Prediction of Parkinson's Disease using Speech Signal with Extreme Learning Machine

[1]Aarushi Agarwal, [1]Spriha Chandrayan and [1]Sitanshu S Sahu
[1]Department of Electronics and Communication Engineering
Birla Institute of Technology, Mesra, Ranchi, India
Email: aarushiagarwal27@gmail.com, spriha.chandrayan@gmail.com, sssahu@bitmesra.ac.in

Abstract— **Speech impairments analysis has been used as an efficient tool for early detection of Parkinson's disease (PD). In this paper, we have proposed an efficient approach using Extreme Learning Machine to predict Parkinson's disease accurately utilising speech samples. The performance of the method has been assessed with a reliable dataset from UCI repository. The proposed method distinguishes Parkinson diseased subjects and healthy subjects with an accuracy of 90.76% and 0.81 MCC for the training dataset. When tested with an independent dataset comprising of Parkinson diseased patients, the proposed method gives 81.55% accuracy. The performance of our method is compared with existing techniques such as Neural Network and Support Vector Machine. The results obtained depict that the proffered method is reliable for identifying the Parkinson's disease.**

Keywords— Parkinson's Disease(PD), Extreme Learning Machines (ELM), dysphonia, telemonitoring.

I. INTRODUCTION

Parkinson's disease (PD) is a degenerative disorder of the central nervous system which mainly affects the motor system [1]. Moreover, it causes voice weakening in approximately 90% of its patients (of over 50 years of age) [2]. The figure of the number of affected people will augment with the aging of the worldwide population and deteriorate with time.

Clinical diagnosis of Parkinson's disease [PD] leads to errors, excessive medical costs and is inconvenient for the patients. Tele-monitoring includes diagnosing, treating, or consultation, via telecommunication for a patient at a distance. This is a solution to the shortcomings of clinical diagnosis of Parkinson's disease. Treatment of PD patients may be carried out using motor-related symptoms, speech processing for voice fluctuations or image processing using CAD [3].

The motor symptoms (such as: postural instability, muscular rigidity, and resting tremor) of Parkinson's disease are due to the death of dopamine-generating cells causing inability in controlling movements. In addition, many PD patients experience non-motor (non-dopaminergic) symptoms such as disorders of mood, behaviour, and characteristic variation in speech. Perceptually, speech of the people with PD is characterized by reduced loudness, breathy, mono pitch, hoarse voice, imperfect articulation, reduced stress, short rushes of speech, and dysfluent speech. Collectively, these speech symptoms are called hypokinetic dysarthria [4].

Effective diagnosis of Parkinson is possible using sustained phonation [2][3]. Phonation is the vibration of the vocal cords to create sound. It needs to fulfil five conditions to be produced: the vocal folds must be close to each other, there must be sufficient respiratory expiration, the vocal folds must be elastic enough to vibrate, vocal fold length and tension must be under volitional control, and the contour of the vocal folds must be suitable.

In literature, several studies have been carried out to analyse the Parkinson's disease using speech impairments. Tsanas *et al.* have used voice samples to forecast the advancement of Parkinson's disease in which they have identified relevant features using signal processing algorithms [1]. Sakar *et al.* investigated Parkinson's dataset using popular machine learning techniques. They have taken the mean and standard deviation of voice features of multiple recordings, showing it to be an effective approach in developing such prognostic models [5]. Little *et al.* have evaluated the proficiency of classic and nonstandard techniques for differentiating healthy people from PD patients using dysphonia [6]. In another work, Noor Afza *et al*. have used the feature selection techniques on sustained vowels using an existing database, to classify the PD by Back-Propagation learning algorithm [7]. Achraf Benba *et al.* have used the compression of MFCCs frames to extract the voiceprints from individuals, and have shown it to be suitable criteria for the identification of Parkinson's disease [2].

Although various tools and techniques have been used on speech signals for accurate identification of Parkinson's disease, still it remains far from achieving the desired results. In this paper, we present a machine learning method using Extreme Learning Machine (ELM) to efficiently predict PD. Various state of art techniques has been compared with the proposed methodology to assess its reliability.

