Audio-Based Cough Classification : Using A Machine Learning Approach

I. Abstract

This Project is aimed at building a machine learning model for cough counting and classification to assist in tuberculosis (TB) diagnosis using audio signal analysis. Our model can differentiate between cough and non-cough sound, hence predict presence of TB allowing early intervention. Diagnosis of TB still continues to be a major threat to public health, especially in areas such as India calling for efficient diagnostic tools. Our approach entails an innovative approach that applies machine learning for possible improvements in current diagnosis procedures. Using audio signals, it constitutes our non-invasive and quick method of finding TB alongside an existing traditional system. What makes our invention unique is the fact that we use technology to improve healthcare outcomes especially in resource constrained settings. Early detection through our model can potentially prevent the spread of TB and decrease sickness caused by it as well as deaths associated with this disease. As part of broader efforts towards controlling.TB, we hope to see a world free from TB one day.

Keywords : *Cough*, *Not Cough*, *Tuberculosis(TB)*, *Convolutional Neural Networks(CNN)*,*Random Forest Classifier*, *Support Vector Machine(SVM)*, *Gradient Boosting Classification*, *Respiratory Diseases*, *Mel Frequency Cepstral Coefficients(MFCC's)*, *Spectral Bandwidth*, *Spectral Centroid*, *Zero Crossing Rate*, *Chroma Features*

II. Introduction

There is a vast increase of disease Tuberculosis. As per recent records there are 24.2 lakh cases in 2022 . For every 100 people 2 people are suffering from the tuberculosis.

Tuberculosis is caused by bacteria called "Mycobacterium". It usually affects the lungs, but it can spread to other parts of the body. It basically starts as When someone with active TB coughs, sneezes, speaks or laughs, they expel tiny droplets containing the bacteria into the air.When we inhale these droplets, the bacteria can enter into our lungs. The bacteria then travel to the small alveoli in our lungs, where they can begin to multiply. So from here the TB disease starts in a person and the person becomes sick. In most people, the body's immune system is able to surround the bacteria with cells called "granulomas", which wall them off and prevent them from spreading. This is called "latent TB infection". People with latent TB infection don't feel sick and cannot spread the bacteria to others.

III. Literature Survey

We have done research work on Tuberculosis detection from acoustic cough signals. As there are very few papers on tuberculosis. We have seen the research papers from 2000 to 2023. We have selected the best 15 research papers from all of them. Every research paper has it own ideas and their own methodologies. In most of the research papers the researches have discussed about the 22 respiratory diseases. In below, we have discussed about all the research papers one by one.

In the "Towards using cough for respiratory disease diagnosis by leveraging Artificial Intelligence: A survey" article, Dr. Aneeqa Ijaz investigates the use of AI in cough sound analysis to diagnose respiratory ailments. The authors discovered that AI is capable of precisely differentiating between coughs after developing a system that classifies coughs from different nations and sickness types. They have developed the smartphone applications to gather information on cough sounds and examined how coughs are produced and vary based on the sickness. The study makes the case that artificial intelligence (AI) can detect diseases fast and precisely from cough noises, but it also stresses the importance of using high-quality data to train these models.

Dr. Liu suggested in the paper "Cough detection using deep neural networks" that researchers could enhance the cough detection system's accuracy and dependability by combining hidden Markov models (HMM) with deep neural networks (DNN) to create a system for cough detection. The researchers have done two types of tests. In both speaker-dependent and speaker-independent tests, the system outperformed a conventional Gaussian Mixture Model (GMM) system in terms of sensitivity, specificity, and F1 measure. This indicates that DNN's are useful for cough detection; nevertheless, for best results, additional noise reduction in audio samples is required.

Dr. Botha proposed in the publication "Detection of tuberculosis by automatic cough sound analysis" that this study looks into the use of cough sounds for low-cost tuberculosis (TB) screening. Statistical methods are applied to coughs to differentiate between TB patients and healthy individuals. The objective is to create a low-cost, non-invasive method for early tuberculosis detection, particularly in places with limited resources. The findings are encouraging, but more testing with bigger sample sizes is required.

