ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 4, No.16 , July 2015
ISSN : 2322-5157
www.ACSIJ.org

Combination of PSO Algorithm and Naive Bayesian Classification for Parkinson Disease Diagnosis

Navid Khozein Ghanad[1],Saheb Ahmadi [2]

[1] **Islamic Azad university Of Mashhad, Faculty of Engineering, Department Of Computer, Mashhad, Iran**
navidghanad@mshdiau.ac.ir

[2] **Islamic Azad university Of Mashhad, Faculty of Engineering, Department Of Computer, Mashhad, Iran**
Sahebahmadi93@gmail.com

Abstract

Parkinson is a neurological disease which quickly affects human's motor organs. Early diagnosis of this disease is very important for its prevention. Using optimum training data and omitting noisy training data will increase the classification accuracy. In this paper, a new model based on the combination of PSO algorithm and Naive Bayesian Classification has been presented for diagnosing the Parkinson disease, in which optimum training data are selected by PSO algorithm and Naive Bayesian Classification. In this paper, according to the obtained results, Parkinson disease diagnosis accuracy has been 97.95% using the presented method, which is indicative of the superiority of this method to the previous models of Parkinson disease diagnosis.

Keywords: *Parkinson disease diagnosis, Naive Bayesian Classification, PSO algorithm*

1. Introduction

Parkinson disease is one of the nervous system diseases, which causes quivering and losing of motor skills. Usually this disease occurs more in people over 60 years old, and 1 out of 100 individuals suffers from this disease. However, it is also observed in younger people. About 5 to 10% of patients are in younger ages. After Alzheimer, Parkinson is the second destructive disease of the nerves. Its cause has not been recognized yet. In the first stages, this disease has significantly low symptoms [1]. It is claimed that 90% of Parkinson patients can be recognized

Through voice disorders[2]. Parkinson patients have a set of voice disorders by which their disease can be diagnosed. These voice disorders have indices whose measurement can be used for diagnosing the disease [3] [4]. In the previous studies, problems of Parkinson disease diagnosis were considered. Using SVM Classification with Gaussian kernel, the obtained result was 91.4% at best [4]. In order to diagnose the Parkinson disease, a new non-linear model based on Dirichlet process mixing was presented and compared with SVM Classification and decision tree. At best, the obtained result was 87.7% [5]. In [6], different methods have been used to diagnose the Parkinson disease, in which the best result pertained to the classification using the neural network with 92.9% accuracy. In [7], the best features were selected for SVM Classification through which 92.7% accuracy could be obtained at best. In [8], using sampling strategy and multi-class multi-kernel relevance vector machine method improvement, 89.47% accuracy could be achieved. In [9], the combination of Genetic Algorithm and Expectation Maximization Algorithm could bring 93.01% accuracy for Parkinson disease diagnosis. In [10], using fuzzy entropy measures, the best feature was selected for classification and thereby 85.03% accuracy could be achieved for classification. In [11], the combination of non-linear fuzzy method and SVM Classification could detect the speaker's gender with 93.47% accuracy. In [12], the combination of RF and CFS algorithms could diagnose the Parkinson disease with 87.01% accuracy. In [13], using parallel forward neural network, Parkinson disease was diagnosed with 91.20% accuracy. In [14], with improvements in OPF Classification, Parkinson disease was diagnosed with 84.01% accuracy. In [15], fuzzy combination with the Nearest Neighbor Algorithm could achieve 96.07% accuracy. In [16] and [17], by focusing on voice analysis, they attempted to gain 94% accuracy. In the previous presented methods, attempts have been made to offer the best classification methods and no attention has been paid to the quality of the training data. In this paper, we presented a new model based on the combination of PSO algorithm and Naive Bayesian Classification for diagnosing the Parkinson disease. This algorithm selects the best training data for Naive Bayesian Classification and this causes no use of non-optimal training data. Due to using optimum training data and not using non-optimal training data, this new model presented in this paper increases the classification accuracy and Parkinson disease diagnosis to 97.95%.

First we consider Naive Bayesian Classification and PSO algorithm. Then, the presented algorithm, results and references will be investigated.

1.1. Naive Bayesian Classification

One very practical Bayesian learning method is naive Bayesian learner which is generally called the Naive Bayesian Classification method. In some contexts, it has been shown that its efficiency is analogous to that of the methods such as neural network and decision tree.

