www.ijsrnsc.org

Available online at www.ijsrnsc.org

IJSRNSC

Volume-5, Issue-3, June 2017

Research Paper

Int. J. Sc. Res. in Network Security and Communication

ISSN: 2321-3256

Diagnosis of Parkinson's Disease using Acoustic Analysis of Voice

Chaitanya Gupte[1*] and Shruti Gadewar[2]

[1*]Department of Electronics and Telecommunication Engineering, SIES Graduate School of Technology, Nerul, India
[2]Department of Electronics and Telecommunication Engineering, SIES Graduate School of Technology, Nerul, India

Received 10th Mar 2017, Revised 25th Apr 2017, Accepted 18th May 2017, Online 30th Jun 2017

Abstract- Acoustic analysis of voice is one of the cost effective method for detecting Parkinson disease. It is also a noninvasive, reliable and easy to use method. Also voice deviation from normal one is the earliest indicator of Parkinson. Voice data of sustained phonation has been collected from 8 healthy and 23 Parkinson subjects. The voice database is analyzed using PRAAT Software and 26 acoustic features are extracted. Some of the features being Jitters, Shimmers, Harmonic to Noise Ratio (HNR), Noise to Harmonic Ratio (NHR), Autocorrelation (AC). The values of these parameters show variation among two groups. A row vector is prepared using these parameters and fed to the classifiers. Classifiers such as Artificial Neural Network (ANN), Support Vector Machine (SVM), k-nearest neighbors (kNN), Adaboost, Decision trees and Random Forest have been tested and it was found that SVM is the best which gives the accuracy of 90%. Performances of classifiers are evaluated in terms of accuracy, precision, recall and total execution time.

Keywords- Parkinson's disease, PRAAT, Acoustic features, Support Vector Machine, Neural Networks

I. INTRODUCTION

Parkinson's disease is the second most common Neurological Disorder and the most common movement disorder. There are over 6 million people worldwide with this disease, but there still isn't much success seen in the development of a cure for the same. It is estimated that around 1 in 500 people are affected by Parkinson's disease and women have one and half time less chance than men to get this disease. The average age for symptoms to start is around 60, although 1 among 20 cases show this under 50. In this control on muscles reduces progressively and results in uncontrolled shaking of the limbs and head, stiffness, slowness and impaired balance. In severe conditions some simple tasks like walking, talking becomes very difficult.

Parkinson's disease is caused by death of Dopamine producing cells in Substantia Nigra and the corpus striatem. These Dopamine neurons are responsible for smooth and controlled movement. When Dopamine gets reduced, the movement becomes disordered. Parkinson results in some motor and non- motor symptoms. Motor symptoms include trembling in fingers, hand, arm, feet, leg, jaw or head and stiffness in the muscles. Non motor symptoms include anxiety, depression, memory loss, diminishing sense of smell etc.

Increasing the amount of Dopamine externally, replacing Dopamine, mimicking Dopamine can treat Parkinson. Treatment with Levodopa is also in practice. But long term of Levodopa produces side effects. Deep brain stimulation surgery is for advanced stages.

For maximizing the effect of treatment, it is necessary that the disease should be diagnosed at an earlier stage. Parkinson patients produce rough, shaky and breathy voice because speech controlling muscles get disturbed. It was found that in 90% cases of Parkinson, voice impairments are there. Also detection of Parkinson using voice is the earliest indicator. Voice Analysis is also used for remote monitoring of patients and thus lowers the inconvenience and cost of physical visits to the hospitals.[1] Parkinson affects the speech in several domains i.e., respiration, phonation, articulation and prosody. Mono pitch, Mono loudness, imprecise articulation, variability of speech rate, Hoarseness, speech disfluencies are the abnormalities in a Parkinson voice.[2] Various types of vocal test are performed for analysis. Sustained vowel, running speech and rapid syllable repetition are used for various vocal tests.

II. DATA

The Parkinson Disease database in the UCI Repository consists of training and test files. The training data belongs to 20 Parkinson (6 female, 14 male) and 20 healthy individuals (10 female, 10 male) who appealed at the Department of Neurology in Cerrahpasa Faculty of Medicine, Istanbul University. From all subjects, multiple types of sound recordings (26 voice samples including sustained vowels, numbers, words and short sentences) are taken. The test data belongs to 28 Parkinson subjects and 5 healthy subjects. During the collection of this dataset, 28 PD patients are asked to say only the sustained vowels 'a' and 'o' three times respectively which makes a total of 168 recordings. A total of 26 features are extracted from voice samples of this dataset [1,2,3,4].

