EVALUATION OF FAST LEARNING MACHINE FOR IDENTIFICATION OF PARKINSON DISEASE

Article · February 2020

CITATIONS
0

READS
73

1 author:

Ahmed Alemran
University of Misan
5 PUBLICATIONS **20** CITATIONS

SEE PROFILE

Journal of Engineering Science and Technology
Vol. 15, No. 1 (2020) 589 - 599

EVALUATION OF FAST LEARNING MACHINE FOR IDENTIFICATION OF PARKINSON DISEASE

AHMED ALEMRAN

Department of Physiology, Facility of Medicine,
Misan University, Al-Amarah, Iraq
E-mail: ahmedalamran2@gmail.com

Abstract

Artificial Intelligence (AI) plays a key part in the timely diagnoses of various degenerative diseases, of which, Parkinson's Disease (PD) is one. Early diagnosis of this degenerative disease is key to controlling its development. A Fast Learning Network (FLN) is a neural network structure that features a single hidden layer and connections among its output and input layers. The approach has been shown to exceed the performance of the comparable Extreme Learning Machine (ELM) method, which has no connections among output and input layers. For this study, voice features were utilised to train FLNs for predicting PD. Numbers of features were changed while 10-fold cross-validation was employed across 10 runs for every model, so as to counter any random weighting effects arising in the FLNs' input hidden layers. A comparison of these results to those for Kernel Extreme Learning Machines (KELM) and conventional Extreme Learning Machines (ELM) shows that FLN outperforms both ELM and KELM in predictive diagnostics. The highest attained accuracy was observed to exceed 90% for FLN.

Keywords: Classification, ELM, FLN, Identification, KELM, Parkinson disease.

1. Introduction

AI is revolutionising many technical applications. Healthcare is a technical field that may benefit from such powerful AI advances. Machine learning is a subfield of AI that can offer the medical field very powerful diagnostic approaches through data gathered from subjects, using models that are trained on key portions of their populations. Among such promising examples would be an advanced diagnosis for Parkinson's Disease (PD) [1].

Parkinson's Disease is among the most prevalent degenerative disorders that afflict the central nervous system. PD has spread markedly in numerous developing nations. The underlying causes remain unknown, and therefore timely diagnoses would be helpful in the alleviation of PD symptoms in its earliest stages. The disease presents several clear symptoms that include bodily rigidity and tremors, slowness of movements, posture instability, and hand asymmetry. Nevertheless, the vocal disorders that arise in almost 90% of PD patients are among the disease's earliest noticeable symptoms that can arise some 5 years prior to clinical diagnosis.

The literature regarding PD diagnoses and Artificial Neural Networks has encompassed a large part of this research. In the study conducted by Berus et al. [2], multiple feed-forward Artificial Neural Network (ANN) models comprising different configurations were employed for PD prediction in tested individuals, in accordance with features extracted from 26 dissimilar voice samples drawn from each person. The results were validated using the Leave-One-Subject-Out (LOSO) procedure. A small number of feature-selection procedures that derive from Pearson's correlation coefficient, Kendall's correlational coefficient, and principal-component analyses as well as self-organising maps, were employed to enhance algorithmic performance and dataset reduction.

Superior test accuracies were achieved with feature selection based on Kendall's correlational coefficient, with the most significant voice samples recognised. The multiple-ANN approach has been shown to be a superior classification method for PD diagnosis that does not rely on feature selection procedures involving raw data. Lastly, the neural network was fine-tuned, with 86.47% test accuracy attained. The work of Morisi et al. [3] combined proton spectroscopy, diffusion tensor imaging, as well as morphometric-volumetric information. These were used to acquire the MR quantitative marker data that Supports Vector Machine (SVM) methods, with the goal of recognising the various PD-related disorders. Procedures for feature selection were also utilised to determine critical classification features. The researchers used a graph-based method derived from the quantitative markers to extract supplementary features from the PD dataset for enhanced classification accuracy.

