Tussis segment classification of acoustic signals using Machine Learning

Abstract

Cough (Tussis) is the most commonly observed indication of many respiratory related disorders. Efficient cough detection is a key for assessing patients with tuberculosis and other various diseases. This paper aims to detect the cough events and their exact boundaries present in an acoustic signal. We treat cough detection as a visual recognition problem by converting them into the patterns of various feature coefficients. The evaluation of these feature coefficients using Machine Learning models find useful in Tussis detection. We trained the Machine Learning classifiers using different feature extraction methods and evaluated them to choose the one with higher efficiency by comparative study. The classification methodology developed in this process should be robust for detecting Tussis in the acoustic signals.

Keywords - Cough Sounds, Automatic Detection, Machine learning, Linear Prediction, FFT

1. Introduction

Cough, sometimes be the root cause or indication of many ailments. In general, cough is a protective reflex which occurs suddenly and also in a repetitive fashion. It's a protective defined mechanism for the lungs and larger airways, which helps to clean out things like foreign bodies, micro bodies, irritants etc. from the respiratory passages. However recurrent and excessive coughing proves harmful to our respiratory passage mucosa, by which it is forced to sleigh of the respiratory pathway.

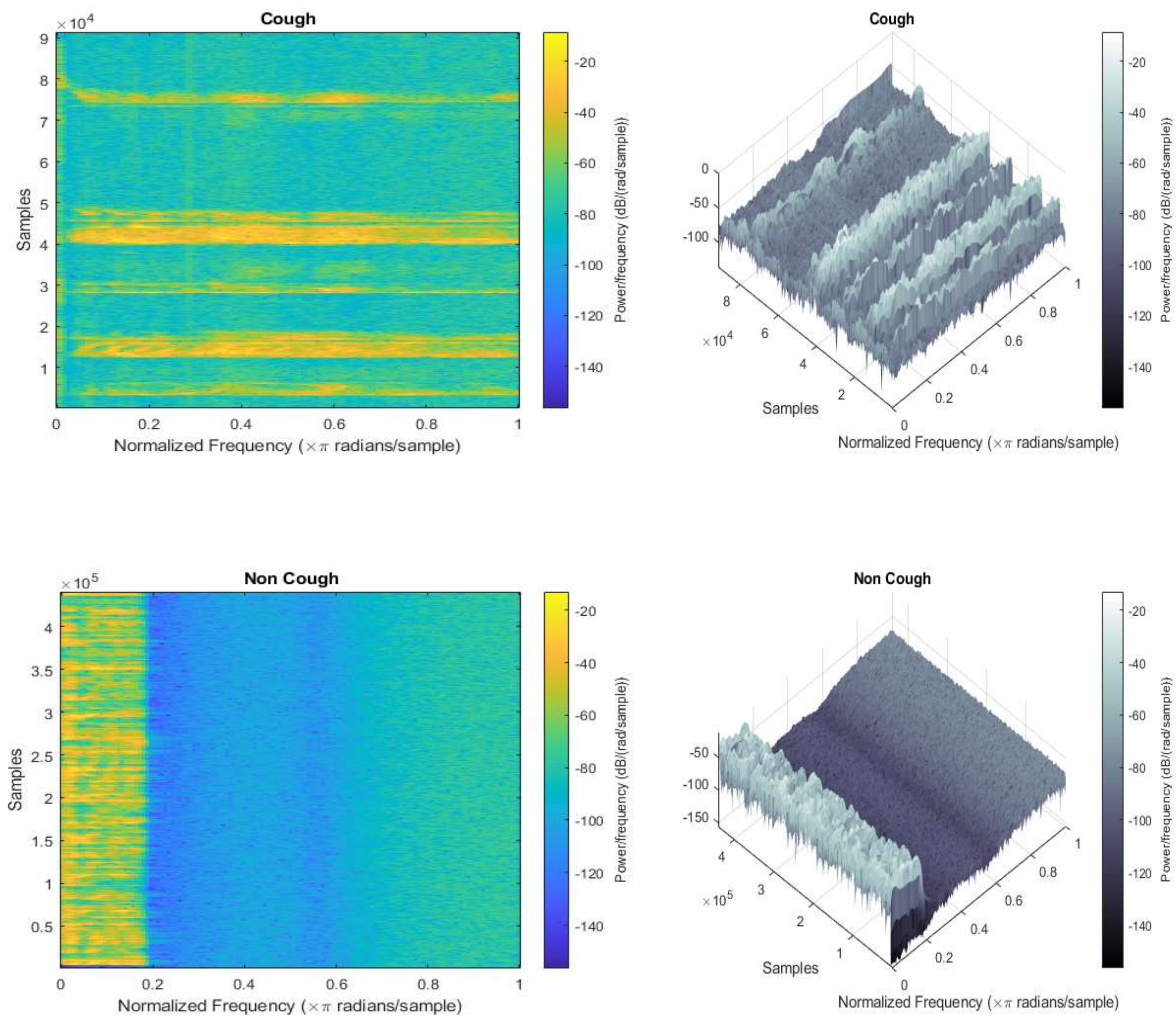
Cough can be categorized by the human ear and can be differentiated easily from other human vocal tract produced like voice, laughing, throat clearing, and snoring. Adding to that, cough can be used to recognize people by the characteristics of their cough and can be diagnosed to get a rough approximation of chest and lung conditions.

Years of research describe that coughs can be categorized with the presence of mucus from the lower airways like trachea and bronchi and without it. But the importance of cough is hugely ignored in the medical assessment of adults and children. Like in the case of asthma and COPD, the initial identification marker is cough. Taking whooping cough as an example, it is associated with trace-esophageal fistula which have well recognized identification features. Even though doctors ask the patients whether the cough is severe or not, studies [1] show that it is advisable to keep it as a marker to access the patients' health. Also even though medical professionals recognize a wheezy quality to cough, identifying a chest disease by their cough characteristics seemed very much limited.

Cough has to be recognized at the earliest as every major respiratory ailment starts from a simple symptom called cough. Proper and fast diagnosis of cough can lead to the fast ailment of many severe diseases which can be fatal at latter stages.

Analysis of voice signals, which are produced by the upper airways gives an interesting way of approach in analysis of cough. First the signal is identified for related features in time or frequency domain. And based on these features a computer model is built with the waveform of the voice signal analyzed comparatively. This leads to interesting differences between both the voice signal and the built features model. These differences can be extracted to identify the presence of the cough in the acoustic signal.

Figure 1 shows the Spectrograms and 3-D bone plots of cough and a non-cough event signal. The spectrogram analysis reveal some transparent features that are present in the voice signals which can be used to differentiate between normal voice signal (non-cough event) and a cough event.



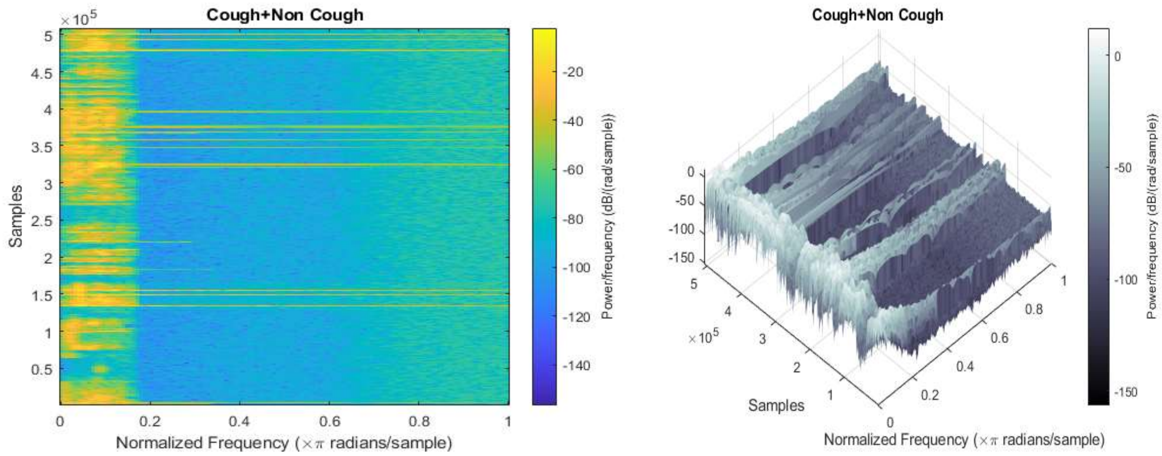


Fig. 1 Spectrograms and 3-D bone plots of cough event signal and a normal signal.