The arrangement of the paper is as follows: In Section II, the methodology comprising of the various features extracted and the classification techniques are described. Section III presents the data specifications and experimental results with discussions. We present the conclusions in Section IV.

978-1-4673-9939-5/16/$31.00 ©2016 IEEE

II. Materials And Method

A. Feature Extraction

Feature extraction of voice samples is an important part of the development of an efficient predictor. Dysphonia measures are well known in identifying speech impairments and helps in diagnosis. In this paper, we have used 6 types of dysphonia parameters comprising of total 26 features. These include frequency (jitter), pulse, amplitude (shimmer), voicing, harmonicity and pitch parameters. The following features are selected because the vocal fold vibrations are periodic in healthy voice samples while are disturbed in pathological cases, taking in account the previous research [6]. The extracted features, arranged into algorithmic "families" sharing common attributes are listed in Table I.

TABLE I. Extracted Features From Speech Recordings

Feature	Group
Shimmer (dda) Shimmer (local) Shimmer (apq3) Shimmer (apq11) Shimmer (apq5) Shimmer (local,dB)	**Amplitude Parameters**
Number of pulses Mean period Number of periods Standard deviation of period	**Pulse Parameters**
Jitter (ddp) Jitter (local) Jitter (rap) Jitter (local, absolute) Jitter (ppq5)	**Frequency Parameters**
Number of voice breaks Fraction of locally unvoiced frames Degree of voice breaks	**Voicing Parameters**
Mean pitch Median pitch Standard Deviation Maximum pitch Minimum pitch	**Pitch Parameters**
Harmonic-to-Noise Noise-to-Harmonic Autocorrelation	**Harmonicity Parameters**

B. Extreme Learning Machine as Classification Technique

Conceptually, machine learning is to generate an algorithm, which learns from the available data and makes predictions on the unknowns. In recent years, Extreme Learning Machine proposed by Huang *et al.* [8] has emerged as a popular technique for classification and regression analysis. Unlike the conventional Neural Network classification technique, in which iterative variation of hidden neurons is required for efficient training, ELM shows that system can be trained without such design [9][10]. The architecture of ELM is depicted in Fig. 1. In this technique, the input weights and biases are arbitrarily chosen and fixed. The output of the network is computed as given in equation 1.

$$O(x) = \sum_{i=1}^{L} \beta_i h_i(x) = h(x)\beta \qquad (1)$$

where, $\beta = [\beta_1, \ldots, \beta_L]$ is the output weight vector directed from the L hidden layers to the output node, and $h(x) = [h_1(x), \ldots, h_L(x)]$ is a nonlinear feature mapping of the input $x = [x_1, \ldots, x_L]$. For determining the weights connecting the hidden layers and the output layers, the approximation error in the squared error sense is minimized. This solution can be obtained by the Moor-Penrose generalized inverse $\beta = H^\dagger T$ (2), where $H^\dagger = (H^T H)^{-1} H^T$ (3) is the Moore-Penrose pseudo-inverse matrix of the hidden layer output matrix H defined as given in equation 4.

$$H(w_1,\ldots w_L, b_1,\ldots b_L, x_1,\ldots x_L) = \begin{bmatrix} g(w_1.x_1 + b_1) & \cdots & g(w_L.x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1.x_N + b_1) & \cdots & g(w_L.x_N + b_L) \end{bmatrix}$$

$$\text{And} \quad \beta = \begin{matrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{matrix} \quad \text{and} \quad T = \begin{matrix} t_1^T \\ \vdots \\ t_N^T \end{matrix} \qquad (4)$$

The ELM classification technique is very fast in training and achieves good generalization performance for infinitely differentiable activation functions in the hidden layers. ELM generates a training model whose output weights can be mathematically calculated using an inverse operation on the hidden layer's weight matrix.

