ACO Feature selection and Novel Black Widow meta-heuristic Learning rate optimized CNN for Early diagnosis of Parkinson's disease

S. Kiruthika[a], Dr. EVRM. Kalaimani[b], Mrs. R.Sudha[c], and Mrs.K.Suganthi[d]

[a]
Assistant Professor, K. Ramakrishnan College of Technology, Trichy
[b]Professor and Head, Arasu Engineering College, Kumbakonam
[c,d]Assistant Professor, Arasu Engineering College, Kumbakonam

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 16 April 2021

Abstract: Parkinson's disease (PD) is a brain cell disease that causes brain cells to interact with a substance called dopamine. The dopamine-producing cells in the brain are answerable for regulating, regulating, and facilitating movement. When 60-80% of these cells are misplaced, dopamine is not formed and Parkinson's motor indicationsseem. The disease is thought to have started years before the onset of motor symptoms. Therefore, investigators are looking for ways to identify non-motorized symptoms that appear at the onset of the disease and avert the development of the disease. In this paper, ACO Feature selection (FS) and Novel Black Widow meta-heuristic Learning rate optimizedConvolutional Neural Network for Early diagnosis of PD is introduced. The proposed diagnostic method includes methods of selection and classification of characteristics. When selecting the characteristics, the meaning of the characteristic and the methods of optimization (OEC) of the ant colony were taken into consideration. The proposed learning rate optimization gives CNN better performance than other methods. Achieved 93.84% accuracy in the diagnosis of PD with minimal noise function.

Keywords: Parkinson's disease, Convolutional Neural Network, ant colony optimization, Novel Black Widow and Learning rate optimization.

1. Introduction

Speaking is an important symptom of Parkinson's disease, therefore, recording and automated analysis of speech signals is the calmest and maximum reasonable way to detect the disease early. [1] Investigators are focused on using this technique to learn more about the nature of the disease and how to examine its signs using data mining procedures. In 2011, he created the Ford Oxford PD Discovery dataset from the Irwin Data Mining (UCI) collection at the University of California and compared some secret collections. [2]. they found that the random tree algorithm classifies the dataset with accuracy and is 100% accurate. R. ArfiShirvan and E. Hami showed the FS using the genetic algorithm and used the K-NN algorithm for the classification [3] thetop accuracy obtained with the top 9 features was 98.2%. In 201 InFS using genetic procedures was applied to the Parkinson's dataset and [4] was used for SVM classification. The uppermost accuracy attained with the top 4 functions was 94.5%. In 2014, Harinish, Gracie Annamari [5] associated the presentation of random plants, M5 rules and ANN algorithms. Demonstrates the highest accuracy of the ANN algorithm. In 201 In, all irrelevant parameters and outlets in the OPD dataset were removed during the study. Comparing the performance of K-NN, Random Forest and Ada-Boost, K-NN found the highest accuracy to be 90.26%.According to the study, the features were selected in 2015 and the datasets were classified using SVM [6]. The maximum accuracy obtained was 100%. In a second study in 2012, two additional target classes were added to Parkinson's dataset [7]. The main features were then screened using main component analysis (PCA) and comparatively different classification rates. Automatic minimum SVM optimization provides maximum ura %%% accuracy. In a second study in 2012, Parkinson's datasets were first generalized, followed by analysis of various methods for selecting and classifying features [8]. It was found that 98.97% accuracy was obtained using the wrapper selection method and K-NN classification. P. Kriparapun and s. United additional NN and more effectual classification of OPD datasets for generalization, amorphous [9].

