A Dimensionality Reduction based Efficient Multiple Voice Disease Recognition Scheme using Mel-Frequency Cepstral Coefficients and K-Nearest Neighbors Algorithm

Shovan Bhowmik[1, *[0000-0001-7218-9875]], Mahedi Hasan[2[0000-0002-1835-9439]] and Muhammad Ataul Hakim[2[0000-0003-3212-0456]]

[1] Bangladesh Army International University of Science and Technology, Cumilla Cantonment, Bangladesh
bhowmik.sshovon5795@gmail.com
[2] Khulna University of Engineering & Technology, Khulna, Bangladesh
mahedi049@gmail.com
akif.ah16@gmail.com

Abstract. Disease diagnosis in medical healthcare leveraging machine learning is an area of significant interest in the whole world. Audio data analysis has facilitated the route of identifying voice disorders in a non-invasive manner. As vocal cord malady can immerse anytime from various bad habits (loud sound, smoking, extra force on vocal cords, etc.) and neurological imbalance, early recognition of voice disorders can save people from causing long-term damage. In this research, the three most injurious voice diseases have been recognized exploiting the LDA based MFCC feature matrix and KNN algorithm. Our empirical study has successfully predicted four voice class labels using this scheme with the foremost 96.49% accuracy and surpassed five other mining algorithms including Artificial Neural Network. Moreover, this model can generate output within 22.37 milliseconds which is also the lowest execution time compared to the traditional clustering methods. Our proposed model can be easily implemented to design end-to-end service for efficient voice disease classification.

Keywords: Voice Pathology, Linear Discriminant Analysis, MFCC, K-Nearest Neighbors, Artificial Neural Network, Voting Classifier.

1 Introduction

With the advent of powerful computational devices and the conception of Machine Learning (ML), researchers have shifted their experiments towards non-invasive disease identification in medical science which can help people to identify malady in a fast, low-cost and trustworthy way. These non-invasive technology dependent modules can assist people not only in early diagnosis of disease but also aid quick recovery from ailments. Voice pathology detection has been a promising area where extracting different acoustic features, for example, Shimmer (%), Jitter (%), Fundamental Frequency (F0), Mel-Frequency Cepstral Coefficients (MFCC), Wavelet Packet Decomposition (WPD), etc. along with ML and Deep Learning (DL) algorithms can produce disorder recognition framework [1].

Generally, voice pathologies are created due to the existence of tissue contagion, breathing disturbance, neurological imbalance, muscular substitution, vocal fold contraction, vocal surface vexation, tissue cell alteration and other factors [2]. There are around 71+ voice diseases associated with voice disorder which include, Dysphonia, Laryngitis, Reinke's Edema, Vocal Fold Nodules and Polyps, Vocal Cord Paralysis, etc. [3][4]. Dysphonia is a voice disease that occurs because of the creation of nodules in the vocal cords, swelling in the larynx, or the sudden traumatic event in the cords. Around 10% people of in the world encounter this type of disease [5]. Another commonly occurring voice disease is Laryngitis which is caused by the swelling of vocal folds. This disease can become acute sometimes if viruses attack the vocal folds [4]. Reinke's Edema can happen by taking overstress or because of immoral habits like smoking, loud shouts, etc. Other voice abnormalities can appear for the aforementioned causes as well. Almost all the voice diseases make the sound scratchy and gruff. As voice is produced from neurological signals, these problems can also hamper brain cells [6]. Thus, careless attitudes obviate severe situations which can be sometimes not curable by surgeries and might lead to gruesome cancer.

Early detection of voice pathology can reduce the risk of grievous circumstances. So far, most of the works related to voice disease recognition using ML and DL are based on the binary classification that predicts a voice sample whether it is healthy or pathological [7][8]. Among traditional ML algorithms, Support Vector Machine (SVM), Gaussian Mixture Model (GMM), Decision Tree (DT), K-Nearest Neighbors (KNN), etc. had been widely used for voice disease detection. Several DL algorithms, e.g., Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) - CNN Hybrid, Bidirectional LSTM, had been considered for this task in previous works too. Previously, multiple public voice datasets were incorporated for this job. For instance, MEEI, SVD and VOICED datasets have been constantly employed for classifying healthy and diseased voices [9]. Unfortunately, the majority of the research works focused on only two class labels and the accuracy was ranged approximately between 70 to 94 percent.

