Kurdistan Journal of Applied Research (KJAR) | Print-ISSN: 2411-7684 – Electronic-ISSN: 2411-7706 | kjar.spu.edu.iq
Volume 2 | Issue 3 | August 2017 | DOI: 10.24017/science.2017.3.121

Feature Selection and Radial Basis Function Network for Parkinson Disease Classification

Ashraf Osman Ibrahim
Faculty of Computer Science,
Future University,
Khartoum, Sudan
Arab Open University, Sudan
ashrafosman2@gmail.com

Walaa Akif Hussien
Faculty of Computer Science,
Future University,
Khartoum, Sudan
walaa.akif@gmail.com

Ayat Mohammoud Yagoop
Faculty of Computer Science,
Future University,
Khartoum, Sudan
ayatmahmoud618@gmail.com

Mohd Arfian Ismail
Soft Computing & Intelligent Systems
Research Group,
Faculty of Computer
System and Software Engineering,
University Malaysia Pahang,
Pahang, Malaysia
arfian@ump.edu.my

Abstract: ***Recently, several works have focused on detection of a different disease using computational intelligence techniques. In this paper, we applied feature selection method and radial basis function neural network (RBFN) to classify the diagnosis of Parkinson's disease. The feature selection (FS) method used to reduce the number of attributes in Parkinson disease data. The Parkinson disease dataset is acquired from UCI repository of large well-known data sets. The experimental results have revealed significant improvement to detect Parkinson's disease using feature selection method and RBF network.***

Keywords: Parkinson's disease, feature selection, artificial neural networks, classification, radial basis function, attributes reduction.

1. INTRODUCTION

Parkinson's disease (PD) is a long term disorder of the central nervous system that affects the motor system . Parkinson's primarily influences neurons in specific zone of the brain called the substantia nigra [1, 2]. Parkinson Disease progresses, amount of dopamine produced in the brain decreases, which lead patient to uncontrollable movement as a normal person. [3]. The key reason of this decline is not recognized yet; nevertheless researchers are conducting many researches to find out a solution.

There are primary symptoms of the PD can be noted as tremors in the hands, legs, arms, jaw and face. Hardness or hardening of limbs and trunk. Slow movement (motion). Positive instability, or poor balance and coordination. These symptoms also become more noticeable [4, 5]. Parkinson's disease can't be diagnosed easily in the early stages since there are many factors to analyze.

In the Initial of the disease, the most noticeable symptoms are shaking, stiffness and slow movement [4]. [4]. Problems of thinking and behavior may also occur. Actually, in the advanced stages of the disease the dementia becomes common. In addition, more than a third of people with PD commonly feel with depression and anxiety [5]. There are further symptoms contain sensory, sleep, and emotional problems [4, 5]. Major motor symptoms are collectively called "parkinsonian syndrome" [6].

Parkinson's disease is thought to include genetic and environmental aspects. A persons with a family members who affected with PD are more likely to get the same disease [7]. There is a higher risk among people exposed to some pesticides and between those who have been injured in the head while there was a lower risk for those who smoking tobacco and drink coffee or tea [7, 8]. The motor symptoms of the disease are caused by cells death in the nigra, a region of the brain's middle. This leads to inadequate dopamine in these areas[4]. Generally, the cause of these death of cells is not understood but contains the growth of proteins in Lewy bodies in neurons [7]. Mainly, the diagnosis of typical cases is depend on symptoms, with the use of some tests such as neuroscience to exclude other diseases. Even though, there is no cure for Parkinson's disease [9, 10]. The Primary treatments are usually with levodopa antiparkensonian, with once can used dopamine stimuli becomes less effective levodopa. With disease progression and nerve cells are still lost, these drugs become less effective, at the same time they produce complex involuntary movements characterized [11].

This paper is organized as follows. Section 2 presents the literature review. Section 3 discusses materials and methods. Section 4 gives brief about the data. Section 5 shows the results of the study. Section 6 concludes the paper.