Even though spectrogram analysis differentiate the signals, these signals were already predetermined as cough and non-cough samples, so it can be easily distinguished. But when it comes to an unknown signal that contains frequencies and areas similar to a cough event, spectrogram analysis of an unknown signal fails.

2. Related Work

Over the course of many years of research several methods have been adopted in training a detecting cough in speech segments. One of the common model is using MFCC feature coefficients as their base model and modifying it based on their number of coefficients and training it using a machine learning model. Detecting can also be applied in a mobile environment. Mobile cough detection approach is implied with features like MFCC are used. But for Training VGG net was used with a change in iterative layers of upto 5 to improve the image recognition accuracy by a significant amount in which they has an average specificity and sensitivity of 91.7% and 90.1% [4].

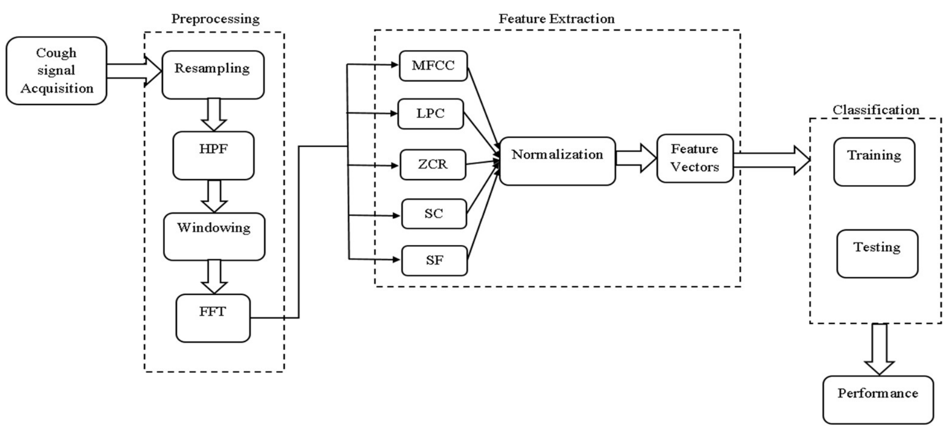
Forward feature group selection method has been used for selecting the best performing feature group within the best group in addition of using MFCC and other baseline features with ISCB[3].A cross fold validation setup was done to analyze the individual features and to rank based on the best selected feature group[5]. Baseline features like ZCR, Kurtosis, MFCC, and Bispectrum Score directly and trained using 1 input layer or 201 neurons. Also a Sigmoid transfer function was used to improve the accuracy along Levenberg-Amrquardt Back-propagation algorithm. This method has a 98% accuracy for prerecorded samples from 3 patients [7].

Also cough has been classified in a direct approach using a Hidden Markov model [8] which gives an accuracy of more than 90%. The audio signal is analyzed by the features that are extracted using MFCC and one of its type Delta-Delta MFCC technique. It is concluded that the standard deviation occurs in Delta-Delta MFCC is minimum and Cepstral coefficients are better when compare to other form of MFCC[9].To classify the cough signals into wet and dry the features like BGS,NGS, formant frequencies, Log energy, ZCR and kurtosis are computed and they are estimated using Logistic

Regression Model classification[10]Also the dataset contains cough samples are dissected by extracting the features like LPC coefficient, tonality index, spectral flatness and spectral centroid and classified as cough and non-cough using Logistic Regression Model[11]. This Method was improved with adding several sets of features like the pitch features (pitch coverage, pitch mean, median, standard deviation, and pitch in harmonicity), spectral features (flux mean, median, and maximum, spectral entropy). They are compared with MFCC, Spectral centroid, spread, skewness, kurtosis, spectral flatness, ZCR and HNR. Even though computations are complex and time consuming compared to the previous paper it gathered significantly higher accuracy using the logistic regression model [12]

Several other Machine Support Vector Machine (SVM) and Bayesian models. Automatic Detection is carried out with the help of Bayesian, SVM, propositional Rule learner and bootstrap Aggregation on two types of features[13] one for each frame and for each relevant segment. A similar work has been carried out using a 10-fold cross validation scheme where the features sets like MFCC, MFC and STFT are classified using DNN, CNN, and LSTM [14].

3. Methodology



The proposed methodology is divided into four sections: Dataset Acquisition**,** Prepossessing, Feature Extraction and Classification. Each process is explained and their simulated results are shown at the result section. Finally the overall performance of the proposed model is discussed and concluded.

3.1 Dataset Acquisition

The cough samples used here have been compiled from various open sources. A total of 1000+ sample recordings were taken and compiled together with specific categories given in Table 1. The cough category is used to train and test the classification model to get close to dead accurate results. Other samples were included to test its efficiency and novelty of the methods used in this paper.