In this work, we have proposed a Linear Discriminant Analysis (LDA) based multiple voice disease distinguishing framework by exploiting MFCC features and the K-NN algorithm. We have introduced three crucial voice disorder classifications, e.g. Dysphonia, Laryngitis and Renkei's Edema along with healthy voices by training our model with the SVD dataset. Moreover, our study has shown that LDA based K-NN algorithm can reach up to 97% accuracy in multiple voice pathology categorization by taking only 13 MFCC attributes. Additionally, four cross-validation techniques have been included in this work for evaluating model performance as our dataset was imbalanced. To benchmark our diagnosis scheme, we have employed ML based supervised polynomial kernel-based SVM, probabilistic GMM, ensemble-based Random Forest (RF), a voting classifier including SVM, KNN and RF, and DL based 3-Layer ANN. Our experiment has found that the accuracy achieved by applying K-NN with Shuffle Split and Stratified Shuffle Split cross-validation techniques can outperform the ANN model. Another highlighted characteristic of this experiment is the intensive comparison among models on execution time.

2 Literature Review

ML and DL models have been devised for voice pathology detection in many previous studies. MFCC and Acoustic Features (AF) had been extracted for applying XGBoost, Isolation Forest and Dense Net to distinguish diseased voices from normal ones in [10]. Four well established public voice databases namely AVPD, MEEI, PDA and SVD was availed with audio files consisting of /a/e/o/ vowels. The highest 0.733 F1-Score and 0.759 precision were achieved from this experiment where dimensionality reduction was not applied for training purposes. Dysphonia was detected by applying SVM, K-NN and RF in [11]. In this paper, Shimmer, Jitter, MFCC, etc. features were considered for three vowel pronunciation recordings from the SVD database. The highest 91.3% classification accuracy was attained for the RF model. Principal Component Analysis (PCA) was applied here for reducing dimensionality that results in a low accuracy ranging from 65% to 80% on average to classify two labels. A comparative analysis was illustrated among Bagging, Boosting and LibD3c ensemble learning models for multiple audio features like Pitch, Intensity, Global Features, MFCC, WPD, CA, etc. in [1]. Although around 96% accuracy was gained in this work, the authors did not validate their model using cross-validation. Moreover, specific voice diseases were not classified in this work. However, K-NN and LDA-based voice ailment recognition was done in [12] which achieved 93% accuracy with a private dataset.

Voice malady identification was established in [13] using Deep Neural Network architecture. A hybrid combination of the CNN-LSTM-Dense layer was proposed in this paper. The validation report demonstrated only 71.36% accuracy in diagnosing pathology in this study. In [14], a CNN framework was devised for the binary classification of voice samples using the SVD database. The authors of this paper achieved 95.41% accuracy in this case with a balanced dataset. However, Cyst, Polyp, Paralysis and healthy voice samples were clustered using VGG16 and CaffeNet models for MEEI and SVD databases in [8]. This work accomplished at most 94.5% accuracy by analysis Fourier Transform of voice spectrums.

The majority of the above-mentioned studies gave priority to only two classes and only several vowel utterances were considered for drawing out voice attributes. Cross-validation was rarely taken into account. Another limitation was the absence of calculating processing time for model execution. Although the speech recognition area has few pieces of research on the distinction between LDA and PCA performance regarding diabetes prediction [15], speech recognition lacks any solid study on it. This research work has shown state-of-art performance in classifying four voice classes utilizing full sentence utterance recordings of the SVD database.

3 The Proposed Multiple Voice Disease Recognition Scheme

In our study, we have emphasized recognizing three types of voice diseases, particularly, Dysphonia, Laryngitis and Reinke's Edema. Several ML and DL algorithms have been selected to fit the MFCC features after reducing the dimensions of the voice attributes using LDA for all the voice samples to finally predict class labels. The overall working principle has been illustrated in Fig. 1.

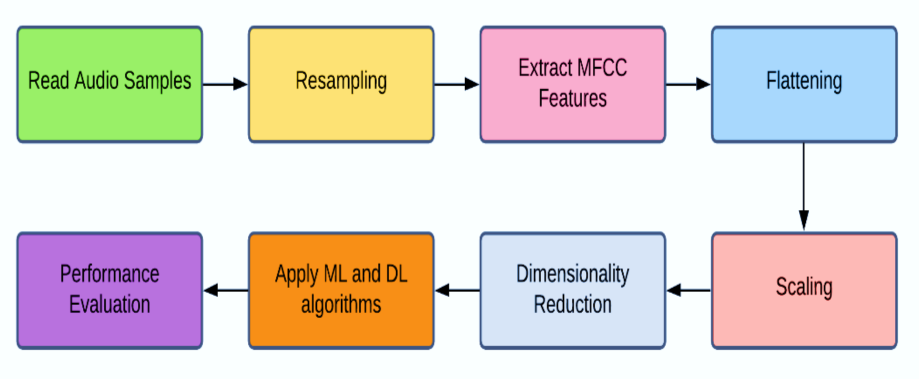

Fig. 1. The Overall Workflow of the Multiple Voice Disease Detection Model

3.1 Dataset Description

We have selected 367 voice samples from a well-known and commonly used audio database namely the "Saarbruecken Voice Database" (SVD) [16]. Whereas most of the works highlighted in the previous section selected few vowels for their research activities, we have taken a full sentence audio clip of ".wav" format to check model performance on lengthy sequence data. Each of the audio clips picked for this work utters "Guten Morgen, wie geht es Ihnen?" and all of the voices are sampled in 50 KHz in this database. Additionally, we have taken only those voices where each voice clip consists of only a single class label like Dysphonia or Laryngitis or Renkei's Edema or Healthy. The total audio dataset statistics are shown in Table 1.

