Medical Hypotheses 138 (2020) 109603

Contents lists available at ScienceDirect

Medical Hypotheses

journal homepage: www.elsevier.com/locate/mehy

Early diagnosis of Parkinson's disease using machine learning algorithms

Zehra Karapinar Senturk

Duzce University, Engineering Faculty, Department of Computer Engineering, 81620 Duzce, Turkey

ARTICLE INFO

Keywords:
Decision support systems
Feature selection
Machine learning
Medical diagnosis
Support Vector Machines

ABSTRACT

Parkinson's disease is caused by the disruption of the brain cells that produce substance to allow brain cells to communicate with each other, called dopamine. The cells that produce dopamine in the brain are responsible for the control, adaptation and fluency of movements. When 60–80%of these cells are lost, then enough dopamine is not produced and Parkinson's motor symptoms appear. It is thought that the disease begins many years before the motor (movement related) symptoms and therefore, researchers are looking for ways to recognize the non-motor symptoms that appear early in the disease as early as possible, thereby halting the progression of the disease. In this paper, machine learning based diagnosis of Parkinson's disease is presented. The proposed diagnosis method consists of feature selection and classification processes. Feature Importance and Recursive Feature Elimination methods were considered for feature selection task. Classification and Regression Trees, Artificial Neural Networks, and Support Vector Machines were used for the classification of Parkinson's patients in the experiments. Support Vector Machines with Recursive Feature Elimination was shown to perform better than the other methods. 93.84% accuracy was achieved with the least number of voice features for Parkinson's diagnosis.

Introduction

Parkinson's is a slow progressing neurodegenerative brain disease [1]. Neurodegenerative means that it causes loss of brain cells. Normally, there are brain cells that produce dopamine in certain regions of the human brain. These cells are concentrated in a certain area of the brain called substantia nigra. Dopamine is a chemical that transmits messages between the substantia nigra and other brain regions that control body movements [1]. Dopamine allows people to make smooth and harmonious movements. When 60–80% of dopamine producing cells are lost, not enough dopamine can be produced and motor symptoms of Parkinson's disease (PD) appear. The earliest symptoms of PD appear in the enteric nervous system, lower brain stem and olfactory tracts. PD spreads from these regions to the higher parts of the brain, namely the substantia nigra and the brain shell [1]. It is thought that the disease begins many years before the motor symptoms such as loss or decrease of sense of smell, sleep disturbances and constipation, tremor and slowing of movement. Besides, 90% of PD patients faces with vocal impairments [2]. Therefore, the researchers are looking for ways to recognize these non-motor symptoms that appear early during the disease as early as possible, thereby halting the progression of the disease.

Machine learning (ML) is frequently used for medical disease diagnosis recently because of its implementation convenience and high accuracy [3]. ML has also been used for the treatment of PD in the literature. [4] reviewed the papers for feature selection (FS) to be used for ML in brain surgery. During brain surgery for PD, the true region to be operated is determined by an ML based approach. This paper focuses on the studies after PD is diagnosed. [5] used ML methods to estimate cognitive consequences of PD. [6] predicted the tremor level of PD patients by an ML application. Stage prediction of PD was also performed by ML [7]. But mostly, the researchers focus on the early diagnosis of PD by this popular approach, ML. [8] tried to predict PD based on the motion data acquired from upper limbs of people. The researchers made the experimental objects (both PD patients and healthy people) wear a device into their upper limbs and directed them to do several performance tasks. Spatiotemporal and frequency data analysis was performed to obtain parameters and then different supervised learning methods were used for classification. Different feature extraction methods and ML methods were used for the detection of PD in [9]. They showed that phonation is the most convenient task for PD detection. K-NN, Multilayer Perceptron (MLP), Optimum Path Forest, and Support Vector Machines (SVM) were the evaluated classifiers in the study. Voice features were reduced using artificial neural networks for the ML based diagnosis of PD in [10]. Classification was performed by SVM. Different from the others an unsupervised method was also used for PD [11]. After dimension reduction by partial least squares, self-organizing map (SOM) was used for clustering and incremental

E-mail address: zehrakarapinar@duzce.edu.tr.