Dr. Tracey stated in the paper "Cough detection algorithm for monitoring patient recovery from pulmonary tuberculosis" that the goal of this work is to investigate a cough analysis approach for monitoring tuberculosis recovery in order to alleviate limited lab access in high-burden regions. An algorithm was developed by researchers to recognize and classify coughs according to their auditory properties. Researchers created an algorithm to identify and categorize coughs based on acoustic characteristics. According to preliminary findings, responding patients' cough frequency has decreased, which may indicate that the treatment is having some effect. To be used with confidence in clinical settings, the algorithm needs more refinement and testing.

The article "TB or not TB? Dr. Frost has predicted the use of acoustic cough analysis for tuberculosis categorization. This paper explores the use of recurrent neural networks, specifically BiLSTMs, for cough classification in order to diagnose tuberculosis (TB). The method works well and finds cough locations that are crucial for tuberculosis identification. The study sheds light on the acoustic properties of tuberculosis coughs by examining cough sounds. It is advised to perform additional validation on larger datasets.

Dr. Alqudaihi has suggested in the paper "Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities" This study looks into the analysis of cough noises and the identification of lung diseases using artificial intelligence (AI) and machine learning (ML). It examines current practices, evaluates effective AI/ML strategies, and talks about potential and problems. In addition to discussing suitable techniques, the authors suggest a method for classifying research according on data size. Their goal is to direct upcoming studies and advancements in the realm of medical technology.

The article "A novel method for wet/dry cough classification in pediatric population" by Dr. Amrulloh suggests To aid in the identification of respiratory disorders, a novel method was developed to categorize children's coughs as dry or wet. It achieves 76% accuracy in cough classification, comparable to doctors' methods, by combining computer analysis, manual labeling, and recording.

Dr. Solinski made a proposal in the publication "A novel method for wet/dry cough classification in pediatric population". This paper introduces an automatic cough detection system for portable spirometry based on airflow signals. It leverages machine learning on the AioCare and NHANES databases to identify coughs during spirometry examinations. The artificial neural network produced the best results out of all the evaluated algorithms (ANN, SVM, random forest, and logistic regression) (accuracy: 0.91, F1 score: 0.88). This airflow signal-based method represents a major breakthrough in spirometry cough detection.

Dr. Orlandic has suggested in the work "A Multimodal Dataset for Automatic Edge-AI Cough Detection" This study includes a public dataset for cough detection algorithms. It emphasizes the advantages of employing wearable sensors and edge-AI—device artificial intelligence—for cough detection while protecting user privacy. The created algorithms distinguished coughs from background noise with a high degree of accuracy.

In his article "Cough classification tool for early detection and recovery monitoring of tuberculosis and asthma" , Dr. Sai wrote This study proposes a smartphone app that could aid in the early detection of respiratory disorders by differentiating between wet and dry coughs. It achieves 82% accuracy by using K-Means clustering and MFCC features. For even more improvement, the study investigates the use of hybrid algorithms and extra features like cough intensity.

The article "Automatic cough classification for tuberculosis screening in a real-world environment" by Dr. Pahar made some recommendations. This paper introduces an automatic cough detection system for portable spirometry based on airflow signals. In order to detect coughs during spirometry testing, it uses machine learning (logistic regression, ANN, SVM, and random forest) on NHANES and AioCare data. The artificial neural network performed the best out of all the methods (F1: 0.88, accuracy: 0.91). This method is a major improvement in cough detection for spirometry.

Dr. Snella makes improvements to the features of coughs, privacy controls, and assessment techniques for audio-based cough detection in the publication "Cough Monitoring Through Audio Analysis". SVM, neural networks, and other techniques are used in current methods to reduce data, provide clarity, and safeguard some privacy.

Dr. De Blasio provides advice on how to properly manage coughs for lung specialists and general care physicians in the paper "Cough management: a practical approach". It examines studies on cough mechanisms, causes, and treatments (including cough suppressants) from large databases (1950 to present). It is emphasized how crucial early intervention and symptom relief—such as using cough suppressants—are, particularly when coughs have a detrimental effect on one's quality of life.

