Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2019 doi:10.20944/preprints201905.0125.v1

Parkinson's Disease Detection Using Biogeography-Based Optimization

Somayeh Hessam [1], Shaghayegh Vahdat[2], Irvan Masoudi Asl[3], Mahnaz Kazemipoor[4], Shahaboddin Shamshirband[5,6] and Timon Rabczuk[7]

[1] Department of Health Services Administration, South Tehran Branch, Islamic Azad University, Tehran, Iran.

[1] Department of Health Services Administration, South Tehran Branch, Islamic Azad University, Tehran, Iran.

[1] PHD of Health Services Management, Associate Professor, Majlis Research Center, Tehran, Iran.

[1] Clinic for Nutrition and Natural Medicine, Karaj, Iran

[1] Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[6]Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam

[7]Institute of Structural Mechanics, Bauhaus-University Weimar, 99423 Weimar, Germany

Abstract: In recent years, Parkinson's Disease (PD) as a progressive syndrome of the nervous system has become highly prevalent worldwide. In this study, a novel hybrid technique established by integrating a Multi-layer Perceptron Neural Network (MLP) with the Biogeography-based Optimization (BBO) to classify PD based on a series of biomedical voice measurements. BBO is employed to determine the optimal MLP parameters and boost prediction accuracy. The inputs comprised of 22 biomedical voice measurements. The proposed approach detects two PD statuses: 0– disease status and 1– reasonable control status. The performance of proposed methods compared with PSO, GA, ACO and ES method. The outcomes affirm that the MLP-BBO model exhibits higher precision and suitability for PD detection. The proposed diagnosis system as a type of speech algorithm detects early Parkinson's symptoms, and consequently, it served as a promising new robust tool with excellent PD diagnosis performance.

Keywords: Parkinson's disease (PD); Biomedical voice measurements; Multi-layer Perceptron Neural Network (MLP); Biogeography-based Optimization (BBO); Medical diagnosis. Bio-inspired computation

1 Introduction

Parkinson's disease (PD) is a type of neurological disorder initiating by the death of cells in the midbrain. There is a lack of a specific method to diagnose PD, but this disease could be typically diagnosed through the medical history, evaluating the signs and symptoms,

 doi:10.20944/preprints201905.0125.v1

and neurological and physical analysis of the patient. At present, there is no specific treatment for this health problem, but it is feasible to alleviate the symptoms and slow down its progress remarkably. Investigations have proven that there is around ninety percent of the individuals with PD exhibit vocal impairment (Ho, Iansek et al. 1999). Subjects with PD frequently suffer from different vocal impairment symptoms recognized as dysphonia. The symphonic signs of PD are important diagnosis measures. Therefore, dysphonic assessments have been considered as the reliable tools for monitoring and detection of PD over the past years (Rahn, Chou et al. 2007, Little, McSharry et al. 2009).

PD diagnosed by clinical features. However, several brain imaging methods comprising positron emission tomography (PET), single photon emission computed tomography (SPECT) and magnetic resonance imaging (MRI) are widely used for PD diagnosis (Pyatigorskaya, Gallea et al. 2014). Mainly, implications of MRI, which provides numerous applicant biomarkers and have the possibility of notifying about the disease process, have primarily been investigated. Zeng et al. (2017) have used an MVPA (Multivariate pattern analysis) method for 45 potential PD patients and 40 healthy subjects as the control group, to investigate the probable alterations in cerebellar gray matter. Based on structural MRI scans, this method combines SVM with voxel-based morphometry to detect morphological abnormalities in the Cerebellum. Also, Cherubini et al. (Cherubini, Morelli et al. 2014) utilized SVMs to distinguish 57 probable PD patients from 21 PSP (Progressive Supranuclear Palsy) patients based on their MRI scans.

Apart from analyzing these conventional biomarkers for PD diagnosis, several studies have explored that speech and gait disorders associated with the PD. Besides, several algorithms and techniques have applied for PD detection. These techniques are mainly classified as gait-based and speech-based methods (Shrivastava, Shukla et al. 2017). Speech and gait disorders are characterized as Axial parkinsonian symptoms (Ricciardi, Ebreo et al. 2016). Gait is signaled as a sensitive indicator for PD progression as PD patients exhibit altered patterns of gait with increased cadence and reduced stride lengths. The specific gait patterns, gait initiation and freezing gait (FOG) characterized as indicators of PD. Gait-based PD detection methods utilize different image and video processing methods for PD detection through the subject’s gait assessment. Speech disorders in PD patients are dissimilar and heterogeneous, comprising hypo-, hyperkinetic and repetitive abnormalities. Recent studies have revealed that some form of vocal impairment detected in more than 90% of PD patients. In general, there are two ways to analyze the speech status: (1) subjective: by speech therapist (perceptive analysis) and (2) objective: by analyzing speech signals through acoustic analysis (Brabenec, Mekyska et al. 2017). Speech-based PD detection methods mainly use the Unified Parkinson’s Disease Rating Scale (UPDRS). Several machine learning models have established for predicting the UPDRS score of the subject by using speech signals. These techniques can provide non-intrusive means of monitoring the onset and development of the PD conditions.