https://doi.org/10.1016/j.mehy.2020.109603
Received 17 December 2019; Received in revised form 22 January 2020; Accepted 24 January 2020
0306-9877/

support vector regression prediction. UPDRS (Unified Parkinson's Disease Rating Scale) was predicted in the study. [12] compared different methods and decided neural networks as the best performing ML method for the problem. [13] used fuzzy C-means clustering for feature weighting and k-NN for classification in PD diagnosis task. Weighted PD dataset was given to k-NN classifier for different k values and the best k value was determined. [14] used Extreme Learning Machines for PD diagnosis. Imbalanced data was enhanced by a weighted strategy and non-linear mapping of kernel function. ABC algorithm was used for FS and optimization of parameters. [15] performed a successful decision for PD diagnosis using PCA (Principle Component Analysis) for dimension reduction, FDR (Fisher Discriminant Ratio) for FS and SVM for classification. Although they obtained very high classification accuracies, they used more than 100 features extracted from brain MRI images. Therefore, their computational cost is high. Abdulhay et al. [16] investigated gait and tremor for the diagnosis of PD. Gait features were extracted from the raw data obtained from PhysioNet database. Peaks and pulses were evaluated from the physical signals obtained via the sensors located into the underneath of the subjects' feet. They achieved 92.7% classification accuracy. [17] used voice features of a common dataset. They applied feature augmentation and obtained 177 features from 44 features on the dataset. After feature augmentation, they used ReliefF to filter out the most effective features and at last 66 features were used for the classification of PD. PD was also diagnosed via handwriting tasks [18] rather than MRI, motion, or voice data using ML methods.

Although high classification rates were obtained in the literature for ML based diagnosis of PD, either they used many features (like [9,14,16]) which increases computation time or the extraction of the features were hard even they use few features. Therefore, indirectly, the computation time is again high. In this paper, decreasing computation time via less amount of effective features, a lightweight feature extraction process and a classifier have been aimed. The features are obtained from the speech signals and so to obtain the features is easier than the other MRI-based [15] or motion-based [7,15] methods in the literature. Although some authors (such as [8,9,12]) used voice features for the diagnosis of PD, they used more number of features than the proposed approach. The least number of features were used in [16] with an acceptable classification accuracy, but they used MRI data for feature extraction and it is harder than obtaining voice features.

The main contributions of the proposed ML based early PD diagnosis method are given below.

- Recursive Feature Elimination (RFE) and Feature Importance (FI) methods were used for the determination of the most relevant features to be used in the classification task.
- The importance of using FS methods in the preprocessing phase of classification of PD patients was proven. The performance of SVM classifier has been improved about 13% by FS.
- The least number of voice features in the literature were used for PD diagnosis and very high detection accuracies (93.84%) were obtained with less effort.
- The performances of different popular classifiers were evaluated and the best classifier was found as SVM for PD diagnosis problem.
- Different optimizers were evaluated for the dataset and the best optimizer was determined.
- The proposed method is better than the other methods with respect to computational cost since few number of voice features were used instead of heavy feature extraction processes such as MRI, motion sensors or handwriting assessments.

Materials

The dataset used in the experiments of this study consists of the features obtained from the speech signals of 31 people at the National Centre for Voice and Speech, Denver, Colorado. The dataset was created by Max Little from University of Oxford and donated to UCI Machine Learning Repository [19]. 23 of the 31 people have PD and 8 of them are the control group. There are 195 biomedical voice measurements in the dataset. Table 1 shows the voice measures used in the experiments. Status column in the database defines the class and gets 0 for healthy, 1 for PD. Class distribution of the dataset is shown in Fig. 1. There are 48 healthy phonetics and 147 PD phonetics that belong to 31 people.

Table 1
The voice measures used in the experiments [19].

Feature no	Voice measure	MEANING
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 (%)	Several measures of variation in
5	MDVP:Jitter (Abs)	fundamental frequency
6	MDVP:RAP	
7	MDVP:PPQ	
8	Jitter:DDP	
9	MDVP:Shimmer	Several measures of variation in amplitude
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
16	HNR	components in the voice
17	RPDE	Two nonlinear dynamical complexity
18	D2	measures
19	DFA	Signal fractal scaling exponent
20	spread1	Three nonlinear measures of fundamental
21	spread2	frequency variation
22	PPE	
23	status	Health status of the subject: (1) Parkinson's, (0) healthy

Methods

FS was performed on the 23 voice features. By performing FS, only the effective features were used and the cost of the analysis was reduced. Different FS algorithms were applied for different classification methods. New feature subsets and classifications were generated by using these algorithms from the original feature set. Performance of the model was evaluated based on multiple criteria. Fig. 2 shows the flow chart of the study.