Table 1. Dataset Information

Category	**Gender**		**No. of .wav file**
	Male	**Female**	
Dysphonia	10	42	52
Laryngitis	76	47	123
Renke's Edema	30	22	52
Healthy	70	70	140
Total No. of Voice Samples			367

3.2 Classification Process

This subsection presents the step-by-step classification process of the voice disease detection using ML and DL algorithms.

Audio Signal Reading and Resampling

The voice samples have been downloaded from the SVD database and read using "Librosa" [17] which is a speech analysis bundle. The usage of this package has facilitated the conversion of all the signals into a "mono" channel by normalizing the whole dataset and resampled all the speeches into 22 KHz. Among all four class labels, the waveform representation for the "Laryngitis" class is displayed in Fig. 2.

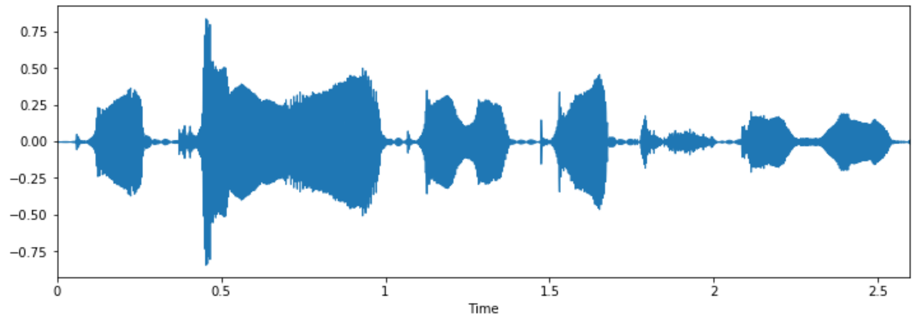

Fig. 2. The Waveform of Laryngitis Category

MFCC Feature Extraction

Mel-Frequency Cepstral Coefficients (MFCC) are computed based on the sample rate and bit value of all the input audio signals to determine the spectral information of the voice [18]. Commonly, 12 to 20 MFCC values are measured for voice characteristic analysis [19]. In our research, we have calculated 13 MFCC features for each voice signal using Eqn. 1.

$$MFCC_m = \sum_{k=1}^{M} (\log(E_k) \times \cos\left[m(k-0.5)\frac{\pi}{M}\right]) \quad (1)$$

In (1), M = Total no. of bands in the signal, k = i-th band number, E_k = Energy value of the k-th band frequency and m = No. of frequency.

The highest number of frames for a single input speech is 213 for our dataset. 13 MFCC feature values have been obtained for all the frames of every individual recording.

Flattening and Scaling

After retrieving the MFCC features for every frame of all 367 voice samples, we have flattened all the MFCC attributes in a single vector. For the highest length audio file, the (213 × 13) matrix has been converted into a (1 × 2769) vector. As the span for all the signals is not the same, we have padded the less no. of frame signals with "0" based on the largest speech vector. Consequently, we have produced a (367 × 2769) MFCC feature matrix after flattening all the sound waves.

As long as the feature values correspond to high numeric numbers, it is required to normalize those numbers into a smaller range for building a cost-effective model. Thus, we have used "Standard Scaler" for scaling the large floating numerals into small decimals.

Dimensionality Reduction

Dimensionality reduction can make the training faster with good accuracy. Moreover, it can visualize the dataset efficiently to resolve inspection [20]. Although PCA works remarkably for category type features and clustering algorithms have gained good performance occupying PCA to label several classes [21], LDA utilization has provided better results for sequential data modeling in our study. When we have visualized our dataset using Scatter Plot and taking the two most explained values of both PCA and LDA, the data points have been placed almost in a similar range for the PCA survey. On the other hand, LDA consideration has distinguished all four labeled data constructively. As a result, we have fit all the data using LDA to reduce the dimensionality before applying ML and DL algorithms. The employment of LDA has converted the (367 × 2769) feature matrix to (367 × 3) MFCC attribute matrix. The characteristic matrix has been reduced to (367 × 367), while we have applied PCA. The Scatter Plot diagram for our dataset is illustrated in Fig. 3.