III. METHODS

The methodology of this study can be broken down into three stages, first is the calculation of features, then comes the preprocessing and preselection of features and the last but not the least is the application of a classification technique to all possible subsets of features for the discrimination of healthy from disordered subjects.

Voice recordings taken from the UCI repository are first converted to .wav files and then fed to PRAAT Software for feature extraction. Feature extraction is done to differentiate Parkinson and healthy voice samples. The 26 features which are fed to the classifiers for further processing of data are jitter

A. Feature Extraction

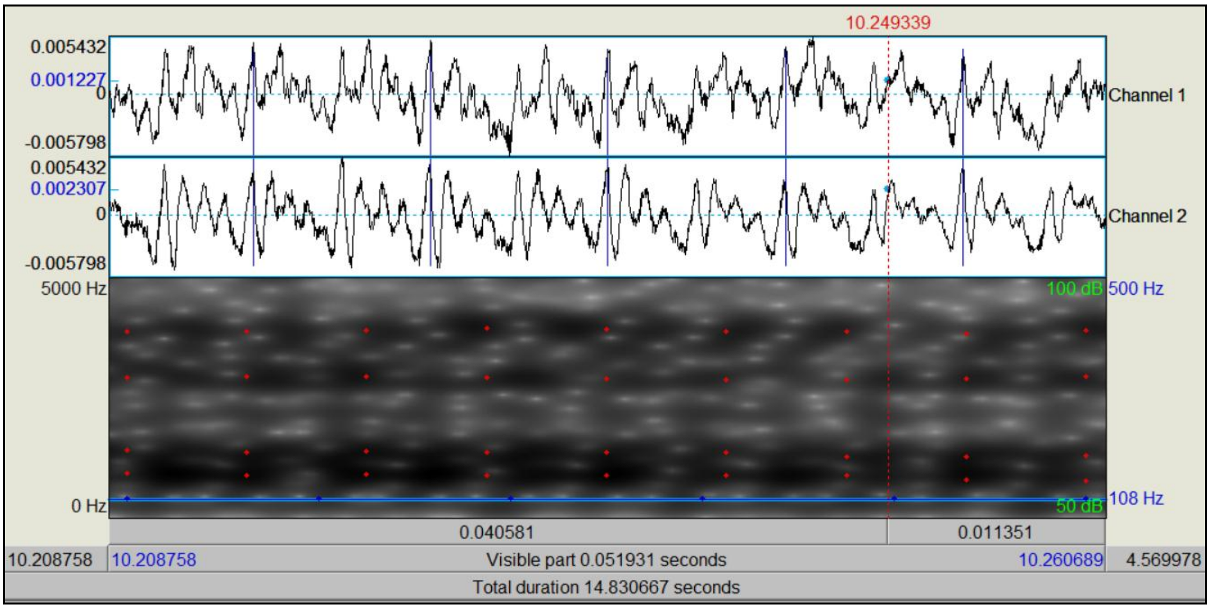

Fig 1: Recorded speech signal of Healthy subject in PRAAT.

