in underdeveloped nations with restricted access to healthcare services. The suggested technique analyses cough sounds and employs machine learning algorithms to discriminate between the coughs of TB positive patients and healthy controls. The classifiers automatically and specifically identify patients who are TB positive using short-term spectral information. These classifiers automatically discriminate between TB-positive patients' coughs and other patients' coughs using short-term spectral information.These classifiers automatically discriminate between the coughs of TB positive patients and healthy controls with an accuracy of 78 percent and an AUC of 0.95 using short-term spectral information. Its accuracy rises to 82 percent when a collection of five clinical measures is also made accessible. The system is capable of achieving a sensitivity of 95% at a specificity of around 72% by selecting an acceptable decision threshold. According to the trials, certain spectrum information that is not perceptible by the human hearing system is used by the classifiers, and some frequencies are more helpful for categorization than others. The paper also discusses the potential applications of this method in remote areas where access to healthcare facilities is limited. Overall, the paper suggests that using machine learning algorithms for analyzing cough sounds can be an effective tool for improving TB diagnosis and management.The automatic classification of cough audio sounds has been tested in real-world environments, showing promising results in meeting WHO triage specifications for identifying patients who should undergo further testing. This makes it a viable means of screening for TB. Further research is needed to improve the accuracy of this method and make it more widely available in low-resource settings where TB is prevalent"

Madhurananda Pahar [3], "This research paper, proposes a deep learning-based automatic cough classifier that can discriminate tuberculosis (TB) coughs from COVID-19 coughs and healthy ones. The paper highlights the importance of early diagnosis of TB and COVID-19, especially in developing countries where access to healthcare facilities is limited. The proposed method uses machine learning algorithms to analyze cough sounds and distinguish between the coughs of TB positive patients, COVID-19 positive patients, and healthy controls. The classifiers use short-term spectral information to automatically detect TB-positive patients with high sensitivity and specificity. The paper also discusses the potential applications of this method in remote areas where access to healthcare facilities is limited. The authors suggest that using machine learning algorithms for analyzing cough sounds can be an effective tool for improving TB and COVID-19 diagnosis and management. Overall, the literature review suggests that there is a need for accurate and timely diagnosis of TB and COVID-19, especially in developing countries where access to healthcare facilities is limited. The proposed method has shown promising results in detecting TB and COVID-19 coughs with high sensitivity and specificity, which can be a valuable tool for improving diagnosis and management of these diseases."

Byron Reeve[4], "The automatic classification of cough audio sounds can be used for tuberculosis (TB) screening in a real-world environment. The experiments are based on a dataset of cough recordings obtained in a developing world clinic from patients with confirmed active pulmonary TB and patients suffering from respiratory conditions suggestive of TB but confirmed to be TB negative. The automatic discrimination between the coughing sounds produced by patients with TB and those produced by patients with other lung ailment is achieved using complex algorithms. The automatic classification of cough audio sounds, when applied to symptomatic patients requiring investigation for TB, can meet the WHO triage specifications for the identification of patients who should undergo expensive molecular downstream testing. This makes it a promising and cost-effective tool for TB screening in resource-limited settings. This system achieves a sensitivity of 93 present at a specificity of 95 percent and thus exceeds the 90 percent sensitivity at 70 percent specificity specification considered by the World Health Organization (WHO) as a minimal requirement for a community-based TB triage test. However, further validation is needed using cohorts of large sample size and diverse populations before any definete conclusion can be made regarding their effectiveness."

Wichern G,[6] "The authors have proposed a method for characterizing sound activity in fixed spaces through segmentation, indexing, and retrieval of continuous audio recordings. Regarding segmentation, we present a dynamic Bayesian network (DBN) that jointly infers on-sets and end times of the most prominent sound events in the space, along with an extension of the algorithm for covering large spaces with distributed microphone arrays. Each segmented sound event is indexed with a hidden Markov model (HMM) that models the distribution of example-based queries that a user would employ to retrieve the event (or similar events). In order to increase the efficiency of the retrieval search, we recursively apply a modified spectral clustering algorithm to group similar sound events based on the distance between their corresponding HMMs. We then conduct a formal user study to obtain the relevancy decisions necessary for evaluation of our retrieval algorithm on both automatically and manually segmented sound clips. Furthermore, our segmentation and retrieval algorithms are shown to be effective in both quiet indoor and noisy outdoor recording conditions. Index Terms—Acoustic signal analysis, acoustic signal detection, Bayes procedures, clustering methods, database query processing"

Yang Ch et.al.,[7]"The authors proposed a novel decentralized feature extraction approach in federated learning to address privacy-preservation issues for speech recognition. It is built upon a quantum convolutional neural network (QCNN) composed of a quantum circuit encoder for feature extraction, and a recurrent neural network (RNN) based end-to-end acoustic model (AM). To enhance model parameter protection in a decentralized architecture, an input speech is first up-streamed to a quantum computing server to extract Mel-spectrogram, and the corresponding convolutional features are encoded using a quantum circuit algorithm with random parameters. The encoded features are then down-streamed to the local RNN model for the final recognition. The proposed decentralized framework takes advantage of the quantum learning progress to secure models and to avoid privacy leakage attacks. Testing on the Google Speech Commands Dataset, the proposed QCNN encoder attains a competitive accuracy of 95.12 percent in a decentralized model, which is better than the previous architectures using centralized RNN models with convolutional features. We also conduct an in-depth study of different quantum circuit encoder architectures to provide insights into designing QCNN-based feature extractors. Neural saliency analyses demonstrate a correlation between the proposed QCNN features, class activation maps, and input spectrograms. We provide an implementation for future studies."

Coppock H et.al.,[10] "The authors of this particular study demonstrated the feasibility of an alternative form of COVID-19 detection, harnessing digital technology through the use of audio biomarkers and deep learning.Since the emergence of COVID-19 in December 2019, multidisciplinary research teams have wrestled with how best to control the pandemic in light of its considerable physical, psychological and economic damage. Mass testing has been advocated as a potential remedy; however, mass testing using physical tests is a costly and hard-to-scale solution. Specifically, they showed that a deep neural network based model can be trained to detect symptomatic and asymptomatic COVID-19 cases using breath and cough audio recordings.Their model, a custom convolutional neural network, demonstrates strong empirical performance on a data set consisting of 355 crowd.sourced participants, achieving an area under the curve of the receiver operating characteristics of 0.846 on the task of COVID-19 classification.This study offers a proof of concept for diagnasing COVID-19 using cough and breath audio signals and motivates a comprehensive follow-up research study on a wiider data sample, gaven the evident advantages of a low-cost, highly scalable digital COVID-19 diagnostic tool."

Qian K et.al.,[12] "Computer audition (CA) has experienced a fast development in the past decades by leveraging advanced signal processing and machine learning techniques. In particular, for its noninvasive and ubiquitous character by nature, CA-based applications in healthcare have increasingly attracted attention in recent years. During the tough time of the global crisis caused by the coronavirus disease 2019 (COVID-19), scientists and engineers in data science have collaborated to think of novel ways in prevention, diagnosis, treatment, tracking, and management of this global pandemic. On the one hand, we have witnessed the power of 5G, Internet of Things, big data, computer vision, and artificial intelligence in applications of epidemiology modeling, drug and/or vaccine finding and designing, fast CT screening, and quarantine management. On the other hand, relevant studies in exploring the capacity of CA are extremely lacking and underestimated. To this end, we propose a novel multitask speech corpus for COVID-19 research usage. We collected 51 confirmed COVID-19 patients' in-the-wild speech data in Wuhan city, China. We define three main tasks in this corpus, i.e., three-category classification tasks for evaluating the physical and/or mental status of patients, i.e., sleep quality, fatigue, and anxiety. The benchmarks are given by using both classic machine learning methods and state-of-the-art deep learning techniques. We believe this study and corpus cannot only facilitate the ongoing research on using data science to fight against COVID-19, but also the monitoring of contagious diseases for general purpose."

Deshmukh S, Al Ismail M et.al.,[13] "In the pathogenesis of COVID-19, impairment of respiratory functions is often one of the key symptoms. Studies show that in these cases, voice production is also adversely affected - vocal fold oscillations are asynchronous, asymmetrical and more restricted during phonation. This paper proposes a method that analyzes the differential dynamics of the glottal flow waveform (GFW) during voice production to identify features in them that are most significant for the detection of COVID-19 from voice. Since it is hard to measure this directly in COVID-19 patients, we infer it from recorded speech signals and compare it to the GFW computed from physical model of phonation. For normal voices, the difference between the two should be minimal, since physical models are constructed to explain phonation under assumptions of normalcy. Greater differences implicate anomalies in the bio-physical factors that contribute to the correctness of the physical model, revealing their significance indirectly. Our proposed method uses a CNN-based 2-step attention model that locates anomalies in time-feature space in the difference of the two GFWs, allowing us to infer their potential as discriminative features for classification. The viability of this method is demonstrated using a clinically curated dataset of COVID-19 positive and negative subjects."

Jing Han, Chloë Brown, Jagmohan Chauhan, et.al.,[14]" The development of fast and accurate screening tools, which could facilitate testing and prevent more costly clinical tests, is key to the current pandemic of COVID-19. In this context, some initial work shows promise in detecting diagnostic signals of COVID-19 from audio sounds. In this paper, we propose a voice-based framework to automatically detect individuals who have tested positive for COVID-19. We evaluate the performance of the proposed framework on a subset of data crowdsourced from our app, containing 828 samples from 343 participants. By combining voice signals and reported symptoms, an AUC of 0.79 has been attained, with a sensitivity of 0.68 and a specificity of 0.82. We hope that this study opens the door to rapid, low-cost, and convenient pre-screening tools to automatically detect the disease. "

Rodriguez CR, Angeles D et.al.,[15] "The objective of the paper is to provide a model capable of serving as a basis for retraining a convolutional neural network that can be used to detect COVID-19 cases through spectrograms of coughing, sneezing and other respiratory sounds from infected people. To address this challenge, the methodology was focused on Deep Learning technics worked with a dataset of sounds of sick and non-sick people, and using ImageNet's Xception architecture to train the model to be presented through Fine-Tuning. The results obtained were a precision of 0.75 to 0.80, this being drastically affected by the quality of the dataset at our availability, however, when getting relatively high results for the conditions of the data used, we can conclude that the model can present much better results if it is working with a dataset specifically of respiratory sounds of COVID-19 cases with high quality"

Kushwaha S, Bahl S, eT.al.,[17] "Machine learning is an innovative approach that has extensive applications in prediction. This technique needs to be applied for the COVID-19 pandemic to identify patients at high risk, their death rate, and other abnormalities. It can be used to understand the nature of this virus and further predict the upcoming issues. This literature-based review is done by searching the relevant papers on machine learning for COVID-19 from the databases of SCOPUS, Academia, Google Scholar, PubMed, and ResearchGate. This research attempts to discuss the significance of machine learning in resolving the COVID-19 pandemic crisis. This paper studied how machine learning algorithms and methods can be employed to fight the COVID-19 virus and the pandemic. It further discusses the primary machine learning methods that are helpful during the COVID-19 pandemic. We further identified and discussed algorithms used in machine learning and their significant applications. Machine learning is a useful technique, and this can be witnessed in various areas to identify the existing drugs, which also seems advantageous for the treatment of COVID-19 patients. This learning algorithm creates interferences out of unlabeled input datasets, which can be applied to analyze the unlabeled data as an input resource for COVID-19. It provides accurate and useful features rather than a traditional explicitly calculation-based method. Further, this technique is beneficial to predict the risk in healthcare during this COVID-19 crisis. Machine learning also analyses the risk factors as per age, social habits, location, and climate."

Jordi Laguarta, Ferran Hueto et.al.,[18]"The authors present the dataset, model architecture, and performance of a zero-cost, rapid, and instantly distributable COVID-19 forced-cough recording AI pre-screening tool achieving 98.5% accuracy, including 100% asymptomatic detection rate. An orthogonal set of biomarkers may be developed to diagnose COVID-19, Alzheimer's, and perhaps other conditions. They hypothesized that COVID-19 subjects, especially including asymptomatics, could be accurately discriminated only from a forced-cough cell phone recording using Artificial Intelligence. To train our MT Open Voice model, we built a data collection pipeline of COVID-19 cough recordings through our website (opensigma.mit.edu) between April and May 2020 and created the largest audio COVID-19 cough balanced dataset reported to date with 5,320 subjects. Methods: We developed an AI speech processing framework that leverages acoustic biomarker feature extractors to pre-screen for COVID-19 from cough recordings and provide a personalized patient saliency map to longitudinally monitor patients in real-time, non-invasively, and at essentially zero variable cost. Cough recordings are transformed with Mel Frequency Cepstral Coefficients and inputted into a Convolutional Neural Network (CNN) based architecture made up of one Poisson biomarker layer and 3 pre-trained ResNet50's."