2. LITERATURE REVIEW

There are many studies in the literature introduced how to help in diagnoses this disease in early stage, these studies used different methods, most of these methods based on application of neural network.

Mehmet Can [12] used neural network system with backpropagation together.

The process of designing neural network system is boosted by filtering. This lead to a significant increase of robustness. In addition, the common voting of 11 parallel networks, recognition rates reached to greater than 90 in spite of 3:1 imbalanced class distribution of the Parkinson's disease data set.

P. Durga , V. Sutha Jebakumari, D. Shanthi [13] used various data mining techniques such as; Naive Bayes, Sequential Minimal Optimization (SMO), J48, Bayesian Network and Multilayer Perceptron and they show a

good accuracy results.

Marius EneS [14], applied three types of probabilistic neural network (PNN), have been used to classification process. The Monte Carlo search (MCS), incremental search (IS) and hybrid search (HS) were use to the smoothing the factor search. The actual model has providing diagnosis accuracy about 79% and 81%.

Indira Rustempasic, Mehmet Can [15] used artificial neural networks, they study the neural network performance using backpropagation along with a majority voting structure. For train samples the authors used boosting by filtering technique with seven committee machines and they used principal component analysis (PCA) for reduction the data.
They concluded in their results the use of proposed techniques had a good results and the ablity of classification the Parkinson's disease is good as well.

In addition, [16] applied four different computational approaches to diagnosis of Parkinson disease and compare the classification results.

Another work in [17] conducted comparative study for the performance of SVM, MLP and RBN on a Parkinson's Disease tremor classification. In addition, [18] were used probabilistic neural network, feed forward, artificial immune system and learning vector quantization and study these methods then got the comparative of the results.

Recent studies employed hybrid methods for Parkinson's Disease, the research conducted by [19] proposed a new hybrid intelligent method to predict of PD progression by using adaptive neuro-fuzzy inference system (ANFIS) and support vector regression (SVR). They used noise removal, clustering and prediction methods. [20] implemented feature dimension reduction technique and developing sequential forward selection algorithm along with the kernel principal component analysis approaches. With accomplish the linear classification from claiming voice records for sound control for healthy and sick people the authors applied the Fisher's linear discriminant analysis (FLDA), maximum a posteriori (MAP) decision rule and SVM with RBF network for classification tasks. [21] displayed the hybridization of the wavelet analysis hybrid and support vector machine can produce efficient classification accuracy for Parkinson's gait identification.

Most of the previous studies focused on using artificial neural networks to find a pattern that can be used to classify the Parkinson's disease. In this paper, mainly we focus on using the feature selection algorithm to reduce the attributes that can help RBF network to give high classification accuracy for the Parkinson's disease.

3. METHODS AND MATERIALS

Recently, the improvements in the area of the artificial intelligence (AI) led to the emanation of the decision support systems and expert systems for medical applications. Artificial Intelligence (AI) are techniques for classification, in this section, we propose using feature selection and the RBF network used as a classifier.

For preprocessing the data set we used Min-Max normalization method. It converts A value to B value which fits between [C, D] values, to transform data set to the range [0.0, 1.0] as in equation 1.

$$\text{Normalization} = (x-\min(x))/ (\max(x)-\min(x)) \quad (1)$$

3.1 Radial Basis Function Network (RBFN)

Is an artificial neural network, mostly used for classification purposes. The classification in RBF is carried out by computing the similarity of inputs to samples from the training set. Each neuron is stored as a "prototype", which is just one sample of a training set. To classify a new input; all neurons calculate euclidean distance between inputs and their model. Figure 1, show the architecture of the RFB network.