Several researchers have applied computational techniques for detection of PD. Little et al. (Little, McSharry et al. 2009) employed a support vector machine (SVM) classifier with Gaussian radial basis kernel functions for PD detection. They also attempted to choose the optimum subset of features. Das (Das 2010) compared various types of classification approaches for effective PD diagnosis, with the prime objective being to discern healthy people. According to the results, the neural network classifier produces the most accurate outcomes. Guo et al. (Guo, Bhattacharya et al. 2010) hybridized genetic programming with the expectation-maximization algorithm to develop the GP-EM approach for detecting healthy individuals and those with PD. The researchers found that GP-EM is highly effective. Hossein et al. (Hossen, Muthuraman et al. 2010) employed wavelet-decomposition with a soft-decision algorithm to diagnose the Parkinson tremor from essential tremor. Luukka (Luukka 2011) applied a feature selection approach based on fuzzy entropy measures together with the similarity classifier for predicting PD and the results indicated a notable prediction enhancement by using the proposed method. Aström and Koker (Åström and Koker 2011) utilized a parallel neural network technique to increase the precision of PD predictions. Based on their results, substantial prediction improvements achieved by using the proposed model. Chen et al. (Chen, Huang et al. 2013) applied the fuzzy k-nearest neighbor (FKNN) technique to develop an efficient model for PD diagnosis. By making a comparison, the researchers demonstrated that FKNN outperforms SVM in PD prediction. Daliri (Daliri 2013) proposed a chi-square distance kernel-based SVM approach to diagnosing PD using gait signals. Based on the assessments of 93 individuals with PD and 73 healthy people, they concluded that the technique could be used successfully for PD diagnosis. Hariharan et al. (Hariharan, Polat et al. 2014) acquired a hybrid intelligent approach comprising feature pre-processing, feature reduction/selection and classification. Their results signified that the proposed scheme is capable of precise classification for PD detection.

Lahmiri et al. (Lahmiri 2017) have also investigated the statistical characteristics and effectiveness of diverse types of dysphonia assessments in PD detection. Results of the statistical tests concluded that all dysphonia assessments usually show diverse variability among PD patients and healthy candidates. The results of classification acquired through SVM classifier, indicated that in contrast to the other dysphonia measures, SVM trained with VFFS produced the maximum accurateness of 88%, while SVM trained with NLDCM resulted in the minimum precision of 80.82%. A three-phase methodology by Travieso et al. (Travieso, Alonso et al. 2017) aimed at automatic detection of voice disease. This study advocates the transformation of the feature space by a Discrete Hidden Markov Model (DHMM) first and then application of RBF-SVM classifier. Wu Y. et al. (Wu, Chen et al. 2017) proposed to use an interclass probability risk (ICPR) technique for the vocal parameter selection. Subsequently, they have compared three different non-linear classifiers including SVM, GLRA (generalized logistic regression analysis) and Bagging ensemble algorithms, to distinguish the voice patterns of PD patients and healthy subjects. The experimental results demonstrated better classification accuracy by SVM and Bagging ensemble classifiers (90.77%) with ICPR. Yang et al. (Yang, Zheng et al. 2014) used two feature dimensionality reduction methods, including kernel principal component analysis (KPCA) and sequential forward selection (SFS). They selected four vocal measures including MDVP: F0, MDVP: Jitter (%), DFA, spread2 and employed MAP (Maximum A Posteriori) for classification. In contrary to little et al. (Little, McSharry et al. 2009), who executed rescaling of feature values from -1 to 1, authors have argued that for such data

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2019 doi:10.20944/preprints201905.0125.v1

set, input data normalization is not required. In their opinion, normalization or rescaling may not be robust for the minor data set, as the full vocal records are less than 200. New recruited voice records may require another rescaling session, and consequently, consuming more computation time.
Moreover, physical magnitude information regarding voice measurements is suspected to be lost after data normalization. Problems of small data set mainly revolve around high variance where overfitting, outliers, and noise emerge considered as significant concerns. To avoid overfitting, Tsanas et al. (Tsanas, Little et al. 2012) suggested using cross-validation for an approximation of the true generalization performance on the unknown cases.

Most of the existing researches on PD detection, primarily focus on the accuracy of prediction and reliability of the diagnosis. However, up to this time, too little attention has been paid to investigate the time efficiency and computational complexity of different classification mechanisms for PD detection. Islam et al. (Islam, Parvez et al. 2014) investigated Feed forward back propagation based on ANN (FBANN), SVM and Random tree classifiers for PD detection using dysphonia measures. Their results signify that FBANN demonstrates higher sensitivity with relatively less execution time. Generally, an appropriate feature selection method can effectively tackle both computation times and cure-of-dimension problems. In the context of Firefly-SVM, Chao et al. (Chao and Horng 2015) advocated that convergence with the most optimal solution within a limited time is possible when firefly-SVM associated with the feature selection.

SVM is known as a machine learning system which has attained considerable significance in applications linked to the environment (Jain, Garibaldi et al. 2009, Ornella and Tapia 2010). SVM is a learning algorithm that applies high-dimensional features. SVM model precision depends on parameter determination (Chapelle, Vapnik et al. 2002). Although structured strategies for parameter selection are vital, model parameter alignment is also required. To choose the SVM model parameters, scientists have utilized several standard optimization algorithms. However, the outcomes are not very efficient due to parameter complexity (Lee and Verri 2003, Friedrichs and Igel 2005, Bao, Hu et al. 2013). The grid search algorithm (Lorena and De Carvalho 2008) and decent gradient algorithm (Chung, Kao et al. 2003, Hsu, Chang et al. 2003) are two algorithms which are applied before. The computational complication is a main disadvantage of the grid search algorithm; therefore, it utilized for selecting a few parameters.
Moreover, the grid search algorithm is commonly disposed to the local minima. Most of the optimization complications have various local solutions, but advanced algorithms appear to be the optimum means of solving these as they offer global solutions. Recently, the optimization techniques applied for classification (Mosavi and Vaezipour 2012), and (Brunato and Battiti 2013).

The Multi-Layer Perceptron (MLP) applied for numerous practical complications. The training on applications required for using MLP, which usually might encounter different complications such as entrapment in local minima, convergence speed, and sensitivity to initialization. In this study, authors propose the Biogeography-Based Optimization (BBO)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2019 doi:10.20944/preprints201905.0125.v1

algorithm for training MLPs to diminish such complications. Their experimental results on several classification datasets such as balloon, iris, breast cancer, heart problems, and several approximating datasets such as sigmoid, cosine, sine, sphere, Griewank, and Rosenbrock demonstrate that BBO has much more ability to escape local minima in comparison with PSO, GA, ACO, ES, and PBIL (Mirjalili, Mirjalili et al. 2014).