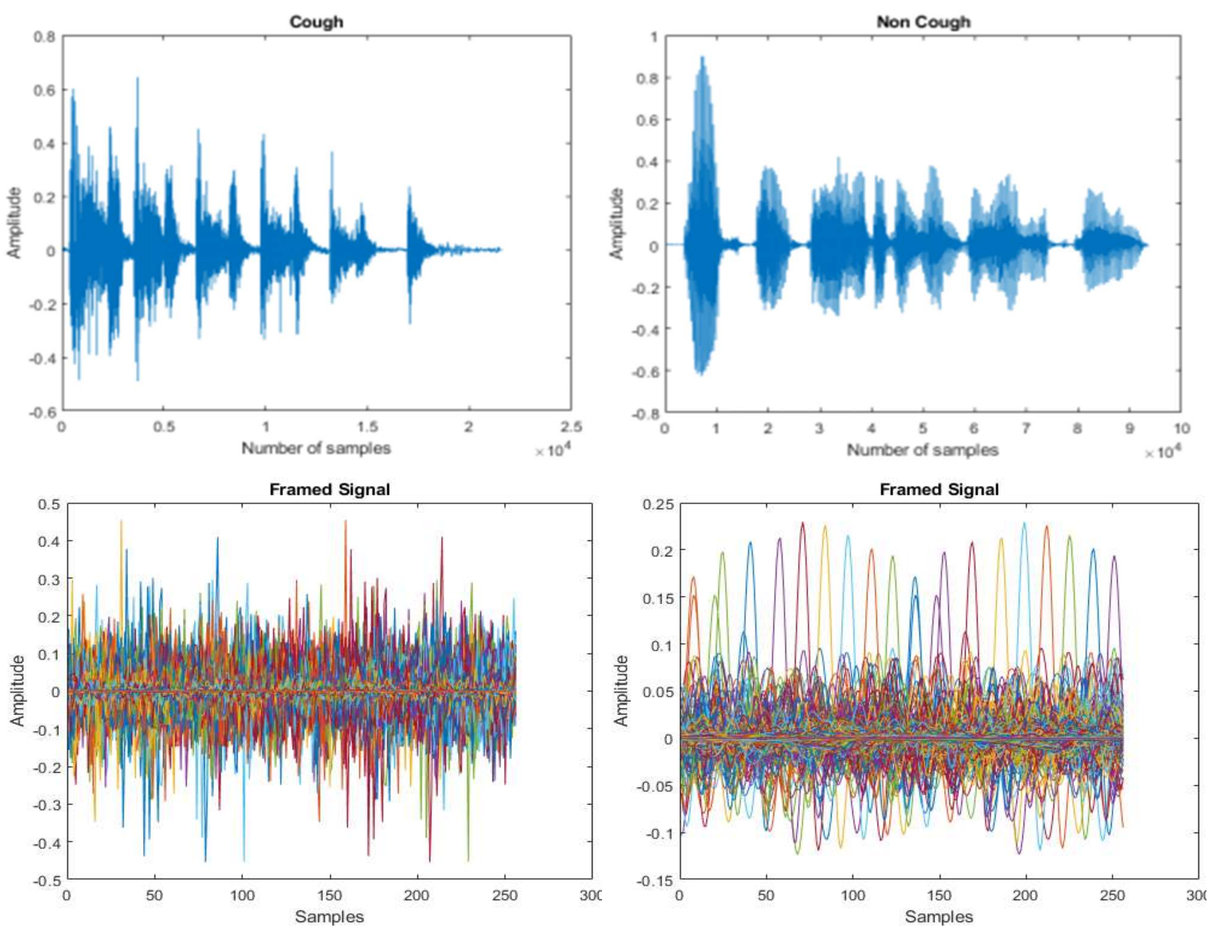
3.2 Preprocessing

Preprocessing is a preliminary process to eliminate the presence of unwanted samples from the signal. Thus, composing the data for further analysis. Preprocessing consists of four steps: Resampling the sample, filtering it to remove the redundant data, framing of the samples and representing the samples in the frequency domain. Figure 2 shows the output of the steps involved in preprocessing the cough and non-cough signals.

The samples taken are resampled to 8000 Hz since most of the features or information regarding cough are observed to contain below 4000 Hz. All the samples were converted to the following format before feature extraction. The amplitude of the samples are then scaled down to -1 to 1 without changing signal distribution to preserve the important features of the signal and also for easy comparison. The samples are then filtered using a high pass filter of a specific low cut off frequency. At very low frequency, the sounds are just pure noise, they are cut off to get accurate results. Also the Amplitudes are cut off at a specific level to prevent the build of noise at very high frequencies.

The silent parts of the samples are identified and removed by fixing a specific threshold and calculating mean deviation for the previous frames. So the processing of signal frames takes place only when the average energy of the frame is above that threshold.

The sampled signals are split into 256ms frames with 50% overlap. These frames are then multiplied with a Black-man window function to smoothen the boundaries of each individual frame. Fast Fourier transform is finally applied to the signal thereby obtaining the spectral components and frequency information of the signal.



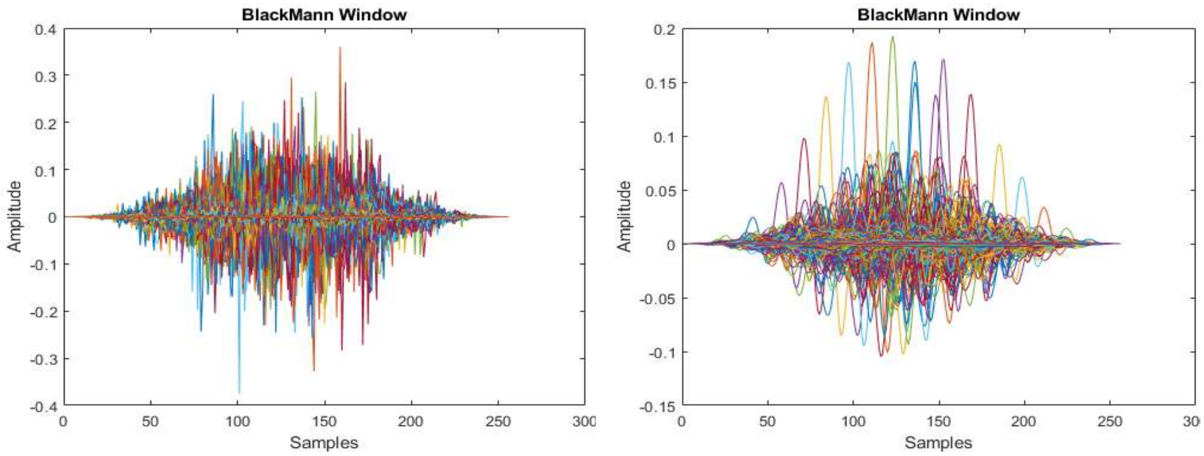


Fig. 2 The preprocessed outputs of sample cough event signal (on the left) and the preprocessed outputs of non-cough event (on the right)

Table 1 Data Acquisition and Preprocessing parameters used in this proposed work

Parameters	
Dataset Acquisition	
No. of cough samples	700
No. of Non-cough samples	800
Audio Format	.wav
Encoding	16Bit PCM
Channels	1-MONO
Preprocessing specifications	
Sampling Frequency	8000Hz
Pre-emphasis filter	[-0.95, 1]
Windowing function	Blackman
Window size	256ms
Overlapping (%)	50% (128ms)

3.3 Feature Extraction

Feature extraction is a process of obtaining the reduced features as it provides pertinent information of the input sample without altering the information of the original sample. It increases the accuracy of the process by using the extracted features.

3.3.1 Mel Frequency Cepstral Coefficients (MFCC)

MFCC are constants obtained by converting an audio signal to the Cepstral domain. The Mel coefficients are derived from Mel scale which places frequency bands equally according to human auditory systems response. The Mel scale aim to mimic non-linear human ear perception of sound.

Human ears are more discriminative at lower frequencies and less discriminative at higher frequencies. Mel filter banks do exactly that by giving a better resolution at low frequencies and less at high. Triangular filter banks help to capture the energy at each critical frequency band and roughly approximates the spectrum shape. This also helps to smooth the harmonic structure.

The MFCC is derived by,

1. Fourier transform of the signal is taken
2. The FFT is converted to Mel scale by multiplying them with the overlapping triangular filters called Mel Filter Bank. We can convert between f (Hertz) and Mel (mm) using the following equations:

$$m = 2595\ \log_{10}\left(1 + \frac{f}{700}\right)$$

$$f = 700(10^{m/2595} - 1)$$

A 20 triangular filter function is applied to the FFT of the audio signals after converting them to the Mel Scale. Figure 3 shows Mel Filter bank that contains 20 Mel-spaced triangular filters. Each filter bank has length N-point FFT used in the transform. Each and every vector in the filter bank is mostly zero until the required frequency of the bank comes and it is filled with non-zero constants. In other words, a process called binning takes place. That is, a coefficient from FFT is multiplied with the respective filter gain and the results are gathered. So each binning has a sum weighted in the gathered results representing the spectral magnitude in that filter bank channel.