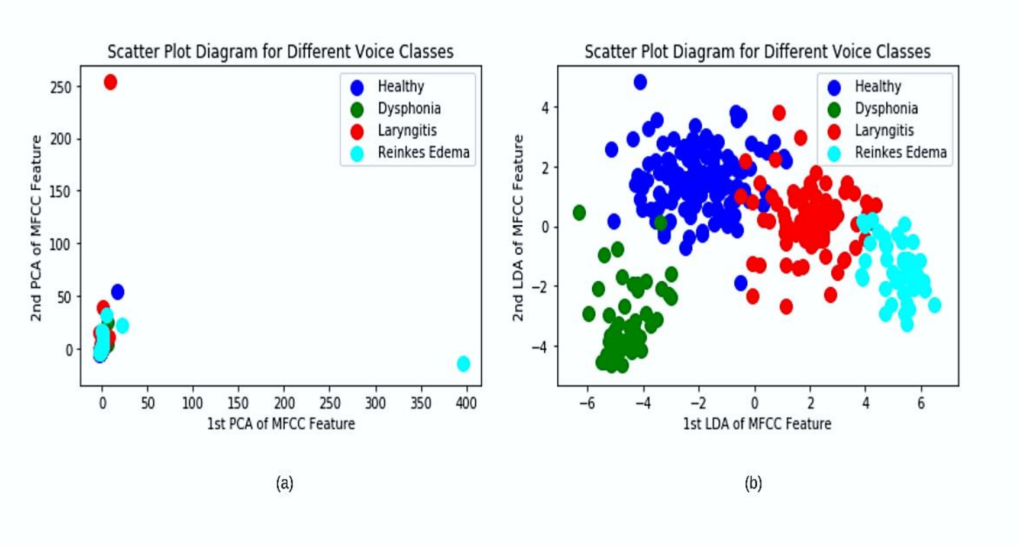

Fig. 3. Scatter Plot Diagram for our Dataset for most the two important values of (a) PCA, (b) LDA

Applying ML and DL Algorithms

After dimensionality reduction with the help of LDA, we have fit the decreased audio feature matrix to some clustering algorithms such as SVM, GMM, KNN, RF and a voting model comprising SVM, KNN and RF. We have chosen one supervised (SVM), one probabilistic (GMM), one unsupervised (KNN) and one ensemble-based (RF) ML algorithm. Polynomial Kernel of degree '3' has been selected when SVM is considered. We have tried varying parameters for GMM, but none has achieved satisfactory performance. But for comparison, we have taken 'Spherical' as the covariance type with 300 as the maximum number of iterations. The default parameters have been chosen for KNN and RF. In the voting classifier, we have selected the same parameters considered in the training of the dataset for each discrete algorithm. For the ANN architecture, we have added three dense layers where the first two layers consist of 256 neurons and both are followed by a dropout layer. "Relu" has been used as the activation function and "Adam" optimizer with the learning rate of 0.0001 has been selected for the ANN model. The batch size for the ANN model has been 10 and 100 epochs have been picked for this three-layer neural network model. The model fitting on the reduced feature matrix is visualized in Fig. 4.