In one of the most recent studies (Pham, Nguyen et al. 2019), the researchers proposed a hybrid machine learning method known as MLP-BBO for estimating the coefficient of consolidation as an essential parameter of soft soil. This technique is according to the Multi-layer Perceptron Neural Network (MLP) and Biogeography-based Optimization (BBO). For comparing the performance of the models applied in their study, standard machine learning methods applied including Backpropagation Multi-layer Perceptron Neural Networks, Radial Basis Functions Neural Networks, Gaussian Process, M5 Tree, and Support Vector Regression. The outcomes of that research model indicated that the recommended MLP-BBO technique has the maximum predictive competency.

In another study by Das et al. (Das, Pattnaik et al. 2014), the researchers have applied Artificial Neural Network (ANN) trained with Particle Swarm Optimization (PSO) for solving the channel equalization problems. According to the proposed method, they used PSO on Artificial Neural Networks (ANN) to find optimal weights of the network on training step, and they tried to consider a suitable network topology and transfer performance of the neuron. The PSO algorithm can optimize the variables, weights and network parameters. Hence, this study emphases on improving the weights, transfer function, and topology of an ANN which made for channel equalization. In the current study, it demonstrated that the equalizer perform better than other ANN equalizer in all noise conditions.

Blum & Socha in 2005 (Blum and Socha 2005) proposed an ACO algorithm for the training of feed-forward neural networks. The algorithm function evaluated by pattern classification complications related to the medical field. They compared their algorithms to several feed-forward neural network training, called BP, LM and genetic algorithm. The functionality of the ACO was as good as the performance of other NN training algorithms. Although the ACO_NN method was initially presented to solve the distinct optimization issues, in recent times, it applied for the improvement of algorithms used for the endless optimization issues.

Moreover, Chandwani et al. (Chandwani, Agrawal et al. 2015) applied hybrid model of Artificial Neural Networks (ANN) and Genetic Algorithms (GA) for modelling slump of Ready Mix Concrete (RMC) related to its design mix constituents viz., cement, fly ash, sand, coarse aggregates, admixture and water-binder proportion. The recommended hybrid approach joined GA to develop the optimum set of first neural network weights and predispositions that were later fine-tuned utilizing Lavenberg Marquardt back-propagation training algorithm. Their research indicated that the hybridizing ANN with GA, the convergence rate of ANN and its estimating accurateness upgraded.

In the current study, the MLP is combined with BBO into a hybrid method (MLP-BBO) to detect PD from 22 biomedical voice assessments. BBO is employed to find out the optimal MLP parameters. The primary objective of this research is to examine the appropriateness of the suggested MLP-BBO approach for PD detection. To verify the MLP-PSO method's precision its capability compared with existing optimization methods.

2 Materials and Methods

2.1 Data Description

For the present research, an investigation was carried out using a PD dataset obtained from the UCI machine learning repository (http://archive.ics.uci.edu/ml/datasets/Parkinsons, last accessed: August 2014). The objective of the data is to diagnose healthy individuals and people suffering from PD, providing the outcomes of several medical examinations performed on the patients. The utilized data includes a collection of biomedical voice assessments related to 31 individuals in which 23 of them suffer from PD. The period from PD diagnosis varies between 0 and 28 years. The subjects are in the 46-85 years old range, with an average of 65.8. Each candidate delivered a middling of six vowel phonations (yielding 195 testers entirely), and the duration of each phonation was 36 seconds. Further information on this dataset presented in the paper published by Little et al. (Little, McSharry et al. 2009). Remarkably, all features are real and no missing and unreliable values exist in the used dataset. The brief explanations about the dataset can be found from the Little et al. (Little, McSharry et al. 2009).

2.2 Biogeography-Based Optimization_ Multi-Layer Perceptron (BBO_MLP)

The basic idea of Biogeography-Based Optimization algorithm was motivated by biogeography, referring to the science of biological creatures related to the geographical spreading over time and space (Simon 2008) . The development of ecosystems to get to a steady condition while making an allowance for diverse species (including predator, prey, etc.), and the influence of migration and mutation was the leading motivation for the BBO algorithm. BBO algorithm uses several search agents known as habitats as chromosomes in Gas, and a Habitat Suitability Index (HSI) states the general fitness of a habitat. The greater the HSI, the higher fit the habitat. The habitats develop over time according to the three principles as below (Ma, Simon et al. 2013).

- Habitants living in environments with more HSI are more probable to immigrate to territories with less HSI.
- Environments with less HSI are more likely to be fascinating for new immigrant habitats from those with more HSI.
- Random alterations may take place in the habitats irrespective to their HSI values.

The BBO algorithm begins with a random set of habitats. Every habitat has dissimilar habitats that represent the number of variables of a particular issue. Emigration (μ_k), immigration (λ_k) and mutation (m_n) for each habitat expressed as functions of the number of habitats as below:

$$\mu_n = \frac{E \times n}{N} \qquad (1)$$

$$\lambda_k = 1 \times \frac{1-n}{N} \quad (2)$$

$$m_n = M \times \left(1 - \frac{p_n}{p_{max}}\right) \quad (3)$$

Where n is the existing number of habitats, N is the acceptable maximum number of habitats which is raised by HSI (the more appropriate the habitat, the greater number of habitats), E is the maximum emigration rate, and I indicates the maximum immigration rate. M is an original value for mutation described by the user, p_n is the mutation possibility of the nth habitat, and p_{max} = argmax(p_n), n = 1,2,. . .,N.
The overall stages of the BBO algorithm is:

1. Initializing step: a random set of habitats
2. do{
3. calculating HIS of each habitat
4. updating the rate of Emigration (μ_k), immigration (λ_k) and mutation (m_n) for each habitat
5. the non_elite habitats are migrated and mutated based on the updated rates
6. selecting the best habitats as elites for next generation}
7. While (non_satisfying the terminated criterion)
8. Returning the best solution (habitats)

For further details about the algorithm refer to (Simon 2008).