P. SecundoDurga, V. Soot Jebkumari and D. Mir, who classified the OPD dataset, showed that MLP has the highest 97.78% purity compared to 48 and NB algorithms [10]. In 201, this study analyzed various methods in Parkinson's data set and then used SVM for classification [11]. The technique of selecting recursive feature exclusion operations using SVM was found to have the highest accuracy with a record of 96.29%. A second study from 201 analyzed the effectiveness of various classification and data segmentation methods in Parkinson's data set [12]. In a 201-ounce study, Oz Zak compared the selection, classification, and segmentation methods for different features in Parkinson's Data [13]. It was found that 99.1% accuracy was achieved when FS was done using PCA, ANN classifier and other 10-fold cross-validity study, Extreme Learning Machine (ELM) comparison. To evaluate on PV [14] was carried out with Neural Network (NN) and support vector machine. It has been found that training and testing of data sets is more accurate than other ELM classifiers. In 2017, a study by Istanbul University used the OPD data set and Parkinson's Speech Data Set (PSD) [15]. Applies to many records and receives a 94% - 98% [16]. The authors associated the presentation of different classifiers on two sets of data. DNN compared other classified mining methods with an accuracy of .068.0% for PSD datasets and 79.79 with %% for OPD datasets. A second study of 201 used a kernel-based EML (KEML) approach with a 5x cross-validation and adaptive artificial colony Bee algorithm (ABC) to diagnose partial discharge with an accuracy of 0.9.9. %%. [17].

2. Literature survey

Ipsita Bhattacharya et.al [18] used a Weka data mining tool, they passed the working dataset and then used one of the classification methods. SVM range from healthy people to people with PD. LibsVM implementation are trying to discovery the best accuracy on different kernel values for a given data set. You downloaded it from the data set [10], which recorded 197 speech samples from 31 people, 23 of whom had PD. You will get better results if you change the split ratio and repeat the test. By randomly dividing the dataset, they concluded that the best accuracy was 65.74%.

Dr. R. GeethaRamani et.al [19] A proposal was made to classify patients with PD and non-PD using the following approaches: linear differential analysis (LDA), RND tree, and SVM. The dataset contains Parkinson's data from the UCI collection. The training data set contained 197 samples from 22 characteristics of patients. The Fisher filter function selection algorithm has proven to be an efficient system for classifying features. The edge tree algorithm achieved a classification accuracy of 100%, while LDA, C4.5, CS-Mac4 and K-NN achieved results with an accuracy of over 90%. The C-plus algorithm achieved the lowest accuracy of 69.74%.

Resul Das et.al [20] Different classification methods were used to identify the partial discharge. Four different methods of class were applied and analyzed, e.g. Satanic Network, Neural Network, Regression, Decision Matching. Various assessmentsystems were used to evaluate the performance of the classification. After evaluating the results, the neural network classifier gave the best results. Adjustable parameters of each security are set. For the neural network classifier, the BPNN algorithm was used on the direct neural network with hidden layers. The algorithm used in the study is the Lewenberg-Marquardt (LM) algorithm. The accuracy of the neural network is 92.9%.

Zahari Abu Bakar et.al [21] The analysis is based on two algorithms. These are neural networks from Lewenbergmarkart (LM) and SCG Multi Layer Perceptron (MLP) in PDI to detect partial discharges. The data in this project is generated from the PD data record. The LM algorithm was found to have achieved better training accuracy and test accuracy with 25 hidden modules compared to other hidden modules, but 97.86 percent during the learning phase and 92.96 percent during the experimental phase. LM. Did a good job with a classification accuracy of 92.95% and SKG achieved an accuracy of 78.21%.

AthanasiosTsanas et.al [22], four partial discharge algorithms were used to selected the feature. And also calculated 132 measurements of dysphony for static tones. Then, using the FS algorithm, four subsets of these steps of dysphonia are selected. These subsets are mapped to two statistical classifications: random forests and binary classification response using SVM. The NCVS database has 263 phonemes on 43 subjects, an extension of the database used on paper [2]. It offers a total classification accuracy of 99% with only ten dysphonia features.

Kiruthika et.al [23], a proposal was made to classify and detect cancer using deep learning techniques like Convolutional neural network which gives the clear clarification about the performance of detection.

Asma Begum et.al[25], proposed a technique to detect heart disease prediction rate using AdaBoost ensemble based machine learning classifier algorithm and achieved the accuracy of about 98%.