Naive Bayesian classification can be applied in problems in which each sample x is selected by a set of trait values and the objective function f(x) from a set like V. Bayesian method for the new sample classification is such that it detects the most probable class or the target value v_{MAP} having trait values $\langle a_1,a_2,\ldots a_n \rangle$, which describes the new sample.

$$v_{map}=\arg_{vi=v}\max p(v_j \text{ I } a_{1,} a_{2,\ldots\ldots,} a_n) \qquad (1)$$

Using Bayesian ' theorem, term (1) can be rewritten as term (2):

$$V_{map}=\arg_{vi=v}\max\frac{p(a1,a2,\ldots.an,Ivj)p(vj)}{p(a1,a2,\ldots..,an)}$$
$$=\arg_{vi=v}\max P(a1,a2,\ldots..,a_n,Iv_{j)})P(v_j) \qquad (2)$$

Now using the training data, we attempt to estimate the two terms of the above equation. Computation based on the training data to find out what is the repetition rate of v_j in the data, is easy. However, computation of different terms $P(a_1,a_2,\ldots a_n \mid V_j)$ by this method will not be acceptable unless we have a huge amount of training data available. The problem is that the number of these terms is equal to the number of possible samples multiplied by the number of the objective function values. Therefore, we should observe each sample many times so that we obtain an appropriate estimation.

Objective function output is the probability of observing the traits $a_1,a_2,\ldots a_n$ equal to the multiplication of separate probabilities of each trait. If we replace it in Equ.2, it yields the Naive Bayesian Classification, i.e. Equ.3:

$$V_{NB=}\arg\max P(V_j)\prod_i P(a\ i \mid V\ j) \qquad (3)$$

Where v_{NB} is Naive Bayesian Classification output for the objective function. Note that the number of terms $P(a_i|v_j)$ that should be computed in this method is equal to the number of traits multiplied by the number of output classes for the objective function, which is much lower than the number of the terms $P(a_1,a_2,\ldots a_n \mid V_j)$

We conclude that naive Bayesian learning attempts to estimate different values of $P(v_j)$ and $P(a_i|v_j)$ using their repetition rate in the training data. This set of estimations corresponds to the learnt assumption. After that, this assumption is used for classifying the new samples, which is done through the above formula. When conditional independence assumption of Naive Bayesian Classification method is estimated, naive Bayesian class will be equal to the MAP class.

1.2. PSO algorithm

Each particle is searching for the optimum point. Each particle is moving, thus it has a speed. PSO is based on the particles' motion and intelligence. Each particle in every stage remembers the status that has had the best result.

Particle's motion depends on 3 factors:

1- Particle's current location
2- Particle's best location so far (pbest)
3- The best location which the wholeset of particles were in so far (gbest)

In the classical PSO algorithm, each particle *i* has two main parts and includes the current location, and *Xi*is the particle's current speed (*Vi*). In each repetition, particle's change of location in the searching space is based on the particle's current location and its updated speed. Particles' speed is updated according to three main factors: particle's current speed, particle's best experienced location (individual knowledge), and particle's location in the best status of group's particles (social knowledge), as Equ.4.

$$V_{i+1} = K(wV_i + C_{1i}(Pbest_i - X_i) + C_{1i}(Gbest_i - X_i)) \quad (4)$$

Where *W* is the i^{th} particle's inertia coefficient for moving with the previous speed. C1i and C2i are respectively the individual and group learning coefficients of the i^{th} particle, which are selected randomly in range {2-0} for the sake of maintaining the algorithm's probabilistic property. Each particle's next speed is obtained by Equ.5:

$$X_{i+1} = X_i + V_{i+1} \quad (5)$$

2. Considering the presented algorithm

In the introduction section, we considered that different methods have been presented for Parkinson disease diagnosis, but no attention has been paid to the quality of the training data. In this paper, we attempt to select the best training data using PSO algorithm for Naive Bayesian Classification. The selection of the best training data is the most important part for training the Naive Bayesian Classification training. This is due to the fact that we observed in our studies that adding or omitting two training data in the whole set of training data caused 4 to 5% more accuracy in disease diagnosis. The suggested method will be introduced in detail in the following.
The diagram below shows the general procedure of the new presented algorithm.