The research by Nilashi et al. [4] covered the advantages of using an incremental machine learning-based method, the Incremental Support Vector Machine, for developing newer approaches to UPDRS prediction. The researchers thereby employed Incremental SVM for Motor-UPDRS and Total-UPDRS predictions, along with the non-linear, iterative partial least-squares method for dataset dimensionality reduction as well as self-organising maps for task clustering. In the selection of optimised feature sets, Shahbakhi et al. [5] employed the Genetic Algorithm (GA). An SVM-based network was utilised to classify healthy subjects and those afflicted with PD. The research dataset comprised a collection of biomedical voice signals drawn from 31 individuals, 23 of whom were afflicted with PD, and 8 healthy subjects. Each individual was requested to pronounce the

letter "A" for 3 seconds. A total of 22 non-linear and linear features were systematically extracted from these voice signals, with 14 derived from F0 (the pitch or fundamental frequency), shimmer, jitter, and the noise-to-harmonics ratio, which comprise the key factors in characterising voice signals. As changes in these aspects are noticeable in individuals afflicted with PD, optimised feature sets were selected from among these subjects. Among the various types of optimised features, data classification methods were evaluated.

The research of Lahmiri and Shmuel [6] emphasised the assessment of performances for 8 different pattern-ranking algorithmic methods (known as feature selection procedures). These were joined with non-linear support vector machines (SVMs) to discriminate healthy control subjects from PD patients. Parameters for the radial basis function kernel in terms of the SVM classifier were optimised with Bayesian optimisation methods.

In other research, Lahmiri et al. [7] assessed the performances of machine learning-based methods for PD diagnoses from dysphonia symptoms. Several machine learning-based methods were examined and thereby trained using a set of 22 voice disorder measurements, in order to classify PD patient and healthy control subjects. These machine learning-based techniques comprised k-Nearest-Neighbours (k-NN), Linear Discriminant Analyses (LDA), Regression Trees (RT), Naive Bayes (NB), Radial-Basis Function Neural Network (RBFNN), Support Vector Machine (SVM) [8], and the Mahalanobis distance classifier. The performances of these techniques were assessed using a 10-fold cross-validation protocol. In other research by Bi et al. [9], certain methods based on randomised SVM clusters were proposed for classifying HC and AD. Drawing from the Alzheimer's Disease Neuroimaging Initiative Database, test subjects that include 25 AD as well as 35 HC patients were recruited. Aich et al. [10] introduced a novel approach that compares performance metrics based on various feature sets, including the originals. Principal-component analysis-based methods for feature reduction were also used to select for feature sets.

Additionally, they employed non-linear-based classification strategies for performance metric comparisons. In the research conducted by Er et al. [11], diverse classification approaches were compared in terms of effective PD diagnoses. These techniques include feedforward, artificial immune systems, learning vector quantisation, and a probabilistic neural-network algorithm. A total of 197 PD data records derived using a 22-voice feature set was utilised in this research. A 10-fold cross-validation procedure was performed to determine algorithmic performances in PD diagnoses. It was found that superior results in classification accuracy are acquired through probabilistic neural-network algorithms. The research results were similarly compared to those from prior studies, based on the same PD dataset.

For their research, a fast learning network was utilised to identify Parkinson's Disease, according to the selection of various numbers of features from the 22-voice feature set. FLN performance was evaluated in comparison to ELM as well as KELM. The rest of the article is structured as follows. In Section 2, we discuss the approach. Section 3 discusses the experimental work and outcomes. Summary and recommendations are provided in Section 4.

2. Methodology

This section introduces the methods for constructing classifications of Parkinson's Disease using FLN approaches. In subsection 2.1, Fast Learning Networks (FLNs)

are introduced. In subsection 2.2, we present our algorithm for constructing characterisation models for Fast Learning Networks based on PD data. In subsection 2.3, we present the assessment measures employed in the evaluation of our method.

2.1. Fast learning network

The Fast Learning Network (FLN) is an ELM variant with added input and hidden-layer connections [12]. FLN training strategy is quite similar to that of ELM. Nevertheless, diverse research studies have shown that FLN yields performance superior to that of ELM. FLN structure is shown in Fig. 1, which comprises a 3-layer schema that includes input, output, and hidden layers. Three connection types are present: connections among hidden and input layers, connections among hidden and output layers, and connections among input and output layers. Matrix or compact representation was utilised as provided for in the equations.

FLN training is conducted based on the estimations of the proper weights, which are presented in Fig. 1 as W^{io}, W^{on} and W^{in}. Each weight corresponds to the connections among different layers W^{in}, with b denoting those weights that represent the connections among the hidden and input layers, which are determined using randomised equations. The terms W^{in}, W^{io}, W^{oh} and c denote the weights that represent the connections among the input-output W^{io}, as well as those connections among the hidden layer and output layer pair, W^{oh} and c. They are depicted in Eq. (1).