Feature selection

FS which is a preprocessing phase in data science identifies the key features of the problem [20]. Satisfactory determination of attributes has a big significance to improve the accuracy of classification. Overall performance of ML methods can be improved by dimensionality reduction [21]. Jain and Singh says "The application of classification

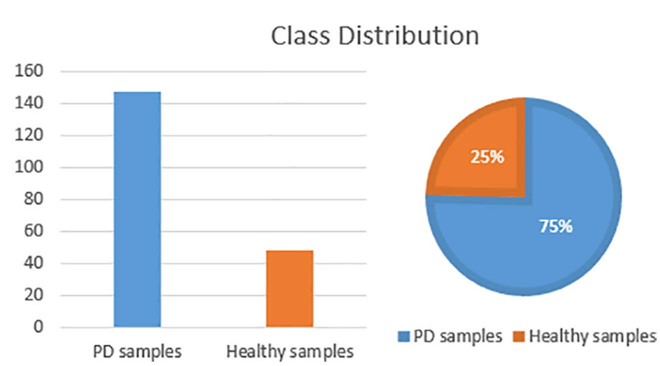

Fig. 1. Class distribution of the recorded phonetics in the dataset.

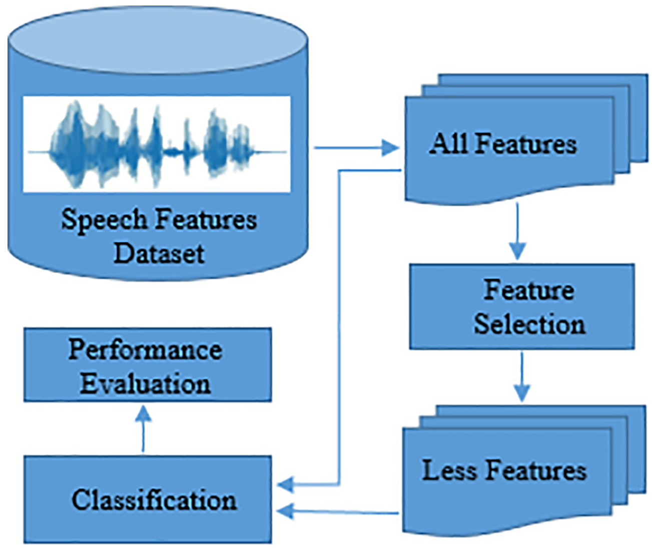

Fig. 2. Flow chart of the proposed decision support system.

algorithms on disease datasets yields promising results by developing adaptive, automated and intelligent diagnostic systems for chronic diseases." [21]. FS has many advantages such as saving time for future data collection, understanding causes of diseases, less computational cost, no degradation in performance [20]. Therefore, in this paper we used FS for the diagnosis of PD. Different FS methods were tried for the problem and they are Univariate Selection (US), RFE, and FI. FI method was determined to be used in Classification and Regression Trees (CART) classification method, and RFE method was used in SVM and Artificial Neural Networks (ANN) classification methods. These FS methods were selected based on the performance of the classifier they were used. FI method gives a score for each feature in the dataset and higher score implies that the feature is more relevant to the output data variable. RFE method fits a model and removes the weakest feature until the desired number of features is reached. Features are ranked and recursively eliminated in loops. Unlike correlation based methods, RFE utilizes SVM. It trains the classifier and optimizes the parameters. Then it computes the ranking criterion for all features and removes the feature with smallest ranking criterion [22].

Classification

CART, SVM and ANN methods were applied to the problem of classifying PD patients and healthy people. CART is a nonparametric statistical methodology and it helps to understand variables or interaction of variables which are responsible for a given problem [23]. If an output variable is continuous, CART produces regression trees; if it is categorical, CART produces classification trees. CART decision tree is a binary tree and has ability to work on noisy data.

SVM maps nonlinear data into a higher dimensional space in which the data is linearly separable. It uses hyper planes to separate two classes and the optimal hyperplane is the one that maximizes the margin between two classes [24]. Its high generalization ability makes it to be used in many fields of classification successfully.

ANN is a very powerful ML method. It inspires from the working principle of human brain. It is a computerized system that is developed to produce new information via learning and to realize this ability automatically [25]. Producing new information or exploring new data is almost impossible by traditional programming methods. ANN is applied to many topics successfully such as learning, association, clustering, generalizing, FS, and optimization because of its properties like parallelism, fault tolerance, easy adaptation to different fields, low computational cost, and self-organization [25].