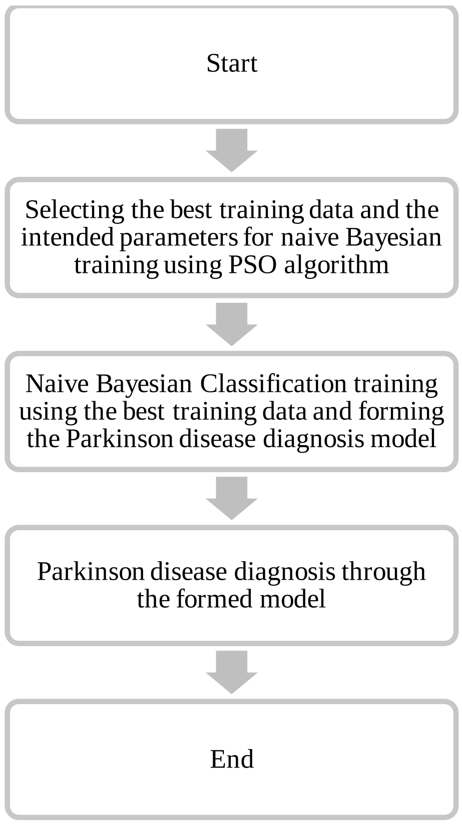

Fig1. The procedure of the suggested method for Parkinson disease diagnosis

The general procedure is very simple. In this paper, first the best data for Naive Bayesian Classification are selected using PSO algorithm, and Naive Bayesian Classification is trained by the optimum training data. Thereby, the Parkinson disease diagnosis model is formed. After the formation of the intended model, the Parkinson disease is diagnosed and identified.

PSO algorithm fitness function for the selection of the optimum training data is expressed in Equ.6:

$$\text{Fitness} = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{x_t - x_f}{x_t}\right| * 100 \qquad (6)$$

where x_t is the real value of the test data, and x_f is the value that has been determined using Naive Bayesian Classification.

Values of the primary parameters of PSO algorithm for the selection of the optimum training data are presented in Table1.

Table1. Primary values given to PSO algorithm parameters

No.	Parameter title	The used parameter value
1	Bird in swarm	50
2	Number of Variable	1
3	Min and Max Range	2-46
4	Availability type	Min
5	Velocity clamping factor	2
6	Cognitive constant	2
7	Social constant	2
8	min of inertia weight	0.4
9	max of inertia weight	0.4

3. Experiments and results

3.1. Dataset descriptions

In this article, we used the dataset of the Parkinson disease belonging to UCI. This dataset is accessible through this link [18]. The number of the items of this dataset is 197, and its features are 23. Features used in Parkinson disease diagnosis are presented in Table2:

ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 4, No.16 , July 2015
ISSN : 2322-5157
www.ACSIJ.org

Table2. Features used in Parkinson disease diagnosis

1	MDVP: FO(HZ)	Average vocal fundamental frequency
2	MDVP: Fhi (HZ)	Maximum vocal fundamental frequency
3	MDVP: Flo(HZ)	Minimum vocal fundamental frequency
4	MDVP: Jitter (%)	
5	MDVP: Jitter (Abs)	
6	MDVP: RAP	
7	MDVP: PPQ	
8	Jitter: DDP	
9	MDVP: Shimmer	Several measures of variation in fundamental frequency
10	MDVP: Shimmer (dB)	
11	Shimmer : APQ3	
12	Shimmer : APQ5	
13	MDVP :APQ	
14	Shimmer :DDA	
15	NHR	Two measures of ratio of noise to tonal components in the voice
16	NHR	
17	RPDE	
18	DFA	
19	Spread 1	Two nonlinear dynamical complexity measure
20	Spread 2	
21	D2	
22	PPE	

3.2. The optimum training data selected for Naive Bayesian Classification using PSO algorithm

As stated in the previous sections, selecting the best training data is the most important part of Naive Bayesian Classification for increasing the accuracy and Parkinson disease diagnosis. In Table3, the number of the optimum training data selected by PSO algorithm can be observed:

Table3. The accuracy of Parkinson disease diagnosis class using the optimum training data selected by PSO algorithm

No.	The number of the optimum training data selected for Naive Bayesian Classification using PSO algorithm	Classification accuracy
1	8	97.95%
2	10	96.93%
3	12	97.95%

In Table3, some of the optimum training data selected using PSO algorithm along with the classification accuracy obtained through the optimum training data can be found. As can be seen in No. 2 of Table3, by adding two training data, classification accuracy has decreased 1.02%. Therefore, it can be concluded that by increasing the training data, there is no guarantee that classification accuracy be increased. The important point in increasing the classification accuracy is the use of optimum training data and no use of noisy training data which decrease the classification accuracy. We increased the number of training data respectively to 50, 60, 70, 80 and 90 training data. The accuracy of the obtained results of this high number of training data can be observed in Table4.

Table4. The relationship between Parkinson disease diagnosis accuracy and training data increase

No.	The number of the training data	Classification accuracy
1	50	88.79%
2	60	77.55%
3	70	76.53%
4	80	69.38%
5	90	67.54%

In Table4, we can see that using Naive Bayesian Classification with increasing the training data will decrease the classification accuracy.

According to the optimum training data selected by PSO algorithm, it is concluded that by having only 8 training data, the highest accuracy in the possible classification can be obtained for Parkinson disease diagnosis.