Thomas F.Quatieri et.al.,[19] "The authors propose a speech modeling and signal processing framework to detect and track COVID-19 through asymptomatic and symptomatic stages. Methods: The approach is based on the complexity of neuromotor coordination across speech subsystems involved in respiration, phonation, and articulation, motivated by the distinct nature of COVID-19 involving lower (i.e., bronchial, diaphragm, lower tracheal) versus upper (i.e., laryngeal, pharyngeal, oral, and nasal) respiratory tract inflammation, as well as by the growing evidence of the virus' neurological manifestations. Preliminary results: An exploratory study with audio interviews of five subjects provides Cohen's d effect sizes between pre-COVID-19 (pre-exposure) and post-COVID-19 (after positive diagnosis but presumed asymptomatic) using the coordination of respiration (as measured through acoustic waveform amplitude) and laryngeal motion (fundamental frequency and cepstral peak prominence), and the coordination of laryngeal and articulatory (formant center frequencies) motion. Conclusions: While there is a strong subject dependence, the group-level morphology of effect sizes indicates a reduced complexity of subsystem coordination. Validation is needed with larger, more controlled datasets and to address confounding influences such as different recording conditions, unbalanced data quantities, and changes in underlying vocal status from pre-to-post time recordings."

Gramming P. et.al.,[21] "Mechanistic hypotheses about airborne infectious disease transmission have traditionally emphasized the role of coughing and sneezing, which are dramatic expiratory events that yield both easily visible droplets and large quantities of particles too small to see by eye. Nonetheless, it has long been known that normal speech also yields large quantities of particles that are too small to see by eye but are large enough to carry a variety of communicable respiratory pathogens. Here, we show that the rate of particle emission during normal human speech is positively correlated with the loudness (amplitude) of vocalization, ranging from approximately 1 to 50 particles per second (0.06 to 3 particles per cm3) for low to high amplitudes, regardless of the language spoken (English, Spanish, Mandarin, or Arabic). Furthermore, a small fraction of individuals behaves as "speech superemitters," consistently releasing an order of magnitude more particles than their peers. Our data demonstrate that the phenomenon of speech superemission cannot be fully explained either by the phonetic structures or the amplitude of the speech. These results suggest that other unknown physiological factors, varying dramatically among individuals, could affect the probability of respiratory infectious disease transmission and also help explain the existence of superspreaders who are disproportionately responsible for outbreaks of airborne infectious disease."

Antonio La Marca et.al.,[23]"The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the resulting coronavirus disease 2019 (COVID-19) pandemic have led to a rapid increase in the need for diagnostic assays to enable mass screening and testing of high-risk groups, as well as to determine the extent of past exposure to the virus at individual and population levels. To meet this demand, there has been a rapid development of both molecular and serological assays across a variety of platforms. This review summarizes the current literature on these testing modalities, including nucleic acid amplification tests, direct viral antigen tests, and laboratory-based and point-of-care serological tests. The combination of these tests can help inform crucial decisions by healthcare providers and policy makers, and understanding their strengths and limitations will be crucial for their appropriate application in the development of treatment algorithms and public health strategies."

Byron Reeve[24], "The automatic classification of cough audio sounds can be used for tuberculosis (TB) screening in a real-world environment. The experiments are based on a dataset of cough recordings obtained in a developing-world clinic from patients with confirmed active pulmonary TB and patients suffering from respiratory conditions suggestive of TB but confirmed to be TB negative. The automatic discrimination between the coughing sounds produced by patients with TB and those produced by patients with other lung ailments is achieved using complex algorithms. The automatic classification of cough audio sounds, when applied to symptomatic patients requiring investigation for TB, can meet the WHO triage specifications for the identification of patients who should undergo expensive molecular downstream testing. This makes it a promising and cost-effective tool for TB screening in resource-limited settings. This system achieves a sensitivity of 93 percent at a specificity of 95 percent and thus exceeds the 90 percent sensitivity at 70 percent specificity specification considered by the World Health Organisation (WHO) as a minimal requirement for a community-based TB triage test. However, further validation is needed using cohorts of large sample size and diverse populations before any definite conclusions can be made regarding their effectiveness."

Wichern G,[25] "The Authors have proposed a method for characterizing sound ac-tivity in fixed spaces through segmentation, indexing, and retrieval of continuous audio recordings. Regarding segmentation, we present a dynamic Bayesian network (DBN) that jointly infers on-sets and end times of the most prominent sound events in the space, along with an extension of the algorithm for covering large spaces with distributed microphone arrays. Each segmented sound event is indexed with a hidden Markov model (HMM) that models the distribution of example-based queries that a user would employ to retrieve the event (or similar events). In order to increase the efficiency of the retrieval search, we recursively apply a modified spectral clustering algorithm to group similar sound events based on the distance between their corresponding HMMs. We then conduct a formal user study to obtain the relevancy decisions necessary for evaluation of our retrieval algorithm on both auto-matically and manually segmented sound clips. Furthermore, our segmentation and retrieval algorithms are shown to be effective in both quiet indoor and noisy outdoor recording conditions. Index Terms—Acoustic signal analysis, acoustic signal detection, Bayes procedures, clustering methods, database query processing."

Madhurananda Pahar [27], "This research paper, proposes a deep learning-based automatic cough classifier that can discriminate tuberculosis (TB) coughs from COVID-19 coughs and healthy ones. The paper highlights the importance of early diagnosis of TB and COVID-19, especially in developing countries where access to healthcare facilities is limited. The proposed method uses machine learning algorithms to analyze cough sounds and distinguish between the coughs of TB positive patients, COVID-19 positive patients, and healthy controls. The classifiers use short-term spectral information to automatically detect TB-positive patients with high sensitivity and specificity. The paper also discusses the potential applications of this method in remote areas where access to healthcare facilities is limited. The authors suggest that using machine learning algorithms for analyzing cough sounds can be an effective tool for improving TB and COVID-19 diagnosis and management. Overall, the literature review suggests that there is a need for accurate and timely diagnosis of TB and COVID-19, especially in developing countries where access to healthcare facilities is limited. The proposed method has shown promising results in detecting TB and COVID-19 coughs with high sensitivity and specificity, which can be a valuable tool for improving diagnosis and management of these diseases."

Zimmer[28] "This research paper, discusses that the use of machine learning algorithms to detect, record, and analyze cough sounds for TB control and individual patient management. The paper proposes a largely automated algorithm for detecting cough sounds, which requires operator input for calibrating the device. The algorithm can be used for TB screening and diagnosis, as well as monitoring treatment progress. The paper highlights the importance of accurate and timely diagnosis of TB, especially in developing countries where access to healthcare facilities is limited. The proposed algorithm has shown promising results in detecting TB coughs with high sensitivity and specificity.One approach is to use artificial intelligence (AI) to detect, record, and analyze coughs. This technology can be used in both TB control and individual patient management[1]. By analyzing coughing patterns, healthcare providers can monitor patients' progress and adjust treatment plans as needed. AI-powered cough detection systems can also help identify patients who may be at risk of developing drug-resistant TB.In addition to AI-powered cough detection systems, other technologies are being developed to improve TB care. For example, researchers are exploring the use of mobile health (mHealth) tools such as text messaging and smartphone apps to support patient self-management and improve treatment adherence. These tools can help patients stay on track with their medication regimens and provide them with information about their condition. Overall, the paper suggests that using machine learning algorithms to analyze cough sounds can be a valuable tool for improving TB care and management.

Thomas F.Quatieri et.al.,[29] "We propose a speech modeling and signal-processing framework to detect and track COVID-19 through asymptomatic and symptomatic stages. Methods: The approach is based on complexity of neuromotor coordination across speech subsystems involved in respiration, phonation and articulation, motivated by the distinct nature of COVID-19 involving lower (i.e., bronchial, diaphragm, lower tracheal) versus upper (i.e., laryngeal, pharyngeal, oral and nasal) respiratory tract inflammation, as well as by the growing evidence of the virus' neurological manifestations. Preliminary results: An exploratory study with audio interviews of five subjects provides Cohen's d effect sizes between pre-COVID-19 (pre-exposure) and post-COVID-19 (after positive diagnosis but presumed asymptomatic) using: coordination of respiration (as measured through acoustic waveform amplitude) and laryngeal motion (fundamental frequency and cepstral peak prominence), and coordination of laryngeal and articulatory (formant center frequencies) motion. Conclusions: While there is a strong subject-dependence, the group-level morphology of effect sizes indicates a reduced complexity of subsystem coordination. Validation is needed with larger more controlled datasets and to address confounding influences such as different recording conditions, unbalanced data quantities, and changes in underlying vocal status from pre-to-post time recordings."

Kushwaha S, Bahl S, eT.al.,[30] "Machine learning is an innovative approach that has extensive applications in prediction. This technique needs to be applied for the COVID-19 pandemic to identify patients at high risk, their death rate, and other abnormalities. It can be used to understand the nature of this virus and further predict the upcoming issues. This literature-based review is done by searching the relevant papers on machine learning for COVID-19 from the databases of SCOPUS, Academia, Google Scholar, PubMed, and ResearchGate. This research attempts to discuss the significance of machine learning in resolving the COVID-19 pandemic crisis. This paper studied how machine learning algorithms and methods can be employed to fight the COVID-19 virus and the pandemic. It further discusses the primary machine learning methods that are helpful during the COVID-19 pandemic. We further identified and discussed algorithms used in machine learning and their significant applications. Machine learning is a useful technique, and this can be witnessed in various areas to identify the existing drugs, which also seems advantageous for the treatment of COVID-19 patients. This learning algorithm creates interferences out of unlabeled input datasets, which can be applied to analyze the unlabeled data as an input resource for COVID-19. It provides accurate and useful features rather than a traditional explicitly calculation-based method. Further, this technique is beneficial to predict the risk in healthcare during this COVID-19 crisis. Machine learning also analyses the risk factors as per age, social habits, location, and climate."

Jing Han, Chloë Brown, Jagmohan Chauhan, et.al.,[31] The development of fast and accurate screening tools, which could facilitate testing and prevent more costly clinical tests, is key to the current pandemic of COVID-19. In this context, some initial work shows promise in detecting diagnostic signals of COVID-19 from audio sounds. In this paper, we propose a voice-based framework to automatically detect individuals who have tested positive for COVID-19. We evaluate the performance of the proposed framework on a subset of data crowdsourced from our app, containing 828 samples from 343 participants. By combining voice signals and reported symptoms, an AUC of 0.79 has been attained, with a sensitivity of 0.68 and a specificity of 0.82. We hope that this study opens the door to rapid, low-cost, and convenient pre-screening tools to automatically detect the disease.