2.3 BBO for MLP

The BBO algorithm used for an MLP with two main phases (Mirjalili, Mirjalili et al. 2014) :

1. Demonstration strategy: the weights and biases must be expressed in the proper format (habitats) for BBO.
   For demonstrating the MLP training problem for BBO, the vector used as habitat formation. This vector contains weights and biases in MLP network. For instance, the last vector of the MLP shown in Fig .1 as below is a sample of this encoding strategy:
   Habitat= [w_{13} w_{23} w_{14} w_{24} …. $w_{22\ 44}$ $w_{22\ 44}$ θ_1 θ_2 θ_3 θ_4…]) w_{ij}: NN weight between neuron I and j and θ_i bias for neuron i

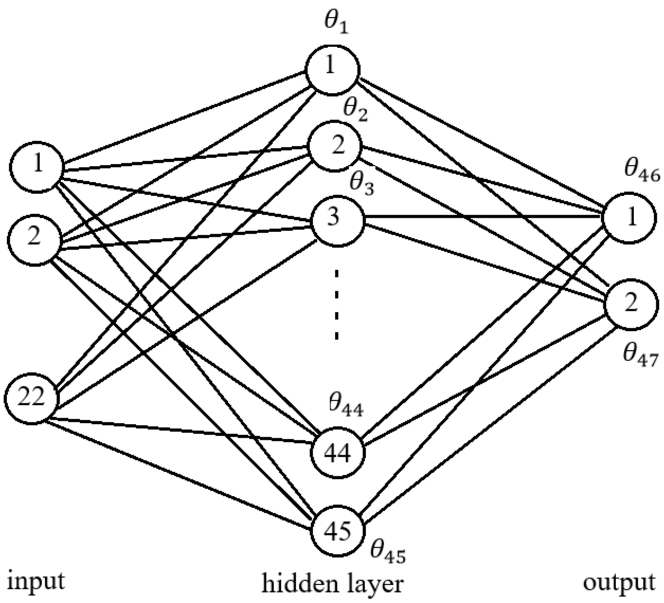


Fig .1 MLP with three layers [3]

2. HSI: a fitness function using the error of the MLP would be described to estimate habitats.

After demonstrating MLPs in the form of habitat vectors, an HSI formulation (fitness function) is prerequisite for calculating each of them. The Mean Square Error (MSE) for all training models used as a fitness function (MSE (habitat)=HIS(habitat)):

$$E = \sum_{k=1}^{q} \frac{\sum_{i=1}^{m}(o_i^k - d_i^k)^2}{q} \qquad (4)$$

where q is the number of training samples, m is the number of outputs, d_i^k is the desired output of the ith input unit when the kth training sample used and o_i^k is the actual output of the ith input unit when the kth training sample appears in the input. The BBO_MLP algorithm explained in Fig .2:

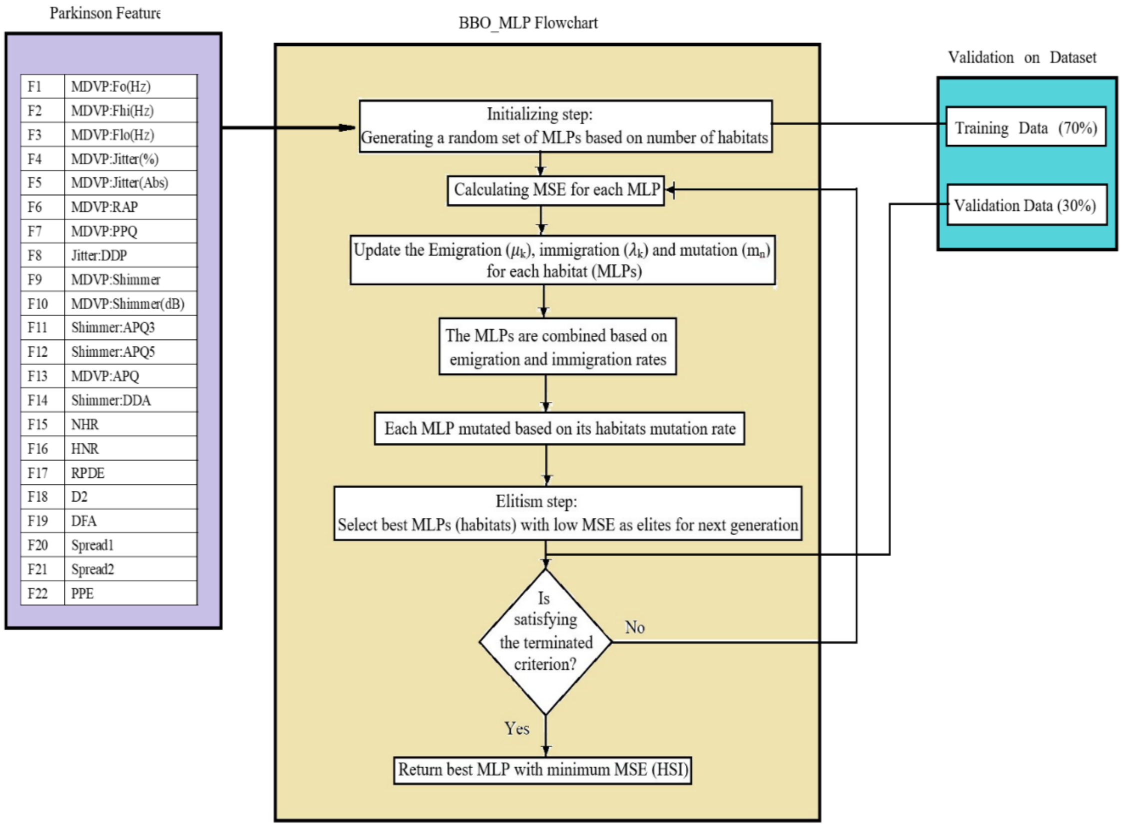


Fig .2 Flow chart of BBO_MLP Method

MLP-BBO is made in six phases as follow:

- Initial step: making a random set of MLPs according to the outlined number of habitats
- Calculating MSE for each MLP
- Update emigration $\mu_n = \frac{E\times n}{N}$, immigration $\lambda_k = 1 \times \frac{1-n}{N}$, and mutation $m_n = M \times \left(1 - \frac{p_n}{p_{max}}\right)$rates
- The MLP are united based on emigration and immigration (create vector [w_{13} w_{23} w_{14} w_{24} …. $w_{22\ 44}$ $w_{22\ 44}$ θ_1 θ_2 θ_3 θ_4…]) w_{ij}: NN weight between neuron I and j and θ_i bias for neuron i
- Each MLP mutated based on its habitat mutation rate
- Elitism step: select the best MLPs with low MSE as elites for next generation
- satisfaction of a termination criterion (if no satisfaction algorithm repeated as flowchart)
- Return best MLP with minimum MSE(HSI)

2.4 Input Parameters

 doi:10.20944/preprints201905.0125.v1

The aptitude of BBO-MLP to produce reliable predictions is reliant on input parameter selection. In the current research, 22 biomedical voice measurements were used to produce the BBO-MLP model. The descriptive statistics including minimum, maximum and mean values, standard deviation and the range of values of the datasets applied in this research presented in Table 1.

Table 1: Descriptive statistics for the data sets used

Variable	Statistics				
	Min	Max	Mean	Standard deviation	Range
In1	88.333	260.105	154.2286	41.3901	171.772
In2	102.145	592.03	197.1049	91.4915	489.885
In3	65.476	239.17	116.3246	43.5214	173.694
In4	0.0017	0.0332	0.0062	0.0048	0.0315
In5	0.000007	0.00026	0.000044	0.0000348	0.000253
In6	0.0007	0.0214	0.0033	0.003	0.0207
In7	0.0009	0.0196	0.0034	0.0028	0.0187
In8	0.002	0.0643	0.0099	0.0089	0.0623
In9	0.0095	0.1191	0.0297	0.0189	0.1096
In10	0.085	1.302	0.2823	0.1949	1.217
In11	0.0046	0.0565	0.0157	0.0102	0.0519
In12	0.0057	0.0794	0.0179	0.012	0.0737
In13	0.0072	0.1378	0.0241	0.0169	0.1306
In14	0.0136	0.1694	0.047	0.0305	0.1558
In15	0.0006	0.3148	0.0248	0.0404	0.3142
In16	8.441	33.047	21.886	4.4258	24.606
In17	0.2566	0.6852	0.4985	0.1039	0.4286
In18	0.5743	0.8253	0.7181	0.0553	0.251
In19	–7.965	–2.434	–5.6844	1.0902	5.531
In20	0.0063	0.4505	0.2265	0.0834	0.4442
In21	1.4233	3.6712	2.3818	0.3828	2.2479
In22	0.0445	0.5274	0.2066	0.0901	0.4829

Table 2 shows the user-defined parameters for MLP-BBO and PSO, ACO, GA and ES for PD detection.

Table 2: Parameter Regularization

Population Size		Maximum number of Generations		Mutation probability		Habitat modification probability [0 1]		Percentage of cross-validation for Train/Test	
Parameter value	Accuracy	Parameter value	Accuracy	Parameter value	Accuracy	Parameter value	Accuracy	Parameter value	Accuracy
20	82 %	50	82 %	0.001	68 %	0.6	74 %	50-50	76.84 %
50	84 %	150	84 %	0.005	84 %	0.8	78 %	**70-30**	**86 %**
150	82 %	**250**	**86 %**	**0.008**	**86 %**	**1**	**86 %**	80-20	80 %
200	**86 %**	350	84 %	0.01	82 %			90-10	84 %
300	76 %	500	84 %	0.05	82 %				
		1000	84 %						

Table 2 shows that the best experimental result achieved with the parameters value 200 for population size, 250 for the maximum number of generation, 0.008 for mutation probability, 1 for habitat modification probability and splitting 70-30 percentage of cross-validation for Train/Test. As seen in table 2, these set of parameters regularization leads to the accuracy of 86 percentage.

Therefore, the best result on multi-layer perceptron based on BBO algorithm was obtained according to the regularization of MLP_BBO parameters.

3. Results and Discussion

3.1 Statistical performance Analysis

The accuracy formula is served as the reliable statistical parameters to appraise the capability of the MLP-PSO model on a more noticeable and individual basis. Table 5 offers the values achieved for accuracy during training and testing. It is evident that the models' performance reduced from training to testing. According to the statistical results presented in Table 4, the proposed hybrid MLP-PSO model naturally exhibits greater PD detection capability and precision compared to the existing optimization model.

The BBO algorithm is equated with PSO, GA, ACO, ES, and PBIL over these benchmark datasets to verify its performance. It is expected that every habitat was randomly adjusted in the range. The population size is 50 for Parkinson dataset. Table 3 shows how the datasets are allocated in terms of training and test sets.

Table 3 dataset in terms of training and testing

Classification Dataset	Number of feature	Number of training samples		Number of testing samples		Number of classes
Parkinson	22	50%	98	50%	97	2
		70%	136	30%	59	
		80%	156	20%	39	
		90%	175	10%	20	

In this study, the researchers have chosen the paramount trained MLP among ten runs, and then they applied it to categorize or estimate the test set. To deliver an unbiased association, the whole algorithms ended when a maximum amount of iterations (250) achieved. Lastly,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2019 doi:10.20944/preprints201905.0125.v1

the merging actions are correspondingly considered in the outcomes to deliver a complete assessment. It reminded that min-max standardization applied for the datasets comprising data with diverse ranges. Finally, the result of MLP_BBO in terms of accuracy rate illustrated in Table 4.