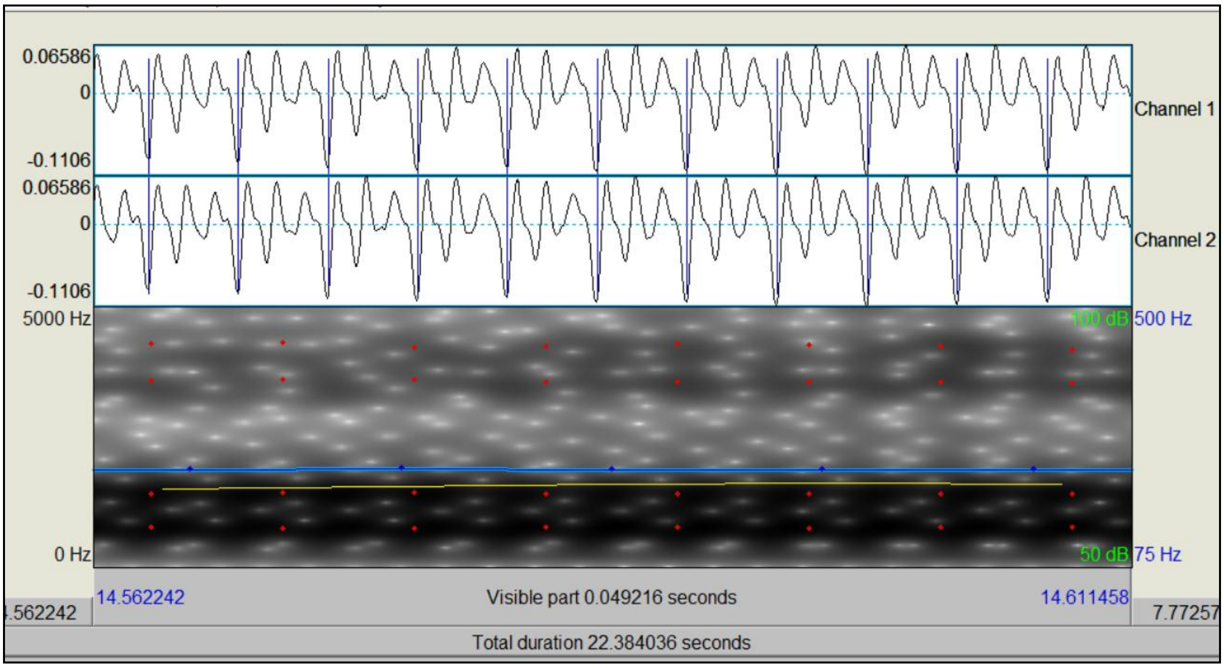

Fig 2: Recorded speech signal of Parkinson subject in PRAAT.

(local),Jitter(absolute),Jitter(rap),Jitter(ppq5),Jitter(ddp), Shimmer(local),Shimmer(dB),Shimmer(apq3),Shimmer(apq5),Shimmer(apq11),Shimmer(dda), Autocorrelation, Noise to Harmonic Ratio, Harmonic to Noise Ratio, Median pitch, Mean pitch, Standard deviation, Minimum pitch, Maximum pitch, Number of pulses, Number of periods, Mean period, Standard deviation of period, Fraction of locally unvoiced frames, Number of voice breaks, Degree of voice breaks.[3][4]

The traditional measures are based on the application of the short-time autocorrelation to successive segments of the signal, with peak picking to determine the frequency of vibration of the vocal folds (pitch period). The jitter and period perturbation measures are derived from the sequence of frequencies for each vocal cycle, by taking successive absolute differences between frequencies of each cycle and averaging over a varying number of cycles, optionally normalizing by the overall average. The shimmer and amplitude perturbation measures are derived from the sequence of maximum extent of the amplitude of the signal within each vocal cycle. The average difference of this sequence is taken as a measure of the deviation between cycle amplitudes. The noise to harmonics (and harmonics-to-noise) ratios are derived from the signal-to-noise estimates from the autocorrelation of each cycle.

B. Artificial Neural Network

Artificial neural networks are based on biological neural system. A neural network can learn and therefore can be trained to find solutions, recognize patterns, classify data, and forecast future events. Neural network consists of neurons and connections between them. Topology of network decides the response. Usually Neural networks include three or more layers and are termed as the input layer, hidden layer and the output layer. In the input layer, number of neurons depends on number of input signals and in output layer neuron number is same as the number of classes. Users can choose number of neurons in hidden layer. Neurons are connected by axons. A weight factor is given to each connection between two neurons. These weights are modified in training iterations where input and output both are given. So, optimum weights are thus selected which give minimized error between estimated output and expected output. [5]