CHAPTER 3

METHODOLOGY

3.1 DATA COLLECTION AND MANIPULATION

The prediction of tuberculosis is a critical task in medical diagnosis. It involves analyzing various features extracted from a patient's cough and determining whether the patient is suffering from tuberculosis or not. In recent years, machine learning has shown great potential in predicting tuberculosis with high accuracy. This is achieved through the use of advanced techniques such as Convolutional Neural Networks (CNN) and deep learning.Loading the data is the first step in any machine learning project. In this code, the data is loaded from a CSV file using the pandas library. The CSV file contains the features extracted from the audio files. Once the CSV file is loaded, the

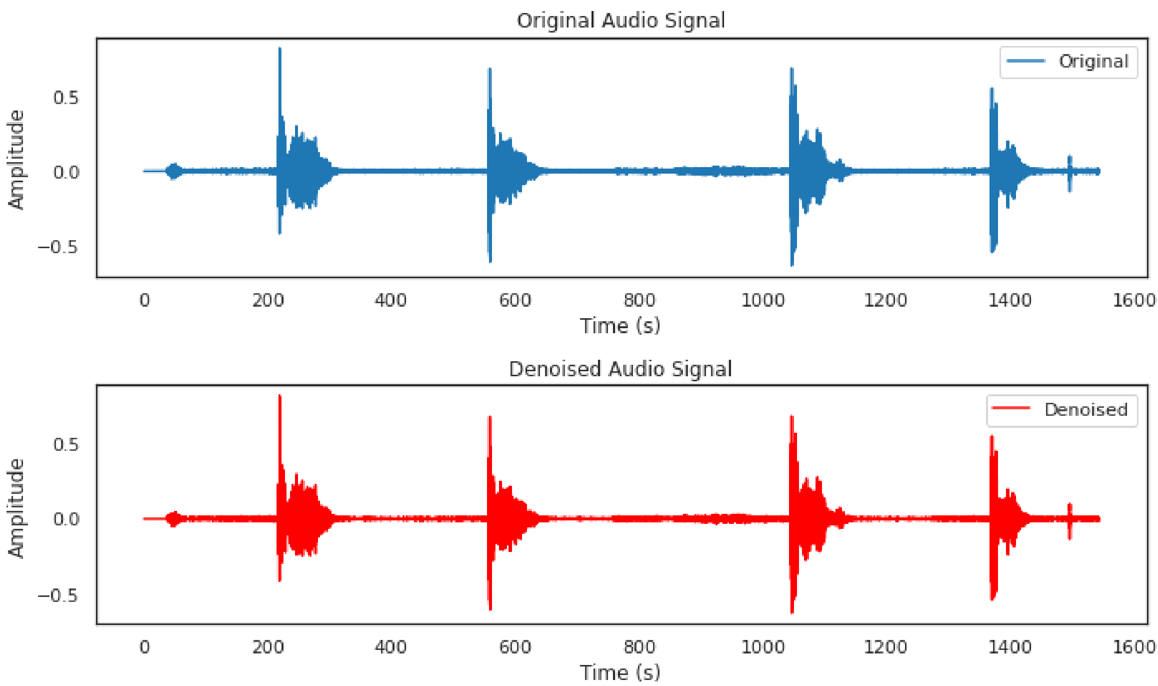

Figure 3.1: Original and Denoised Audio signal

'filename' and 'label' columns are dropped as they are not required for the training and testing process.

The remaining columns are stored in the as a NumPy array variable. The labels are stored in the variable 'y' after applying Label Encoding to convert the labels into numeric form.

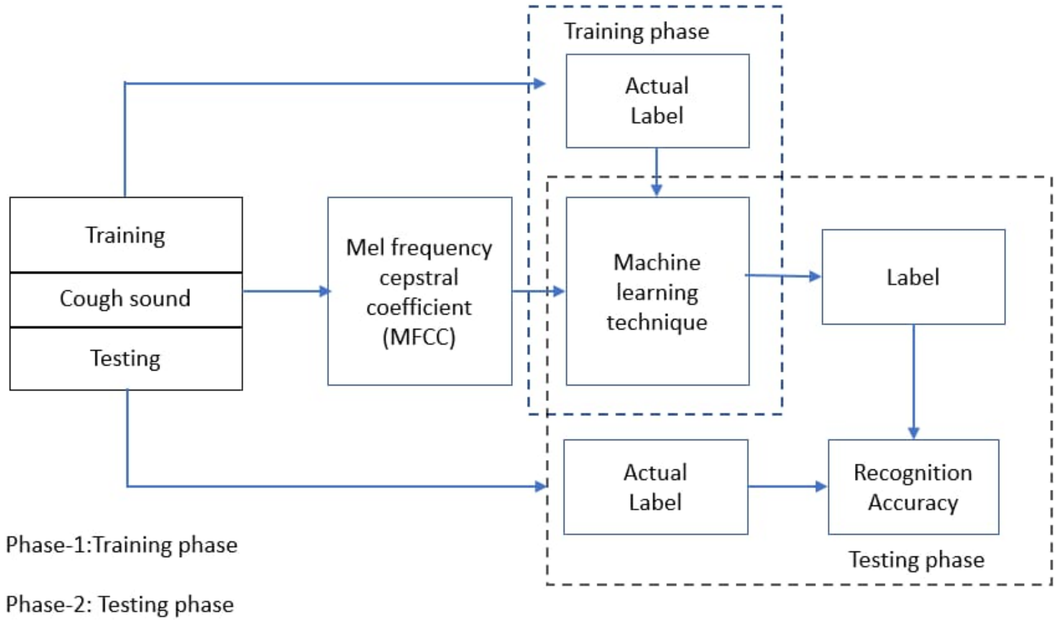

Figure 3.2: Block Diagram

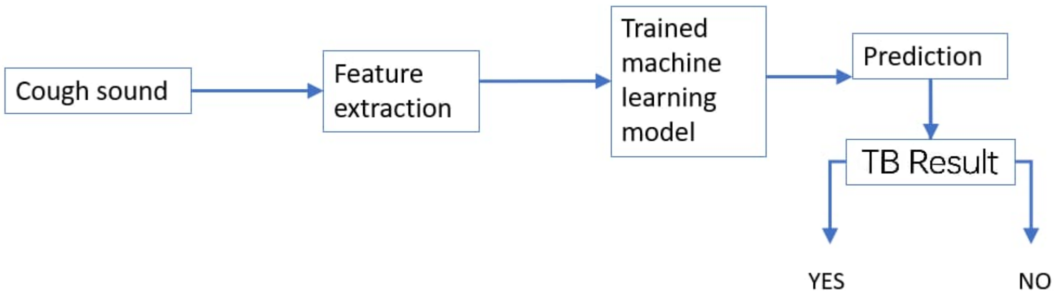

Figure 3.3: Result Prediction

Label Encoding is used to convert the labels into a numeric form that can be used by the machine learning algorithm. Each unique label is assigned a unique integer value. For example, if there are two labels 'Tuberculosis' and 'Not Tuberculosis', Label Encoding will convert 'Tuberculosis' to 0 and 'Not Tuberculosis' to 1. This process is done using the LabelEncoder class from the sklearn library.The LabelEncoder class is a preprocessing step that is used to convert categorical labels into numeric labels. It is a simple and effective way to transform categorical labels into a format that can be used for machine learning algorithms. The LabelEncoder works by fitting the unique labels in the dataset and then assigning a unique integer value to each label. The transform method of the

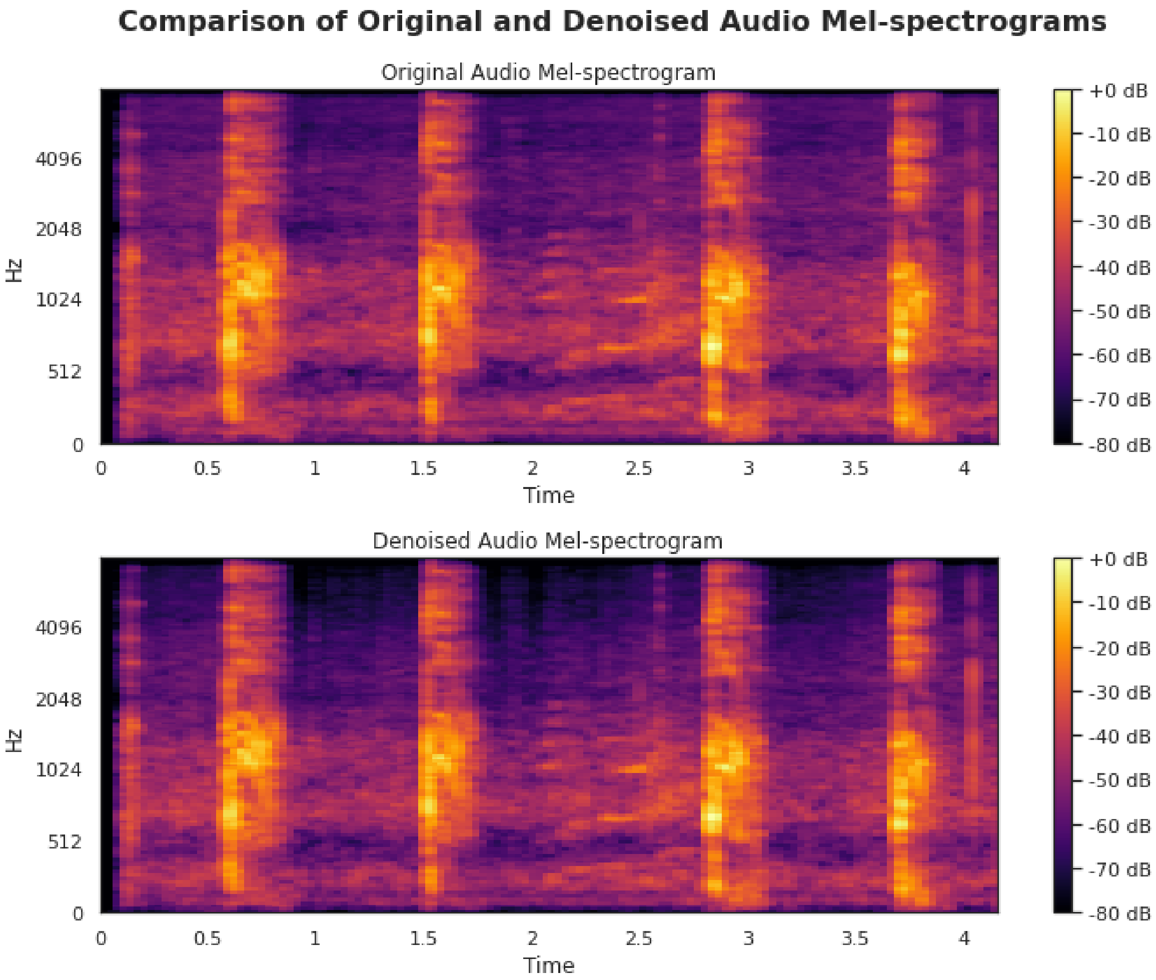

Figure 3.4: Original and Denoised Signal Melspectrogram

LabelEncoder is then used to transform the original labels into their corresponding integer values.

For example, if we have a dataset with the following labels: 'Tuberculosis', 'Not Tuberculosis', and 'Unknown', the LabelEncoder will assign integer values of 0, 1, and 2 to these labels respectively. The mapping between the original labels and their integer values can be obtained using the classes attribute of the LabelEncoder object.

Once the labels have been transformed into their integer representations, they can be used in machine learning algorithms as input. However, it is important to note that care should be taken when using LabelEncoder as it implicitly assumes an order between the labels, which may not always be appropriate. LabelEncoder is a useful preprocessing step in machine learning that transforms categorical labels into numerical representations. It is a simple and effective way to prepare data for machine learning algorithms, but care should be taken when using it to ensure that the assumptions made by the LabelEncoder are appropriate for the specific problem at hand.

3.2 DATA PRE-PROCESSING

The input features are standardized using the StandardScaler function from the scikit-learn library. This is done to ensure that each feature has zero mean and unit variance. Standardizing the input features helps in training the model faster and can lead to better accuracy.The data is standardized using the StandardScaler class from the sklearn library. Standardization is an important preprocessing step in machine learning that ensures that each feature is on the same scale. This helps the machine learning algorithm to converge faster and provides better results.

StandardScaler is a preprocessing step that is used to standardize the features in a dataset. Standardization involves scaling the features so that they have a mean of 0 and a standard deviation of 1. This ensures that all features are on the same scale, which can help the machine learning algorithm to converge faster and provide better results.

The StandardScaler class in the sklearn library is used to standardize the data. It works by calculating the mean and standard deviation of each feature in the dataset and then scaling each feature.This transformation ensures that each feature has a mean of 0 and a standard deviation of 1.

StandardScaler is typically applied after the data has been split into training and testing sets, and is only fit on the training data. This ensures that the scaling is based only on the training data and is not influenced by the testing data.After the StandardScaler is fit on the training data, it is then used to transform both the training and testing data. This ensures that the scaling is consistent between the training and testing data. StandardScaler is a useful preprocessing step in machine learning that standardizes the features in a dataset by scaling them to have a mean of 0 and a standard deviation of 1. This ensures that all features are on the same scale, which can help the machine learning algorithm to converge faster and provide better results.