3. Proposed system

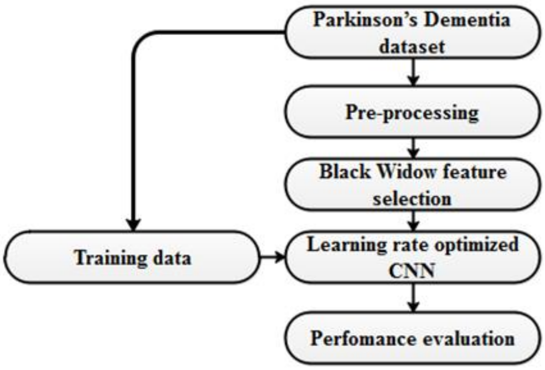

Fig.1. Proposed flow diagram.

In this paper, for early diagnosis of PD a novel approach is introduced which is defined in the fig.1. The proposed method consist of three major blocks such as pre-processing, feature selection and classification. In the first stage the data set data's are readied and the null values are removed in the preprocessing stage. To improve the classification accuracy in this paper a ACO FS technique is used and finally, to improve the CNN performance in this work Learning rate has been optimized with the help of Novel Black Widow meta-heuristic the brief discussion of the proposed system is defined in the below section.

3.1. Data Source

The dataset used in this research experiment includes features obtained from speech signals from 31 people at the Voice Speech Center in Max Little of Oxford University created the dataset and presented it at the UCI Machine Learning Repository. Of the 31, 23 are PDs and 8 are control groups. The dataset contains 195 biomedical voice measurements. Table 1 displays the voice solutions used in the experimentations. The status in the filedescribes the class of the column and gets 0 for health, 1 for PD. The distribution of the classes of the dataset is exposed in Figure 2. There are 48 healthy phonetic and 147 phonological PDs for 48 people.

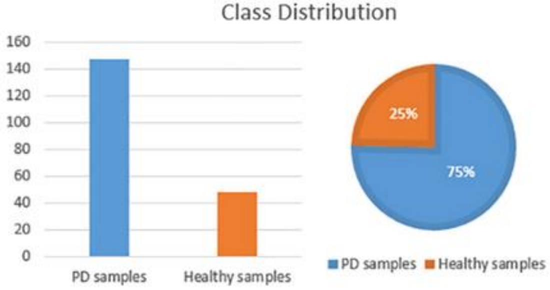

Fig.2.Recorded phonetics Class distribution in the dataset.

Table. 1. The voice actions used in the experimentations [19].

Feature no	Voice measure	MEANING
1	MDVP:Fo (Hz)	Average frequency
2	MDVP:Fhi (Hz)	Maximum frequency
3	MDVP:Flo (Hz)	Minimum frequency
4	MDVP:Jitter (%)	Numerous measures of difference in
5	MDVP:Jitter (Abs)	fundamental frequency
6	MDVP:RAP	
7	MDVP:PPQ	
8	Jitter:DDP	
9	MDVP:Shimmer	Numerous measures of variation in amplitude
10	MDVP:Shimmer (dB)	
11	Shimmer:APQ3	
12	Shimmer:APQ5	
13	MDVP:APQ	
14	Shimmer:DDA	
15	NHR	Two ratio measures of noise to tonal
16	HNR	components in the voice
17	RPDE	Two nonlinear dynamical difficulty
18	D2	measures

3.2. Feature selection by using ACO

Function selection activities can be converted into an ACO-eligible task. ACO must present the problem graphically. Here the nodes represent existence and the edges between them indicate the next element to be selected. The search for a subset of optimal properties consists of the intersection of "ants" in the diagram visiting the minimum nodes that meet the cross-stop criteria. The ant finishes its journey and creates a subset of these functions as candidates for data processing. A suitable genetic trait for performance analysis can be any subgroup scoring function, such as: B. Entropy measurements, fixed dependency measures or Fisher's Discrimination Rate (FDR). The heuristic benefits of lateral and borderline pheromone levels together form what is known as the potential transition rule, which assumes the likelihood that an ant on the property will decide to enter the area at a certain point in time.:

$$P_{IJ}^{K} = \begin{cases} \frac{[\tau_{IJ}(t)]^{\alpha}.[\eta_{ij}]^{\beta}}{\sum_{\iota \epsilon J_I^K}[\tau_{IJ}(t)]^{\alpha}.[\eta_{ij}]^{\beta}}, \ if\ j\epsilon J_I^K \\ 0, \ \ otherwise \end{cases} \qquad (1)$$