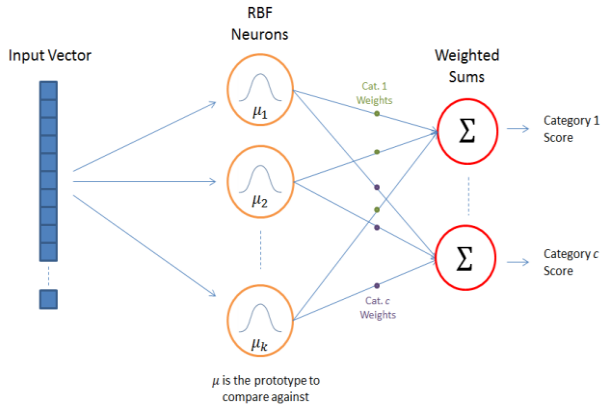

Figure 1 RBF network architecture [22]

3.2 Feature selection algorithm

Is the procedure of select a subset of related features for use in model construction. There are three reasons for using feature selection methods:

- Simplify prototypes to make them easier to interpret by users.
- To reduce the training times.
- To enhance the generalization via reduce over fitting.

FS method is like a combination of a search method for offering new feature subsets, alongside an evaluation measure which scores the different feature subsets. However, the simplest method is to test every potential subset of features finding the particular case which minimizes the error rate. This may be an exhaustive search of the space, also will be computationally ungainly for everything except the lowest of feature sets.

We used feature selection algorithms to minimize the number of features and choose the best feature that gives

high accuracy.

Four feature selection algorithms were used in this study; Cfs Subset Eval, Info Gain Attribute Eval, principal components and Wrapper Subset Eval and as search method we used ranker and best first.

3.3 Evaluate model

Usually to evaluate the performance of the classifiers need one of the evaluation measures. We have classified Parkinson's disease data set to classify the patient either healthy or sick. Sensitivity, specificity and accuracy were used to evaluate the model.

The correct positive samples that generate the classifier are called sensitivity (SEN). On the other hand, the correct negative samples which depend on the number of true negatives and false positives is called specificity (SPE). The equations 2-4 show the calculation of the evaluation method. The sensitivity (SEN) is given by:

$$SEN = {t - pos}/{pos} \qquad (2)$$

Where, t-pos is the number of true positives correctly classified as healthy and pos is the number of positive healthy samples. The specificity (SPE) is given by:

$$SPE = {t - neg}/{neg} \qquad (3)$$

True positive, false positive, true negative and false negative are suitable to calculate the accuracy. The classification the accuracy is given by:

$$Accuracy = {n0,0 + n1,1}/{n} \qquad (4)$$

With these equations we calculate the accuracy of our classifier with the feature selection algorithm.

4. DATA SET

The data set was made by the University of Oxford, in collaboration with the National Centre for Voice and Speech, Denver, Colorado, who recorded the speech signals. The original study published the feature extraction methods for general voice disorders.

This dataset is composed of a range of biomedical voice measures from 31 people, 23 with Parkinson's disease (PD). In the table data, every column is a particular voice measure, and each row corresponds one of 195 voice recording from these individuals ("name" column). The key purpose of this data is to discriminate healthy people from those with PD. According to "status" column in the table, 0 values set for healthy and 1 values set for sick people. Table 1 describes the dataset attribute in details.

Table 1: Data set attributes description

Attribute	Description
MDVP:Fo(Hz)	Average vocal fundamental frequency
MDVP:Fhi(Hz)	Maximum vocal fundamental frequency
MDVP:Flo(Hz)	Minimum vocal fundamental frequency
MDVP:Jitter(%) **MDVP:Jitter(Abs)** **MDVP:RAP** **MDVP:PPQ** **Jitter:DDP**	Several measures of variation in fundamental frequency
MDVP:Shimmer **MDVP:Shimmer(dB)** **Shimmer:APQ3** **Shimmer:APQ5** **MDVP:APQ** **Shimmer:DDA**	Several measures
RPDE **D2**	Two nonlinear dynamical complexity measures
DFA	Signal fractal scaling exponent
spread1 **spread2** **PPE**	Three nonlinear measures of fundamental frequency variation
NHR **HNR**	Two measures of proportion of noise to tonal ingredient in the voice

5. RESULT AND DISCUSSION

This section provides the experimental results along with some discussion about the results. Table 2 shows the classification accuracy for all feature subsets A, B, C and D, (including the data set before feature selection) for cross validation and training set test.