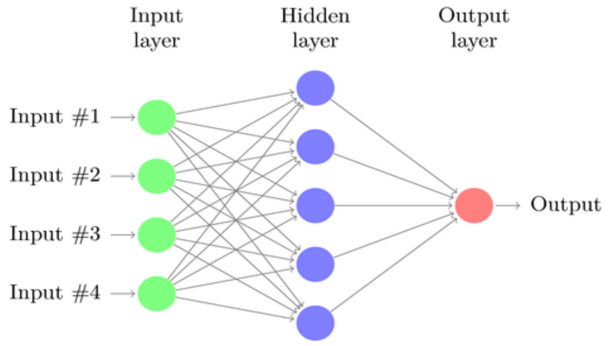

Fig 3: Artificial Neural Network

C. Support Vector Machine

In Machine Learning, Support Vector Machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. A support vector machine constructs a hyperplane or set of hyperplanes in a high or infinite-dimensional space. A good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class. This hyperplane is called the Functional Margin. SVM is effective in high dimensional spaces. It is effective in cases where number of dimensions is greater than the number of samples. It uses a subset of training points in the decision function (called support vectors), so it is also memory efficient. If the number of features is much greater than the number of samples, the method is likely to give poor performances. SVMs do not directly provide probability estimates, these are calculated using an expensive five-fold cross-validation.

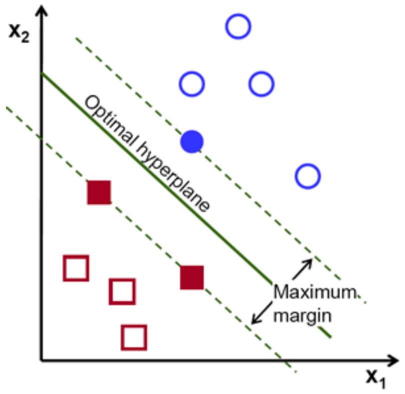

Fig 4: Support Vector Machine

D. Other Classifiers

Other Classifiers used in this paper are k-Nearest neighbors, Decision Trees, Random Forest, Adaboost and Naïve Bayes. The performance of all these classifiers along with SVM and ANN have been compared. [6]

K-Nearest Neighbor (*k*NN) classifier operate on the premises that classification of unknown instances can be done by relating the unknown to the known according to similarity function. Classification using *k*NN is a matter of locating the nearest neighbor in instance *space* and labelling the unknown instance with the same class label as that of the known neighbor. An odd value of *k* is preferred for better performance. The high degree of local sensitivity makes *k*NN classifiers highly susceptible to noise in the training data.

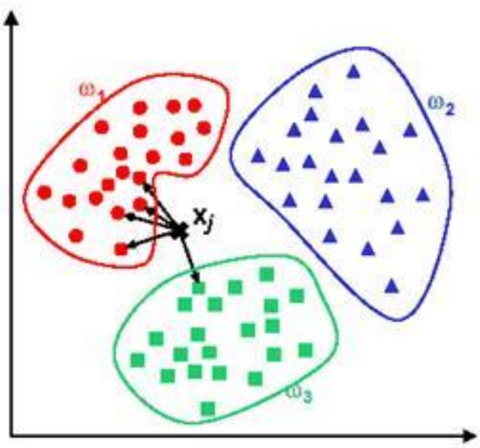

Fig 5: *k* - Nearest Neighbors

Decision tree learning uses a decision tree as a predictive model which maps observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves). It is one of the predictive modelling approaches used in statistics, data mining and machine learning. Tree models where the target variable can take a finite set of values are called classification trees; in these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. A tree can be "learned" by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called recursive partitioning.

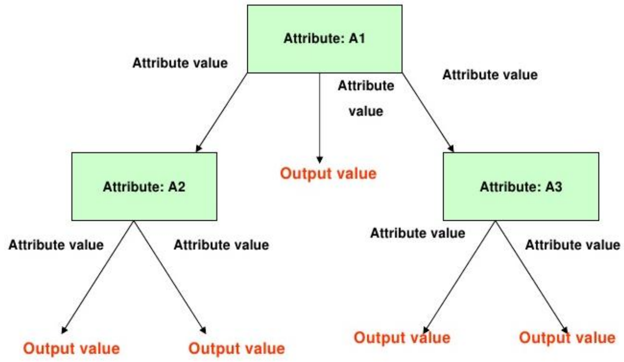

Fig 6: Building Decision Tree

Random Forest classifier operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes or mean prediction of the individual trees. Basically Random forests are a way of averaging multiple deep decision trees, trained on different parts of the same training set, with the goal of reducing the variance. This comes at the expense of a small increase in the bias, some loss of interpretability and overfitting of the training dataset, but generally greatly boosts the performance of the final model. Also Random forests can be used to rank the importance of variables in a regression or classification problem in a natural way.