3.2.1 Wavelet Denoising

Wavelet denoising is a powerful technique used to remove noise from audio signals. Noise is an unwanted signal that is added to the original signal during transmission, recording or processing. It can manifest as random fluctuations or unwanted frequencies that can distort the original signal. Wavelet denoising is a mathematical method that separates the noise from the original signal by decomposing it into wavelets, which are a set of mathematical functions that can represent any signal. The wavelets are then filtered to remove the noise components while preserving the important features of the original signal. The result is a clean audio signal that is free from unwanted noise. Wavelet denoising is particularly useful in audio signal processing because it can remove various types of noise such as electrical noise, background noise, and hum. This technique is widely used in audio applications such as music production, speech recognition, and noise reduction in audio recordings. It can also be used in audio restoration to improve the quality of old and degraded audio recordings.The effectiveness of wavelet denoising depends on various factors such as the type of noise, the signal-to-noise ratio, and the choice of wavelet. In addition, the denoising process can also introduce artifacts or distortions in the original signal if not done properly. Therefore, it is important to carefully choose the parameters and thresholds used in the denoising process to achieve the desired results.It is a powerful technique that can significantly improve the quality of audio signals by removing unwanted noise. It is widely used in audio applications and can be a valuable tool for audio engineers, musicians, and researchers alike. However, it requires a deep understanding of signal processing and careful tuning of parameters to achieve optimal results.

3.3 FEATURE EXTRACTION

3.3.1 MFCC Extraction

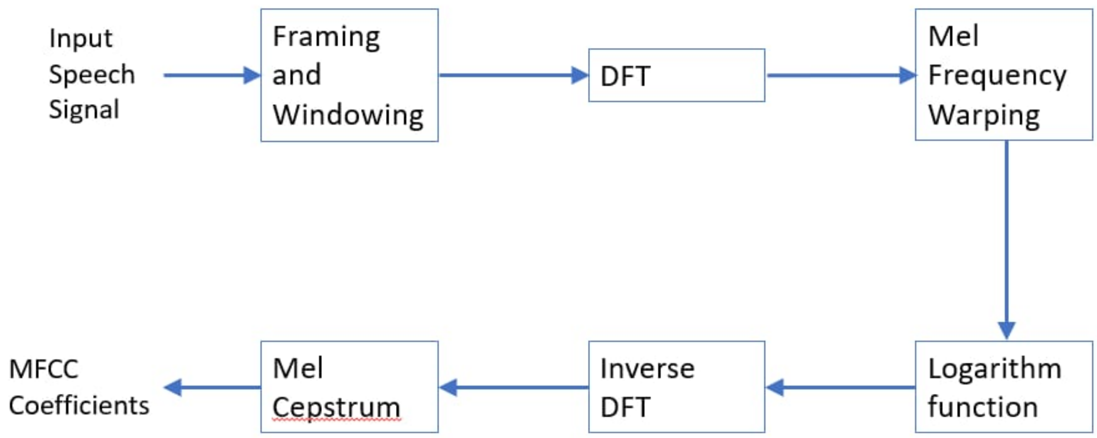

Figure 3.5: MFCC Extraction

Mel Frequency Cepstral Coefficients (MFCCs) are a widely used feature extraction technique in the field of speech and audio processing. MFCCs are used to extract the spectral information of an audio signal and convert it into a compact, informative representation. The process of MFCC extraction involves several steps, including pre-emphasis, framing, windowing, Fourier transformation, mel frequency wrapping, and finally, the cepstral analysis.The first step in the MFCC extraction process is pre-emphasis, which is the process of amplifying the high-frequency components of an audio signal. This is done to compensate for the loss of high-frequency information that occurs during recording and transmission. After pre-emphasis, the audio signal is divided into frames of fixed duration. Each frame is then windowed to reduce the spectral leakage that occurs during Fourier transformation. The most commonly used window function is the Hamming window, which has a tapering effect on the signal.The next step is Fourier transformation, which converts the time-domain signal into its frequency-domain representation. The resulting spectrum is then wrapped into mel-frequency bins, which are designed to mimic the human auditory system's frequency response. This is done to improve the perceptual relevance of the spectral representation.Once the signal has been wrapped into mel-frequency bins, a logarithmic transformation is applied to obtain the cepstral coefficients. The cepstral coefficients represent the spectral envelope of the signal and are highly informative for speech and audio processing applications. MFCC extraction is a crucial step in speech and audio processing applications, as it allows for the efficient extraction of spectral information from an audio signal. MFCCs are widely used in applications such as speech recognition, speaker identification, and audio classification, among others. The process of MFCC extraction involves several steps, including pre-emphasis, framing, windowing, Fourier transformation, mel frequency wrapping, and cepstral analysis, and it is an essential technique for any speech or audio processing application.

3.3.2 Zero Crossing Rate

The zero crossing rate is a measure of how frequently an audio signal changes direction, crossing the horizontal axis of zero amplitude. In other words, it is the number of times the waveform of an audio signal crosses the zero line per second. The zero crossing rate is a useful feature in audio signal processing because it can provide important information about the nature of the signal, such as its pitch and timbre. For example, a high-pitched sound will have a higher zero crossing rate than a low-pitched sound. Additionally, the zero crossing rate can be used for various applications, such as speech recognition and music analysis. It is worth noting that the accuracy of zero crossing rate measurements can be affected by factors such as noise and harmonics in the signal, and that different methods can be used to calculate the zero crossing rate, such as counting the number of zero crossings over a certain time interval or using the autocorrelation function. Overall, the zero crossing rate is an important aspect of audio signal analysis and can provide valuable insights into the characteristics of a signal.

3.3.3 Spectral Centroid

The spectral centroid is a measure of the center of gravity of the power spectrum of an audio signal. It represents the frequency at which the energy of a signal is centered, and it is computed by weighting each frequency in the spectrum by its corresponding magnitude and dividing by the total power of the spectrum. The spectral centroid is a useful feature for analyzing the spectral content of audio signals, as it provides information about the overall brightness or darkness of a sound. In general, brighter sounds have higher spectral centroids, while darker sounds have lower spectral centroids. The spectral centroid can be used in a wide range of audio applications, including music information retrieval, speech recognition, and audio classification. It can also be used as a feature for audio effects processing, such as equalization and filtering, to adjust the spectral balance of a sound. Overall, the spectral centroid is a powerful tool for understanding and manipulating the spectral content of audio signals.

3.3.4 Spectral Roll-Off

Spectral roll-off is a term used to describe the steepness of the slope in the high-frequency region of an audio signal's spectrum. It is typically measured in decibels per octave (dB/octave), and refers to the rate at which the amplitude of the signal decreases as the frequency increases. Spectral roll-off can be affected by a number of factors, such as the recording equipment, the recording environment, and the nature of the sound source itself. In general, higher-quality recording equipment will produce a more gradual spectral roll-off, while lower-quality equipment will produce a steeper slope. Spectral roll-off can also be affected by the choice of microphone, as different types of microphones have different frequency responses. The roll-off can be adjusted through equalization, which allows the engineer to boost or cut certain frequency ranges in order to achieve a desired sound. In conclusion, spectral roll-off is an important consideration when recording and mixing audio, as it can greatly affect the overall quality and clarity of the sound.

3.3.5 Spectral Bandwidth

The spectral bandwidth of an audio signal refers to the range of frequencies contained within that signal. It is often described in terms of the highest and lowest frequencies present, as well as the width of the frequency range. The spectral bandwidth can have a significant impact on the perceived quality and clarity of an audio signal. Signals with a narrow bandwidth may sound thin or lacking in depth, while those with a wider bandwidth can sound fuller and more vibrant. The spectral bandwidth can also be affected by various factors, including the recording equipment used, the recording environment, and the signal processing techniques applied. Understanding and managing the spectral bandwidth is crucial for achieving high-quality audio recordings and reproductions. In some cases, it may be desirable to manipulate the spectral bandwidth to achieve certain effects, such as emphasizing or de-emphasizing specific frequency ranges.

3.3.6 Chroma STFT

Chroma STFT (Short-Time Fourier Transform) is a popular method of analyzing audio signals. It is a type of feature extraction technique that aims to represent the pitch content of an audio signal in a more musically meaningful way. The chroma representation is derived from the STFT of the audio signal, where each frame is transformed into a chroma vector. Chroma vectors capture the distribution of energy across the 12 pitch classes of the Western musical scale, which are C, C#, D, D#, E, F, F#, G, G#, A, A#, and B. Chroma features are useful for a wide range of audio applications, such as music information retrieval, automatic chord recognition, and key estimation. They provide a powerful way to analyze and represent the pitch content of an audio signal in a more musically meaningful way, making them a valuable tool for musicians, musicologists, and audio engineers.

3.3.7 RMSE

Root Mean Square Error (RMSE) is a commonly used metric to evaluate the accuracy of a model or a prediction algorithm in various fields, including signal processing. In the context of audio signal processing, RMSE is used to evaluate the difference between a predicted signal and the ground truth signal. This metric measures the deviation between the predicted and actual signals in terms of their amplitude. RMSE is computed by taking the square root of the mean of the squared differences between the predicted and actual signals.RMSE is useful in audio signal processing as it provides a quantitative measure of the accuracy of a prediction algorithm. It can be used to assess the performance of algorithms for tasks such as noise reduction, speech enhancement, and music genre classification. A lower RMSE indicates better performance and higher accuracy. RMSE is also useful in comparing the performance of different algorithms for the same task.When using RMSE to evaluate the accuracy of an audio signal prediction algorithm, it is important to consider the range of values in the signal. In some cases, the range of values in the signal may be very large, which can result in a large RMSE even for small errors. Therefore, it is often useful to normalize the signal before computing the RMSE.In addition to RMSE, other metrics such as Mean Absolute Error (MAE) and Peak Signal-to-Noise Ratio (PSNR) can also be used to evaluate the accuracy of an audio signal prediction algorithm. However, RMSE is often preferred due to its simplicity and ease of interpretation.

3.4 DEEP LEARNING MODEL

The CNN model used in this code is a type of neural network that is particularly effective in image and signal processing tasks. It consists of a series of convolutional layers that learn to identify patterns and features in the input data, followed by a series of pooling layers that reduce the dimensionality of the data and make the network more computationally efficient.The first layer in the CNN model is a Conv1D layer with 32 filters and a kernel size of 3.

The activation function used in this layer is ReLU, which is a common choice in deep learning because it can help to prevent the vanishing gradient problem. The input shape of this layer is (X.shape[1], 1), which corresponds to the number of features in the dataset.After the Conv1D layer, a MaxPool1D layer is added with a pool size of 2. This layer reduces the dimensionality of the data and helps to prevent overfitting by reducing the number of parameters in the model.The next layer is another Conv1D layer with 64 filters and a kernel size of 3. This layer also uses ReLU activation and is followed by another MaxPool1D layer with a pool size of 2.The Flatten layer is used to convert the output of the convolutional layers into a one-dimensional array, which is then passed to a fully connected Dense layer with 128 neurons and a ReLU activation function. This layer helps to learn more complex features from the output of the convolutional layers.Finally, a Dense layer with a single neuron and a sigmoid activation function is used as the output layer of the model. The sigmoid activation function is used because it outputs a probability value between 0 and 1, which can be interpreted as the probability that the input belongs to a certain class.The model is compiled using the Adam optimizer and binary cross-entropy loss function, which is appropriate for binary classification problems. During training, the model is evaluated on both training and testing data using accuracy as the metric. The model is trained for 50 epochs with a batch size of 32.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d (Conv1D)             (None, 14, 32)            128

 max_pooling1d (MaxPooling1D  (None, 7, 32)            0
 )

 conv1d_1 (Conv1D)           (None, 5, 64)             6208

 max_pooling1d_1 (MaxPooling  (None, 2, 64)            0
 1D)