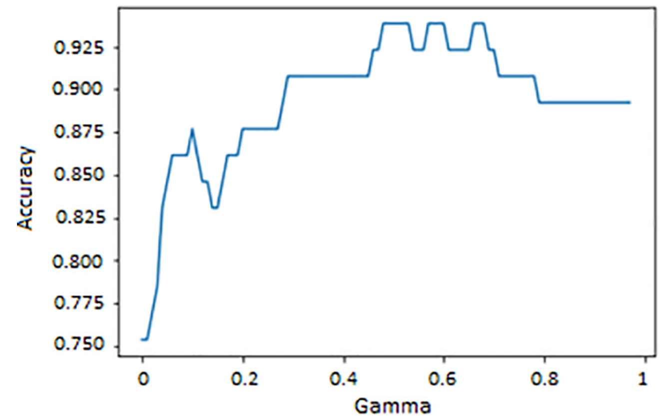

Fig. 3. Variation of Accuracy with respect to Gamma (γ).

Parameter adjustments of classifiers

Binary tree established in CART method had six levels, i.e. root was in level 0 and leafs were in level 6. Seven phonetic based features were used and the tree was branched according to these seven features. In SVM method, the most important issue is to determine the best [c,γ] pair. For this purpose, we realized a grid search for the best choice of these parameters. The performance of classification highly depends on these parameters. Fig. 3 shows the accuracy of classification with respect to gamma (γ). As for ANN, the network topology is determined by trial and error. Activation functions in layers, number of hidden layers, and the number of neurons in hidden layers must be determined. For the Parkinson disease diagnosis problem, hundreds of different experiments were performed to determine the best model. The network topology consisted of single hidden layer with 8 neurons after trials. Tanh, Relu, and sigmoid activation functions were used in input, hidden and output layers respectively. Optimal epoch number was 150 and batch size was 50. Different optimization methods were also evaluated for the dataset as given in Table 3 and nadam was determined as the best optimizer.

Experimental results

FS was performed in this study for the diagnosis of PD via the phonetic features. There were 22 phonetic features extracted from the speech signals of PD patients and the healthy people. The present study not only aimed to diagnose PD patients but also to evaluate the performances of FS algorithms on the classification performance. The dataset was rearranged to include less number of columns (features) for an efficient classification. 7 features were used for CART and 13 features were used for SVM and ANN methods. Table 2 shows the features used in classifiers. The experiments were performed using Python programming language and its libraries Keras, Tensorflow, and sci-kit learn. Accuracy given in equation (1) was used as performance metric. TP in the equation stands for the number of PD patients classified correctly as having PD, TN means the number of healthy subjects classified correctly as healthy, FP stands for the number of healthy people misclassified as PD patient, and FN is the number of PD patients misclassified as healthy. Accuracy is the ratio of the number of correct assessments to the number of overall assessments [26]. The success of different classifiers before and after FS are shown in Table 4.

As seen from Table 4, SVM with FS showed the highest classification performance (93.84%) and pure SVM showed the lowest performance (79.98%). By these results it is proven that the determination of the most relevant features increases the classification performance in accordance with the literature. RFE method made SVM the most successful classifier for PD diagnosis while it was the worst before FS.

Table 2
Selected features based on classifiers.

Classifier	FS method	Determined features
CART	FI	Shimmer:APQ3, spread2, D2, Shimmer:APQ5, DFA, RPDE, PPE
SVM, ANN	RFE	Shimmer:APQ3, spread1, Shimmer:APQ5, RPDE, DFA, spread2, PPE, D2, Jitter:DDP MDVP:APQ, HNR, NHR, MDVP:Shimmer

Table 3
Optimization methods used.

No	Optimizer	No	Optimizer
1	Adam	5	Adagrad
2	Adamx	6	Adadelta
3	Adamax	7	rmsprop
4	Nadam	8	TfOptimizer

Table 4
Accuracy of classification before and after FS.