Boosting is a general ensemble method that creates a strong classifier from a number of weak classifiers. AdaBoost is best used to boost the performance of decision trees on binary classification problems. It is used to boost the performance of any machine learning algorithm, mostly the weak learners. Weak learners are models that achieve accuracy just above random chance on a classification problem.

$$H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$

Where ht(x) is the weak learner function.

IV. RESULTS

Using the algorithms like SVC, Neural Network, *k*NN, Decision Tree, Random Forest and Adaboost different parameters like Accuracy, Precision, Recall, F1-score, Area under the ROC curve and total execution time in minutes have been calculated. Scatter plots or decision boundaries, confusion matrix, learning curves and receiver operating characteristics for every algorithm have been plotted. It has been found that Adaboost gives the best accuracy of 96.94%.

The figures shown below are the decision boundaries and the confusion matrix plotted using the above stated algorithms.

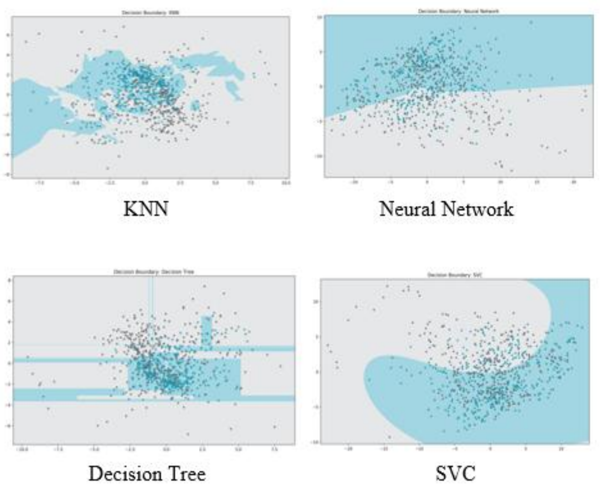

Fig 7: Decision Boundaries

Confusion matrix is a table used to describe the performance of a classifier on a set of test data for which the true values are already known. The confusion matrix of various classifiers are shown below.

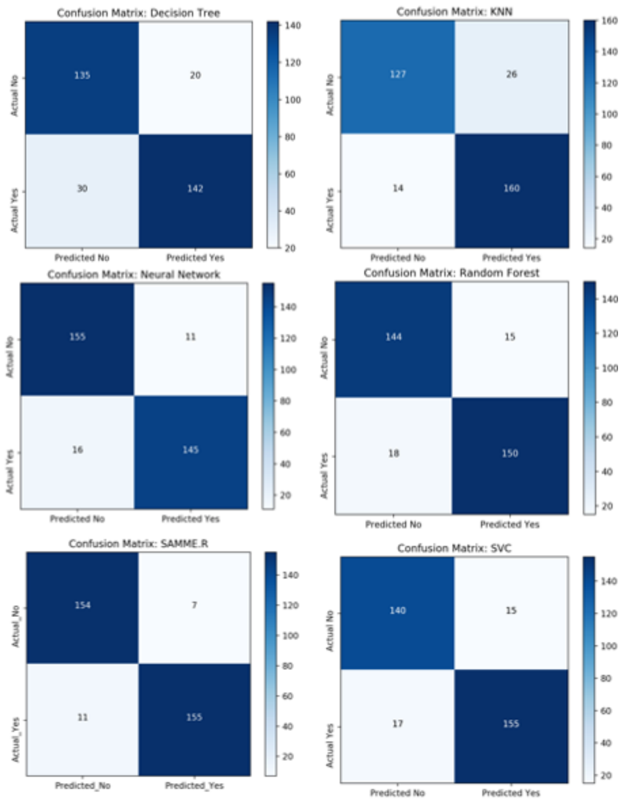

Fig 8: Confusion Matrices

Accuracy is defined as the ratio of all predictions that are correct to all the predictions made. It basically measures how good a model is.

$$accuracy = \frac{TP + FN}{TP + FN + TP + TN} = \frac{\text{correct predictions}}{\text{all predictions}}$$

$$Precision = \frac{TP}{TP + FP} = \frac{\text{positive predicted correctly}}{\text{all positive predictions}}$$

Precision is a measure of how many positive predictions were actual positive observations.

$$Recall = TPR = \frac{TP}{TP + FN} = \frac{TP}{P} = \frac{\text{predicted to be positive}}{\text{all positive observations}}$$

Recall is defined as the proportion of all positive observations made that are predicted incorrectly.