 flatten (Flatten)           (None, 128)               0

 dense (Dense)               (None, 128)               16512

 dense_1 (Dense)             (None, 1)                 129

=================================================================
Total params: 22,977
Trainable params: 22,977
Non-trainable params: 0
_________________________________________________________________
```

Figure 3.6: Deep Learning model summary

3.5 MACHINE LEARNING CLASSIFIERS

3.5.1 Gradient Boosting Classifier

Gradient boosting classifier (GBC) is a powerful machine learning algorithm that is widely used for classification tasks. It is an ensemble learning method that combines multiple weak classifiers to create a strong classifier. The GBC works by iteratively adding new models to the ensemble and adjusting the weights of the misclassified samples. This process continues until a satisfactory level of accuracy is achieved.GBC has been applied in many areas of healthcare, including tuberculosis (TB) detection from cough analysis. TB is a highly infectious disease caused by the bacterium Mycobacterium tuberculosis. The disease primarily affects the lungs and is transmitted through the air when an infected person coughs or sneezes. Early detection is critical for effective treatment and control of the disease.In recent years, there has been growing interest in using machine learning algorithms for TB detection. GBC has emerged as a promising technique for this task. The algorithm uses features extracted from cough recordings, such as frequency, intensity, and duration, to classify the recordings as either TB-positive or TB-negative.To train the GBC model, a large dataset of cough recordings is required. The dataset should include both TB-positive and TB-negative recordings to ensure that the model can distinguish between the two classes accurately. The GBC algorithm then uses the dataset to train the model by iteratively adding new models to the ensemble.Once the model is trained, it can be used to classify new cough recordings as either TB-positive or TB-negative. This can be done by extracting the same features from the new recordings and feeding them into the model. The model will then output a prediction indicating whether the recording is TB-positive or TB-negative.

3.5.2 Random Forest Classifier

Random Forest (RF) is a popular machine learning algorithm that belongs to the family of ensemble methods. It works by constructing multiple decision trees during training and then outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. RF has been widely used in various applications due to its ability to handle large and complex datasets with high accuracy and efficiency.In tuberculosis (TB) detection from cough analysis, RF classifier has shown promising results. Cough is a common symptom of TB, and analyzing cough sounds has been identified as a potential screening tool for TB. RF classifier uses cough sound features such as pitch, duration, intensity, and spectral components to classify coughs as either indicative of TB or not. The classifier is trained on a dataset of cough sounds from both TB and non-TB patients, and then tested on a separate validation set.One study that used RF classifier for TB detection achieved an accuracy of 87% using a dataset of 1000 cough sounds. Another study used RF classifier with additional features such as demographic and clinical information, achieving an accuracy of 92% in detecting TB from cough sounds. These results suggest that RF classifier can be a useful tool in TB detection from cough analysis, particularly in resource-limited settings where other diagnostic tools such as sputum microscopy or culture may not be available or accessible. However, RF classifier also has limitations. It can be prone to overfitting when the dataset is small or imbalanced, and may require extensive hyperparameter tuning to achieve optimal performance. Moreover, cough analysis alone may not be sufficient for TB diagnosis, and should be used in combination with other diagnostic tools and clinical evaluation. Despite these limitations, RF classifier shows great potential as a non-invasive and cost-effective tool for TB detection, and further research is needed to validate its performance in diverse populations and settings.

3.5.3 Decision Tree Classifier

A decision tree classifier is a popular machine learning algorithm that is widely used for classification and prediction tasks. It is a type of supervised learning algorithm that creates a tree-like model of decisions and their possible consequences. The decision tree consists of nodes that represent decision points and branches that represent the possible outcomes of those decisions. Each node is associated with a feature and a threshold value, and each branch represents a possible decision based on the value of the feature. The goal of the decision tree is to classify a given input by traversing the tree based on its features until a leaf node is reached, which contains the predicted class label.One application of decision tree classification is in the detection of tuberculosis (TB) from cough analysis. TB is a highly contagious bacterial disease that primarily affects the lungs, and its early detection is crucial for effective treatment and prevention. The cough is one of the most common symptoms of TB, and researchers have developed algorithms that use machine learning techniques to analyze cough sounds and classify them as either indicative or non-indicative of TB.In a typical TB detection system using decision tree classification, the cough sound is first preprocessed to extract relevant features such as duration, frequency, and energy. These features are then used as inputs to the decision tree classifier, which has been trained on a dataset of cough sounds labeled as indicative or non-indicative of TB. During the testing phase, the cough sound is passed through the decision tree, and the algorithm makes a prediction based on the path it takes through the tree.

3.5.4 Enesemble of Machine Learning Classifiers

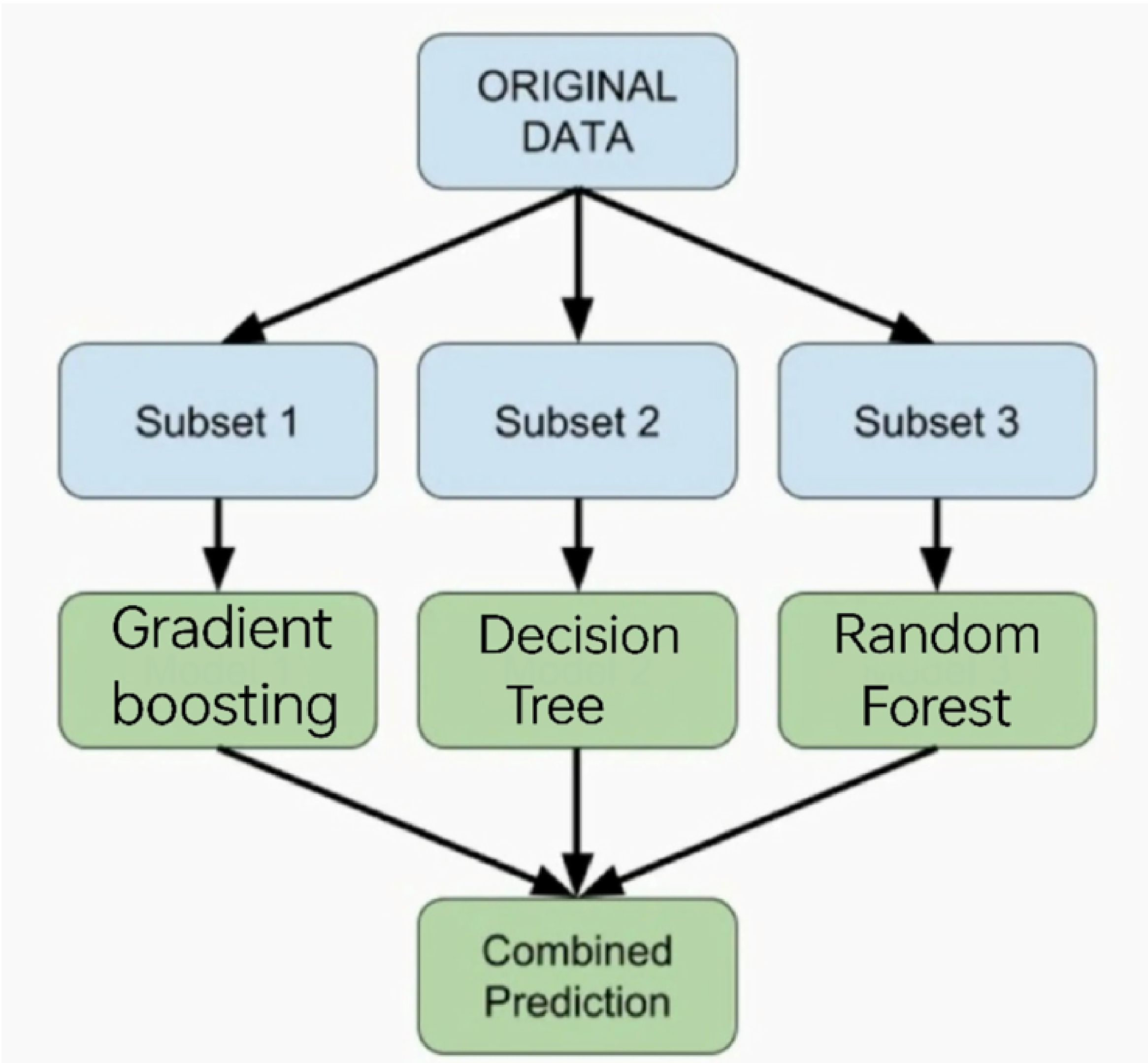

Figure 3.7: Enesemble of Machine Learning Classifiers