Table 2: Classification accuracy %

	Dataset	Subset A	Subset B	Subset C	Subset D
Cross validation	68.20	76.41	71.80	65.64	79.49
Training set	78.97	75.90	77.43	69.23	83.59

As we can notice from the table 2 that the highest accuracy is from the subset D, from this result, we can understand that the features which subset D contains it has the best prediction, but at the same time we can't consider D as the best subset, because it has only four features which are not the best representation of the data set.

Moreover, we can see that the subset C has the lowest accuracy value, which means that subset C's features don't contain the best feature for prediction.

We can see that the accuracy when we use subset becomes higher than the accuracy before the feature selection. Table 3, shows the measurement criteria for the model evaluation.

Table 3: Measurement criteria for the model

Measurement Criteria	Partition	Dataset	Subset A	Subset B	Subset C	Subset D
Correctly Classified Instances	**Validation cross**	133	149	140	128	155
	Training set	154	151	128	135	163
Incorrectly Classified Instances	**Validation cross**	62	46	55	67	40
	Training set	41	44	67	60	32
Kappa statistic	**Validation cross**	0.34	0.48	0.38	0.26	0.56
	Training set	0.54	0.49	0.26	0.36	0.65
Mean Absolute error	**Validation cross**	0.39	0.34	0.36	0.43	0.31
	Training set	0.32	0.32	0.43	0.41	0.29
Root mean squared error	**Validation cross**	0.46	0.43	0.45	0.47	0.41
	Training set	0.40	0.40	0.47	0.45	0.30
Relative absolute error	**Validation cross**	80.88	70.64	76.19	89.09	65.18
	Training set	67.14	65.99	89.09	84.79	60.72
Root relative squared error	**Validation cross**	94.11	87.11	91.03	95.92	82.84
	Training set	81.92	81.33	95.92	92.18	77.87

The results in table 3, show the measurement criteria of the models. The results show how the feature selection algorithms make the classification more reliable. In addition, the best result comes from subset D which its output occurred after using Wrapper Subset Evaland as a feature selection algorithm.

Table 4, shows the confusion matrices for the model of training and cross validation test data partition after and before feature selection.

Table 4: The confusion matrices for the training and cross validation test

Model	Desired output	Training set data		Cross validation set test	
		sick	**Healthy**	**sick**	**healthy**
Dataset	**Sick**	108	10	85	33
	Healthy	31	46	29	48
Subset A	**Sick**	113	5	104	14
	Healthy	41	36	36	41
Subset B	**Sick**	106	12	91	27
	Healthy	41	36	39	38
Subset C	**Sick**	95	23	92	26
	Healthy	37	40	41	36
Subset D	**Sick**	105	13	102	16
	Healthy	19	58	24	53

All these measurements proofs that the accuracy after feature selection has higher value than the accuracy before feature selection. In addition, the best result from the previous measurement comes from subset D which was selected after using wrapper Subset Evaland as a feature selection algorithm. Wrapper Subset Evaland feature selection algorithm and RBF network when used to classify the Parkinson Disease shows better accuracy results.

6. CONCLUSION

One of the most significant challenges is choosing the right classifier algorithm for the classification the medical data. In the present study, we choose the RBF network as a classifier approach, we used a feature selection algorithm to reduce the attribute of the Parkinson Disease that can help RBF network to increase the accuracy results. Four algorithms of the feature selection were used and divided the dataset to four subsets according to these algorithms, after the classification we compared between the results of the dataset and the four subsets, which a proof that feature selection helps improving the classification results.

7. REFERENCE

[1] W. Dauer and S. Przedborski, "Parkinson's disease: mechanisms and models," *Neuron*, vol. 39, pp. 889-909, 2003.

[2] M. Fjodorova, E. M. Torres, and S. B. Dunnett, "Transplantation site influences the phenotypic differentiation of dopamine neurons in ventral mesencephalic grafts in Parkinsonian rats," *Experimental neurology*, vol. 291, pp. 8-19, 2017.