Where the number of ant is signify by k and choosing feature j heuristic desirabilityη_ij and τ_(IJ(t)) is the sum of virtual pheromone on edge (i,j). The process begins by generating a sum of ants, k, which are then randomly placed on the graph (that is, each ant begins with a random characteristic). Instead, the sum of ants to be placed on the graph can be set equal to the sum of functions within the data; each ant begins the construction of the path with a different characteristic. From these starting positions, they will likely cross the edges until a cross stop criteria is met. The resulting subsets are collected and then appraised. If an optimal subset was found or the algorithm was run multiple times, the process stops and generates the best subset of functionality found.

$$\boldsymbol{\tau_{IJ}}(t+1) = (1-\rho).\tau_{ij}(t) + \rho.\Delta\tau_{ij}(t)(2)$$

Where

$$\Delta\tau_{ij}(t) = \sum_{k=1}^{n} {(\gamma(S^k)} \big/ {(|S^K|)}$$

Novel Black Widow meta-heuristic Optimization

Like any other algorithm, the proposed algorithm can develop with the initial population of spiders, so each spider is a solution. This first pair of spiders is trying to recreate a novel generation. The black widow eats the man during or after mat. He then replaces the semen kept in his sperm with egg yolk. After 11 days the spiders lay eggs. They get pregnant numerous days a week. During this time the brothers become cannibals. Then they are carried through the air.

Initial population

To resolve an optimization problematic, the values of the difficult must be formulated as an appropriate construction to solve the present problem. In black widow meta-heuristic optimization terminology, this structure is referred to as "chromosome" or "particle position", but here in the BWO algorithm, "widow". In the BWO algorithm, the Black Widow Spider is seen as a possible solution to any problem. Every one black widow spider represents the value of a variable problem. In this article, this structure should be seen as an array for performing test functions.

In a N_var–dimensional optimization problem, a widow is an array of 1 × N_varr signifying the solution of the problem. This array is distinct as shadows:

Widow= [x_1 ,x_2 ,…,x_(N_var)],

Every variable values (x_1 ,x_2 ,…,x_(N_var)) is floating-point numeral. The fitness of widow is attained by evaluation of fitness function f at a widow of (x_1 ,x_2 ,…,x_(N_var)). So Fitness = (widow) =(x_1 ,x_2 ,…,x_(N_var)),

To run the optimization algorithm, a entrant widow matrix of size N_pop× N_var is generated with an initial spider population. Parent pairs are then randomly selected to conduct the mating stage.

Procreate

As the sets of each other, they begin to mate to breed a fresh generation. At the same time, each pair mates separately from the others in their network. In the real world, about 1,000 eggs are produced each time they mate, but in the end some of the stronger spider chicks survived. Well, here, in this algorithm, a matrix called Alpha must be created to reproduce when the widow's matrix contains random numbers. Then, the offspring are created using α with the following equation (Equation 1), where x1 and x2 are parents, y1 and y2 are children..

$$\begin{cases} y_1 = a \times x_1 + (1-a) \times x_2 \\ y_2 = a \times x_2 + (\grave{}1-a) \times x_1 \end{cases} \qquad (3)$$

This process is repetitive Nvar / 2 times and there is no need to duplicate randomly selected numbers. In the end, children and mothers are additional to the ranks and arranged by fitness scores. Based on the cannibal rating, some of the best people have now been added to the afresh formed population. These measures relate to all couples.

c) Cannibalism

Here we have three types of male predators. The initial is sexual racism, in which a black widow eats her partner during or after intercourse. With these algorithms, we were able to identify women and men based on their fitness values. The second type is related to cannibals, in which strong spiders eat their pathetic allies. In these algorithms we set the Cannibalism Score (CR), which determines the number of survivors. In particular circumstances, a third species of ogre is often found, with small spiders eating their mother. We use fitness values to identify weak or strong spiders.

d) Mutation

At this point, we arbitrarily select Mute Pop individuals from the population. As shown in Figure 2, each of the selected solutions arbitrarily changes two basics in the array. Intended the mute pop based on the mutation rate.