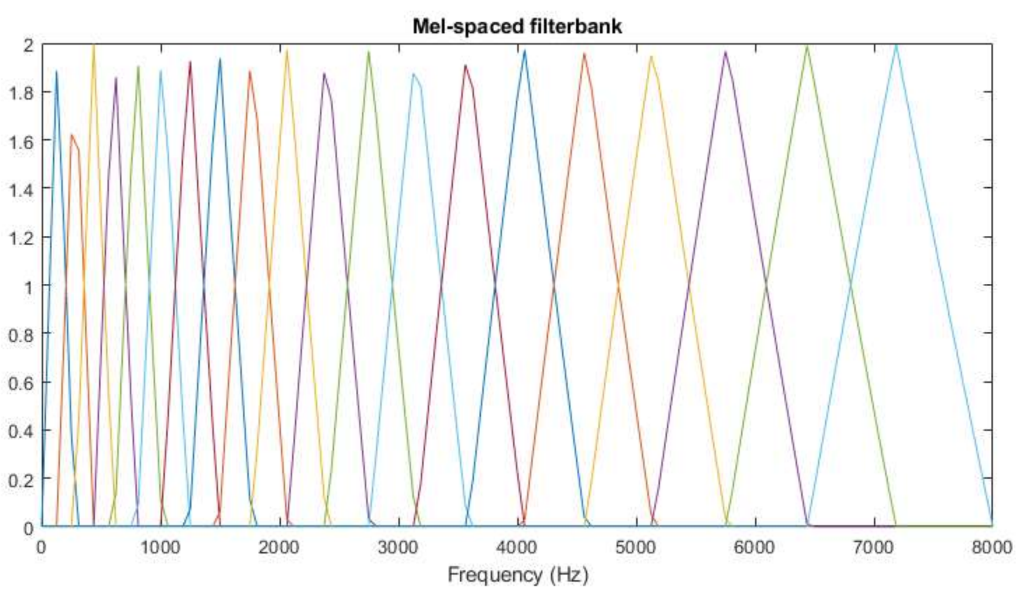


Fig. 3 Mel Filter Bank of 20

3. The Computed values from the Bank is taken log and is converted to cepstral domain as shown in figure 4, by taking discrete cosine transform.

The Fourier transform gives corresponding frequency spectrum which signifies any periodic component in the given signal. This analysis is used for converting a time domain signal to frequency domain. This is the process that takes place in formulating the Mel-Frequency Coefficients. Figure 5 represents MFCC feature extracted from the preprocessed sample signals of both cough and non-cough events.

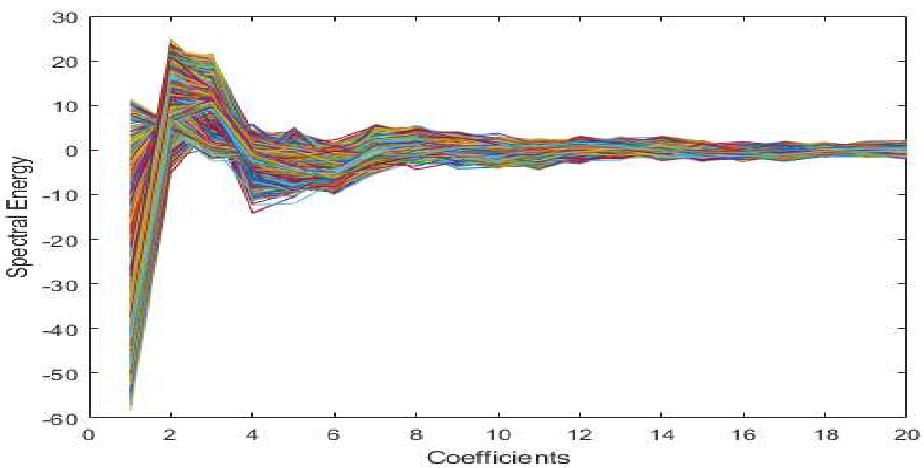


Fig. 4 Cepstral Map of an Audio Signal

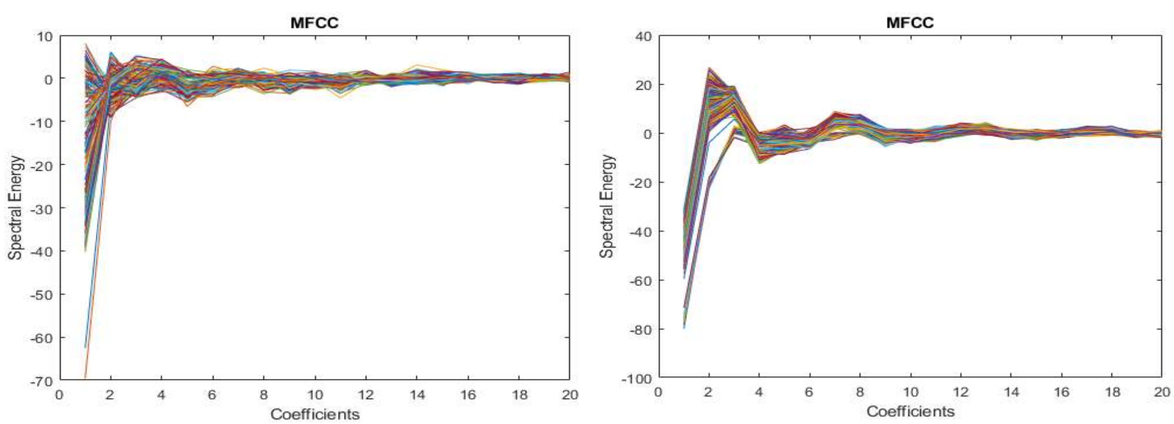


Fig. 5 Extracted Mel Frequency Cepstral Coefficients of a cough event (on the left) and a non-cough event (on the right)

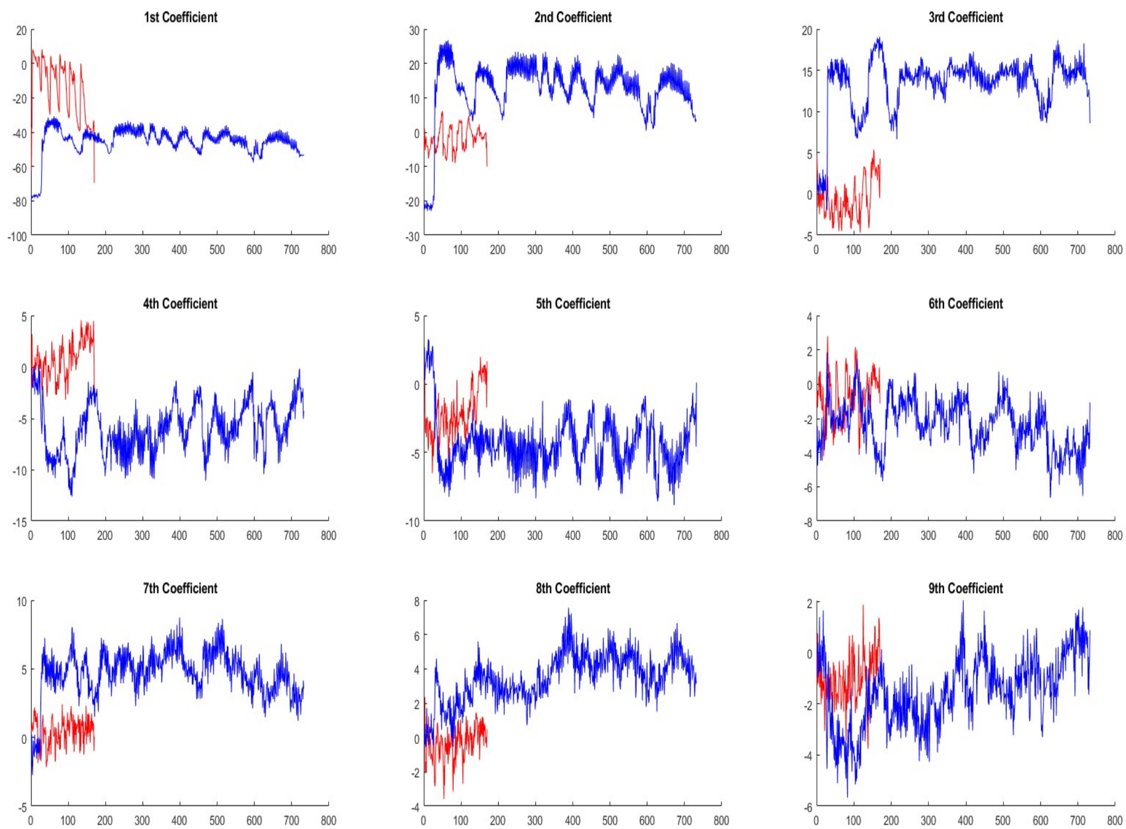


Fig. 6 This compilation of graphs represent each MFCC coefficient of signal with cough event (red) and without a cough event (blue) w.r.t sampling rate.