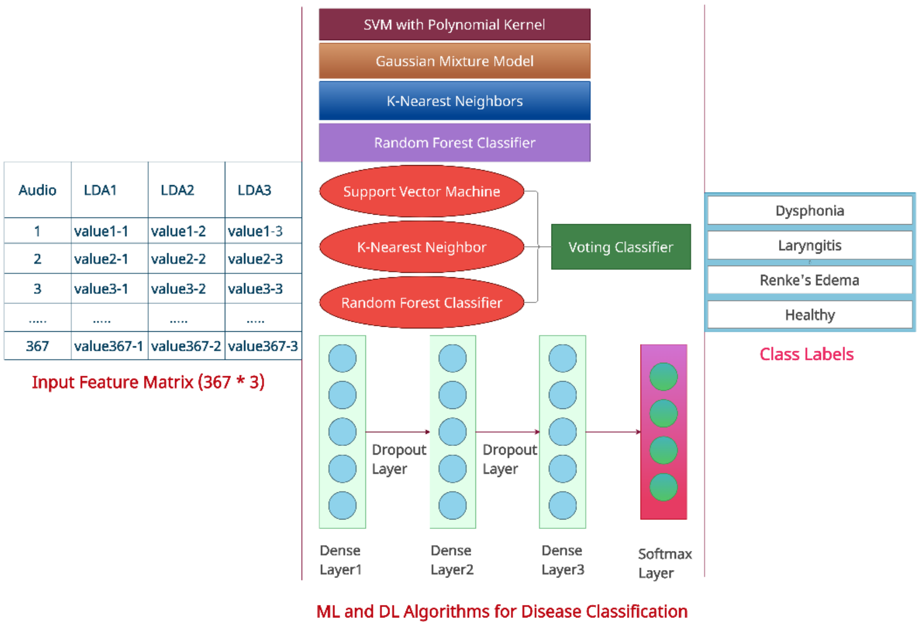

Fig. 4. Applying ML and DL Algorithms to determine four Class Labels

Performance Evaluation

Because of the imbalanced dataset, we have measured accuracy, F1-Score and AUC area for the whole dataset using a cross-validation approach. Four popular cross-validation techniques have been employed namely “K-Fold”, “Stratified K-Fold”, “Shuffle Split” and “Stratified Shuffle Split”. We have calculated the accuracy of the recognition framework for all the mentioned techniques and computed the other two performance evaluation metrics only for the best cross-validation models. In our study, we have placed the voice features into 10-folds for the efficacy assessment. Moreover, we have also determined the total time needed for executing all four cross-validation models by each algorithm.

4 Experimental Results

This research has dwelled on the performance of LDA based dimension minimization for sound waves. The finding of this work implies significant results when KNN has been applied to detect different voice diseases. Other ML and DL algorithms have also performed saliently on LDA based features.

This work has mainly aimed to find the performance of ML and DL algorithms accompanying dimensionality reduction techniques to detect multiple audio classes. Therefore, we have taken only the voice samples which are related to a single category. Due to the lack of a single labeled dataset available in the public domain, we have to rely on a small bunch of audio samples. Moreover, according to the “Hopkins Medicine Organization” [3], people suffer more from Laryngitis, Dysphonia and Renkei’s Edema that occur because of vocal cord trouble.

In this proposed model, the dimensionality of the MFCC feature matrix has been brought into (367 × 367) for PCA based classification and (367 × 3) for LDA based prediction. PCA works on the correlation between the features by placing the data points in the subspace which has the maximum variance [22]. This has resulted in an unsupervised separation of features in the orthogonal axes and a covariance matrix has been calculated to find the eigenvectors which are multiplied with the features to reduce the dimension. Consequently, we have found 367 principal components (PC) for individual audio. These PCs have explained the majority of the characters in each audio sample. In contrast, LDA always tries to sum up the vector components. Because of the known class labels, 1-D mean vectors are generated for each class from where scatter matrices are formed. Each scatter matrix computed from the mean vectors is added together to create a within-class matrix. Between-class matrices are also devised from the whole dataset by subtracting the overall mean from each feature value of the dataset. Finally, the within-class and between-class metrics are multiplied to get the reduced feature matrix [22][23]. By applying these steps, LDA has identified three essential features for each sample that can explain the distinct characters of each audio file. As a result, we have achieved a smaller matrix compared to PCA based reduction.

Table 2 shows the mean accuracy for the several cross-validation methods while taking MFCC voice features without decreasing dimensions, applying PCA and considering LDA respectively.

Table 2. Performance Analysis of Several ML and DL Algorithms for Four Cross-Validation Techniques

Dimensionality of the Voice Features	ML and DL Algorithms	Cross-Validation Methodology Accuracy (%)			
		K-Fold	Stratified K-Fold	Shuffle Split	Stratified Shuffle Split
No Dimensionality Reduction (367 × 2769)	SVM	39.46 (±15)	38.19 (±2)	35.41 (±12)	37.84 (±1)
	GMM	47.22 (±13)	55.56 (±19)	**56.77 (±2)**	35.14 (±18)
	KNN	31.40 (±18)	39.45 (±11)	41.08 (±15)	38.92 (±16)
	Random Forest	39.38 (±19)	48.60 (±13)	46.49 (±18)	46.22 (±16)
	Voting Classifier	39.20 (±10)	41.41 (±8)	42.16 (±20)	39.46 (±4)
	ANN	44.50 (±12)	48.27 (±9)	48.92 (±14)	47.84 (±12)
Dimensionality Reduction using PCA (367 × 367)	SVM	45.36 (±14)	43.10 (±6)	44.33 (±11)	44.60 (±10)
	GMM	48.77 (±4)	33.33 (±17)	37.84 (±13)	32.43 (±10)
	KNN	41.40 (±19)	39.45 (±11)	39.46 (±7)	40.54 (±5)
	Random Forest	38.76 (±8)	34.89 (±9)	41.08 (±10)	38.64 (±14)
	Voting Classifier	45.73 (±12)	43.22 (±5)	47.84 (±14)	44.89 (±11)
	ANN	**59.06 (±16)**	44.96 (±6)	45.95 (±14)	43.78 (±13)
Dimensionality Reduction using LDA (367 × 3)	SVM	91.56 (±8)	93.86 (±6)	94.87 (±3)	94.87 (±3)
	GMM	50.32 (±19)	69.44 (±10)	74.38 (±3)	62.97 (±10)
	KNN	93.73 (±5)	95.14 (±3)	**96.49 (±3)**	95.14 (±4)
	Random Forest	91.83 (±8)	93.28 (±6)	92.16 (±5)	95.41 (±4)
	Voting Classifier	91.82 (±8)	94.64 (±5)	**95.68 (±4)**	95.41 (±3)
	ANN	92.37 (±5)	94.25 (±5)	93.78 (±6)	**96.22 (±3)**