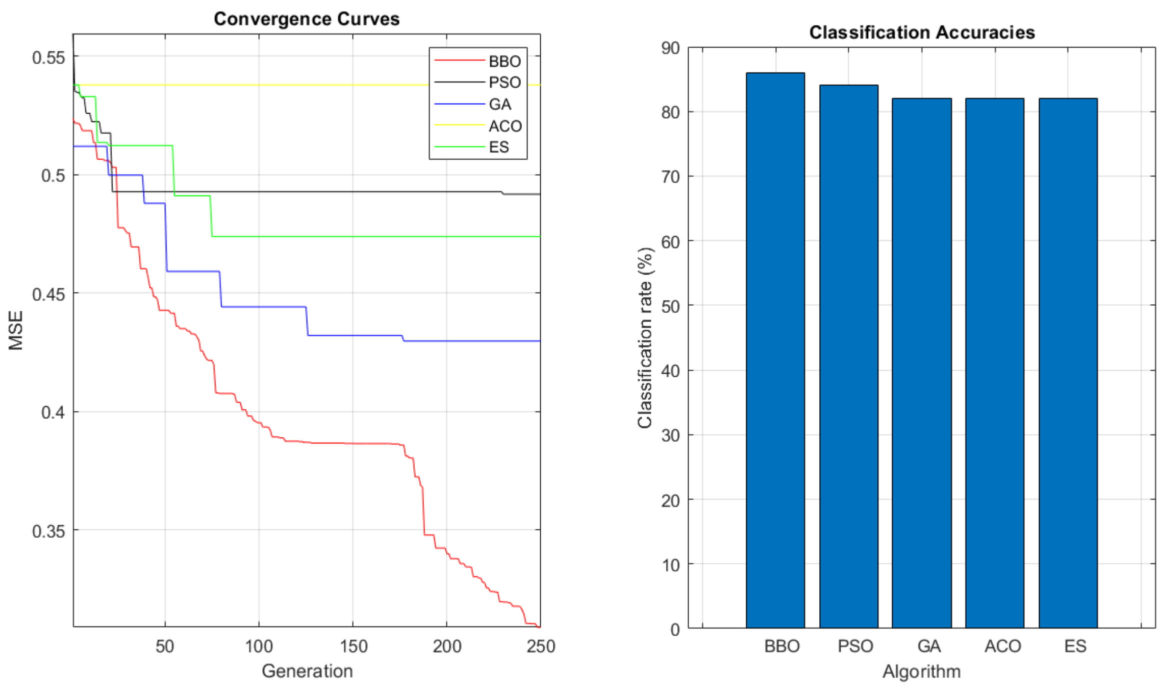
Table 4: Comparable Algorithms

Method	**BBO_MLP**	PSO_MLP	GA_MLP	ACO_MLP	ES_MLP	SVM_MLP
Accuracy	**86**	84	82	82	82	81

Table 4 compares the six-optimization algorithm in terms of accuracy of the multi-layer perceptron (MLP). The above table indicates that the accuracy of BBO_MLP is more than the other five optimization MLP algorithm. The accuracy calculated as follow:

$$\text{Accuracy} = \frac{(\text{TP}+\text{TN})}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} * 100 \qquad (5)$$

where TP, TN, FP, and FN are true positive, true negative, false positive and false negative respectively.



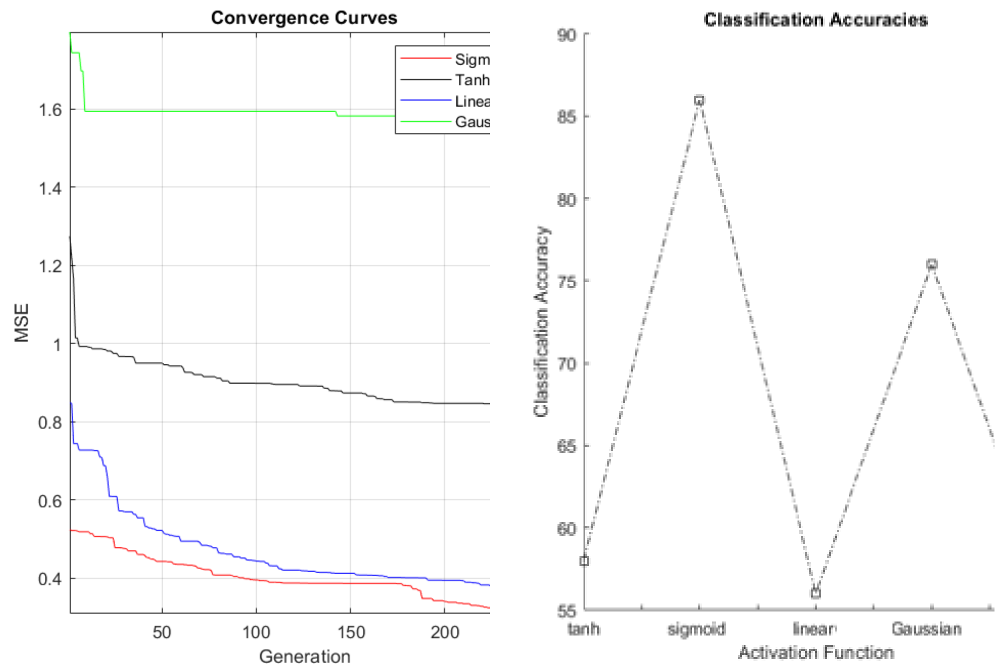
a b

Fig 3a: The rate of error (RMSE) convergence and b. The rate of accuracy for several MLP based algorithm