3.3.2 Zero Crossing Rate:

Zero crossings are said to occur at the points where the signal attains zero amplitude. The rate at which the signal changes its amplitude from positive value to negative value and vice versa is called zero crossing rate. It is a measure of how many times the signal crosses the zero value in a particular time period. Zero crossing rate is highly used to distinguish between the silent and non-silent sounds based on their smoothness. Smoothness in non-silent sounds provides characteristic information of the cough sample. The smoothness in the sample is identified using the zero crossing rate by analyzing the sign changes within each frame.

The energy of the silent parts are mostly focused on the higher frequencies whereas for non-silent parts, the signal energy is found at lower frequencies. For example, as the signal oscillates, a 50Hz Non-silent sound may contain 50 zero crossings while a silent fricative sound may have 1000 zero crossings. Thus, a low frequency non-silent fricative signal will have low zero crossing rate while a silent signal will have high zero crossing rate. Figure 7 shows the zero crossing rate of sample signal with the cough event.

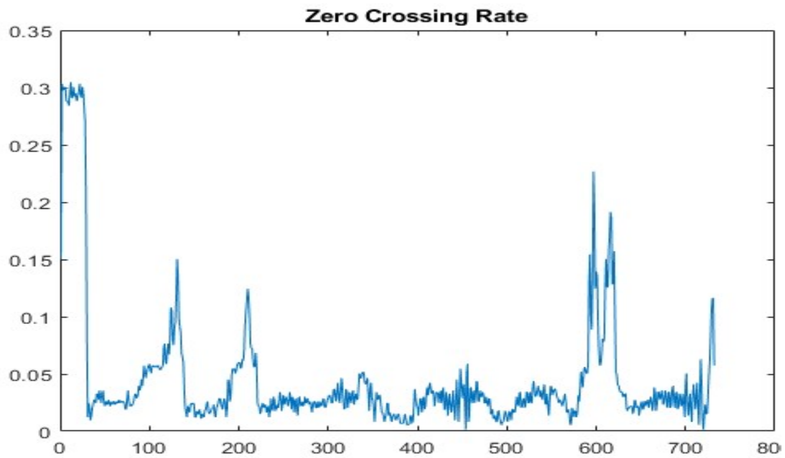


Fig. 7 Zero Crossing rate of sample audio signal

3.3.3 Linear prediction coefficient:

LPC is a format estimation technique which analyses the speech signal by evaluating the formants which correspond to the formant regions of the speech signal. The output of the LPC model includes residue where the residual sound is very close to the vocal tract input signal. Thus, the residual error is obtained by calculating the difference between the predicted and actual value.

The output of linear predictor model

$$\hat{s}(n) = \sum_{k=1}^{p} a_k s(n-k)$$

Where $\hat{s}$ is the predicted sample, p is the order and a_k is the prediction coefficient.

Transfer function system,

$$H(z) = \frac{G}{1 - \sum_{k=1}^{p} a_k z^{-k}}$$

Where the denominator indicates the inverse filter

After windowing each windowed frame is auto correlated and the order of LPC analysis is given by the highest auto correlation value. These auto correlated frames can be converted into LPC parameter set using Levinson-Durbin recursion method. This process is called LPC analysis. Figure 8 represents Linear Prediction coefficients of pure cough signal and pure non-cough signal with a 8th order FIR filter. Since the predictor coefficients show high variance they are never used in recognition. In each frame where LPC analysis can be done is answerable to the process of voiced or unvoiced. LPC yields better estimation of the spectral envelope of the vocal tract whereas in unvoiced regions it is less effective.

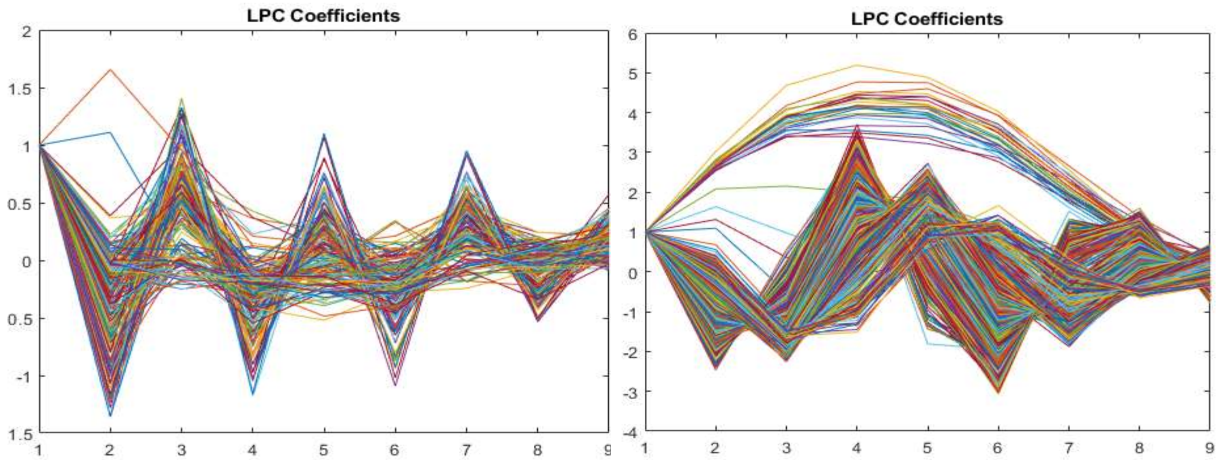


Fig. 8 LPC coefficients of sample cough event signal (left) and a non-cough event signal (right)

3.3.4 Spectral Flatness:

Spectral Flatness helps to evaluate how a sound is too close to being a noise which reflects the flatness properties of the power spectrum of noise. The standard definition of spectral flatness is that

$$Spectral\ Flatness = \frac{Geometric\ mean\ of\ signal\ magnitude\ of\ power\ spectrum}{Arithmetic\ mean\ of\ signal\ magnitude\ of\ power\ spectrum}$$

$$= \frac{exp(\frac{1}{N}\sum_{n=0}^{N-1} ln(x(n))}{\frac{1}{N}\sum_{n=0}^{N-1} x(n)}$$

Thus, the flatness of the signal is calculated per frequency bins and is equal to the average of the flatness values of each frequency bins. Spectral Flatness of an audio signal is shown in figure 9. It is measured in decibels (dB) ranging from -∞ to 0dB.

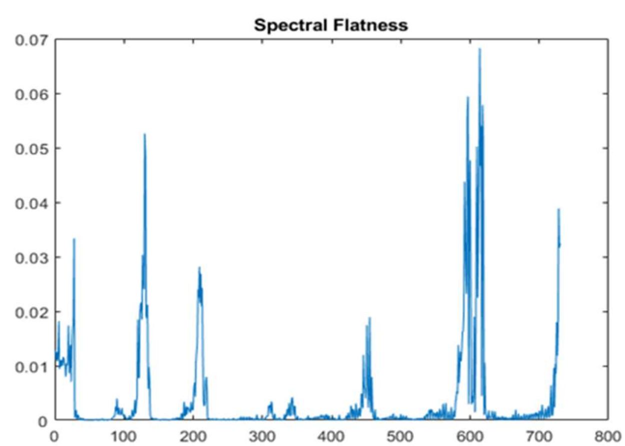


Fig. 9 Spectral flatness graph of sample audio signal

Spectral Flatness is also known as Wiener Entropy. The spectrum having high spectral flatness is said to have uniform and flat spectrum which is similar to white noise spectrum showing uniform/same amount of power in all frequency bins. Similarly, the spectrum with low spectral flatness will not have the same amount of power in all frequency bins. The spectrum will have abrupt changes in the power. The spectrum with high spectral flatness indicates the presence of noise while the spectrum of low spectral flatness value represents pure sound.