[3] D. M. Vogt Weisenhorn, F. Giesert, and W. Wurst, "Diversity matters–heterogeneity of dopaminergic neurons in the ventral mesencephalon and its relation to Parkinson's Disease," *Journal of neurochemistry*, vol. 139, pp. 8-26, 2016.

[4] A. Jamak, A. Savatić, and M. Can, "Principal component analysis for authorship attribution," *Business Systems Research*, vol. 3, pp. 49-56, 2012.

[5] M. Can, "Neural Networks to Diagnose the Parkinson's Disease," 2013.

[6] L. V. Kalia, S. K. Kalia, and A. E. Lang, "Disease-modifying strategies for Parkinson's disease," *Movement Disorders*, vol. 30, pp. 1442-1450, 2015.

[7] J. P. Iannotti and R. Parker, *The Netter Collection of Medical Illustrations: Musculoskeletal System, Volume 6, Part III-Musculoskeletal Biology and Systematic Musculoskeletal Disease E-Book*: Elsevier Health Sciences, 2013.

[8] J. L. Barranco Quintana, M. F. Allam, A. S. Del Castillo, and R. F. n.-C. Navajas, "Parkinson's disease and tea: a quantitative review," *Journal of the American College of Nutrition*, vol. 28, pp. 1-6, 2009.

[9] N. Singh, V. Pillay, and Y. E. Choonara, "Advances in the treatment of Parkinson's disease," *Progress in neurobiology*, vol. 81, pp. 29-44, 2007.

[10] C. Camara, P. Isasi, K. Warwick, V. Ruiz, T. Aziz, J. Stein, *et al.*, "Resting tremor classification and detection in Parkinson's disease patients," *Biomedical Signal Processing and Control*, vol. 16, pp. 88-97, 2015.

[11] S. Sveinbjornsdottir, "The clinical symptoms of Parkinson's disease," *Journal of neurochemistry*, vol. 139, pp. 318-324, 2016.

[12] M. Can, "Diagnosis of parkinson's disease by boosted neural networks," 2013.

[13] P. Durga, V. S. Jebakumari, and D. Shanthi, "Diagnosis and Classification of Parkinsons Disease Using Data Mining Techniques," *International Journal of Advanced Research Trends in Engineering and Technology*, vol. 3, pp. 86-90.

[14] M. Ene, "Neural network-based approach to discriminate healthy people from those with Parkinson's disease," *Annals of the University of Craiova-Mathematics and Computer Science Series*, vol. 35, pp. 112-116, 2008.

[15] I. Rustempasic and M. Can, "Diagnosis of Parkinson's disease using principal component analysis and boosting committee machines," 2013.

[16] R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Systems with Applications*, vol. 37, pp. 1568-1572, 2010.

[17] S. Pan, S. Iplikci, K. Warwick, and T. Z. Aziz, "Parkinson's Disease tremor classification–A comparison between Support Vector Machines and neural networks," *Expert Systems with Applications*, vol. 39, pp. 10764-10771, 2012.

[18] O. Er, O. Cetin, M. S. Bascil, and F. Temurtas, "A Comparative study on parkinson's disease diagnosis using neural networks and artificial immune system," *Journal of Medical Imaging and Health Informatics*, vol. 6, pp. 264-268, 2016.

[19] M. Nilashi, O. Ibrahim, and A. Ahani, "Accuracy improvement for predicting Parkinson's disease progression," *Scientific reports*, vol. 6, 2016.

[20] S. Yang, F. Zheng, X. Luo, S. Cai, Y. Wu, K. Liu, *et al.*, "Effective dysphonia detection using feature dimension reduction and kernel density estimation for patients with Parkinson's disease," *PloS one*, vol. 9, p. e88825, 2014.

[21] D. Joshi, A. Khajuria, and P. Joshi, "An automatic non-invasive method for Parkinson's disease classification," *Computer Methods and Programs in Biomedicine*, vol. 145, pp. 135-145, 2017.

[22] pp. http://mccormickml.com/2013/08/15/radial-basis-function-network-rbfn-tutorial/, 2013.