Figure 3a shows the MSE for each method of BBO, PSO, GA, ACO, and ES based on MLP. As it is evident in the figure, BBO method significantly decreases errors in comparison with other approaches. Also, the bar chart of the above figure (Fig 3b) indicates

doi:10.20944/preprints201905.0125.v1

that the MLP-BBO technique with an accuracy rate of 86% has offered better results compared with other developing methods. According to Fig 3b, the MLP-ACO, MLP-GA, and MLP-ES with the 82% had the same percentage of accuracy. Furthermore, in this study the recommended approach is examined on different activation functions such as sigmoid, linear, tanh, sin and Gaussian and the results are observed in Fig 4.



a b

Fig 4a: The rate of error (RMSE) convergence and **b:** The rate of accuracy for several activation functions on BBO_MLP algorithm

Figure 4 shows that which activation functions have a better result in terms of high accuracy and low rate of RMSE error in BBO-MLP classification method. As it is evident in the Fig 4.a, sigmoid method significantly decreases errors in comparison with other activation functions in MLP. As it observed from the Fig 4.b, sigmoid activation functions with the 86% have better performance in comparison with other activation functions (Tanh: 58%, Linear:56%, Gaussian:76% and Sin: 56%).

4. Conclusion

In this study, a hybrid approach proposed for the detection of Parkinson's disease (PD) determined from biomedical voice measurements. To achieve this purpose, the MLP was combined with the BBO to develop the hybrid MLP-BBO method. MLP essentially achieves structural minimization, whereas other traditional optimization approaches focus on error minimization and are much less efficient.

 doi:10.20944/preprints201905.0125.v1

As mentioned above, due to the lack of performance in MLP, a set of processes known as "Meta-heuristic algorithms" could reach to a solution by frequently bringing up to date the applicant solution and assessing to an optimal result to a problematic issue, through improving the objective function. In this research, the MLP parameters are optimized utilizing BBO that through calculating its performance, it is inferred to outperform the MLP performance.

Through the method of merging BBO with MLP, the flashing actions of the fireflies could be conveyed to form an objective function that could be useful to adjust the parameters of MLP. By BBO, it recognized that the higher frequency of comparisons between the BBO to find the optimum location in the swarm, the superior the outcomes would be.

The principal aim was to identify the suitability of the MLP-BBO method developed for detecting two PD statuses: 0 – disease status and 1 – reasonable control status. The accuracy of MLP-BBO with 86 percentage verified against the lower accuracy of PSO, GA, ACO and ES method. Accuracy was served to assess the MLP-BBO models' PD detection performance statistically. The findings indicate that the MLP-BBO model developed in this study is more precise than PSO, ACO, GA, and ES in PD detection. Consequently, the proposed diagnosis system exhibits favorable precision and is supposed as a promising and appealing tool for detecting PD.

References

Åström, F. and R. Koker (2011). "A parallel neural network approach to the prediction of Parkinson's Disease." Expert systems with applications **38**(10): 12470-12474.

Bao, Y., Z. Hu and T. Xiong (2013). "A PSO and pattern search based memetic algorithm for SVM parameters optimization." Neurocomputing **117**: 98-106.

Blum, C. and K. Socha (2005). Training feed-forward neural networks with ant colony optimization: An application to pattern classification. Hybrid Intelligent Systems, 2005. HIS'05. Fifth International Conference on, IEEE.

Brabenec, L., J. Mekyska, Z. Galaz and I. Rektorova (2017). "Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation." Journal of Neural Transmission **124**(3): 303-334.

Brunato, M. and R. Battiti (2013). "Learning and intelligent optimization (LION): one ring to rule them all." Proceedings of the VLDB Endowment **6**(11): 1176-1177.

Chandwani, V., V. Agrawal and R. Nagar (2015). "Modeling slump of ready mix concrete using genetic algorithms assisted training of Artificial Neural Networks." Expert Systems with Applications **42**(2): 885-893.

Chao, C.-F. and M.-H. Horng (2015). "The construction of support vector machine classifier using the firefly algorithm." Computational intelligence and neuroscience **2015**: 2.

Chapelle, O., V. Vapnik, O. Bousquet and S. Mukherjee (2002). "Choosing multiple parameters for support vector machines." Machine learning **46**(1): 131-159.

Chen, H.-L., C.-C. Huang, X.-G. Yu, X. Xu, X. Sun, G. Wang and S.-J. Wang (2013). "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach." Expert systems with applications **40**(1): 263-271.

Cherubini, A., M. Morelli, R. Nisticó, M. Salsone, G. Arabia, R. Vasta, A. Augimeri, M. E. Caligiuri and A. Quattrone (2014). "Magnetic resonance support vector machine discriminates between Parkinson disease and progressive supranuclear palsy." Movement Disorders **29**(2): 266-269.

Chung, K.-M., W.-C. Kao, C.-L. Sun, L.-L. Wang and C.-J. Lin (2003). "Radius margin bounds for support vector machines with the RBF kernel." Neural computation **15**(11): 2643-2681.

Daliri, M. R. (2013). "Chi-square distance kernel of the gaits for the diagnosis of Parkinson's disease." Biomedical Signal Processing and Control **8**(1): 66-70.

Das, G., P. K. Pattnaik and S. K. Padhy (2014). "Artificial neural network trained by particle swarm optimization for non-linear channel equalization." Expert Systems with Applications **41**(7): 3491-3496.

Das, R. (2010). "A comparison of multiple classification methods for diagnosis of Parkinson disease." Expert Systems with Applications **37**(2): 1568-1572.

Friedrichs, F. and C. Igel (2005). "Evolutionary tuning of multiple SVM parameters." Neurocomputing **64**: 107-117.

Guo, P.-F., P. Bhattacharya and N. Kharma (2010). Advances in detecting Parkinson's disease. International Conference on Medical Biometrics, Springer.