In the paper "Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions", Dr. Vizel examined the precision of PulmoTrack-CCTM, a paper cough detection technology, across various coughing situations in their study titled "Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions." The device's ability to distinguish between coughs and other sounds and vibrations in the chest was tested against that of human experts. The findings demonstrated that while PulmoTrack-CCTM performed well in general—even when walking—it requires considerable refinement in some scenarios, such as when ascending stairs.

Dr. Vizel evaluated the accuracy of PulmoTrack-CCM, a paper cough detection technology, in a study titled "Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions." The study looked at the precision of the technology in a variety of coughing scenarios. The device was tested against human experts to determine how well it could differentiate between coughs and other sounds and motions in the chest. The results showed that although PulmoTrack-CCTM functioned well overall—including when walking—it has to be significantly improved in certain situations, such climbing stairs.

The Hull Automatic Cough Counter (HACC), a cutting-edge method created by Dr. Barry for the automated identification and counting of coughs in audio recordings, is presented in the publication ”The automatic recognition and counting of cough”. When comparing HACC to manual methods, counting time was greatly reduced. It produced consistent results (100 percent repeatability) and good accuracy (80 percent sensitivity, 96 percent specificity).

IV. Methodology Used

The project's work flow is depicted in the flow chart below. The dataset will first be balanced, according to the flow chart. Then, using Librosa's assistance, we extract the required features and save the mean values in a.csv file. Subsequently, various strategies are implemented for model training, enabling the model to categorize signals as cough or non-cough.

The flow chart of our project goes as :

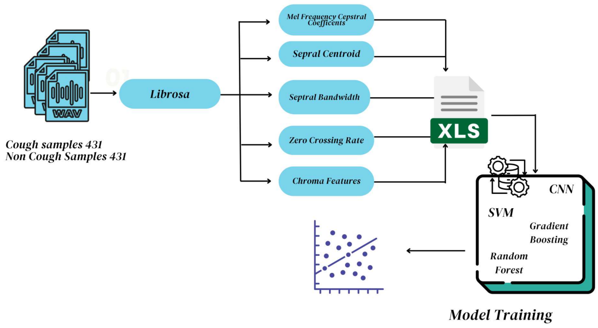

Fig. 1. FLOW CHART

We have split the workload into two tasks:

- Task 1: Predicting the given audio signal is cough or not cough.
- Task 2: If the given audio signal is cough we will predict the given audio signal has TB or not TB.

Right now we have started working on the task 1 .

We have got an data set of 432 cough and 432 not cough signals. With these we have started working on them .

We have imported the required libraries. We have used librosa library because the librosa library works on the stereo audio signals. So that we can get an best result by using the librosa library . We have used pandas,sklearn,matplot and other libraries.

Firstly, we have trimmed the audio signals at start and end with an threshold frequency so that few of the silence will get removed from the audio signals. Next we have pre-processed our audio signals and we have extracted the required features such as zero crossing rate , spectral centroid, spectral bandwidth, chroma features, Mel Frequency Cepstral Co-efficients(MFCC's) for each and every audio signal and we have stored all the data in an .csv(comma separate variable) file. While extracting these features we have extracted 20 frames of MFCC's for each and every signal and we have calucated the mean for each frame of the MFCC's and we have calculated the mean vaue for the Spectral Centriod, Spectral Bandwidth, Zero Crossing Rate and Chroma Features . So that it will make comfortable to develop a model using these values. So that we can use this file to develop a model.

1) Zero Crossing Rate (ZCR): How often the signal crosses the zero amplitude line, reflecting rapid audio changes (e.g., clicks, transients).

Cough vs. Non-Cough ZCR Characteristics:

Cough: Coughs typically have a rapid succession of high and low-pressure sounds due to the forceful expulsion of air. This translates to a higher ZCR compared to speech or other ambient sounds.