Classification Method	Accuracy before FS	Accuracy after FS
CART	85.23%	90.76%
SVM	79.98%	**93.84%**
ANN	80.25%	91.54%

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{1}$$

Table 5 shows a brief summary and comparison of the state-of-the-art. The number of features, the way that the features were obtained, the classification method, and the accuracies are shown in the table. Although high classification rates were obtained in the literature for ML based diagnosis of PD, either they used many features (like [9,14]) which increases computation time or the extraction of the features were hard even they use few features. Therefore, indirectly, the computation time is again high. In the proposed approach, computation time was decreased via less number of effective features, a lightweight feature extraction process and a classifier. The features were obtained from the speech signals and so to obtain these features is easier and less costly than the other MRI-based [15], motion-based ([7,15]) and hand-writing based ([18]) methods in the literature. Although some authors (such as [8,9,12]) used voice features for the diagnosis of PD, they used more number of features than the proposed approach. Besides using more features, a costly FS process used in [10]. The least number of features were used in [16] with an acceptable classification accuracy, but they used MRI data for feature extraction and it is harder than obtaining voice features. [14] used similar number of voice features but they determined the features by a heavier FS method. A version of artificial bee colony algorithm was utilized to reduce number of features. Also, their classification method has greater algorithmic complexity than the proposed approach.

Discussion

The present study developed an FS based decision support system by using the features extracted from speech signals of PD patients and healthy people. Different FS methods were applied to different classifiers and the one with the highest performance was determined.

FS has a strong effect on the performance of the classifier. It is shown in [20] that FS is a beneficial preprocessing tool and it not only reduces the number of inputs of classifiers, but also helps people understand the underlying causes of diseases. Three FS methods were evaluated in the present study and they showed different performance for different classifiers. The best combination of FS method and classification method was determined and used for the diagnosis of PD. Table 4 shows the effect of FS methods on the classification performance. Remarkable results were obtained using FS. It provided about 13% performance improvement for SVM, about 11% for ANN, and about 5% improvement for CART. In addition, classification performance also depends on the parameters of the classifiers. Determination of the optimal parameters for certain classification method is a key point. For SVM, c and gamma must be properly determined. In case of ANN based classification, the topology must be determined properly. That is the number of hidden layers, the number of neurons in hidden layers, the activation functions, the parameters like learning rate and momentum coefficient, normalization of data, epoch number, etc. must be investigated in detail.

PD diagnosis system proposed in this study differs from the literature in terms of FS method, a lightweight feature extraction process and a classifier. A high enough classification performance has been achieved. Using voice features in the diagnosis of PD helped very much. Obtaining voice features are both easier and cheaper when compared to MRI-based or motion based diagnosis methods. In this study, SVM with RFE gave the best classification accuracy. These findings suggest that using certain subset of voice features help researchers classify PD patients more accurately and less efforts can be made to extract features

Table 5
Comparison with the state-of-the-art.

Ref. No	Number of features	FS method	Source of features	Method	Accuracy (%)
[8]	28	Correlation	Motion-based	RF	95
[9]	18	–	Voice-based	k-NN	94.55
[10]	20	MLP	Voice-based	LSVM	100%
[11]	10	NIPALS	Voice-based	ISVR-SOM	0.4656 (MAE)
[12]	23	–	Voice-based	NN	92.9
[13]	23	FCMFW	Voice-based	k-NN	97.93
[14]	14	AABC	Voice-based	KWELM	98.97
[15]	100+	FDR	MRI-based	SVM	100
[16]	4	–	Gait and tremor-based	SVM	92.7
[17]	66	ReliefF	Voice-based	SVM	91.25
[18]	9	Wrapper	Hand-writing	NB	91.00
Proposed	13	RFE	Voice-based	SVM	93.84

Abbreviations in the table are: LSVM: Lagrangian Support Vector Machines, RF: Random Forests, k-NN: k Nearest Neighbor, AABC: Adaptive Artificial Bee Colony, OPF: Optimum-Path Forest, ISVR: Incremental Support Vector Regression, KWELM: Kernel-based Weighted Extreme Learning Machine, MAE: Mean Absolute Error, NIPALS: Non-linear iterative partial least squares, FCMFW: Fuzzy C-Means Feature Weighting.

from voice signals of candidate PD patients. Besides, the classification can be realized by less computational cost.