Hariharan, M., K. Polat and R. Sindhu (2014). "A new hybrid intelligent system for accurate detection of Parkinson's disease." Computer methods and programs in biomedicine **113**(3): 904-913.

Ho, A. K., R. Iansek, C. Marigliani, J. L. Bradshaw and S. Gates (1999). "Speech Impairment in a Large Sample of Patients with Parkinson’s Disease." Behavioural Neurology **11**(3).

Hossen, A., M. Muthuraman, J. Raethjen, G. Deuschl and U. Heute (2010). "Discrimination of Parkinsonian tremor from essential tremor by implementation of a wavelet-based soft-decision technique on EMG and accelerometer signals." Biomedical Signal Processing and Control **5**(3): 181-188.

Hsu, C.-W., C.-C. Chang and C.-J. Lin (2003). "A practical guide to support vector classification."

Islam, M. S., I. Parvez, H. Deng and P. Goswami (2014). Performance comparison of heterogeneous classifiers for detection of Parkinson's disease using voice disorder (dysphonia). Informatics, Electronics & Vision (ICIEV), 2014 International Conference on, IEEE.

Jain, P., J. M. Garibaldi and J. D. Hirst (2009). "Supervised machine learning algorithms for protein structure classification." Computational biology and chemistry **33**(3): 216-223.

Lahmiri, S. (2017). "Parkinson's disease detection based on dysphonia measurements." Physica A: Statistical Mechanics and its Applications **471**: 98-105.

Lee, S. W. and A. Verri (2003). "Support vector machines for computer vision and pattern recognition." International Journal of Pattern Recognition and Artificial Intelligence **17**(3): 331-332.

Little, M. A., P. E. McSharry, E. J. Hunter, J. Spielman and L. O. Ramig (2009). "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease." IEEE transactions on biomedical engineering **56**(4): 1015-1022.

Lorena, A. C. and A. C. De Carvalho (2008). "Evolutionary tuning of SVM parameter values in multiclass problems." Neurocomputing **71**(16): 3326-3334.
Luukka, P. (2011). "Feature selection using fuzzy entropy measures with similarity classifier." Expert Systems with Applications **38**(4): 4600-4607.
Ma, H., D. Simon, M. Fei and Z. Xie (2013). "Variations of biogeography-based optimization and Markov analysis." Information Sciences **220**: 492-506.
Mirjalili, S., S. M. Mirjalili and A. Lewis (2014). "Let a biogeography-based optimizer train your multi-layer perceptron." Information Sciences **269**: 188-209.
Mosavi, A. and A. Vaezipour (2012). "Reactive search optimization; application to multiobjective optimization problems." Applied Mathematics **3**(10): 1572.
Ornella, L. and E. Tapia (2010). "Supervised machine learning and heterotic classification of maize (Zea mays L.) using molecular marker data." Computers and electronics in agriculture **74**(2): 250-257.
Pham, B. T., M. D. Nguyen, K.-T. T. Bui, I. Prakash, K. Chapi and D. T. Bui (2019). "A novel artificial intelligence approach based on Multi-layer Perceptron Neural Network and Biogeography-based Optimization for predicting coefficient of consolidation of soil." Catena **173**: 302-311.
Pyatigorskaya, N., C. Gallea, D. Garcia-Lorenzo, M. Vidailhet and S. Lehericy (2014). "A review of the use of magnetic resonance imaging in Parkinson’s disease." Therapeutic advances in neurological disorders **7**(4): 206-220.
Rahn, D. A., M. Chou, J. J. Jiang and Y. Zhang (2007). "Phonatory impairment in Parkinson's disease: evidence from nonlinear dynamic analysis and perturbation analysis." Journal of Voice **21**(1): 64-71.
Ricciardi, L., M. Ebreo, A. Graziosi, M. Barbuto, C. Sorbera, L. Morgante and F. Morgante (2016). "Speech and gait in Parkinson's disease: When rhythm matters." Parkinsonism & related disorders **32**: 42-47.
Shrivastava, P., A. Shukla, P. Vepakomma, N. Bhansali and K. Verma (2017). "A survey of nature-inspired algorithms for feature selection to identify Parkinson's disease." Computer Methods and Programs in Biomedicine **139**: 171-179.
Simon, D. (2008). "Biogeography-based optimization." IEEE transactions on evolutionary computation **12**(6): 702-713.
Travieso, C. M., J. B. Alonso, J. R. Orozco-Arroyave, J. Vargas-Bonilla, E. Nöth and A. G. Ravelo-García (2017). "Detection of different voice diseases based on the nonlinear characterization of speech signals." Expert Systems with Applications **82**: 184-195.
Tsanas, A., M. A. Little, P. E. McSharry, J. Spielman and L. O. Ramig (2012). "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease." IEEE Transactions on Biomedical Engineering **59**(5): 1264-1271.
Wu, Y., P. Chen, Y. Yao, X. Ye, Y. Xiao, L. Liao, M. Wu and J. Chen (2017). "Dysphonic Voice Pattern Analysis of Patients in Parkinson’s Disease Using Minimum Interclass Probability Risk Feature Selection and Bagging Ensemble Learning Methods." Computational and Mathematical Methods in Medicine **2017**.

 doi:10.20944/preprints201905.0125.v1

Yang, S., F. Zheng, X. Luo, S. Cai, Y. Wu, K. Liu, M. Wu, J. Chen and S. Krishnan (2014). "Effective dysphonia detection using feature dimension reduction and kernel density estimation for patients with Parkinson's disease." PloS one **9**(2): e88825.

Zeng, L.-L., L. Xie, H. Shen, Z. Luo, P. Fang, Y. Hou, B. Tang, T. Wu and D. Hu (2017). "Differentiating Patients with Parkinson's Disease from Normal Controls Using Gray Matter in the Cerebellum." The Cerebellum **16**(1): 151-157.