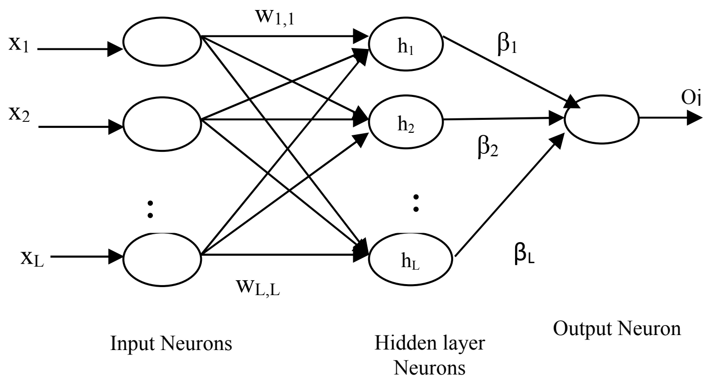

Fig. 1. Architecture of Extreme Learning Machine Classifier

For the training dataset: $N = \{(x_i,t_i) | x_i \in R^n\ t_i \in R^m,\ i = 1,\ldots,N\}$, activation function $g(x)$ and L hidden neurons the algorithm for implementing ELM in feed-forward networks is as follows:

Step 1: Randomly choose the input weights w_i and biases b_i for $i = 1,\ldots, L$.

Step 2: Compute the matrix H as described in equation 4.

Step 3: Determine the output weights b_i from equation 2, where β and T are described in equation 4.

C. Evaluation Parameters

The efficacy of the various classification techniques has been assessed using the following parameters:

a) Average Accuracy: It is the fraction (expressed in percentage) of correctly determined (both healthy and PD).

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \qquad (5)$$

b) Specificity: It is the fraction (expressed in percentage) of PD correctly determined as PD.

$$Specificity = \frac{TN}{FP+TN} \quad (6)$$

c) Sensitivity: It is the fraction (expressed in percentage) of healthy individuals correctly determined as healthy individuals.

$$Sensitivity = \frac{TP}{TP+FN} \quad (7)$$

d) MCC: Matthews's correlation coefficient is 1 for perfect prediction while 0 for extremely arbitrary prediction.

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (8)$$

True positive (TP) is the count of healthy patients predicted accurately as healthy; false negative (FN) is the count of healthy patients predicted to be diseased; false positive (FP) is the count of diseased patients predicted as healthy and true negative (TN) is the count of diseased subjects accurately predicted diseased.

The evaluation parameters of ELM have been compared with other well-known classifiers: Support Vector Machine (SVM) and Artificial Neural Networks (ANN). All these classifiers have been implemented by varying individual associated parameters in MATLAB.

III. Results And Discussion

A. Data Source

The dataset used in this study was collected from UCI Machine learning Repository created by Sakar *et al.* from Istanbul University. The database comprises of training data belonging to 20 PD patients (6 female, 14 male) and 20 healthy individuals (10 female, 10 male) who applied to the department of Neurology in Cerrahpasa Faculty of Medicine, Istanbul University. Multiple voice recordings (26 speech samples including short sentences, words, numbers and sustained vowels) were taken from all applicants. A collection of 26 linear and time–frequency based features have been extracted from each sample [5]. The training data set consists of 1040 voice samples, out of which 40 random samples (one from each individual) have been separated for a self-consistency check of the developed predictor.

By the same examination procedure and under the same conditions, an independent test set from Parkinson diseased people was also collected. It contains 168 recordings comprising of only the sustained vowels 'a' and 'o' recorded thrice each by 28 PD patients.

B. Experimental Results

The speech samples have been normalized between 0 and 1 to maintain uniformity in all the features. The proposed ELM method is compared with the existing classification techniques such as Artificial Neural Network and SVM. In ELM method, from an empirical study, we have found that 500 neurons with log sigmoid kernel function is best in the discriminating between PD and healthy subjects. For SVM, radial basis function (RBF) kernels with σ=10 and C=50 was found best for prediction. In Neural network, 50 neurons provided better results. The performance of all the methods for the training dataset is listed in Table II.