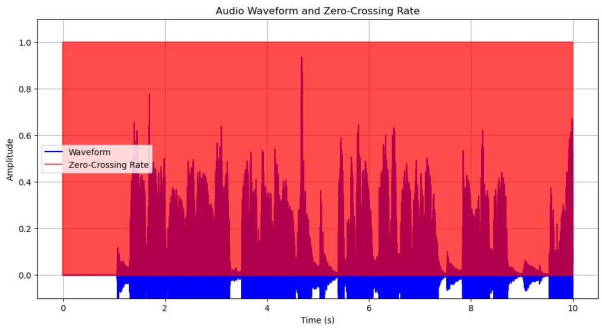

Fig. 2. Zero Crossing Rate of Cough

Non-cough: Speech and other ambient sounds often have a wider range of frequencies but with less rapid transitions. This results in a lower ZCR compared to coughs.

Fig. 3. Zero Crossing Rate of Non-Cough

Plotting and Analysis for ZCR:

Obtain Audio Data: Record audio containing both coughs and non-cough sounds (speech, background noise).

Segment the Audio: Divide the audio into small time windows (e.g., 10 milliseconds).

Calculate ZCR for Each Segment: Within each window, count the number of times the signal crosses the zero amplitude line.

Plot the ZCR: Create a time-series plot where the x-axis represents time (window segments) and the y-axis represents the calculated ZCR values for each segment.

2) Spectral Centroid: The ”center of mass” of the frequency spectrum, indicating the average pitch of the sound.

The formula to calculate the Spectral centriod is :

$$\text{Centroid} = \frac{\sum f(n) \cdot x(n)}{\sum x(n)}$$

where,
Centroid: Represents the spectral centroid value.
n: Iterates over each frequency bin (0 to N-1).
f(n): Represents the center frequency of bin n.
x(n): Represents the magnitude (weight) of bin n, often obtained from the magnitude spectrum of the signal using a Fourier transform.

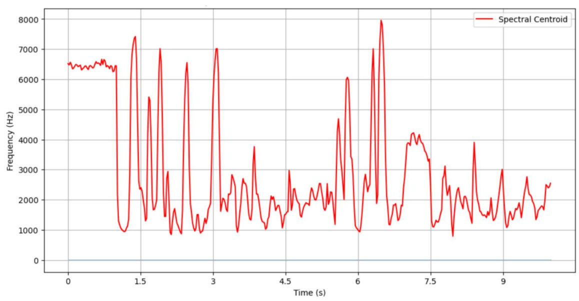

Fig. 4. Spectral Centroid of Cough

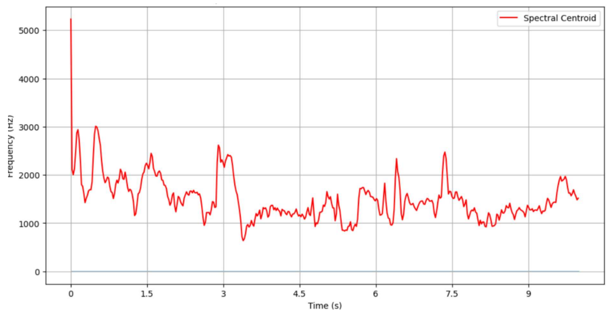

Fig. 5. Spectral Centroid of Non-Cough

3) Spectral Bandwidth: The spread of frequencies in the spectrum, representing the ”tonal richness” of the sound (wide for instruments, narrow for vocals).
The formula we have used for the Spectral bandwidth calculation is :

$$B = \sqrt{\frac{\sum ((f - f_c)^2 \cdot P(f))}{\sum P(f)}}$$

where:
B is the spectral bandwidth (Hz)
f is the frequency (Hz)
f_c is the spectral centroid (center of mass of the spectrum, Hz)
P(f) is the power spectral density at frequency f

Cough vs. Non-Cough Spectral Bandwidth Characteristics:

Cough: Coughs are characterized by a sudden expulsion of air, resulting in a wide range of frequencies being produced. This broad frequency range translates to a wider spectral bandwidth compared to speech or other non-cough sounds.