From Table 2, it can be easily observable that exploiting LDA has helped to far outweigh the accuracy of the detection models using PCA and without dimensionality reduction schemes. We have achieved the best accuracy of 96.49% when KNN has been applied to LDA based reduced attribute matrix and when shuffle split has been considered as the cross-validation approach. For the same validation technique, 95.68% accuracy has been achieved for the voting classifier as well. Moreover, the Stratified Shuffle Split method has shown more than 96% accuracy for the same feature vector. Other algorithms have also obtained good accuracy of more than 91% except for the GMM model which has conveyed a lower performance in multiple voice disease recognition processes. However, our findings exhibit poor output for the other two criteria. For instance, direct feed of MFCC values in ML and DL algorithms has provided accuracy ranging from 35% to 57%. GMM has achieved the largest accuracy of only 56.77% which is not a satisfactory value. Furthermore, PCA-based diagnosis models have performed poorly too. In this case, 59.06% accuracy has been reached by ANN when the K-Fold method has been realized. Because of splitting our dataset into 10 folds, the accuracy has varied obviously, let alone the extent of which is mentioned in the table.

Since LDA is a supervised algorithm that incorporates Bayesian theory to discriminate multiple classes by maintaining linear distance between class labels, it can outperform unsupervised PCA algorithm in the case of multi-label clustering with a smaller dataset [24]. This probabilistic modeling might have helped to classify three diseases by separating categories depending on the highest probability of the feature values with an upper success rate than PCA.

As LDA depended clustering framework has outperformed the other two criteria by a massive margin, we have observed the F1 Score, AUC area and execution time for the best cross-validation results for each model. Table 3 provides the same.

Table 3. F1-Score, AUC area and Execution Time of the ML and DL Algorithms

Name of the Algorithm	**F1-Score (%)**	**AUC**	**Execution Time (milliseconds)**
SVM	94.32	0.9506	23.99
GMM	63.28	0.6639	69.96
KNN	**95.83**	**0.9709**	**22.37**
Random Forest	95.13	0.9572	141.17
Voting Classifier	95.05	0.9583	199.14
ANN	95.54	0.9669	103135.49

Our findings from Table 3 imply that LDA based approach has carried out salient achievement for F1-Score, AUC and Execution time accordingly. Excluding the probabilistic GMM model, all the other models have gained more than 95% F1-Score value which indicate significant precision and recall values for this purpose. Among all the models, KNN provides the foremost outcome with extraordinary F1-Score and AUC value along with only 22.37 milliseconds execution time. ANN has been praiseworthy too, but it is costly with an enormous execution time. Even though the mean accuracy for multiple cross-validation techniques is almost similar for SVM, VC, KNN and ANN, the AUC result is much more perfect for KNN with just over 97%. This suggests the balance between the prediction of true positive rate (TPR) and false positive rate (FPR) is more satisfactory. Moreover, the time for model training is also lower than even SVM for the KNN algorithm. As a result, we have placed KNN on the priority list in the multiple voice disease recognition scheme.

Fig. 5 portrays the average accuracy of all the cross-validation methods accomplishing the recognition task.

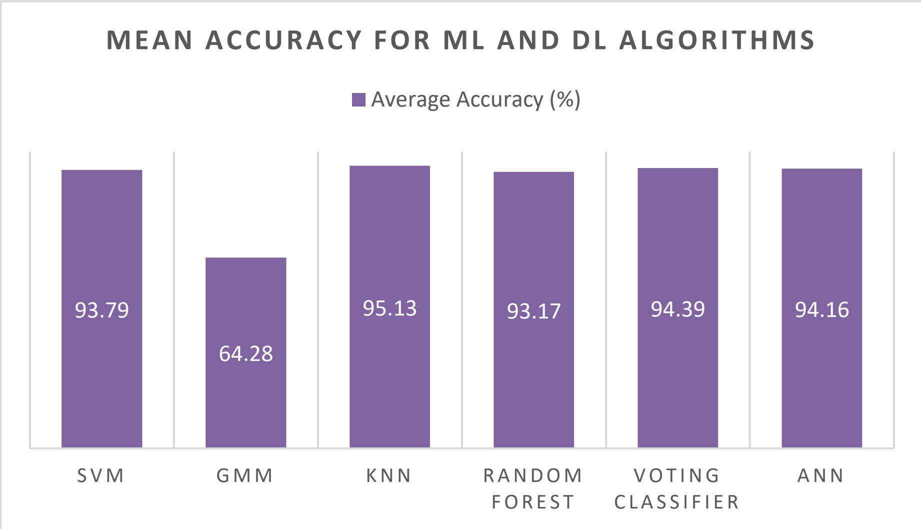

Fig. 5. Mean Accuracy of the Cross-Validation Approaches for LDA based Disease Recognition

From Fig. 5, it can be interpreted that, KNN functions better in measuring the mean accuracy among all the algorithms. As execution time of a model is also an important parameter as computationally expensive models often face resource constraints [25], the execution time comparison in model fitting for the KNN algorithm is illustrated in Fig. 6.