TABLE II. Performance Comparison Of The Classifiers On The Training Data Set

Method	Result	Overall Accuracy	S_p	S_n	MCC
Neural Network	*Avg.*	65.18	64.39	66.33	0.305
	Best	71.80	70.10	73.90	0.437
SVM	*Avg.*	68.66	72.28	65.04	0.374
	Best	71.40	75.00	68.20	0.428
ELM	*Avg.*	89.95	89.54	90.64	0.802
	Best	**90.76**	**90.58**	**91.80**	**0.815**

Avg.= Average of 500 runs, Best= The best result out the 500 runs, S_P= Specificity, S_n=Sensitivity

It shows that the ELM is superior amongst the existing methods by providing 90.76 % overall accuracy with MCC of 0.815. In the best case, it is able to identify the healthy subjects with an accuracy of 91.80 % and PD subjects of 90.58%. The consistency of the three methods is tested with a dataset consisting 40 voice samples extracted from PD and healthy people. The performance results are presented in table III. Again, ELM provides better result with an overall accuracy of 87.50 % and MCC of 0.75.

TABLE III. Performance Comparison Of 40 Random Samples (One From Each Individual)

Method	Overall Accuracy	S_p	S_n	MCC
Neural Network	67.50	64.00	73.30	0.361
SVM	72.50	69.56	76.41	0.450
ELM	**87.50**	**89.47**	**85.71**	**0.750**

Avg.= Average of 500 runs, Best= The best result out the 500 runs, S_P= Specificity, S_n=Sensitivity

TABLE IV. Prediction Accuracy Of The Independent Dataset (168 Voice Samples)

Classifier	Accuracy
Neural Network	72.61
SVM	82.14
ELM	**81.55**

Also, the reliability of the proposed method is assessed by an independent set of data and the results are reported in Table IV. The ELM method predicts the PD subjects with an accuracy of 81.55% whereas Neural Network and SVM provides 72.61% and 82.14% respectively. From the results, it is inferred that the ELM method is superior to the neural network and comparable to the SVM. From the analysis, it is concluded that the sustained vowels carry sufficient information to distinguish between the PD and normal subjects, which is in agreement with the earlier findings.

IV. CONCLUSION

In this paper, we have proposed an effective approach to generate an accurate predictive model for telemonitoring of Parkinson's disease using Extreme Learning Machine (ELM). This method is able to identify the PD subjects with an accuracy of 81.55%. From our extensive study, it is evident that the sustained vowels carry sufficient information to predict Parkinson's disease. The simple architecture and inherent learning of the data make ELM a reliable method for prediction. Finally, future studies can incorporate different data sets with new features on this technique to explore its proficiency for varied medical applications.

References

[1] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel Speech Signal Processing Algorithms for High-Accuracy Classification of Parkinson's Disease", *IEEE transactions on biomedical engineering*, vol. 59(5), 2012.

[2] A. Benba, A. Jilbab, and A. Hammouch, "Detecting Patients with Parkinson's disease using Mel FrequencyCepstral Coefficients and Support Vector Machines", *International Journal on Electrical Engineering and Informatics*, vol. 7(2), 2015.

[3] A.Valli and Dr.G.Wiselin Jiji, "Parkinsons Disease Diagnosis Using Image Processing Techniques A Survey", *International Journal on Computational Sciences & Applications*, vol.4(6), 2014.

[4] K. M. Rosen, R. D. Kent, A. L. Delaney, and J. R. Duffy, "Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers", *J. Speech. Lang. Hear. Res.*, vol. 49, pp. 395–411, 2006.

[5] E. Sakar, B. Isenkul, M. Sakar, C.O. Sertbas, A. Gurgen, F. Delil, S. Apaydin and H. Kursun, "Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound Recordings", *IEEE Journal of Biomedical and Health Informatics*, vol. 17(4), pp. 828-834, 2013.

[6] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of Dysphonia Measurements for Telemonitoring of Parkinson disease", *IEEE Transactions on biomedical engineering*, vol. 56(4), 2009.

[7] N. Afza, M. Challa and J. Mungara, "Speech Processing Algorithm For Detection of Parkinson's Disease", *International Journal of Engineering Research & Technology*, vol. 2(4), 2013.

[8] G. Huang, Q. Zhu and C. Siew, "Extreme learning machine: Theory and applications", *Neurocomputing*, vol. 70, pp. 489-501, 2006.

[9] G. Huanga, G. Huang, S. Songa and K. Youa, "Trends in extreme learning machines: A review", *Neural Networks*, vol. 61, pp. 32-48, 2015.

[10] G. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification", *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, vol. 42, pp. 513-529, 2012.