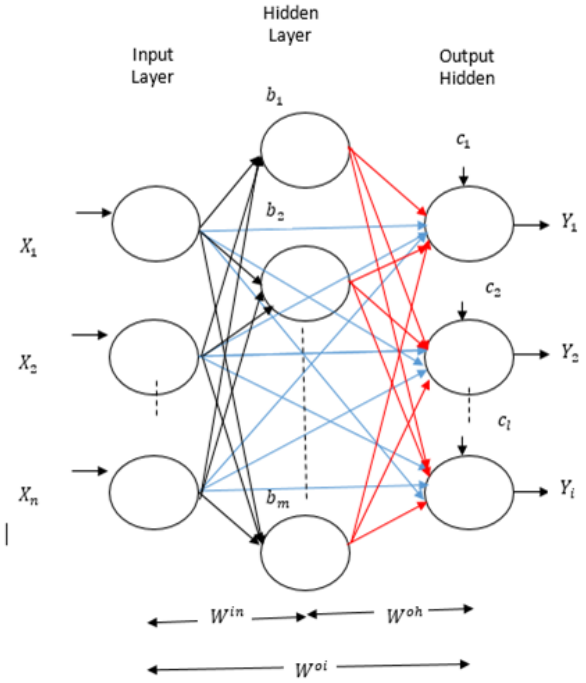

Fig. 1. Topological representation of the fast learning network.

The output matrix is computed on the basis of the input hidden weights and the activation function as displayed in Eq. (2). The other weights are calculated by using the Eqs. (3) and (4).

$$w = [W^{io}\ W^{oh}c\ ] \tag{1}$$

$$G = G(W^{in}, b, x) \tag{2}$$

$$\widehat{W} = Y\begin{bmatrix}X\\G\end{bmatrix}^{T}\left(\begin{bmatrix}X\\G\end{bmatrix}^{in}\begin{bmatrix}X\\G\end{bmatrix}^{T}\right)^{-1} \tag{3}$$

$$= YH^{T}(HH^{T})^{-1} \tag{4}$$

2.2. Algorithm

With the aim to employ Fast Learning Machine on the dataset of PD, we create a descriptive model of the FLN performance by using 2 kinds of factors: Activation function type as well as the amount of neurons as shown in Fig. 2. The precision of every case of the factors is acquired and at last, the best precision is considered as the model precision. It is noteworthy that every model was assessed on the basis of 10 runs and 10 folds. 10 runs are used in order to counter the performance change that takes place because of the arbitrary weights between hidden and input layer. Each run will generate different weights of values between the hidden and the input layer by seed changing.

```
Input
maxNumOfNeurons = 200
numberofActivationFncs = 5;
numberOfFolds = 10
numberOfRuns = 10
Data

Output
BestFLNCof

Start
Folds{1:numberOfFolds}=DivideData(Data)
a = 1
n = 1
r = 1
TestingAccruacy = []
accuraciesMatrix = zeros(maxNumOfNeurons,numberofActivationFncs)
for a = 1:numberofActivationFncs
    for n = 1:maxNumOfNeurons
            for f = 1:numberOfFolds
               for r = 1:numberOfRuns
                   FLNTrained=TrainFLN(activationFunction = a,
                                numberOfNeurons= n,seed = r,Folds except f)
                   accuracy=TestFLN(FLNTrained,activationFunction = a,
                            numberOfNeurons = n,seed = r, Folds except f)
            TestingAccruacyRuns=[TestingAccruacy accuracy]
          avgAccuracyRuns=mean(TestingAccruacy)
              end
            TestingAccruacyRuns=[TestingAccruacyRuns avgAccuracyRuns]
            avgAccuracyFolds=mean(TestingAccruacyRuns)
          end
         accuraciesMatrix(a,n)=avgAccuracyFolds
    end
end
[aMax,nMax]=FindMax(accuraciesMatrix)
BestFLNCof=[aMax,nMax]
```

Fig. 2. Pseudocode of building characterizing FLN based on two parameters: Type of activation function and number of hidden neurons.

2.3. Evaluation measures

The measures of assessment that are going to be employed for our approach's performance evaluation are provided in the equations. These measures are computed on the basis of the components of a binary confusion matrix, given in Table 1.

Table 1. Binary confusion matrix for Parkinson disease.

