

A Modified Brain Storm Optimization Based Feature Selection For Parkinson Disease Classification

^[1] Mrs. B.Sathyabama, ^[2] Dr. N. Revathy

^[1] Research Scholar, Department of Computer Science,
Hindusthan College of Arts & Science,
Coimbatore, Tamilnadu.

^[2] Professor, PG and Research Department of Computer Applications,
Hindusthan College of Arts & Science,
Coimbatore, Tamilnadu.

Abstract: Parkinson's disease (PD) is an advancing neurological condition chiefly resulting in sequential motor problems. Therapy necessitates the robust and describable diagnosis pertaining to the severity level for PD. However, very less data is available with respect to serious PD patients, however there is massive amount of data available for moderately critical PD patients, and due to this unbalanced distribution, the diagnosis accuracy is reduced. During the early phases of the disease, PD patients primarily experience vocal impairments. Therefore, diagnosis systems that are founded on vocal conditions lead the current studies on PD detection. This study developed a feature selection approach based on modified brain storm optimization and a classification model that uses a discrete wavelet transform (DWT) for signal modification. Features such as wavelet Shannon entropies, energies, zero-crossing rates (ZCR), Mel frequency cepstral coefficients (MFCC), and linear predictive coding (LPC) are derived. The Modified Brainstorm Optimisation, which is used to reduce the features, is then used to pick the features. Lastly, the PD data is classified using a classifier based on Support Vector Machines (SVM). The outcomes of simulation shows that the proposed technique provides the optimal accuracy results so that treatment and therapy of PD patients are empowered.

Keywords: Parkinson's disease (PD), Mel frequency cepstral coefficient MFCC, wavelet Shannon entropy, Modified brain storm optimization, Support Vector Machine (SVM).

1. Introduction

Based on the fact sheets that the World Health Organization (WHO) has published, chronic diseases form the primary reason for the disability and mortality worldwide [1]. PD are chronic neurodegenerative conditions of nervous systems where motor functionalities are primarily affected. It is categorized to be a movement disorder, exhibits characteristics of disability of voluntary movement (akinesia), reduced and slower mobility (bradykinesia), raised muscle tonus (rigidity), and shaky movement during the resting position (Parkinson's tremor) [2]. Few other features are limited facial expression, issues with balance and definitive variations in speech and voice. People suffering with PD can also have loss in sense of smell (anosmia) and experience sleep problems during the rapid eye movement sleep (REMs) stage. It has been estimated that around 1% of those over 60 suffer with PD. Dopamine, a neurotransmitter that is also a hormone, is essentially a substance that is produced by brain cells to carry signals that control muscle actions. This hormone degenerates due to PD [3]. Hand tremors and thick, sluggish, incomprehensible speech are among the signs of PD. These symptoms can progress to uncontrollable tremors throughout the body, as well as altered memory and cognition. Voice problems impact about 90% of PD patients..

PD is generally related to the increased usage of vowel during speeches, redundancies in syllable, pauses between sentences (separated sentences) and conversations [4]. The reduction in the intensity during speech can be noticed via redundant syllables. PD constitutes the second highly prevalent neurological condition, and all over the world, 6.3 million people are affected with it. People who are affected with PD stay alive for quite long, however, those affected with PD can go on living, though there is a reduction in quality of the life led by them [5]. Lately, the tracking of postural gait and swaying in the throat during speech is done and therefore, the identification of anomalies is done while recording the speech of patients. This second approach may be found to

be extensible. Owing to the extended lifespan of the patients who are diagnosed early, improved accuracy and robust health informatics systems are necessary for identifying the PD patients [6-8]. These systems are also useful in decreasing the difficulties of clinicians.

Information may be found and extracted from medical databases using data mining techniques. When utilising data and analytics in the healthcare business, data mining greatly aids health systems in identifying discrepancies, resource waste, and even completely different best practises that contribute to healthcare improvement and cost reduction[9]. The central nervous system is impacted by PD, which makes mobility challenging. The neuro-pathologic and histo-pathologic criteria can be followed to validate the clinical diagnosis of PD [10]. Depending on the sensitivity and specificity of the defining clinical criteria, a thorough study of the literature data and selection process may be used to construct a clinical diagnostic classification of PD. Examining the clinical, pathologic, and nosologic research according to the frequency of occurrence, characteristics, and risk factors in patients is required in accordance with clinic-pathologic analysis in the distinctive community of individuals affected with PD [11]. Regression, Decision Trees, Neural Networks, and DMeural were previously used to determine the classifiers' performance score, allowing for the accurate diagnosis of PD. Vocal impairment brought on by PD affects speech, motor abilities, behaviour, mood, sensory, and thinking processes. Tele monitoring of the disease with the help of voice measurement plays a significant part to make an early PD diagnosis. This study presents a feature selection and classification model based on modified brain storm optimisation, which employs a DWT for signal transformation. In order to construct a classification for a predictive model and evaluate the significance and statistical importance of the PD relations with regard to attributes, classic bootstrapping or leave-one-out validation procedures using SVM are utilised.

The remaining research papers are arranged as follows: in Section 2, the most recent methods for diagnosing PD are examined. The procedure for the suggested approach is covered in section 3. The findings and discussion are given in section 4. The conclusion and suggestions for further study are given in section 5.

2. Literature Review

In this section reviews few recent approaches for PD detection employing artificial intelligence techniques.