Conclusion

In the present study, a FS based decision support system was developed using the features of voice signals of both PD patients and the healthy people for the early diagnosis of Parkinson's. Different FS methods and different classification methods were used in the experiments. The primary objective in doing so was to improve the performance and the accuracy of the model and also to reduce the computational cost of classification task. Accuracies of the classification methods were evaluated with and without FS and the remarkable effect of FS was shown. The results indicate that using FS methods together with classification methods is quite advantageous especially when dealing with speech signals in which hundreds of phonetic features can be obtained. By the help of the developed early diagnosis system, PD can be diagnosed with a high accuracy rate in its early stages and the worse symptoms of the disease can be stopped.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Karabudak R. Parkinson's disease. 2014.
[2] Sakar CO, Kursun O. Telediagnosis of Parkinson's disease using measurements of dysphonia. J Med Syst 2010;34(4):591–9.
[3] Rayan Z, Alfonse M, Salem A-BM. Machine learning approaches in smart health. Procedia Comput Sci 2019;154:361–8. https://doi.org/10.1016/j.procs.2019.06.052.
[4] Wan KR, Maszczyk T, See AAQ, Dauwels J, King NKK. A review on microelectrode recording selection of features for machine learning in deep brain stimulation surgery for Parkinson's disease. Clin Neurophysiol 2019;130(1):145–54.
[5] Salmanpour MR, et al. Optimized machine learning methods for prediction of cognitive outcome in Parkinson's disease. Comput Biol Med 2019;111.
[6] Pedrosa TÍ, Vasconcelos FF, Medeiros L, Silva LD. Machine learning application to quantify the tremor level for Parkinson's disease patients. Procedia Comput Sci 2018;138:215–20.
[7] Prashanth R, Dutta Roy S. Novel and improved stage estimation in Parkinson's disease using clinical scales and machine learning. Neurocomputing 2018;305:78–103.
[8] Cavallo F, Moschetti A, Esposito D, Maremmani C, Rovini E. Upper limb motor preclinical assessment in Parkinson's disease using machine learning. Park Relat Disord 2019;63:111–6.
[9] Almeida JS, et al. Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. Pattern Recogn Lett 2019;125:55–62.
[10] Parisi L, RaviChandran N, Manaog ML. Feature-driven machine learning to improve early diagnosis of Parkinson's disease. Expert Syst Appl 2018;110:182–90.
[11] Nilashi M, Ibrahim O, Ahmadi H, Shahmoradi L, Farahmand M. A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. Biocybern Biomed Eng 2018;38(1):1–15.
[12] Das R. A comparison of multiple classification methods for diagnosis of Parkinson disease. Expert Syst Appl 2010;37(2):1568–72.
[13] Polat K. Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering. Int J Syst Sci 2012;43(4):597–609.
[14] Wang Y, Wang AN, Ai Q, Sun HJ. An adaptive kernel-based weighted extreme learning machine approach for effective detection of Parkinson's disease. Biomed Signal Process Control 2017;38:400–10.
[15] Singh G, Vadera M, Samavedham L, Lim ECH. Machine learning-based framework for multi-class diagnosis of neurodegenerative diseases: a study on Parkinson's Disease. IFAC-Papers OnLine 2016;49(7):990–5.
[16] Abdulhay E, Arunkumar N, Narasimhan K, Vellaiappan E, Venkatraman V. Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease. Fut Gener Comput Syst 2018;83:366–73.
[17] Yaman O, Ertam F, Tuncer T. Automated Parkinson's disease recognition based on statistical pooling method using acoustic features. Med Hypotheses 2019:109483.
[18] Kotsavasiloglou C, Kostikis N, Hristu-Varsakelis D, Arnaoutoglou M. Machine learning-based classification of simple drawing movements in Parkinson's disease. Biomed Signal Process Control 2017;31:174–80.
[19] UCI machine learning repository: Parkinsons data set. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/parkinsons. [Accessed: 10-Dec-2019].
[20] Remeseiro B, Bolon-Canedo V. A review of feature selection methods in medical applications. Comput Biol Med 2019;112.
[21] Jain D, Singh V. Feature selection and classification systems for chronic disease prediction: a review. Egypt Inform J 2018;19(3):179–89.
[22] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines.
[23] Yohannes Y, Webb P. Classification and Regression Trees, CART: a user manual for identifying indicators of vulnerability to famine and chronic food insecurity Pap/Dsk. Intl Food Policy Research Inst; 1999.
[24] Hsu C-W, Chang C-C, Lin C-J. A practical guide to support vector classification. 2003.
[25] Öztemel E. Yapay Sinir Ağları. 3rd ed. İstanbul: Papatya Yayıncılık; 2012.
[26] Zhu W, Zeng N, Wang N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS ® implementations. 2010.