	Prediction of Parkinson is positive	**Prediction of Parkinson is negative**
Parkinson is positive	*TP*	*FN*
Parkinson is negative	*FP*	*TN*

The measures are found in the equations:

$$\text{A}ccuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$\text{Recall} = sensitivity = TP/(TP + FN) \quad (6)$$

$$Specificity = TN/(FP + TN) \quad (7)$$

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

$$F - measure = \frac{2 \times precision \times Recall}{Precision + Recall} \quad (9)$$

$$G - mean = \sqrt{sensitivity \times precision} \quad (10)$$

3. Experimental Work

This portion provides the empirical assessment of the Fast Learning Machine for recognising Parkinson's Disease. The technique has been verified on the given dataset in 3.1 sub-section and the measures of assessment were found and evaluated in 3.2 sub-section.

3.1. Dataset

This research is carried out on the dataset of PD obtained from the UCI repository of machine learning [13]. This dataset was developed by Max Little belonging to the Oxford University, in association with the National Centre for Voice and Speech situated in Denver, Colorado. He recorded the signals of the speech. The actual research published the techniques for feature extraction for common voice disorders. This dataset consists of several measurements of biomedical voice obtained from 21 individuals, out of whom 23 had PD. Every table column is a specific measure of voice, and every row signifies one of the 195 recordings from these people ("name" column). The primary objective of the data is to segregate the healthy individuals from individuals with PD, as per the column "status", wherein 1 stand for PD and 0 stands for healthy.

3.2. Result and discussion

For assessing FLN and identifying PD, the assessment measures have been determined and compared between three approaches: Kernel Extreme Learning

Machine, Extreme Learning Machine, and Fast Learning Machine. The measures were computed for several numbers of chosen features 1, 5, 10, 15, 20, and 22.

Figure 3 shows the accuracy. It can be noted that irrespective of the amount of features, FLN shows superiority in precision as compared to other techniques, KELM and ELM. Moreover, the best-obtained accuracy is at the time when the amount of features was 5, in which, case the accuracy was found to be more than 90 percent for FLN.

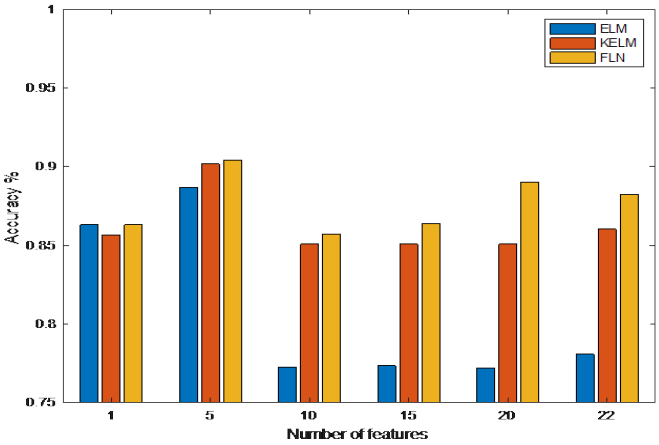

Fig. 3. Accuracy of FLN, ELM, and KELM according to the number of selected features.

Besides the accuracy, we also display the sensitivity and its quantities in accordance with the amount of chosen features in Fig. 4. Likewise, we discovered that the FLN specificity is better compared to other techniques KELM and ELM. Also, we discovered that the highest sensitivity was for 15 features.

The 3 models' specificity was found with various amounts of features as given in Fig. 5. Likewise, we discovered that the models' specificity changes as per the changes in the amount of the features and we discovered that the KELM technique has obtained the biggest specificity values compared to FLN and ELM. Nonetheless, for each of the cases of the amount of features, FLN was better than ELM. The 3 models' accuracy was observed to display an equivalent behaviour to the specificity where KELM was superior to ELM and FLN and FLN over ELM as shown in Fig. 6. Nevertheless, in case the amount of features was 20, it was found that FLN had more accuracy than KELM and ELM. Moreover, in every case of the amount of features, FLN was better than ELM.

F-measure was computed for all the 3 models in case of multiple features amounts, and it was shown that we get distinct *F*-measure values for a different amount of features as shown in Fig. 7. We note that FLN was better in terms of *F*-measure in comparison to KELM and ELM.

Lastly, the *G*-mean was computed for these 3 models for multiple feature amounts and it was found that FLN was superior to KELM and ELM as shown in Fig. 8.