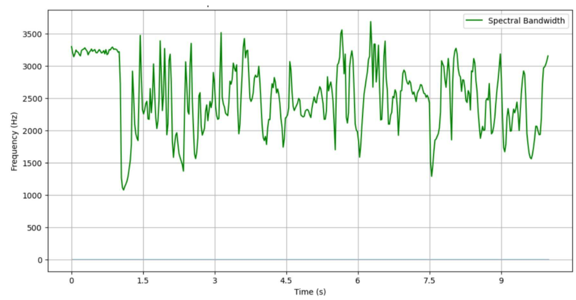

Fig. 6. Spectral Bandwidth of Cough

Non-Cough Sounds: Speech typically occupies a lower spectral bandwidth than coughs, concentrating on the vocal range (around 300 Hz to 3 kHz). Other non-cough sounds might have varying bandwidths depending on their source, but generally tend to be narrower than coughs.

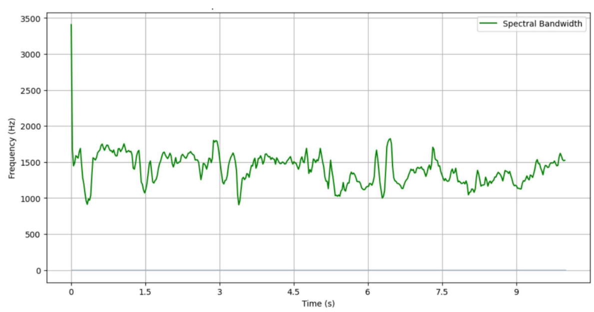

Fig. 7. Spectral Bandwidth of Non-Cough

Spectral Bandwidth Analysis for Cough Detection:

Short-Time Fourier Transform (STFT): Divide the audio signal into short segments (windows) and compute the magnitude spectrum for each window using the STFT. This provides a time-frequency representation of the signal. *Spectral Bandwidth Calculation:* For each window, calculate the spectral bandwidth. There are different ways to do this, but common methods include: *Full Width at Half Maximum (FWHM):* This is the range of frequencies within which the spectral magnitude falls to half of its maximum value. *Interquartile Range (IQR):* This is the range of frequencies that contains the middle 50% of the spectral magnitude distribution. *Cough vs. Non-Cough Classification:* Based on the calculated spectral bandwidth for each window, classify the segment as cough or non-cough. You can establish a threshold based on the observed differences in bandwidth between coughs and other sounds. For example, segments with bandwidth exceeding a certain value could be classified as coughs.

4) Mel-Frequency Cepstral Coefficients (MFCC): MFCCs are a set of coefficients that represent the short-term power spectrum of an audio signal. They record essential data regarding the signal's spectral content . This is the is the important feature as:
    - MFCCs are less sensitive to noise and environmental variations compared to raw audio features.
    - They capture essential information while minimizing noise effects.
    - MFCCs mimic the human auditory system's just like Our ears are more sensitive to certain frequency ranges, and MFCCs reflect this.

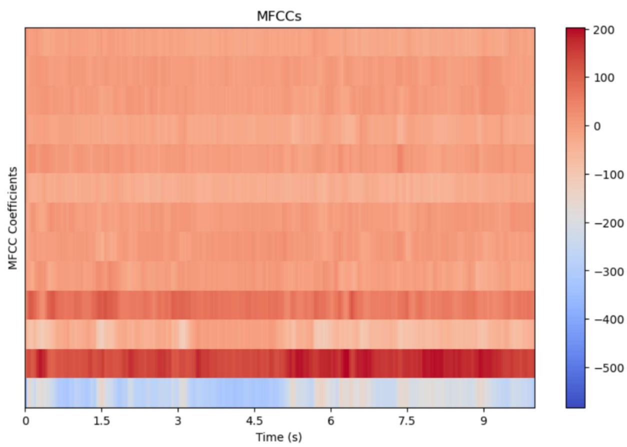

Fig. 8. MFCC of Cough

The coefficients having lower frequencies (where cough energy lies) would likely be maximally red. These red areas indicate that the MFCC values are high, capturing the characteristic features of the cough sound.
The coefficients having the higher frequencies (where non-cough energy lies) would likely be maximally blue. These blue areas indicate that the MFCC values are low, capturing the characteristic features of the non-cough sound.