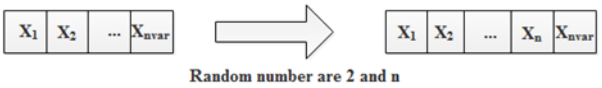

Fig. 2. Mutation.

e) Convergence

Similar to other algorithms, three closing state of affairs can be deliberated: (a) a predetermined sum of repetitions. (B) The best widow compliance rating for most delegates does not change. (C) Specified accuracy. The pseudocode shown in the figure summarizes the main phases of BWO. The next section discusses some of the issues with optimizing the BWO test using. Since the best solutions for testing activities are already known, the availability of a certain level of accuracy is taken into account to determine the level of accuracy of the experimental algorithm. Also, Section 4 sets the maximum repetitions in the experiments as stop conditions.

f) Parameter setting

The projected scheme has some parameters that are required to get the best results. These factors comprise purchase rate (PP), cannibalism rate (CR) and conversion rate (PM). The parameters should be adjusted accordingly so that the algorithm can find the best solutions. Better handling of a number of parameters, higher navigation capability on any local um platform and higher ability to locate globally. Therefore, the exact number of parameters can allow to control the balance between the absorption phase and the inspection phase. The proposed algorithm has three important control factors, including PP, CR and PM. PP is the ratio of ownership that regulates how many people are involved in the product. By regulatory the production of different offspring, this parameter delivers more variety and more possibilities for a clear definition of research location. The control parameter of the CR cannibal operator is to exclude unqualified people from the population. The Prime Minister is the percentage of people who have changed. The fair value of these parameters can strike a balance between use and search. This factor can control the migration of agents in the local phase from global and lead them to better resolution.

Table 2. Pseudo code for BWO algorithm.

```
Input: Maximum nu
mber of iteration, rate of procreating, rate of cannibalism, rate of mutation
Output: near-optimal solution for the objective function
// Initialization
    1. The initial population of BWS
       Each pop is a D-dimensional array of chromosomes for a D-dimensional problem
// Loop until the terminal condition
    1. Based on procreating rate, calculation the number of reproduction "nr"
    2. Select the best nr solution in pop and save them in pop1:
// procreating and connbalism
    3. For 1=l to nr do
    4. Randomly select two solution as parents from pop1;
    5. Generate D children use equation 1:
    6. Destroy father;
    7. Based on the cannibalism rate, destroy some for the children (new achieved solution)
    8. Save the remain solution into pop 2
    9. End for
// Mutation
```

```
    10. Based on the mutation rate, calculate the number of mutation Children “nm”
    11. For i=1 to nm do
    12. Select a solution from pop 1;
    13. Mutate randomly one chromosome fo the solution and generate a new solution;
    14. Save the new one into pap3;
    15. End for
// Updating
    16. Update Pop =po2+pop3;
    17. Returning the best solution;
    18. Return the best solution from pop;
```

3.4. Convolutional Neural Network (CNN)

CNN is based on a backpropagation algorithm for training information [21] and uses vector X of trained samples as input for the corresponding target class y. Learning is achieved by comparing the desired goal relative to the output of each CNN, and the difference between them gives a learning error. Assuming the role of the next CNN tuition fee mathematically,

$$E(\omega) = \frac{1}{2}\sum_{\rho=1}^{p}\sum_{j=1}^{N_l}(o_{j,p}^{l} - y_{j,p})^2 \quad (4)$$