3.3.5 Spectral Centroid:

Spectral Centroid is expressed as "center of gravity" or "center of mass" since it locates the center of the mass of the spectrum or the spectral energy of each sub-band. Spectral Centroid gives the principal frequency of a signal and can be measured over the power spectrum which illustrates the sharpness of the sound. This is associated with the brightness of the sound.

The specialty of the spectral centroid is that it provides details about the peak positions of spectrum in each sub-band with the typical sub band power so that the robustness of cough recognition can be increased upon the included additional noises coming from the surroundings. Spectral centroid will mainly focus on large peaks. It will also find the location of the formants. Figure 10 represents the Spectral Centroid of sample signal.

In the spectral centroid frequency (SCF), the weighted average frequency is calculated using:

$$\text{Centroid} = \frac{\sum_{m-0}^{N-1} f(m)x(m)}{\sum_{m=0}^{N-1} x(m)}$$

Where $f(m)$ - Center frequency of the bin

$x(m)$ - Value of the weighted frequency and

m – no. of bins.

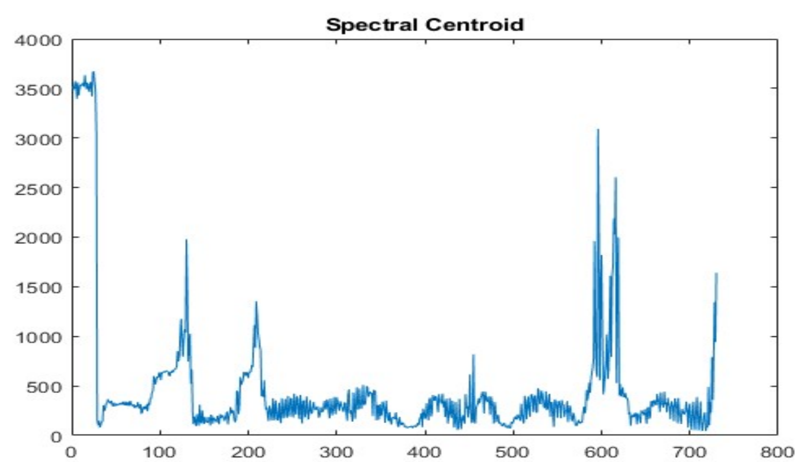


Fig. 10 Spectral Centroid graph of sample audio signal

3.4 Training and Validation

The first and foremost thing is to study or understand about the various acoustic features of the cough signal before attempting for classification. The characteristics features of the cough sounds rely on the physical size as well as the resonance (vibration) that occurs on the vocal tract. The study of cough features will lead to the detection or classification of Tussis in an acoustic signal. In order to detect and classify the cough sounds, the Machine Learning models are built to predict the output as either cough or non-cough sounds. The model is trained with the extracted acoustic feature and tested to observe its performance.

Machine Learning Algorithms used in this proposed work is: Support Vector Machine, K-Nearest Neighbor and Stochastic Gradient Descent with L1 and L2 regularization, Naive Bayes classifier and AdaBoost classifier.

3.4.1 Algorithms:

I. Support Vector Machine (SVM):

SVM is a supervised learning model which is useful for classification and regression. SVM has the unique attribute to increase the margin and reduce the classification error. SVM works great when a smaller number of features are considered.

A set of data is given to the SVM as input. For each input, there is a possibility that the input belongs to one of the two classes. This led the SVM to be a linear classifier. New samples are then mapped into the same space. This will predict to which category the data belongs to and it is based on which side of the gap they occupy. The class of each data point is specified using vectors. The hyperplane is established by using the support vectors and margins.

The advantage of Support Vector Machine (SVM) is that Overfitting can be prevented and the risk is less as it has good generalization capabilities. SVM is effective in cases where the number of samples is less than the number of dimensions and also in high dimensional spaces. SVM is stable since altering the feature extracted data will not much affect the hyperplane.

The disadvantages of SVM include overlapping classes when data is scattered which affects the performance. Also choosing the kernel function of SVM is highly difficult including difficulty in interpretation of SVM model with extended training time.

II. K Nearest Neighbor (KNN):

KNN classification is an intense based learning which is practiced for both classification and regression. KNN algorithm depends on a polling scheme which categorizes the group of which the data point belongs to. KNN classification is a simple, effective method and accuracy is better. The way of discovering K is favorable. K is also discovered by running the algorithm with various values of K and selecting the K value which provides best execution.

The accuracy of the algorithm will not get affected when new data are added which is a huge advantage in terms of efficiency. Other advantages of KNN include easy interpretation and neglected training step since there is not any discriminative function from the data set which is under training.

For higher dimensional data the complexity of prediction increases also with increase in calculated distance outliers will be greatly affected.

III. **Stochastic Gradient Descent (SGD)**:

Stochastic Gradient Descent is an iterative algorithm that is used to optimize the machine learning model. It randomly chooses the batch with a single sample for performing each iteration to attain the minimum value instead of using the whole dataset for computation. Each time the sample is reorganized before executing the next iteration. It therefore minimizes the value of the cost function.

Regularization is the process used to tune the function in order to overcome the error that occurs by regularizing the function.

There are many regularization techniques introduced to improve the efficiency of the model. Here, L1 and L2 regularization techniques are used to avoid the overfitting problem that occurs in the Stochastic Gradient Random algorithm.

SGD is incredibly flexible and loss function can be highly enhanced. It also doesn't require data pre-processing. As it need many trees it is expensive analytically. Also due to this interpretation of the model is difficult.

IV. Naive Bayes:

This classifier is simple and really fast classification method based on the Bayes theorem with every pair of features under classification assuming naïve independence. Even though they are fast their bad estimators lead to limited usage of them. The Gaussian Naive Bayes implements the algorithm as follows:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi{\sigma_y}^2}} exp\left(-\frac{(x_i - \mu_y)^2}{2{\sigma_y}^2}\right)$$

The parameters σ_y and μ_y are estimated using maximum likelihood.

The advantages of Naïve Bayes include easy estimation of test data set, easy implementation and better performance of with respect to other models with lesser training data.

But Naïve bayes considers all the features as mutually independent which is not ideal for many training tasks and also the model results in poor prediction as it assumes priori probability.

V. AdaBoost classifier:

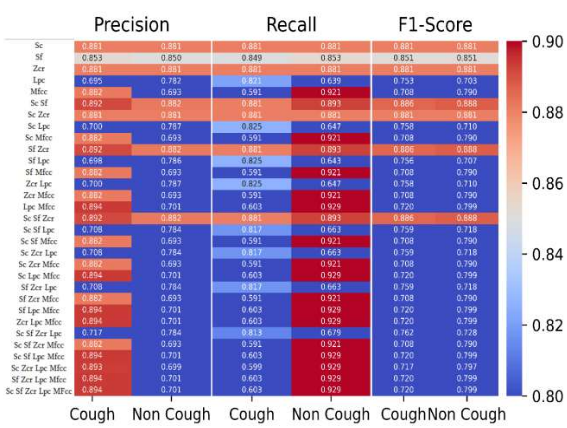
AdaBoost or Adaptive boost is an ensemble meta- learning algorithm and boosting technique which boosts prediction by trying to reduce the error or unpredictability based on the previous model. It corrects the previous model until maximum predictability is reached on the current model.