Mel Frequency Cepstral Coefficients Analysis for Cough Detection

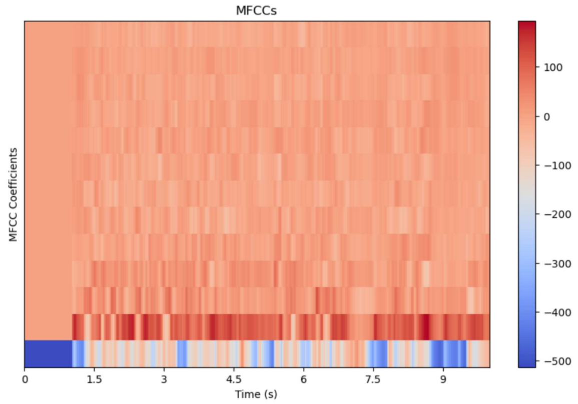

Fig. 9. MFCC of Non-Cough

I. Frame Segmentation: We divided the audio signal into frames to analyze short segments. Hence, we used 20 frames.
II. Pre-Emphasis: To emphasize high-frequency components, we applied a pre-emphasis filter to the signal. This step helps improve sensitivity to higher frequencies.
III. Windowing: Each frame is multiplied by a window function (e.g., Hamming window) to reduce spectral leakage.
IV. Fast Fourier Transform (FFT): We computed the FFT of each windowed frame to obtain the power spectrum. [10]

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-i2\pi kn/N}$$

V. Mel Filterbank: The power spectrum is passed through a set of triangular filters in the Mel scale. These filters are spaced according to human perception of pitch. [10]

$$H_m(k) = \begin{cases} 0 & \text{if } k < f(m-1) \\ \frac{k-f(m-1)}{f(m)-f(m-1)} & \text{if } f(m-1) \leq k < f(m) \\ \frac{f(m+1)-k}{f(m+1)-f(m)} & \text{if } f(m) \leq k < f(m+1) \\ 0 & \text{if } k \geq f(m+1) \end{cases}$$

VI. Logarithm and Discrete Cosine Transform (DCT): We took the logarithm of the filterbank energies to compress the dynamic range. Finally, we applied the DCT to obtain the MFCCs.

$$C_i = \sum_{j=1}^{M} \log(E_j) \cdot \cos\left[\frac{\pi i}{M}\left(j - \frac{1}{2}\right)\right]$$

where:
- $X(k)$ is the Fourier transform of the signal,
- $x(n)$ is the input signal,
- N is the number of samples in the frame,
- $H_m(k)$ is the mth Mel filter,
- $f(m)$ is the frequency corresponding to the mth Mel filter,
- E_j is the energy in the jth filterbank,

- C_i is the ith MFCC coefficient, and
- M is the number of Mel frequency filters.

5) Chroma Features: Represent the presence of musical notes (pitches) across 12 pitch classes, useful for analyzing music harmony and chord progressions.