Our goal is the minimalizing of cost function E(ω), finding a minimizer ω̃=ω̃^1,ω̃^2,…,ω̃^v ϵ R^v,where $\mathrm{v} = \sum_{\mathrm{k}=1}^{\mathrm{L}} \mathrm{WeightNum\ (k)}$ and signify that the space of weight R^vis equal to the number of weights (WeigtNum(.)) at each k layer of total L layers of the CNN network.

$$\nabla \mathrm{E_i}(\omega_\mathrm{i}) = \left(\frac{\partial \mathrm{E_i}}{\partial \omega_\mathrm{i}^1}, \ldots, \frac{\partial \mathrm{E_i}}{\partial \omega_\mathrm{i}^\mathrm{v}}\right) \quad (5)$$

$$\omega_{i+1} = \omega_i - n\nabla E_i(\omega_i)(6)$$

Where nis the learning rate (step) value.The n has been selected with the help of BWS optimization techniques.

4. Results and Discussion

To evaluate the various performance metrics such as accuracy, class error, specificity, and sensitivity were used to evaluate the performance of the classifier. The time required to implement the model was also measured. All calculations were done in Python using an Intel (R) Core ™ i5-2400CPU at 3.10 GHz. The main results of the proposed research are the following:

4.1. Performance Evaluation Metrics

In this study, various presentation metrics were used to test the effectiveness of classification. We used a random matrix, evaluating each observation in the test sample, even in one field. As a repo class of 2, it is a 2 × 2 matrix. In addition, it offers two types of accurate predictions and two types of false predictions. Table 3 presents the confusion matrix.

Table 3: Confusion matrix.

	Predicted HD patient (1)	Predicted healthy person (0)
Actual HD patient (1)	TP	FN
Actual healthy person (0)	FP	TN

From confusion matrix, we calculate the subsequent terms as:

TP: predicted output as true positive (TP)

TN: predicted output as true negative (TN)

FP: predicted output as false positive (FP)

FN: predicted output as false negative (FN)

Classification accuracy: accuracy indicationsthat the complete performance of the classification system as follows

$$classification\ Accracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100(7)$$

Classification error: it is the overall improper classification of the classification ideal which is considered as follows:

$$\text{Classification error} = \frac{FP+FN}{TP+TN+FP+FN} \times 100\ (8)$$

Sensitivity: This is the proportion of newly classified patients with heart disease out of the total number of patients with heart disease.

$$\text{Sensitivity(Sn) or recall or true positive rate} = \frac{TP}{TP+FN} \times 100(9)$$

Specificity:aanalytical test is negative and the being is healthy and is accurately written as follows:

$$\text{specificity(Sp)} = \frac{TN}{TN+FP} \times 100(10)$$

***Precision*:** the equation of precision is given as follows:

$$Precision = \frac{TP}{TP+FP} \times 1(11)$$

4.2 Result analysis

In this section, the proposed system has compared in two different ways such as without hybrid embedded feature and with feature selection. In this both the cases we have considered with the Convolutional Neural Network. The results are mentioned in the table.4.

Table 4:*Result analysis.*

Method	**Sensiti**	**specifici**	***Precisio***	**Classific**
Without feature selection	93.34	83.34	*15.18*	3.47
With feature selection	94.47	84.47	*12.64*	1.83

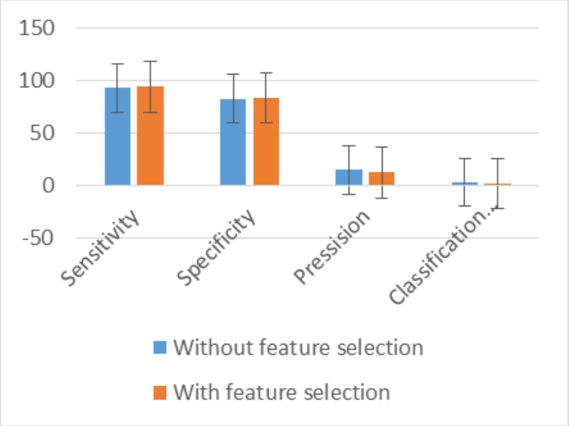

Fig.3. Result analysis.