When LDA features are selected, KNN takes only around 23 seconds to complete the prediction as stated in Fig. 6. The other two criteria take a long time than LDA. The less execution time for LDA based feature extraction can be realized in a way that the feature matrix for LDA is (367×4) in size whereas PCA based attribute matrix is off (367×367) dimensions and which is extended massively for no reduction scheme with (367×2769) area. Thus, LDA has provided more correct predictions with a shorter amount of time for this sequential data.

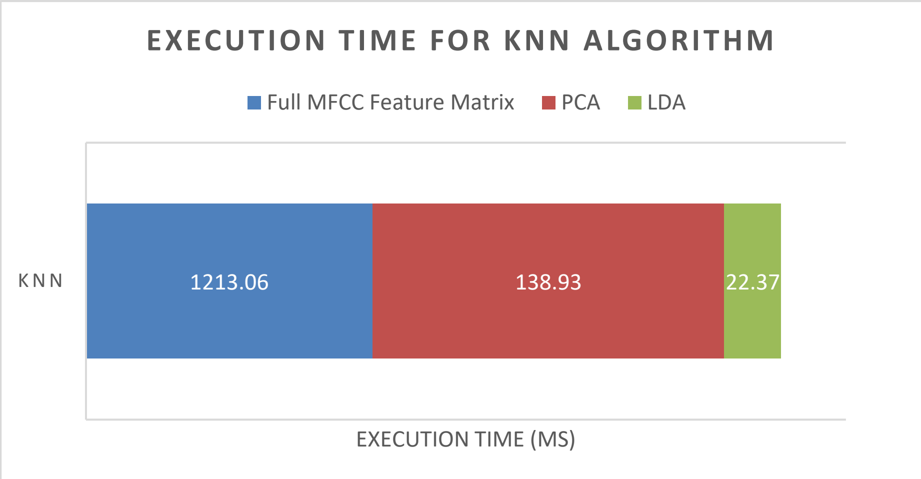

Fig. 6. Execution time comparison among three criteria for KNN Algorithm

Finally, it can be interpreted from the above result analysis is that LDA based feature minimization has been useful for detecting four labeled voice categories and the ML based KNN algorithm has even surpassed accuracy, F1-Score and AUC area of even DL based ANN architecture.

5 Conclusion and Future Works

Voice diseases can cause brain stroke or cancer which can lead to death if proper precaution is not taken. In this work, we have proposed LDA based multi-label voice disease recognition model with ML and DL algorithms and shown a remarkable rise in accuracy while using LDA as a dimensionality reduction system instead of PCA or without dimension exclusion strategy. Besides, we have wonderfully found that KNN can even outweigh ANN when LDA is considered predicting disorder within a less amount of time. Our multiple voice disease detection models can be deployed in various web and mobile healthcare applications which will be fruitful in the early detection and treatment of voice diseases. In the future, we will try to compare more ML and DL algorithms like Decision Tree, CNN, LSTM, Bidirectional LSTM, etc. for vigorously validating our work. Additionally, the dataset we have utilized in this study has a shortage in the number of audio samples. To ensure performance stability and for achieving more unbiased results, we will incorporate more labeled speech signals. Moreover, we have taken only 13 MFCC features for each voice sample which will be extended with more MFCC features and other voice characteristics mentioned in various studies previously. Furthermore, there is a plan to evaluate the model's robustness by testing audio recordings of general people in real-time. We will essentially apply our model to more audio related works for discovering more robust methods and analysis in the field of speech recognition.