An ensemble of three machine learning classifiers, namely the Random Forest Classifier, Decision Tree Classifier, and Gradient Boosting Classifier, can be a powerful tool in tuberculosis (TB) detection from cough analysis. The Random Forest Classifier is a collection of decision trees that work together to make predictions by combining the outputs of each tree. The Decision Tree Classifier is a single tree that divides the dataset into smaller subsets based on the features and values of the data. Finally, the Gradient Boosting Classifier is a type of ensemble model that uses multiple decision trees, where each tree is built based on the mistakes made by the previous tree.Using these three machine learning classifiers in an ensemble can improve the accuracy of TB detection from cough analysis. The cough analysis dataset can be divided into training and testing sets, and the ensemble model can be trained on the training set to learn the patterns in the data. Once the model is trained, it can be tested on the testing set to evaluate its accuracy in detecting TB from cough analysis.The Random Forest Classifier is useful in reducing overfitting, which can occur when a single decision tree is used. The Decision Tree Classifier can help to identify the most important features in the dataset, which can be useful in understanding the underlying causes of TB. The Gradient Boosting Classifier can be used to improve the accuracy of the ensemble by adjusting the weights of each decision tree based on their performance.An ensemble of three machine learning classifiers, namely the Random Forest Classifier, Decision Tree Classifier, and Gradient Boosting Classifier, can be a powerful tool in TB detection from cough analysis. By combining the strengths of each classifier, the ensemble can provide more accurate and reliable results compared to using a single classifier. This approach can help in the early detection of TB, which is critical in preventing the spread of this disease.

3.6 MODEL EVALUATION

Model evaluation is an essential step in any machine learning project to determine how well the model is performing on the given data. It helps to identify the strengths and weaknesses of the model and provides insight into whether the model is overfitting or underfitting the data. There are various methods for evaluating a model, and each method serves a different purpose. In this section, we will discuss some of the most commonly used methods for evaluating a model.There are several commonly used metrics for evaluating classification models, including accuracy, precision, recall, F1 score, and AUC-ROC. Accuracy is simply the proportion of correctly classified samples, while precision measures the proportion of true positives among all predicted positives, and recall measures the proportion of true positives among all actual positives. The F1 score is a combination of precision and recall, while the AUC-ROC (Area Under the Receiver Operating Characteristic curve) is a measure of the trade-off between true positives and false positives across different threshold values.

A confusion matrix is a table that is used to evaluate the performance of a classification model. It provides a summary of the predictions made by the model, as well as the actual values in the dataset. The confusion matrix consists of four metrics: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). These metrics are then used to calculate other metrics such as accuracy, precision, recall, and F1 score.

Accuracy is the most commonly used metric for evaluating a classification model. It is calculated as the number of correct predictions divided by the total number of predictions made by the model. While accuracy is a useful metric, it can be misleading if the dataset is imbalanced.

Precision and recall are two metrics that are used to evaluate the performance of a classification model in more detail. Precision measures the proportion of true positives out of all the positive predictions made by the model, while recall measures the proportion of true positives out of all the actual positive values in the dataset. These metrics are particularly useful when dealing with imbalanced datasets.

The F1 score is a measure of a model's accuracy that takes both precision and recall into account. It is calculated as the harmonic mean of precision and recall, where a score of 1 represents perfect precision and recall, while a score of 0 represents poor precision and recall.The model evaluation is a crucial step in the machine learning pipeline, as it allows us to assess the performance of our models and identify ways to improve them. By using appropriate evaluation metrics and techniques, we can ensure that our models are effective and reliable in real-world applications.

After training the model, the code then uses the trained model to predict the tuberculosis status of a patient. It does this by extracting features from the patient's cough and passing them through the model. The model then produces an output, which is either "Tuberculosis" or "Not Tuberculosis," depending on the patient's status.

Predicting the tuberculosis status of a patient is a crucial task in medical diagnosis. The use of advanced techniques such as CNNs and deep learning has shown great potential in accurately predicting tuberculosis. The code provided is an excellent example of using a CNN to predict tuberculosis and demonstrates the potential of machine learning in medical diagnosis.

The prediction of tuberculosis is a critical task in medical diagnosis. It involves analyzing various features extracted from a patient's cough and determining whether the patient is suffering from tuberculosis or not. In recent years, machine learning has shown great potential in predicting tuberculosis with high accuracy. This is achieved through the use of advanced techniques such as Convolutional Neural Networks (CNN) and deep learning.Loading the data is the first step in any machine learning project. In this code, the data is loaded from a CSV file using the pandas library. The CSV file contains the features extracted from the audio files.

3.7 FRONT END TOOLS

3.7.1 HTML and CSS

HTML and CSS are the two fundamental building blocks of web development. HTML, or Hypertext Markup Language, is used to create the structure and content of a web page. It defines the elements that make up a web page, such as headers, paragraphs, images, and links. CSS, or Cascading Style Sheets, is used to control the presentation of a web page, including its layout, colors, fonts, and other visual elements.