It is a very slow algorithm but also very efficient. It works based on a set of weak learners. Basically it sets example weight on the weak learner based on the ensemble predictions to improve the overall learning.

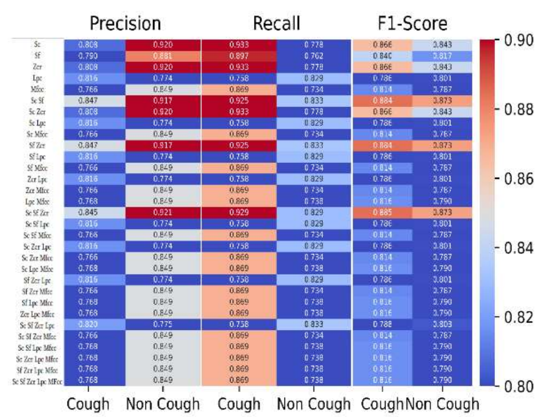
As it is an ensemble algorithm its high flexibility results higher efficiency. But it is highly sensitive to noisy data and outliers. So quality of data is highly important. Also due to its ensemble form, it is very much slower compared to other models and even Xgboost.

4. Training Results and Analysis

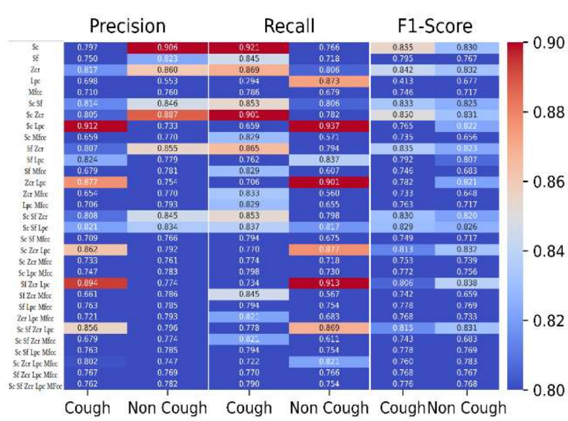
Each algorithm was separately trained on dataset with all the combination of features for more detailed observation. Each feature set was separately trained and then trained with combination of features.



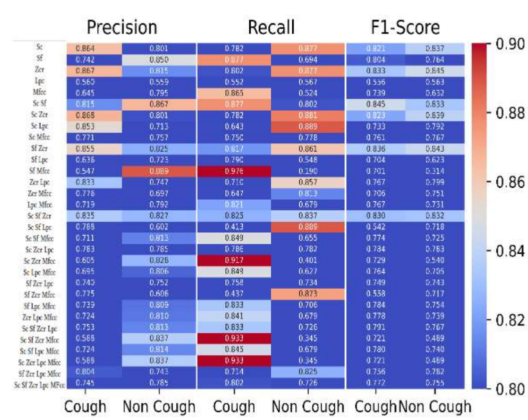
(a)



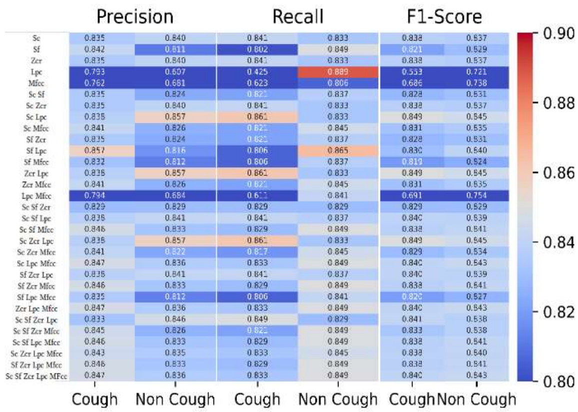
(b)



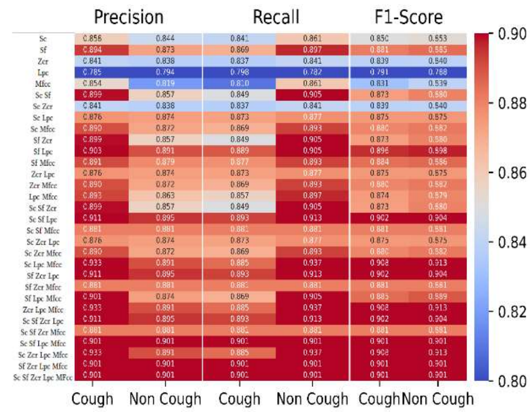
(c)



(d)



(e)



(f)

Fig. 11 The figures represent validated results for following algorithms (a) SVM, (b) KNN, (c) SGDCL1, (d) SGDCL2, (e) Naive Bayes and (f) Adaboost

The Support Vector Machine has a cough predictability with maximum precision in LPC and MFCC with 89.4% and recall of 88.1%. Same in non-cough the precision is highest with spectral Features 88.2% and maximum recall value of 92.9% is with MFCC and its combination. It has a high recall rate with non-cough for almost all the features signifying that SVM works best for evaluating speakers based on their vocal features as shown in figure 11(a).

Again spectral features show high precision and recall rate comparatively with other features while using this classifier for both cough and non-cough. Cough has highest precision and recall percent of 84.7% and 93.3% respectively while non cough has maximum 92.1% precision with spectral features and zero crossing rate, recall rate is at 83.3% again with spectral features.

As shown in figure 11(c), in the SGDC-L1 algorithm, when the features extracted from the combined SC, SF, LPC technique are taken into consideration, the model obtain the recall of 83.7% for cough and 81.7% for non-cough while the precision of the model is 82.1% for cough and 83.7% for non-cough. In SGDC-L2 algorithm(figure 11(d)), when SC, SF, ZCR are combined and their features are taken into consideration, the precision of the model is 82.5% for cough and 83.7% for non-cough while the model gives recall of 83.5% for cough and 82.7% for non-cough. When both the regularization techniques are compared, L1 and L2 show comparatively equal results as both the techniques have equal but different properties**.**

The combination of Spectral Flatness and Linear Prediction Coefficient show the maximum precision of 85.7% and the LPC with SC and ZCR show the maximum recall value of 86.1% for cough prediction. As Naive Bayes is an independent event classifier it shows comparatively lower output in dependent features such as MFCC where 1-5 coefficients are needed to classify a cough efficiently.
For Non-Cough there is a maximum of 88.9% recall value with Linear Prediction Coefficient alone and SC, ZCR shows best results with LPC with a maximum of 85.7%. Overall LPC works best with Naive Bayes algorithm projecting its independent feature quality.

For cough Prediction with AdaBoost (in figure 11(f)), LPC and MFCC shows high predictability rate with up to 93.3% for precision and 90.1% for recall. For non-cough again LPC and MFCC with ZCR

show high recall value where highest precision is offered by LPC and MFCC combined with spectral features centroid and flatness.
As AdaBoost is an ensemble method the results are pretty high compared to other classifiers but it also takes too much time for classification. The base estimator used here is a Decision Tree classifier. For each of these features the no of estimators is kept at 50 even though when working with spectral features it stops early at 30 considering a perfect fit.

5. Summary

Cough is a protective reflex, a strong expulsion of air from the lungs through airway which helps to clean out things like foreign bodies, micro bodies etc. from the respiratory passages. Even though cough is a protective mechanism that helps to remove any unhealthy particles in the respiratory tract, Cough is found to be the most common symptom of illness that occurs on human body.