In Table5, the result of the algorithm presented in this paper is compared with the results of the previous works:

Table5. Comparison of the suggested method's accuracy and previous models of Parkinson disease diagnosis

No.	Presented works	Result and accuracy of the presented model
1	[9]	93.01%
2	[11]	93.01%
3	[13]	91.20%
4	[15]	96.01%
5	[16][17]	94%
6	Proposed Algorithm	**97.95%**

According to the comparison made between the suggested method and the previous models of Parkinson disease diagnosis in Table5, it is shown that the suggested method is superior to the previous models of Parkinson disease diagnosis. Based on the comparison it can be concluded that in order to increase the classification accuracy, it is not always necessary to present a new classification method; rather by selecting the best training data and omitting the inappropriate training data, classification accuracy can be significantly increased.

4. Conclusion

In this paper, we suggested a new model for Parkinson disease diagnosis based on the combination of PSO algorithm and Naive Bayesian Classification. Using PSO algorithm, the best training data were selected for Naive Bayesian Classification. Due to the fact that this presented algorithm selects the best training data and avoids choosing those that cause drop and decrease in classification accuracy, it gets the classification accuracy and Parkinson disease diagnosis to 97.95%. This classification accuracy shows the superiority of the suggested method to the previous models of Parkinson disease diagnosis. Also, according to the result obtained in the paper, it can be reminded that in order to increase the classification accuracy, it is not always necessary to present a new classification method; rather by selecting the best training data and omitting the inappropriate training data, classification accuracy can be significantly increased.

References

[1] Singh, N., Pillay, V., &Choonara, Y. E. (2007). Advances in the treatment of Parkinson's disease. Progress in Neurobiology.81,29-44

[2] Ho, A. K., Iansek, R., Marigliani, C., Bradshaw, J. L., & Gates, S. (1998).Speechimpairment in a large sample of patients with Parkinson's disease. Behavioural Neurology, 11,131-138

[3] Little, M. A., McSharry, P. E., Hunter, E. J., Spielman, J., &Ramig, L. O (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Transactions on Biomedical Engineering, 56(4), 1015–1022

[4] Rahn, D. A., Chou, M., Jiang, J. J., & Zhang, Y. (2007). Phonatory impairment in Parkinson's disease: evidence from nonlinear dynamic analysis and perturbation analysis. Journal of Voice, 21, 64–71.

[5] Shahbaba, B., & Neal, R. (2009). Nonlinear models using Dirichlet process mixtures.The Journal of Machine Learning Research, 10, 1829–1850.

[6] Das, R. (2010). A comparison of multiple classification methods for diagnosis of Parkinson disease. Expert Systems with Applications, 37, 1568–1572.

[7] Sakar, C. O., &Kursun, O. (2010). Telediagnosis of Parkinson's disease using measurements of dysphonia. Journal of Medical Systems, 34, 1–9

[8] Psorakis, I., Damoulas, T., &Girolami, M. A. (2010). Multiclass relevance vectormachines: sparsity and accuracy. Neural Networks, IEEE Transactions on, 21,1588–1598.

[9] Guo, P. F., Bhattacharya, P., &Kharma, N. (2010). Advances in detecting Parkinson's disease. Medical Biometrics, 306–314.

[10] Luukka, P. (2011). Feature selection using fuzzy entropy measures with similarity classifier. Expert Systems with Applications, 38, 4600–4607.
[11] Li, D. C., Liu, C. W., & Hu, S. C. (2011). A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets. Artificial Intelligence in Medicine, 52, 45–52.
[12] Ozcift, A., &Gulten, A. (2011). Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. Comput Methods Programs Biomed, 104, 443–451.
[13] AStröm, F., &Koker, R. (2011). A parallel neural network approach to prediction of Parkinson's Disease. Expert Systems with Applications, 38, 12470–12474.
[14] Spadoto, A. A., Guido, R. C., Carnevali, F. L., Pagnin, A. F., Falcao, A. X., & Papa, J. P. (2011). Improving Parkinson's disease identification through evolutionarybased feature selection. In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE (pp. 7857–7860).
[15] Hui-Ling Chen a, Chang-Cheng Huang a, Xin-Gang Yu b, Xin Xu c, Xin Sun d, Gang Wang d, Su-Jing Wang(2013). An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach. In Expert Systems with Applications 40 (2013) 263–271
[16] Yoneyama, M.;kurihara, y.;watanabe,k;mitoma, h. accelerometry-Based Gait Analysis and Its Application to Parkinson's Disease Assessment— Part 1: Detection of Stride Event(Volume:22 , Issue: 3)page:613-622, May 2014
[17] Yoneyama, M.;kurihara, y.;watanabe,k;mitoma, h.Accelerometry-Based Gait Analysis and Its Application to Parkinson's Disease Assessment— Part 2: New Measure for Quantifying Walking Behavior (Volume:21 , Issue: 6)page:999-1005,Nov. 2013
[18] UCI machine learning repository. (http://archive.ics.uci.edu/ml/datasets/Parkinsons)