The tabulation 3. And fig.3 shows that the proposed feature selection and CNN for Machine Learning archive better results in terms of Sensitivity, Specificity precision and Classification error. In that particularly 94.47% of Sensitivity and 1.83% error rate has archived. Compared to without proposed feature selection classification error is less in our proposed system.

4.3. Accuracy Analysis

In this section, the proposed system has compared in two different ways such as without hybrid embedded feature and with feature selection. In terms of accuracy. The measurements are defined in the table 5 and figure 4.

Table 5: Accuracy analysis

Method	**Accuracy**
Without feature selection	**96.53**
With feature selection	**98.17**

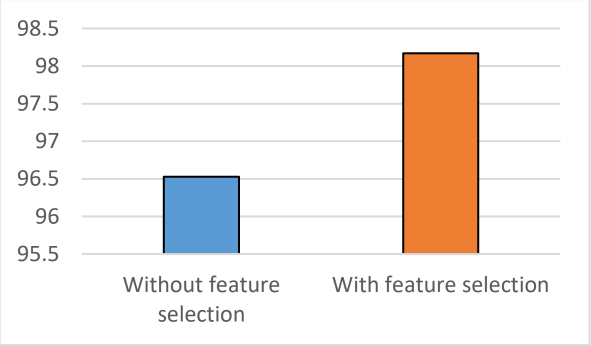

Fig.4. Accuracy analysis.

In tabulation 4 and figure 4 shows that the comparison analysis of accuracy by using two techniques such as without feature selection and with feature selection. Technique with feature selection achieved the accuracy of 96.53%, but classification algorithm with feature selection technique attained the better accuracy of 98.17%. Compared by two technique, with feature selection method achieved better accuracy than without feature selection technique.

5. Conclusion

According to the current study, FS-based decision support systems have been developed that can detect Parkinson's disease as soon as possible using features of speech signals from PD patients and healthy people. Different FS approaches and different classification approaches were used in the trials. The main impartial was to improve the efficiency and accuracy of the model, as well as to reduce the computation cost of the classification problem. The accuracy of the classification methods was evaluated with and without FS, showing a significant effect of FS. Results show that FS methods can be used in combination with classification methods, particularly speech signals that can provide hundreds of phonetic features. With the help of advanced early detection mechanisms, Parkinson's disease can be diagnosed as early as possible and poor symptoms can be avoided.

Reference

1. “Parkinson’s Foundation: Better Lives. Together.”.
2. R. G. Ramani, “Parkinson Disease Classification using Data Mining Algorithms,” vol. 32, no. 9, pp. 17–22, 2011.
3. R. A. Shirvan and E. Tahami, “Voice Analysis for Detecting Parkinson’s Disease Using Genetic Algorithm and KNN Classification Method,” no. December, pp. 14–16, 2011.
4. M. Shahbakhi, D. T. Far, E. Tahami, and M. Shahbakhi, “Speech Analysis for Diagnosis of Parkinson’s Disease Using Genetic Algorithm and Support Vector Machine,” J. Biomed. Sci. Eng., vol. 7, no. 7, pp. 147–156, 2014.
5. S. Hariganesh and G. Annamary, “A Survey of Parkinson’s Disease Using Data Mining Algorithms,” Int. J. Comput. Sci. Inf. Technol., vol. 5, 2014.
6. Saloni, R. K. Sharma, and A. K. Gupta, “Detection of Parkinson Disease Using Clinical Voice Data Mining,” vol. 9, pp. 320–326, 2015.
7. Caesarendra W, Putri FT, Ariyanto M, Setiawan JD. Pattern recognition methods for multi stage classification of parkinson's disease utilizing voice features. In2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM) 2015 Jul 7 (pp. 802-807). IEEE.
8. M. S. Wibawa, H. A. Nugroho, N. A. Setiawan, J. Grafika, and Y. Indonesia, “Performance Evaluation of Combined Feature Selection and Classification Methods in Diagnosing Parkinson Disease Based on Voice Feature,” Int. Conf. Sci. Inf. Technol., pp. 126– 131, 2015.
9. Kraipeerapun P, Amornsamankul S. Using stacked generalization and complementary neural networks to predict Parkinson's disease. In2015 11th International Conference on Natural Computation (ICNC) 2015 Aug 15 (pp. 1290-1294). IEEE.