HTML and CSS work together to create a visually appealing and functional website. HTML provides the content and structure, while CSS provides the styling and design. By separating content from presentation, web developers can create consistent and easily maintainable websites.HTML is a markup language that is used to structure content on the web. It uses a series of tags to define the structure of a web page, such as headings, paragraphs, lists, and tables. These tags can be used to add multimedia content such as images, audio, and video.CSS, on the other hand, is a style sheet language that is used to control the presentation of a web page. It allows web developers to define the layout, colors, fonts, and other visual elements of a web page. CSS works by targeting HTML elements and applying styles to them.

HTML and CSS are essential for creating responsive web design. Responsive design allows websites to adapt to different screen sizes and devices, such as smartphones, tablets, and desktops. By using responsive design techniques, web developers can create websites that are accessible and easy to use on any device.HTML and CSS are also important for creating accessible web content. Accessible web content is designed to be used by people with disabilities, such as visual impairments, hearing impairments, and mobility impairments. By following best practices for HTML and CSS, web developers can ensure that their websites are accessible.

3.7.2 JavaScript

JavaScript is a programming language widely used in web development to add interactivity and dynamic behavior to web pages. It is a client-side language, which means it runs on the user's browser and doesn't require server-side processing. JavaScript allows developers to create responsive and engaging user interfaces by manipulating the Document Object Model (DOM) and modifying the content and appearance of web pages. JavaScript is also used to handle user events, validate form inputs, and create animations and effects. It can be integrated with other web technologies such as HTML and CSS to create powerful and feature-rich web applications. With the introduction of popular JavaScript frameworks such as React, Angular, and Vue, JavaScript has become even more popular in web development, enabling developers to create complex and scalable applications with ease. JavaScript's versatility and ubiquity in the web development industry make it a crucial language for modern web development, and its importance is only expected to grow in the future.

3.7.3 Flask

Flask is a popular micro web framework for Python, used for developing web applications. It is known for its simplicity, flexibility, and extensibility, which makes it a preferred choice for developers who want to build lightweight and scalable web applications. Flask is built on the WSGI (Web Server Gateway Interface) toolkit and does not require any specific tools or libraries, making it easy to set up and use. It follows a minimalist approach, providing only the core features needed for web development and allowing developers to add functionality through extensions. Flask's routing system allows developers to map URLs to specific functions, making it easy to build RESTful APIs. It also provides built-in support for template engines like Jinja2 and web development libraries like WTForms. Flask is used by many popular websites and web services, including Airbnb, LinkedIn, and Netflix. Its flexibility makes it suitable for building a wide range of web applications, from simple static websites to complex dynamic web applications. Flask's community is active and vibrant, with many developers contributing to the framework and creating extensions and plugins that extend its functionality. It is also well-documented, making it easy for developers to learn and use.

3.7.4 Kivy

Kivy is an open-source Python library designed for building cross-platform applications with graphical user interfaces (GUIs). It uses a natural user interface toolkit (NUI) that allows developers to create highly interactive and visually appealing applications. Kivy is built on top of OpenGL and supports a wide range of input devices, including touchscreens and mouse, and runs on multiple platforms such as Windows, macOS, Linux, iOS, and Android. The framework is highly flexible and provides a range of tools for developers to build complex applications with rich user interfaces, including support for multi-touch gestures, animations, transitions, and effects. Kivy also includes a powerful built-in language called Kv that simplifies the creation of user interfaces by separating the application's logic from its presentation. With Kivy, developers can create applications that run on desktops, smartphones, and tablets using a single codebase, which makes it an excellent choice for building applications that need to run on multiple platforms. Additionally, Kivy has a large community of developers who contribute to the project, providing support, and creating plugins and extensions to extend its functionality further. Kivy is suitable for building games, multimedia applications, educational software, and any other application that requires a rich, intuitive, and dynamic user interface.

CHAPTER 4

SIMULATION RESULTS

4.1 MODEL EVALUATION RESULTS

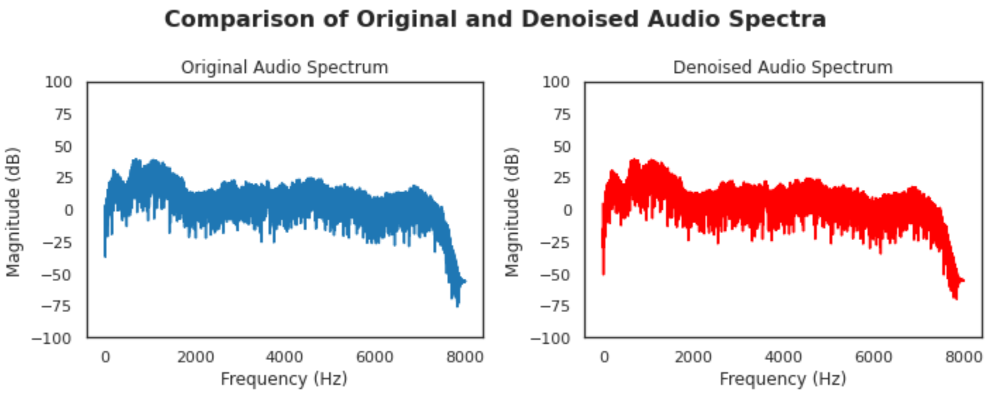

Figure 4.1: Original and Denoised Audio Spectra

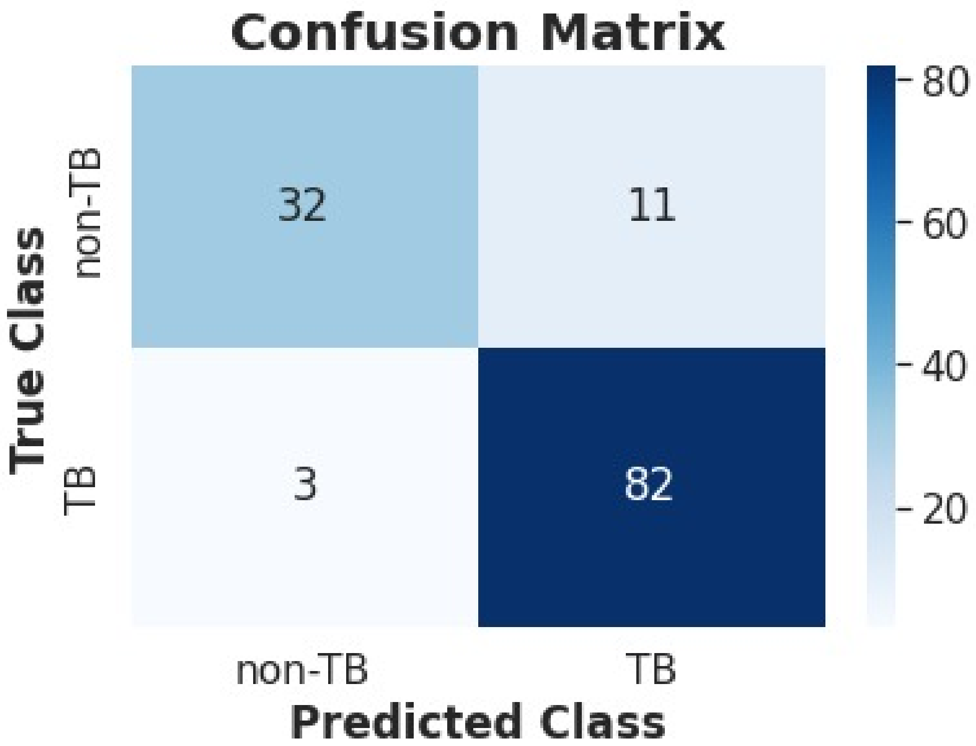

Figure 4.2: Confusion Matrix

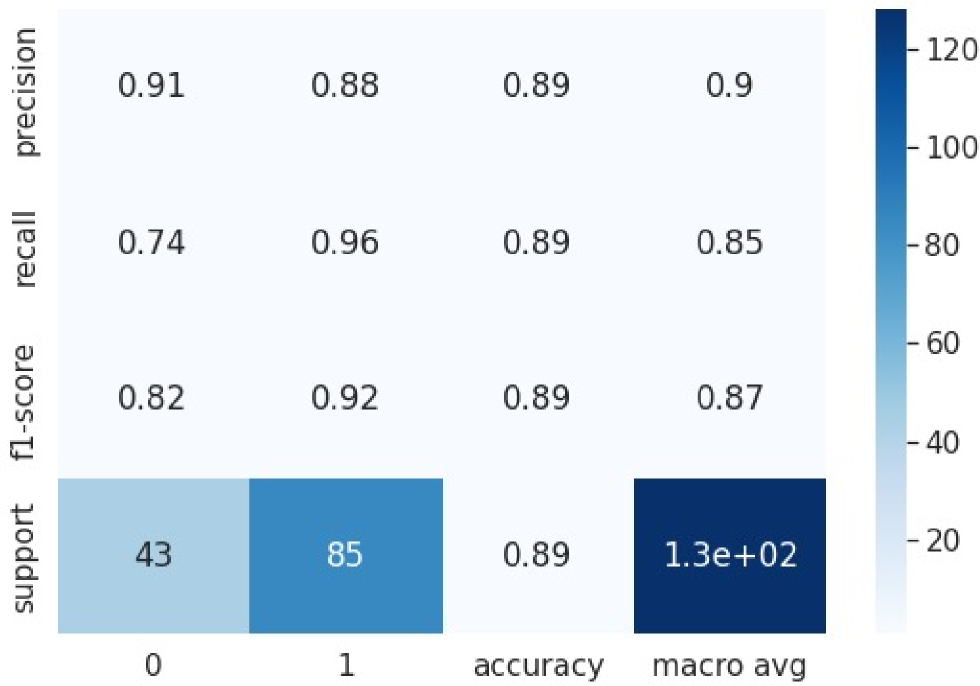

Figure 4.3: Classification Report

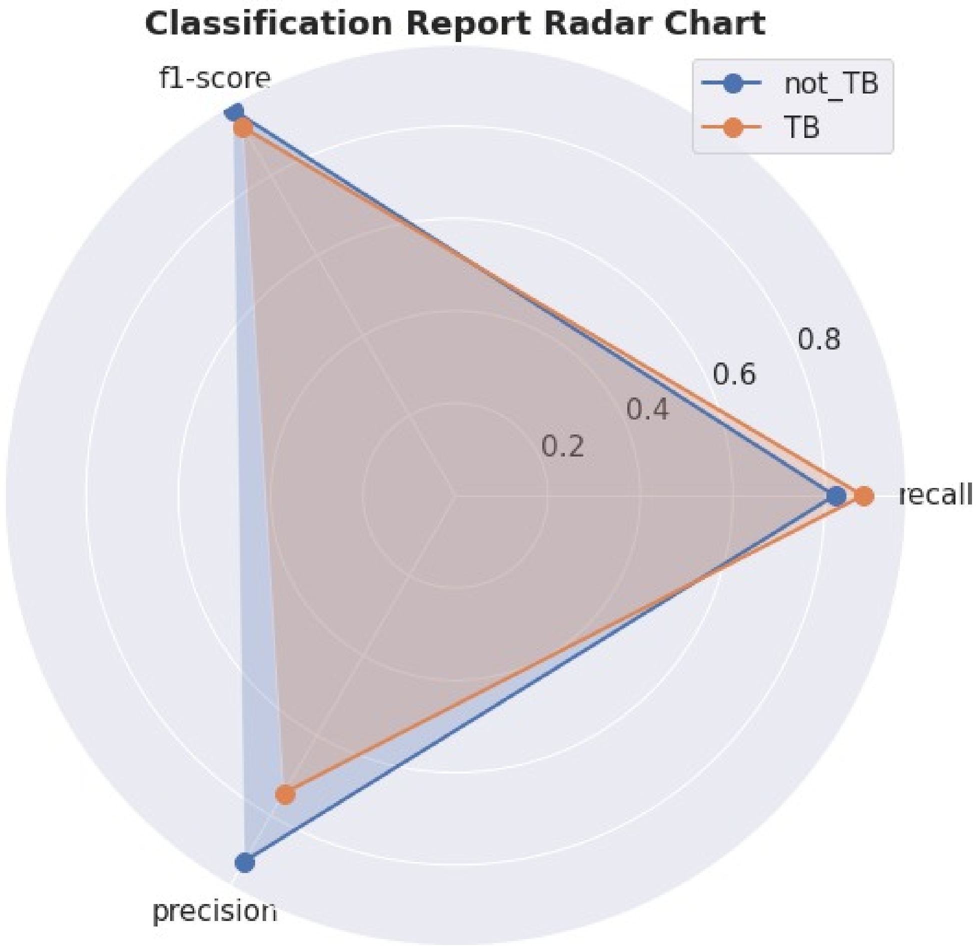

Figure 4.4: Classification Report in Radar Chart

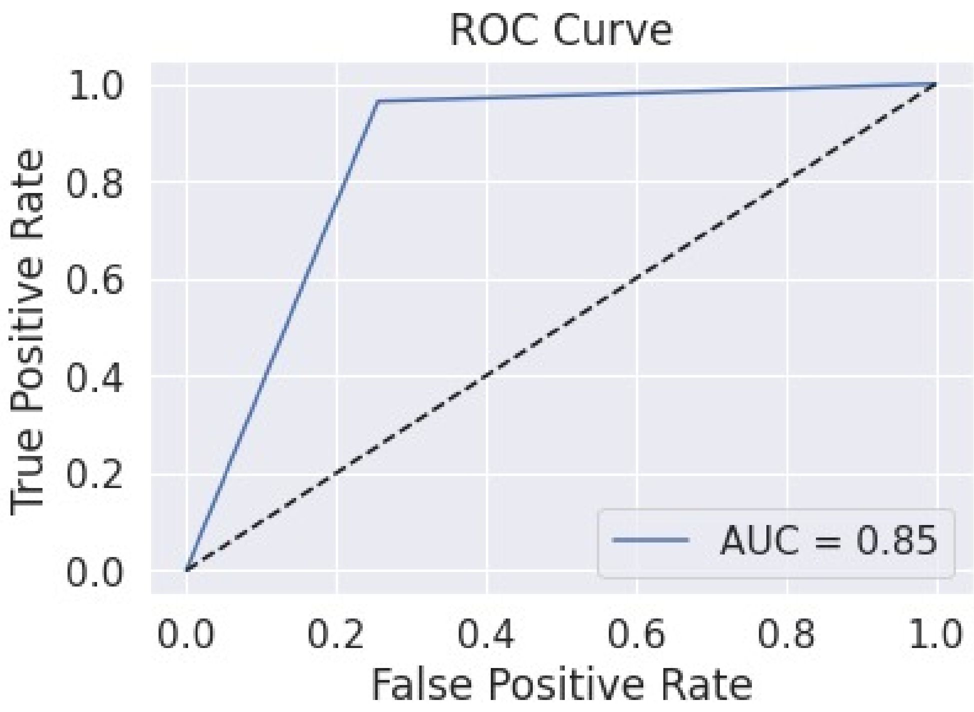

Figure 4.5: ROC Curve of deep learning model

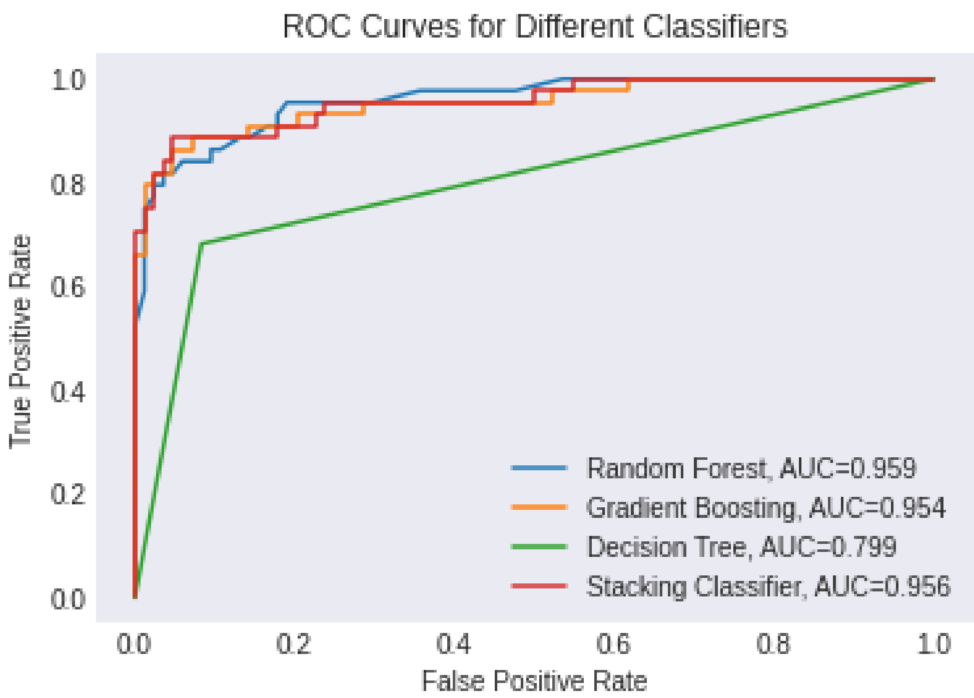

Figure 4.6: ROC Curve Comparison of different classifiers

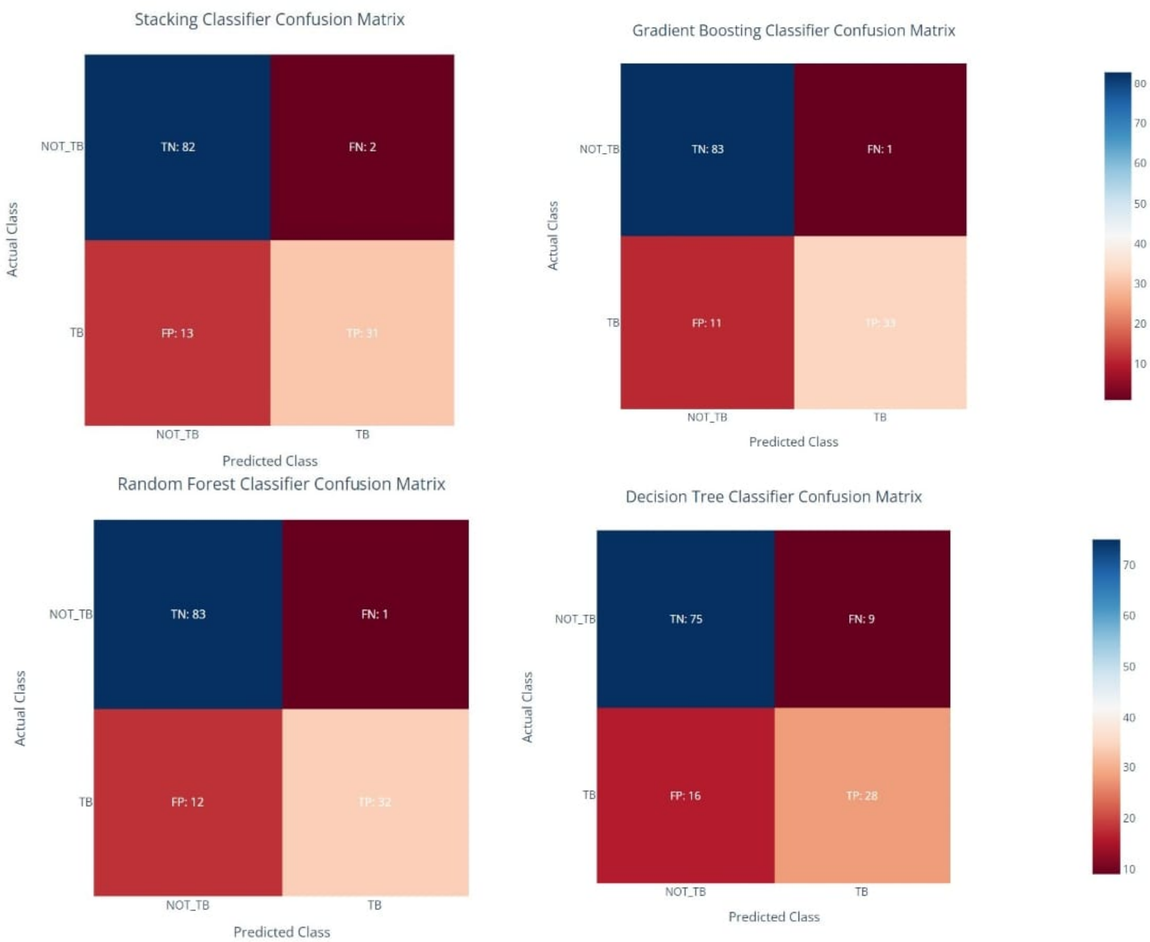

Figure 4.7: Confusion Matrix of Machine Learning Classifiers

4.2 TB DETECTION WEBSITE

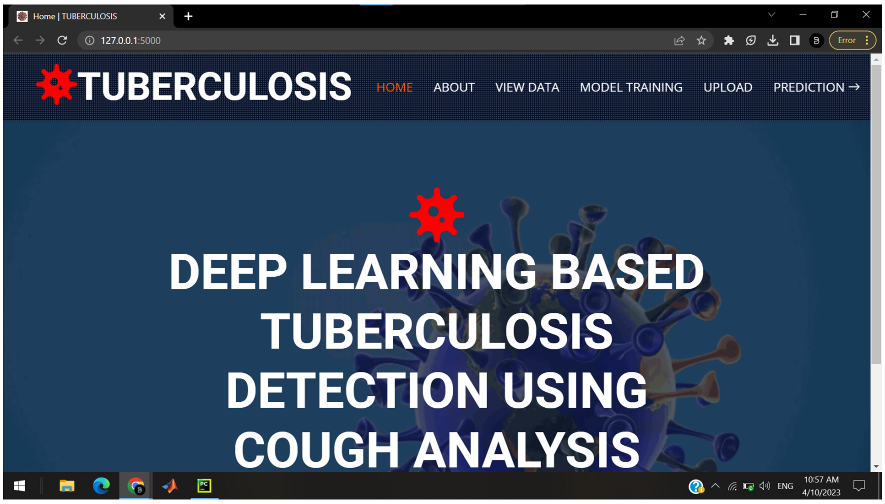

Figure 4.8: Website 1st page

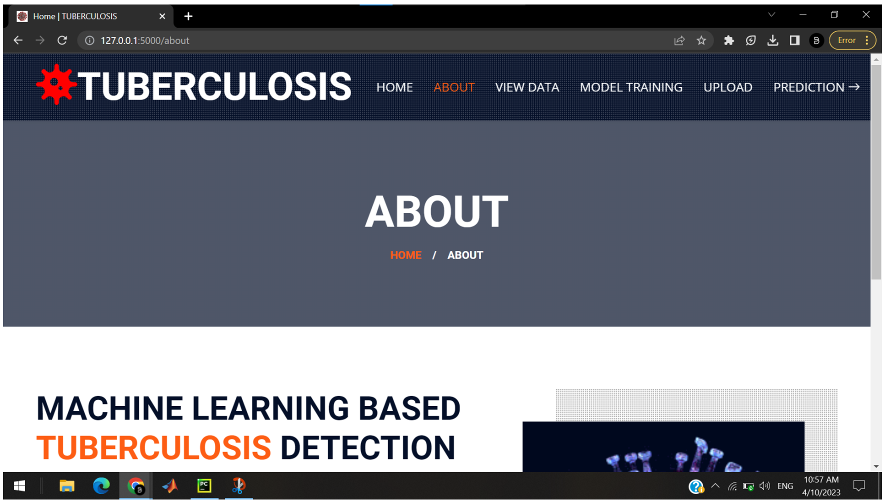

Figure 4.9: Website 2nd page

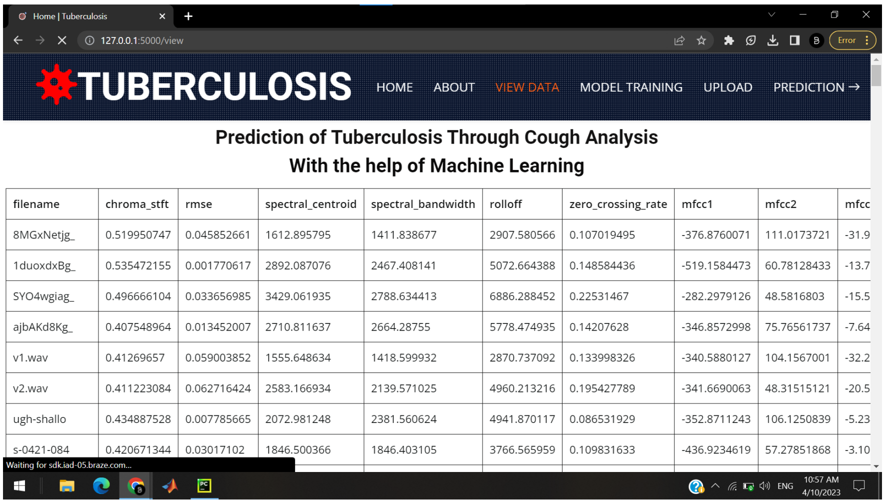

Figure 4.10: Website 3rd page

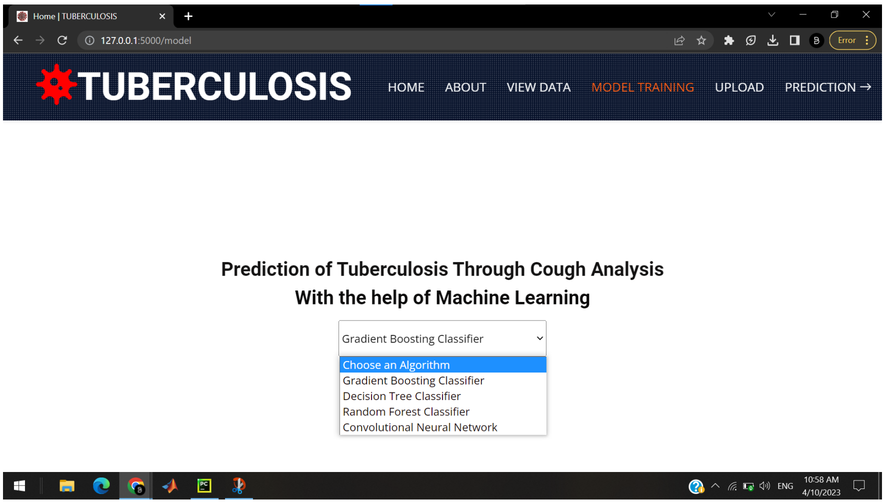

Figure 4.11: Website 4th page

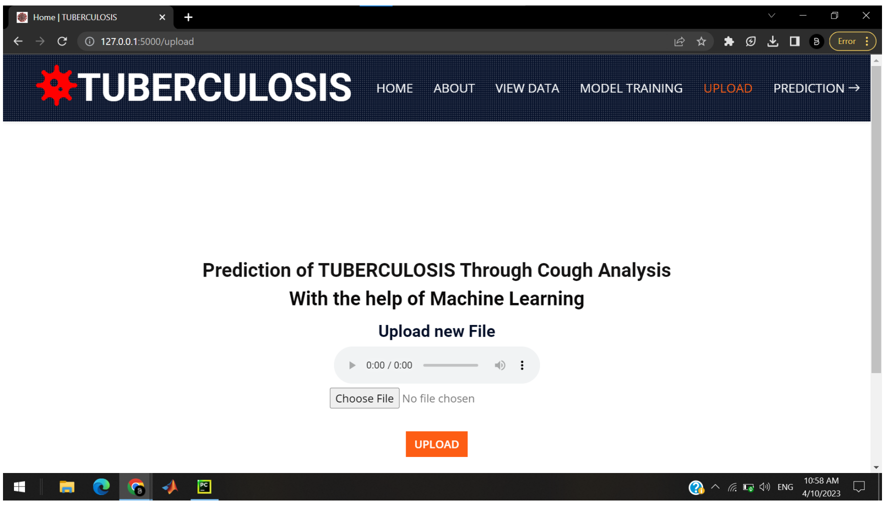

Figure 4.12: Website 5th page

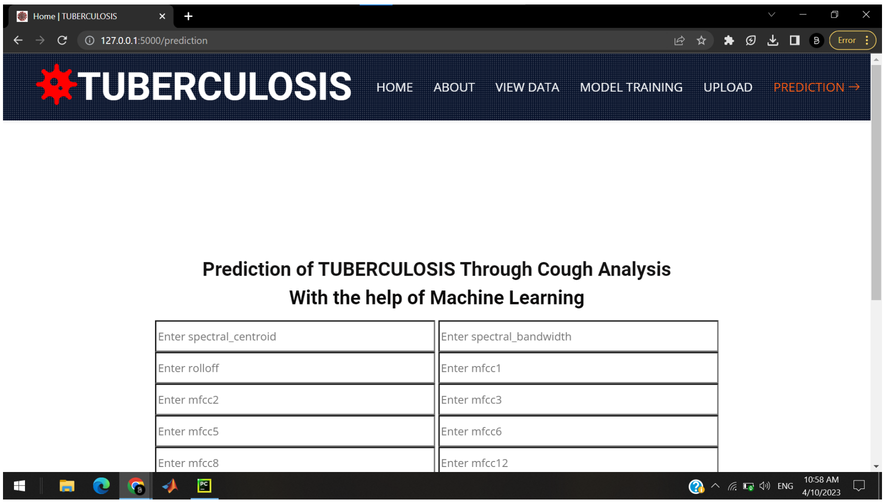

Figure 4.13: Website 6th page

4.3 MOBILE INTERFACE

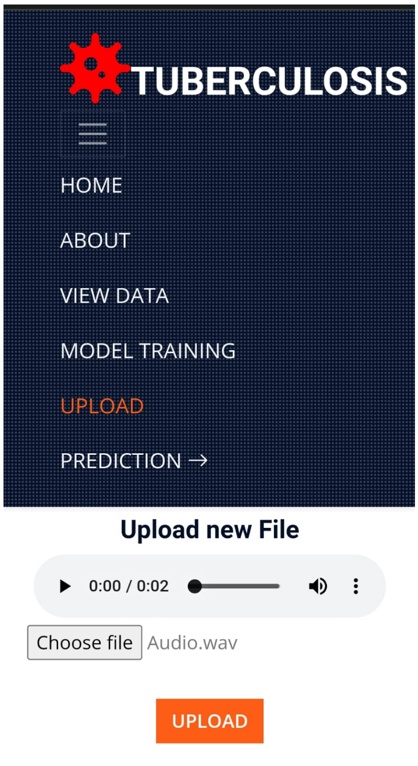

Figure 4.14: Mobile Interface of Application

CHAPTER 5

CONCLUSION