Here,
TP: True Positives
FP: False Positives
TN: True Negatives
FN: False Negatives

Table 1: Comparison of Accuracy, Precision and Recall

Parameters	SVM	ANN	*k*NN	Decision Tree	Random Forest	Adaboost
Accuracy	0.9	0.917	0.88	0.8471	0.89908	0.96941
Precision(PD)	0.91	0.93	0.86	0.88	0.91	0.98
Precision(H)	0.89	0.91	0.9	0.82	0.89	0.96
Recall(PD)	0.9	0.9	0.92	0.83	0.89	0.96
Recall(H)	0.9	0.93	0.83	0.87	0.91	0.98
f1-score(PD)	0.91	0.91	0.89	0.85	0.9	0.97
f1-score(H)	0.9	0.92	0.86	0.84	0.9	0.97

Where,
PD: Parkinson's disease
H: Healthy

Harmonic mean of Precision and Recall is called the F1-score. It can be interpreted as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and worst at 0. It can be calculated using the formula given below,

$$F1 = 2\frac{Precision * Recall}{Precision + Recall}$$

Total Execution time is the time required for the classifier to learn and then classify the test data set into categories. Accuracy and the execution time taken by the classifier are the best indicators of how good a classifier is. The table below shows the F1-score and the total execution time in minutes.

Table 2: Comparison of F1-Score and Execution time

Parameters	SVM	ANN	*k*NN	Decision Tree	Random Forest	Adaboost
f1-score(PD)	0.91	0.91	0.89	0.85	0.9	0.97
f1-score(H)	0.9	0.92	0.86	0.84	0.9	0.97
Execution Time(mins)	0.81	0.444	0.43	0.4071	0.5271	0.4724

V. DISCUSSIONS

Nowadays, a lot of research is going on in detecting diseases using voice analysis. In this paper, voice analysis has been performed successfully for differentiating between healthy and Parkinson subjects on the basis of acoustic features. Few literatures show less classification accuracy than the proposed method due to use of only two parameters (ECP and ASR). ECP and ASR parameters show good deviations among healthy and Parkinson subjects. ECP parameters have lower values for Parkinson's subjects than the healthy one. Slew rates are higher for the healthy subjects. SO in order to increase the classification accuracy all the 24 features have been considered in the proposed system. Among ANN, SVM, Adaboost, Decision trees, Random Forest and KNN, Adaboost gives a very good classification accuracy of 96.4%. The database can be enhanced for better classification.

VI. ACKNOWLEDGMENT

The authors are grateful to Mr. Vikram Patil, Dr. Atul Kemkar and Mrs. Preeti Hemnani at SIES Graduate School of Technology, Nerul for their guidance throughout the project and also for the comments on early drafts of the paper that prompted improvements to the paper.

REFERENCES

[1]. A. K. Ho, "*Speech impairment in a large sample of patients with Parkinson's disease*", Behav. Neurol., Vol.11, Issue.3, pp. 131-137, 1997
[2]. S. Saloni, R. K. Sharma, Anil K. Gupta, "*Disease detection using voice analysis: A review*", International Journal of Medical Engg. and Informatics, Vol.6, Issue.3, pp.189-209, 2014.
[3]. M. Shahbakhti, D. Taherifar, A. Sorouri, "*Linear and non-linear speech features for detection of Parkinson's disease*", Biomedical Engineering International Conference, Thailand, pp.1-3, 2013.
[4]. J.A. Logemann, H.B. Fisher, B. Boshes, E.R. Blonsky, "*Frequency and co-occurrence of vocal-tract dysfunctions in speech of a large sample of Parkinson patients*", Journal of Speech and Hearing Disorders, Vol.43, Issue.1, pp,.47-57, 1978.
[5]. A. Dastanpour, S. Ibrahim, R. Mashinchi, "*Effect of Genetic Algorithm on Artificial Neural Network for Intrusion Detection System*", International Journal of Computer Sciences and Engineering, Vol.4, Issue.10, pp.10-18, 2016.
[6]. A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, L. O. Ramig, "*Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease*", IEEE Transactions on Biomedical Engineering, Vol.59, Issue.5, pp.1264-1271, 2012.