References

1. Mythili, J., Vijaya, M. S.: Pathology Voice Detection and Classification Using Ensemble Learning. International Journal of Engineering Science Invention (IJESI), 7(8), 1-8 (2018).
2. Titze, I. R., Martin, D. W.: Principles of voice production. Prentice-Hall, Englewood Cliffs (1998).
3. Voice Disorders. https://www.hopkinsmedicine.org/health/conditions-and-diseases/voice-disorders, last accessed 2021/05/14.
4. Internal Medicine. https://www.medstarsouthernmaryland.org/our-services/internal-medicine/conditions/ear-nose-and-throat-conditions/voice-and-swallowing-disorders/, last accessed 2021/05/14.
5. Martins, R. H. G., do Amaral, H. A., Tavares, E. L. M., Martins, M. G., Gonçalves, T. M., Dias, N. H.: Voice disorders: etiology and diagnosis. Journal of voice, 30(6), 761-e1 (2016).
6. Uma Rani, K., Holi, M. S.: A hybrid model for neurological disordered voice classification using time and frequency domain features. Artificial Intelligent Research, 5(1), 87-94 (2016).
7. Verde, L., De Pietro, G., Sannino, G.: Voice disorder identification by using machine learning techniques. IEEE access, 6, 16246-16255 (2018).
8. Alhussein, M., Muhammad, G.: Voice pathology detection using deep learning on mobile healthcare framework. IEEE Access, 6, 41034-41041 (2018).
9. Verde, L., De Pietro, G., Alrashoud, M., Ghoneim, A., Al-Mutib, K. N., Sannino, G.: Leveraging Artificial Intelligence to Improve Voice Disorder Identification Through the Use of a Reliable Mobile App. IEEE Access, 7, 124048-124054 (2019).
10. Harar, P., Galaz, Z., Alonso-Hernandez, J. B., Mekyska, J., Burget, R., Smekal, Z.: Towards robust voice pathology detection. Neural Computing and Applications, 32, 15747–15757 (2020).
11. Dankovičová, Z., Sovák, D., Drotár, P., Vokorokos, L.: Machine learning approach to dysphonia detection. Applied Sciences, 8(10), 1927 (2018).
12. Boyanov, B., Hadjitodorov, S.: Acoustic analysis of pathological voices. A voice analysis system for the screening of laryngeal diseases. IEEE Engineering in Medicine and Biology Magazine, 16(4), 74-82 (1997).
13. Harar, P., Alonso-Hernandezy, J. B., Mekyska, J., Galaz, Z., Burget, R., & Smekal, Z. Voice pathology detection using deep learning: a preliminary study. In: 2017 international conference and workshop on bioinspired intelligence (IWOBI), pp. 1-4. IEEE, Funchal, Portugal (2017).
14. Mohammed, M. A., Abdulkareem, K. H., Mostafa, S. A., Ghani, M. K. A., Maashi, M. S., Garcia-Zapirain, B., Alhakami, H., Al-Dhief, F. T.: Voice pathology detection and classification using convolutional neural network model. Applied Sciences, 10(11), 3723 (2020).
15. Choubey, D. K., Kumar, M., Shukla, V., Tripathi, S., Dhandhania, V. K.: Comparative analysis of classification methods with PCA and LDA for diabetes. Current diabetes reviews, 16(8), 833-850 (2020).
16. Pützer, M., Koreman, J.: A German database of patterns of pathological vocal fold vibration. Phonus, 3, 143-153 (1997).
17. McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., Nieto, O.: librosa: Audio and music signal analysis in python. In: Proceedings of the 14th python in science conference, 8, 18-25 (2015).

18. Gupta, S., Jaafar, J., Ahmad, W. W., Bansal, A.: Feature extraction using MFCC. Signal & Image Processing: An International Journal (SIPIJ), 4(4), 101-108 (2013).
19. Poorjam, Hossein, A.: Why we take only 12-13 MFCC coefficients in feature extraction? https://rb.gy/2mimzc, last accessed: 2018/05/31.
20. Arjmandi, M. K., Pooyan, M., Mohammadnejad, H., Vali, M.: Voice disorders identification based on different feature reduction methodologies and support vector machine. In: 2010 18th Iranian Conference on Electrical Engineering, pp. 45-49, IEEE, Isfahan, Iran (2010).
21. Bhowmik, S., Reno, S., Sultana, S., Ahmed, M.: Clusterization of Different Vulnerable Countries for Immigrants Due to Covid-19 Using Mean Probabilistic Likelihood Score and Unsupervised Mining Algorithms. In: 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), pp. 285-290, IEEE, Dhaka, Bangladesh (2021).
22. Ottensen, C.: Comparison between PCA and LDA. https://dataespresso.com/en/2020/12/25/comparison-between-pca-and-lda/, last accessed: 2020/12/25.
23. Sarkar, P.: What is LDA: Linear Discriminant Analysis for Machine Learning. https://www.knowledgehut.com/blog/data-science/linear-discriminant-analysis-for-machine-learning, last accessed: 2019/09/30.
24. Mehta, A.: Everything You Need to Know About Linear Discriminant Analysis. https://www.digitalvidya.com/blog/linear-discriminant-analysis/, last accessed: 2020/01/04.
25. Priya, R., de Souza, B. F., Rossi, A. L., de Carvalho, A. C.: Predicting execution time of machine learning tasks for scheduling. International Journal of Hybrid Intelligent Systems, 10(1), 23-32 (2013).