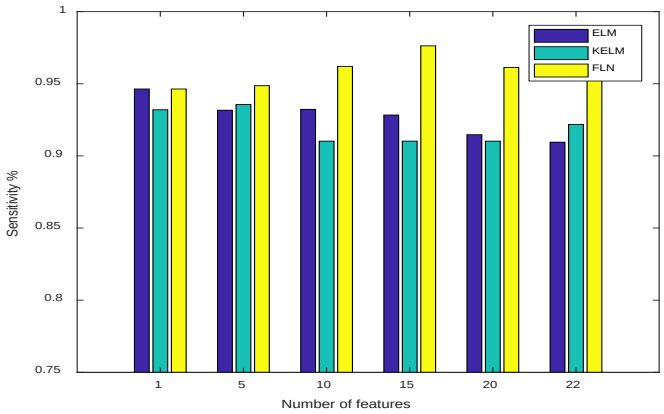

Fig. 4. The sensitivity of FLN, ELM, and KELM according to the number of selected features.

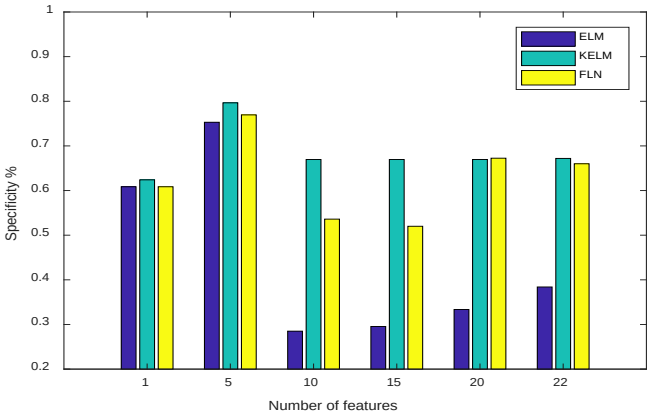

Fig. 5. The specificity of FLN, ELM, and KELM according to the number of selected features.

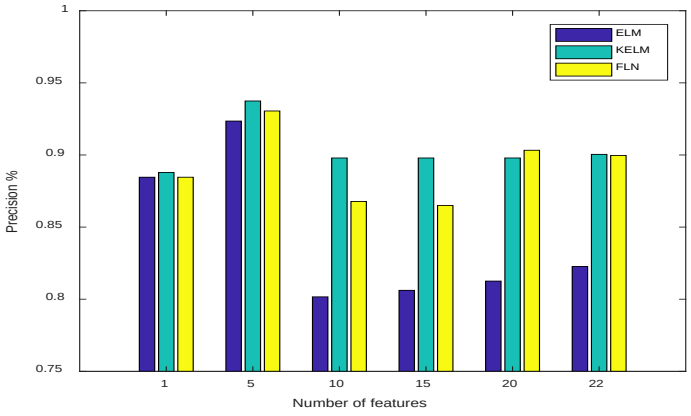

Fig. 6. The precision of FLN, ELM, and KELM according to the number of selected features.

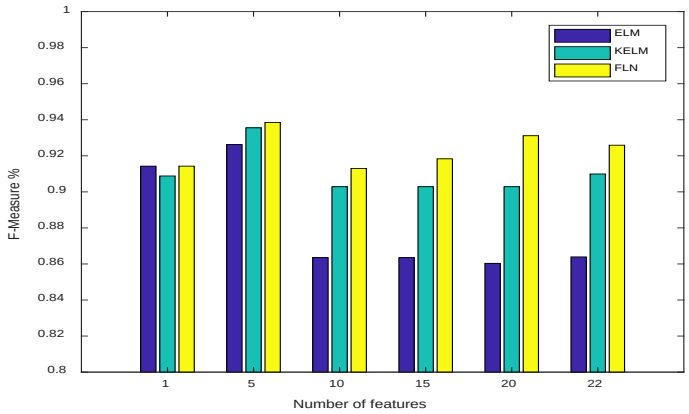

Fig. 7. *F*-measure of FLN, ELM, and KELM according to the number of selected features.

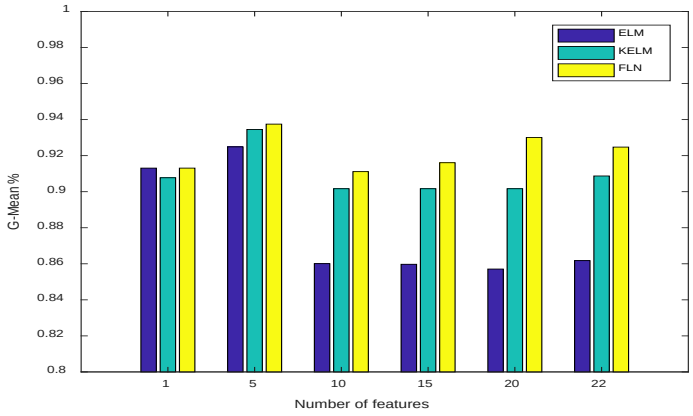

Fig. 8. *G*-mean of FLN, ELM, and KELM according to the number of selected features.