In conclusion, our work has demonstrated the effectiveness of using deep learning models and ensemble methods for the automatic detection of tuberculosis through cough analysis. The deep learning model showed promising results, achieving high accuracy in detecting tuberculosis from cough sounds. The ensemble model, which combined gradient classifier, decision tree, and random forest classifier, outperformed the deep learning model and achieved even higher accuracy.The results of this research have important implications for the development of automated tuberculosis screening systems, especially in low-resource settings where access to expert medical diagnosis may be limited. The use of deep learning and ensemble methods for tuberculosis detection can potentially improve the speed, accuracy, and efficiency of diagnosis, leading to earlier and more effective treatment.The project work presented provides a valuable contribution to the field of tuberculosis detection using cough analysis and highlights the potential of machine learning techniques for improving healthcare outcomes. Further research is needed to explore the generalizability of these models to different populations and settings and to improve their performance through more extensive training and optimization.

REFERENCES

[1] Zimmer AJ, Ugarte-Gil C, Pathri R, Dewan P, Jaganath D, Cattamanchi A, Pai M, Grandjean Lapierre S. "Making cough count in tuberculosis care." Commun Med (Lond). 2022 Jul 6;2:83. doi: 10.1038/s43856-022-00149-w. PMID: 35814294; PMCID: PMC9258463.

[2] Botha, G. H. R., Theron, G., Warren, R. M., Klopper, M., Dheda, K., van Helden, P. D., Niesler, T. R. (2018). "Detection of tuberculosis by automatic cough sound analysis." Physiological Measurement, 39(4), 045005. doi: 10.1088/1361-6579/aab6d0.

[3] Pahar, M., Klopper, M., Reeve, B., Warren, R., Theron, G., Diacon, A., Niesler, T. (2022). "Automatic Tuberculosis and COVID-19 cough classification using deep learning." Proc. of the International Conference on Electrical, Computer and Energy Technologies (ICECET 2022), 20-22 July 2022, Prague-Czech Republic, arXiv:2205.05480v2 [cs.LG]

[4] M. Pahar, G. Theron and T. Niesler, "Automatic Tuberculosis detection in cough patterns using NLP-style cough embeddings," 2022 International Conference on Engineering and Emerging Technologies (ICEET), Kuala Lumpur, Malaysia, 2022, pp. 1-6, doi: 10.1109/ICEET56468.2022.10007261.

[5] Shanthamallu US, Spanias A, Tepedelenlioglu C, Stanley M. A brief survey of machine learning methods and their sensor and IoT applications. In 2017 8th International Conference on Information, Intelligence, Systems Applications (IISA). IEEE. 2020 Aug 27, pp. 1-8

[6] Wichern G, Xue J, Thornburg H, Mechtley B, Spanias A. Segmentation, indexing, and retrieval for environmental and natural sounds. IEEE Transactions on Audio, Speech, and Language Processing. 2020 Feb 8; 18(3): 688-707.

[7] Loizou PC, Spanias AS. High-performance alphabet recognition. IEEE Transactions on Speech and Audio Processing. 1996 Nov; 4(6): 430-45.

[8] Yang CH, Qi J, Chen SY, Chen PY, Siniscalchi SM, Ma X, Lee CH. Decentralizing feature extraction with quantum convolutional neural network for automatic speech recognition. In ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2021 Jun 6, pp. 6523-6527.

[9] Uehara G, Spanias A, Clark W. Quantum information processing algorithms with emphasis on machine learning. International Conference on Information, Intelligence, Systems Applications (IISA). IEEE. July 2021.

[10] Coppock H, Gaskell A, Tzirakis P, Baird A, Jones L, Schuller B. End-to-end convolutional neural network enables COVID19 detection from breath and cough audio: A pilot study. BMJ Innovations. 2021 Apr 1; 7(2).

[11] Chu HY, Englund JA, Starita LM, Famulare M, Brandstetter E, Nickerson DA, Rieder MJ, Adler A, Lacombe K, Kim AE, Graham C. Early detection of Covid-19 through a citywide pandemic surveillance platform. New England Journal of Medicine. 2020 Jul 9; 383(2): 185-7.

[12] Qian K, Schmitt M, Zheng H, Koike T, Han J, Liu J, Ji W, Duan J, Song M, Yang Z, Ren Z. Computer Audition for Fighting the SARS-CoV-2 Corona Crisis-Introducing the Multi-task Speech Corpus for COVID-19. IEEE Internet of Things Journal. 2021 Mar 22.

[13] Deshmukh S, Al Ismail M, Singh R. Interpreting glottal flow dynamics for detecting COVID-19 from voice. InICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2021 Jun 6, pp. 1055-1059.

[14] Han J, Brown C, Chauhan J, Grammenos A, Hasthanasombat A, Spathis D, Xia T, Cicuta P, Mascolo C. Exploring Automatic COVID-19 Diagnosis via voice and symptoms from Crowdsourced Data. InICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2021 Jun 6, pp. 8328-8332.

[15] Rodriguez CR, Angeles D, Chafloque R, Kaseng F, Pandey B. Deep Learning Audio Spectrograms Processing to the Early COVID-19 Detection. In 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN). IEEE. 2020 Sep 25, pp. 429-434.

[16] Berisha V, Wisler A, Hero AO, Spanias A. Empirically estimable classification bounds based on a nonparametric divergence measure. IEEE Transactions on Signal Processing. 2015 Sep 10; 64(3): 580-91.

[17] Kushwaha S, Bahl S, Bagha AK, Kulwinder Singh P, Javaid M, Haleem A, Singh RP. Significant applications of machine learning for COVID-19 pandemic. Journal of Industrial Integration and Management. 2020.

[18] Laguarta J, Hueto F, Subirana B. COVID-19 artificial intelligence diagnosis using only cough recordings. IEEE Open Journal of Engineering in Medicine and Biology. 2020 Sep 29; 1: 275-81

[19] Quatieri TF, Talkar T, Palmer JS. A framework for biomarkers of COVID-19 based on coordination of speech-production subsystems. IEEE Open Journal of Engineering in Medicine and Biology. 2020 May 29; 1: 203-6.

[20] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016, pp. 770-778.

[21] P. Gramming, J. Sundberg, S. Ternström, R. Leanderson, and W. H. Perkins, “Relationship between changes in voice pitch and loudness,” J. Voice, vol. 2, no. 2, pp. 118–126, Jan. 1988, doi: 10.1016/S0892- 1997(88)80067-5.

[22] B. J. Tromberg et al., ”Rapid scaling up of covid-19 diagnostic testing in the United States—The NIH radx initiative”, New Engl. J. Med., vol. 383, no. 11, pp. 1071-1077, 2020.

[23] A. La Marca, M. Capuzzo, T. Paglia, L. Roli, T. Trenti and S. M. Nelson, ”Testing for SARS-CoV-2 (COVID-19): A systematic review and clinical guide to molecular and serological in-vitro diagnostic assays”, Reprod. BioMed. Online, vol. 41, no. 4, pp. 483-499, Sep. 2020.