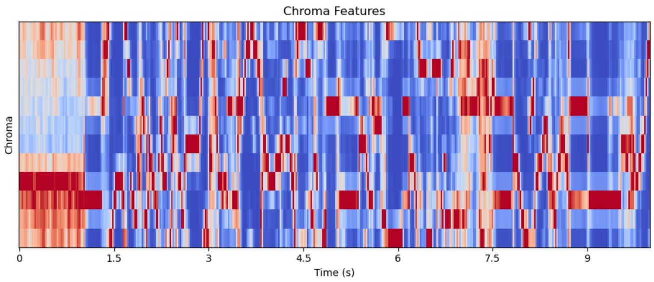

Fig. 10. Chroma Features of Cough

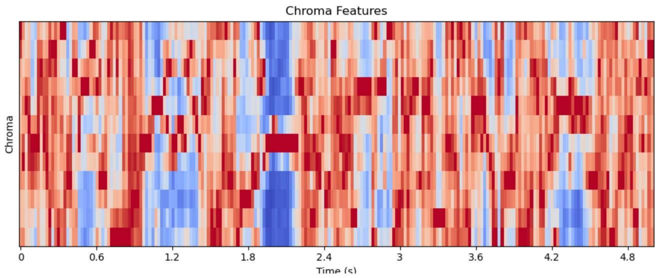

Fig. 11. Chroma Features of Non-Cough

Next, we have started with creating a model. We have used different types of supervised and Neural Network techniques. We have used Support Vector Machine(SVM), Random Forest Classifier, Gradient Boosting and Convolutional Neural Networks(CNN) techniques in our model. We have got different accuracies, fl scores to these techniques . From the 4 techniques we have choosen one technique.

V. User Interface

We have made an simple user friendly interface i.e,Fig. 11 using the python streamlit library. From this user interface we can detect the given audio signal is cough or not cough.

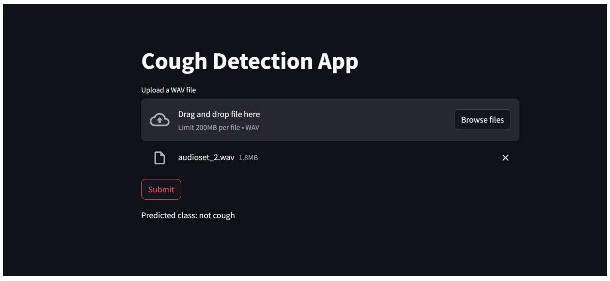

Fig. 12. User Interface

VI. Results and Discussion

As we discussed above we have used 4 techniques of supervised an unsupervised learning . Let's see the table 1 i.e, comparision of accuracies of the four techniques in the plotted graph from fig 12.

MODEL	Accuracy	Precision	Recall	F1 Score
Convulution Neural Networks	0.83	0.80	0.88	0.84
Random Forest Classifier	0.76	0.77	0.76	0.76
Gradient Boosting Classifier	0.80	0.80	0.80	0.80
Support Vector Machine	0.81	0.82	0.81	0.81

TABLE I
Comparision of accuracies

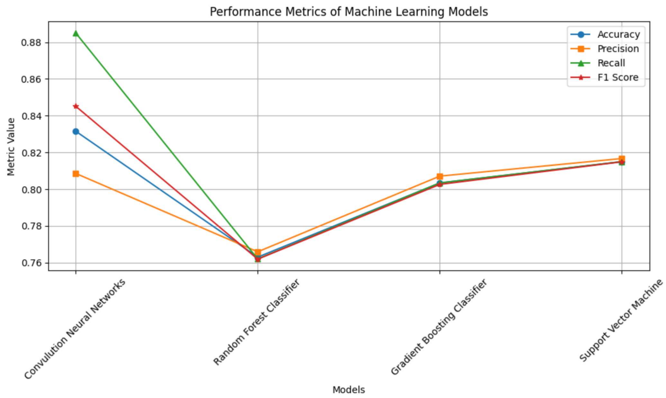

Fig. 13. Performance Metrices

‘

VII. Conclusion and Future Work

By using Convolutional Neural Networks we were getting an accuracy of 0.83 which is going to predict the given acoustic signal is cough or not cough . In future we want to develop this model that which will predict the person having TB or not TB once the signal is predicted as cough signal.

In future we want to develop our user interface as it going to show the MFCC's graph of the qaudio signal of the patient.

If this happens we can decrease the rapid increase of TB from the world. Furtherly , we will develop this project as from the acoustic signal it will perdict the person is suffering from any of the 22 respiratory diseases.

References

[1] Ijaz, Aneeqa, Muhammad Nabeel, Usama Masood, Tahir Mahmood, Mydah Sajid Hashmi, Iryna Posokhova, Ali Rizwan, and Ali Imran. "Towards using cough for respiratory disease diagnosis by leveraging Artificial Intelligence: A survey." Informatics in Medicine Unlocked 29 (2022): 100832.