4. Conclusions

This study proposed the application of the Fast Learning Machine for Parkinson's Disease diagnosis and evaluating its performance with 2 techniques: KELM (Kernel Extreme Learning Machine) and ELM (classical Extreme Learning Machine). FLN was chosen to owe to its better performance that can be obtained from the link between the layers of the input and the output. After FLN was trained on several features amounting from 1 to 22, the measures of the performance have confirmed that FLN has a better performance compared to KELM and ELM as far as precision and other factors are concerned. The best accuracy that was obtained was above 90 percent. We observed that the change in the amount of features plays a significant role in the measurement of accuracy. Nonetheless, FLN still remains better compared to KELM and ELM irrespective of the different number of features. In future, FLN's kernel variant can be developed and tested for the classification of PD.

Abbreviations

ANN	Artificial Neural Network
ELM	Extreme Learning Machine
FLN	Fast Learning Networks
GA	Genetic Algorithm
KELM	Kernel Extreme Learning Machines
PD	Parkinson's Disease
SVM	Supports Vector Machine

Reference

1. Vashistha, R.; Yadav, D.; Chhabra, D.; and Shukla, P. (2019). Chapter 5 - Artificial intelligence integration for neurodegenerative disorders. *Leveraging Biomedical and Healthcare Data*, 77-89.
2. Berus, L.; Klancnik, S.; Brezocnik, M.; and Ficko, M. (2018). Classifying parkinson's disease based on acoustic measures using artificial neural networks. *Sensors*, 19(1), 15 pages.
3. Morisi, R.; Manners, D.N.; Gnecco, G.; Lanconelli, N.; Testa; C.; Evangelisti; S.; Talozzi, L.; Gramegna, L.L.; Bianchini, C.; Calandra-Buonaura, G.; Sambati, L.; Giannini, G.; Cortelli, P.; Tonon, C.; and Lodi, R. (2018). Multi-class Parkinsonism disorders classification with quantitative MR markers and graph-based features using support vector machines. *Parkinsonism and Related Disorders*, 47, 64-70.
4. Nilashi, M.; Ibrahim, O.; Ahmadi, H.; Shahmoradi, L.; and Farahmand, M. (2018). A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. *Biocybernetics and Biomedical Engineering*, 38(1), 1-15.
5. Shahbakhi, M.; Far, D.T.; and Tahami, E. (2014). Speech analysis for diagnosis of Parkinson's disease using genetic algorithm and support vector machine. *Journal of Biomedical Science and Engineering*, 07(04), 147-156.
6. Lahmiri, S.; and Shmuel, A. (2019). Detection of Parkinson' s disease based on voice patterns ranking and optimized support vector machine. *Biomedical Signal Processing and Control*, 49, 427-433.
7. Lahmiri, S.; Dawson, D.A.; and Shmuel, A. (2017). Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures. *Biomedical Engineering Letters*, 8(1), 29-39.
8. Alemran, A.; Rahmatullah, B.B.; and Hadi, A. (2018). Systematic review on ear identification. *International Journal of Engineering & Technology*, 7(4.31), 251-259.
9. Bi, X.-a.; Shu, Q.; Sun, Q.; and Xu, Q. (2018). Random support vector machine cluster analysis of resting-state fMRI in Alzheimer's disease. *PLoS ONE*, 13(3), 1-17.
10. Aich, S.; Hui, K.L.; Al-Absi, A.A.; and Sain, M. (2018). A nonlinear decision tree based classification approach to predict the Parkinson' s disease using different feature sets of voice data. *Proceedings of the 20th International Conference on Advanced Communications Technology* (ICACT). Chuncheon-si Gangwon-do, South Korea, 638-642.

11. Er, O.; Cetin, O.; Bascil, M.S.; and Temurtas, F. (2016). A comparative study on Parkinson's Disease diagnosis using neural networks and artificial immune system. *Journal of Medical Imaging and Health Informatics*, 6(1), 264-268.
12. Li, G.; Niu, P.; Duan, X.; and Zhang, X. (2014). Fast learning network: A novel artificial neural network with a fast learning speed. *Neural Computing and Applications*, 24(7-8), 1683-1695.
13. Little, M.A.; McSharry, P.E.; Roberts, S.J.; Costello, D.A.E.; and Moroz, I.M. (2007). Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering Online*, 6(23), 1-19.