10. P. Durga, V. S. Jebakumari, and D. Shanthi, “Diagnosis and Classification of Parkinsons Disease Using Data Mining Techniques,” ISSNOnline) Int. J. Adv. Res. Trends Eng. Technol., vol. 3, no. 14, pp. 2394–3777, 2016.
11. H. Ma, T. Tan, and H. Z. and T. Gao, “Support Vector Machine-Recursive Feature Elimination for the Diagnosis of Parkinson Disease based on Speech Analysis,” IEEE, 2016.

Mohamed GS. Parkinson ’s disease Diagnosis: Detecting the Effect of Attributes Selection and Discretization of Parkinson’s Disease Dataset on the Performance of Classifier Algorithms. Open Access Library Journal. 2016 Nov 2;3(11):1-1.

H. Ozkan, “A comparison of classification methods for telediagnosis of Parkinson’s disease,” 2016.

A. Agarwal, S. Chandrayan, S. S. Sahu, and C. Engineering, “Signal with Extreme Learning Machine,” Int. Conf. Electr. Electron. Optim. Tech.,

Y. N. Zhang, “Can a Smartphone Diagnose Parkinson Disease? A Deep Neural Network Method and Telediagnosis System Implementation,” vol. 2017, no. 1, 2017.

ÇALIŞKAN A, BADEM H, BAŞTÜRK A, YÜKSEL M. DIAGNOSIS OF THE PARKINSON DISEASE BY USING DEEP NEURAL NETWORK CLASSIFIER.

Y. Wang, A. N. Wang, Q. Ai, and H. J. Sun, “An adaptive kernel-based weighted extreme learning machine approach for effective detection of Parkinson’s disease,” Biomed. Signal Process. Control, vol. 38, pp. 400–410, 2017.

Ipsita Bhattacharya, M. P. S. Bhatia, “SVM Classification To Distinguish Parkinson Disease Patients”, Proceedings Of The 1st Amrita Acm-W Celebration On Women In Computing In India, 2010.

Dr. Geetha R Ramani, G Sivagami,”Parkinson Disease Classification Using Data Mining Algorithms”, International Journal Of Computer Applications 32(9):17-22, October 2011.

Resuldas, “A Comparison Of Multiple Classification Methods For Diagnosis Of Parkinson Disease”, Expert Systems With Applications, Volume 37, Issue 2, March 2010, Pages 1568-1572.

Zahari Abu Bakar, NooritawatiMd Tahir, Ihsan M Yassin, "Classification Of Parkinson's Disease Based On Multilayer Perceptrons Neural Network', Ieee Colloquium In Signal Processing And Its Applications (Cspa), 2010.

A. Tsanas, M.A. Little, P.E. Mcsharry, J. Spielman, L.O. Ramig, "Novel Speech Signal Processing Algorithms For High-Accuracy Classification Of Parkinson's Disease", Ieee Transactions On Biomedical Engineering, Vol. 59, No. 5, Pp. 1264-1271, 2012.

S.Kiruthika, S.Rahmath Nisha, “Automated Oral Cancer Detection and Classification using very Deep Convolutional Neural Network Algorithm”, Test Management and Engineering Volume 83 Page Number: 20019 – 20027 Publication Issue: March - April 2020.

C Shyamala, S Rahmath Nisha, K ShyamPrasad, S Sankar, G Siddharth, “Seeking information using online and offline q&a system”,

M.Asma Begum, S. Abirami, R. Anandhi, K. Dhivyadharshini and R.Ganga Devi, “Prediction of Heart Disease Using Machine Learning”, Bioscience Biotechnology Research Communications Volume 13 No (4) 2020