[2] Liu, Jia-Ming, Mingyu You, Zheng Wang, Guo-Zheng Li, Xianghuai Xu, and Zhongmin Qiu. "Cough detection using deep neural networks." In 2014 IEEE international conference on bioinformatics and biomedicine (BIBM), pp. 560-563. IEEE, 2014.

[3] Botha, G. H. R., Grant Theron, R. M. Warren, Marisa Klopper, Keertan Dheda, P. D. Van Helden, and T. R. Niesler. "Detection of tuberculosis by automatic cough sound analysis." Physiological measurement 39, no. 4 (2018): 045005.

[4] Tracey, Brian H., Germán Comina, Sandra Larson, Marjory Bravard, José W. López, and Robert H. Gilman. "Cough detection algorithm for monitoring patient recovery from pulmonary tuberculosis." In 2011 Annual internatiTonal conference of the IEEE engineering in medicine and biology society, pp. 6017-6020. IEEE, 2011.

[5] Frost, Geoffrey, Grant Theron, and Thomas Niesler. "TB or not TB? Acoustic cough analysis for tuberculosis classification." arXiv preprint arXiv:2209.00934 (2022).

[6] Alqudaihi, Kawther S., Nida Aslam, Irfan Ullah Khan, Abdullah M. Almuhaideb, Shikah J. Alsunaidi, Nehad M. Abdel Rahman Ibrahim, Fahd A. Alhaidari et al. "Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities." Ieee Access 9 (2021): 102327-102344.

[7] Amrulloh, Y.A., Wati, D.A., Pratiwi, F. and Triasih, R., 2016, May. A novel method for wet/dry cough classification in pediatric population.In 2016 IEEE region 10 symposium (TENSYMP) (pp. 125-129). IEEE

[8] Solinski, M., Łepek, M. and Kołtowski, Ł., 2020. Automatic cough ´ detection based on airflow signals for portable spirometry system. Informatics in medicine unlocked, 18, p.100313.

[9] Orlandic, L., Thevenot, J., Teijeiro, T. and Atienza, D., 2023, July. A Multimodal Dataset for Automatic Edge-AI Cough Detection. In 2023 45th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC) (pp. 1-7). IEEE.

[10] Sai, P., Rao, N., Kumar, N., Brahmaiah, P. and Ajay, D., 2015. Cough classification tool for early detection and recovery monitoring of tuberculosis and asthma. In 4th International Conference on Computing, Communication and Sensor Network, CCSN.

[11] Pahar, M., Klopper, M., Reeve, B., Warren, R., Theron, G. and Niesler, T., 2021. Automatic cough classification for tuberculosis screening in a real-world environment. Physiological Measurement, 42(10), p.105014

[12] Snella, T., 2021. Cough Monitoring Through Audio Analysis.

[13] De Blasio, F., Virchow, J.C., Polverino, M., Zanasi, A., Behrakis, P.K., Kilinc¸, G., Balsamo, R., De Danieli, G. and Lanata, L., 2011. Cough management: a practical approach. Cough, 7, pp.1-12.

[14] Vizel, E., Yigla, M., Goryachev, Y., Dekel, E., Felis, V., Levi, H., Kroin, I., Godfrey, S. and Gavriely, N., 2010. Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions. Cough, 6(1), pp.1-8.

[15] Yellapu, G.D., Rudraraju, G., Sripada, N.R., Mamidgi, B., Jalukuru, C., Firmal, P., Yechuri, V., Varanasi, S., Peddireddi, V.S., Bhimarasetty, D.M. and Kanisetti, S., 2023. Development and clinical validation of Swaasa AI platform for screening and prioritization of pulmonary TB. Scientific Reports, 13(1), p.4740.

[16] Barry, S.J., Dane, A.D., Morice, A.H. and Walmsley, A.D., 2006. The automatic recognition and counting of cough. Cough, 2(1), pp.1-9.