The proposed methodology is divided into four sections: Dataset Acquisition, Prepossessing, Feature Extraction and Classification. A total of 1000+ sample recordings were taken and compiled together as signals with and without cough event.The samples taken are resampled to 8000 Hz since most of the features or information regarding cough are observed to contain below 4000 Hz. Then they are filtered using a high pass filter of specific low cut off frequency to remove noise. Later the silent parts are removed by a threshold method and split into frames for fast fourier transform.The FFT is applied after passing the signal through the Blackman window.Feature extraction is a process of obtaining the reduced features as it provides pertinent information of the input sample without altering the information of the original sample. It increases the accuracy of the process by using the extracted features. In this paper we have extracted MFCC, spectral centroid, spectral flatness, zero crossing rate and linear prediction coefficient.

In order to detect and classify the cough sounds, the Machine Learning models are built to predict the output as either cough or non-cough sounds based on the extracted features. The following machine learning algorithms were used to test their efficiency in detection of cough event in this paper. Support vector Machine(SVM), K-nearest neighbour(KNN), Stochastic Gradient Descent(SGD), Naive bayes and Adaboost.

Each Algorithm has different sets of efficiency in different sets of features as included in the paper. The overall highest efficiency was obtained from adaboost with all the sets of extracted features which is at 90% recall, F1 score and precision. This proposed work is further planned to be improved on cough detection in real time unknown signals and also by offering more precision and lesser error rate.

6. Conclusion

With diseases increasing day by day, medical technology should also keep up with it to save human lives. Efficient and automatic cough segment detection is one such method to diagnose diseases related to respiratory systems to avoid complications. With this as the motivation we have focused on improving the cough detection system very simple and efficiently enough for real time diagnosis. Five different set features have been extracted from cough events after all the preprocessing stages. These features are classified with different types of classifiers to test for its efficiency. Each showed a different set of results according to their characteristics. The overall training results of commonly used MFCC algorithm have been tabulated against all the extracted features in the table shown below.

Table 2 Comparison between MFCC and other features on the basis of classifiers

Classification Method	**Features Selection**	**Detected Class**	**Precision**	**Recall**	**F1 score**
SVM	MFCC	Tussis	0.88	0.59	0.71
		Non Tussis	0.69	0.92	0.79
	MFCC+ZCR+LPC+SC+SF	Tussis	0.89	0.60	0.72
		Non Tussis	0.70	0.93	0.80
k-NN	MFCC	Tussis	0.77	0.87	0.81
		Non Tussis	0.85	0.73	0.79
	MFCC+ZCR+LPC+SC+SF	Tussis	0.77	0.87	0.82
		Non Tussis	0.85	0.74	0.79
SGDC l1	MFCC	Tussis	0.76	0.53	0.62
		Non Tussis	0.64	0.84	0.73
	MFCC+ZCR+LPC+SC+SF	Tussis	0.67	0.85	0.75
		Non Tussis	0.80	0.58	0.67
SGDC L2	MFCC	Tussis	0.77	0.69	0.73
		Non Tussis	0.72	0.80	0.76
	MFCC+ZCR+LPC+SC+SF	Tussis	0.61	0.91	0.73
		Non Tussis	0.83	0.42	0.55
Naïve Bayes	MFCC	Tussis	0.76	0.62	0.69
		Non Tussis	0.68	0.81	0.74
	MFCC+ZCR+LPC+SC+SF	Tussis	0.85	0.83	0.84
		Non Tussis	0.84	0.85	0.84
AdaBoost	MFCC	Tussis	0.85	0.81	0.83
		Non Tussis	0.82	0.86	0.84
	MFCC+ZCR+LPC+SC+SF	Tussis	0.90	0.90	0.90
		Non Tussis	0.90	0.90	0.90

Overall the AdaBoost classifier gave the best result with all the five features combined for both the cough and non-cough when compared to all the other ML models with difference set of combination. So this preferred method gives 90% precision and recall in both cough and non-cough. This consistent performance across all the parameters markers makes it an excellent framework for detection of a cough event in an audio signal.

7. Future Work

The Machine learning Model in spite of its high efficiency is still in its development stage. Even though it can identify a cough event efficiently, it has to be an unknown signal without prior calibration to the type of signal it is given. It needs more tuning and development. Also the recent development of transfer learning has significantly boosted the efficiency of Machine oriented learning. So for future development we focus on combining Machine learning models with deep learning architecture to make a more efficient and compact framework for efficient detection of cough events in an unknown signal. Also the preprocessing is done for quality samples as some machine models like AdaBoost are heavily sensitive to it. But preprocessing a real world signal without knowing its complete nature can result in degradation of signal more. So noise sensitivity should be further reduced in our model.

8. References

[1] Mello, C. J., Irwin, R. S., & Curley, F. J. (1996). Predictive values of the character, timing, and complications of chronic cough in diagnosing its cause. Archives of internal medicine, 156(9), 997-1003.

[2] Ashurst, L., Smith, J. A., Jack, S., Woodcock, A. A., & Earis, J. E. (2003). Subjective recognition of cough sounds by respiratory professionals. Eur Respir J, 22(suppl 45), 172s.

[3] Barata, F., Kipfer, K., Weber, M., Tinschert, P., Fleisch, E., & Kowatsch, T. (2019). Towards device-agnostic mobile cough detection with convolutional neural networks. In 2019 IEEE International Conference on Healthcare Informatics (ICHI) (pp. 1-11). IEEE.

[4] Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Scherer, K., Ringeval, F., ... & Mortillaro, M. (2013). The INTERSPEECH 2013 computational paralinguistics challenge: Social signals, conflict, emotion, autism. In Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association, Lyon, France.

[5] Yadav, S., Keerthana, M., Gope, D., & Ghosh, P. K. (2020). Analysis of Acoustic Features for Speech Sound Based Classification of Asthmatic and Healthy Subjects. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 6789-6793). IEEE.

[6] Digital Signal Processing: Principles, Algorithms, and Applications Book by Dimitris Manolakis and John G Proakis., Pearson Publication

[7] Swarnkar, V., Abeyratne, U. R., Amrulloh, Y., Hukins, C., Triasih, R., & Setyati, A. (2013). Neural network based algorithm for automatic identification of cough sounds. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1764-1767). IEEE.

[8] Matos, S., Birring, S. S., Pavord, I. D., & Evans, H. (2006). Detection of cough signals in continuous audio recordings using hidden Markov models. IEEE Transactions on Biomedical Engineering, 53(6), 1078-1083.

[9] Ranjan, R., & Thakur, A. (2019). Analysis of feature extraction techniques for speech recognition system. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(7C2).

[10] Swarnkar, V., Abeyratne, U. R., Amrulloh, Y. A., & Chang, A. (2012). Automated algorithm for Wet/Dry cough sounds classification. In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 3147-3150). IEEE.

[11] Pramono, R. X. A., Imtiaz, S. A., & Rodriguez-Villegas, E. (2019). Automatic Identification of Cough Events from Acoustic Signals. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 217-220). IEEE.

[12] Rocha, B. M., Mendes, L., Couceiro, R., Henriques, J., Carvalho, P., & Paiva, R. P. (2017). Detection of explosive cough events in audio recordings by internal sound analysis. In 2017 39th Annual international conference of the IEEE engineering in medicine and biology society (EMBC) (pp. 2761-2764). IEEE.

[13] Rocha, B. M., Pessoa, D., Marques, A., Carvalho, P., & Paiva, R. P. (2020). Personalized Detection of Explosive Cough Events in Patients With Pulmonary Disease. In 2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON) (pp. 249-301). IEEE.

[14] Miranda, I. D., Diacon, A. H., & Niesler, T. R. (2019). A comparative study of features for acoustic cough detection using deep architectures. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2601-2605). IEEE.