Khemphila et al [12] created a Multi-Layer Perceptron (MLP) using the Back-Propagation learning method for the purpose of classification in order to effectively diagnose PD. For PD diagnosis, a feature selection algorithm in conjunction with the results of biological tests is taken into account. This article uses a small set of characteristics to classify the incidence of PD. The categorization process makes use of original 22,22 characteristics. In this case, the qualities that assisted in lowering the quantity of attributes that must be gathered from patients are determined using information gain. Artificial neural networks (ANNs) are useful for patient diagnosis and categorization. There are sixteen qualities instead of twenty-two. The accuracy attained was 83.333% in the validation data set and 82.051% in the training data set. A novel Multiple Feature Evaluation Approach (MFEA) of a multi-agent system was proposed by Mostafa et al. [13]; (ii) five different classification schemes, namely Decision Tree, Naïve Bayes, Neural Network, Random Forests, and SVM, are implemented on the diagnosis of PD before and after using the MFEA; and (iii) the diagnosis accuracy in the outcomes is evaluated. The classifiers' average rate of improvement in diagnostic accuracy is 10.51% for Decision Trees, 15.22% for Naïve Bayes, 9.19% for Neural Networks, 12.75% for Random Forests, and 9.13% for SVM. These findings show that the MFEA significantly enhances the classifiers' diagnostic outcomes. Wavelet analysis was introduced by Joshi et al. [14] as an alternative approach, and it was discovered that integrating wavelet analysis with SVM might improve classification accuracy. In order to obtain computationally straightforward information, wavelet modification is applied, and SVM is then used to identify Parkinson's gait. To determine which gait characteristic is optimal for this type of categorization, a variety of gait parameters, including stride interval, swing interval, and stance interval (from both legs), have been assessed.

Bhattacharya et al [15] used SVM that distinguished people with PD from healthy people and attempted for optimum accuracies using different kernel values. As the number of cross validation folds increases, the ROC curve shifts and the manner in which the true positive and false positive rates vary is also examined. This goal can be achieved by the early diagnostic model of PD utilising speech signals, which Oung et al. [16]

presented. In this technological effort, speech behaviour is used as a biomarker for speech data gathering in order to distinguish between individuals with mild and severe PD and healthy individuals. MFCC, LPC, Linear Prediction Cepstral Coefficients (LPCC), and Weighted Linear Prediction Cepstral Coefficients (WLPPC) are among the feature extraction techniques that were taken into consideration. Two different classifier types are used for the classification: Probabilistic Neural Networks (PNN) and K-Nearest Neighbours (KNN). The tests' outcomes demonstrate that, when employing 10-fold cross-validation measures, the PNN and KNN classifiers produce higher average classification performance of 92.63% and 88.56%, respectively. Thankfully, the suggested methods indicate the possibility of a unique tool preference developing that might enable PD detection with outstanding results. Sharma et al. [17] created three different classifier types that rely on MLP, KNN, and SVM to help the experts diagnose PD. The biological speech signals from 31 people—23 of whom have PD and 8 of whom are healthy—make up the dataset for this technological endeavour. The PD data set, which was taken from the UCI machine learning database, was used to achieve this goal. The findings demonstrate that an enhanced accuracy of over 85.294% was attained.

Polat et al [18] presented a novel method for PD identification that takes advantage of the qualities that speech signals produce. There are two stages to the suggested hybrid machine learning technique: (i) pre-processing the data (over-sampling), and (ii) categorization. There are two classes in the PD dataset (PD dataset). Here, 192 data come from individuals who are normal (healthy), and 564 data are from the sick class (PD). There is an uneven distribution of classes in the data collection. Using the SMOTE (Synthetic Minority Over-Sampling approach) approach, the imbalanced dataset is transformed into a balanced dataset. The PD dataset was later classified using the Random Forests classification algorithm after the class distribution was adjusted to a balanced state. Bhardwaj et al.'s [19] early PD diagnostics used conventional characteristics based classifications. After the various characteristics from the aforementioned classes were examined, the participants under consideration for this study were divided into four groups according to their Unified PD Rating Scale (UPDRS) scores. Additionally, KNN (Binary Data) and Decision Tree Classifier (Multiclass Data) yield accuracy of 87.83 and 98.63%, respectively. Regression analysis (RA) and ANNs, two deep learning techniques, are hybridised by Sahuet et al. [20] to efficiently diagnose diseases using probability estimates. In this regard, the benefits of the various available techniques' approaches promise reliable probability estimate. RA performs data preprocessing and probability computation on preprocessed data. The next approach, which is already accessible, aids in the identification of Parkinson's disease patients by comparing the neuron's threshold value. The estimate is based on a group of people's speech recognition, iron content, and pulse rate data sets. The proposed approach is compared to current methods such as SVM and k-NN classifiers. The estimated results show how much better the proposed technique is, with an accuracy of 93.46%.

Naseer et al [21] augmented data and applied transfer learning techniques into deep convolutional neural networks using ImageNet and MNIST dataset. Two methods namely freezing and fine-tuning of transfer learning were examined. When utilising a fine-tuning-based approach with ImageNet and PaHaW dataset, a trained network produces an accuracy of 98.28%. When compared to benchmark methodologies, the suggested technique enhances PD identification, according to the findings of studies conducted on a standard dataset. A DWT was suggested by Soumaya et al. [22] as classification models utilising signal processing and using approximations a_3 , properties of LPC, energies, ZCR, MFCC, and wavelet Shannon entropies were extracted. Later, the best accuracy of 91.18% was with GA and SVM classifications.

3. Proposed Methodology

In this work, Modified brain storm optimization based feature selection and a model of classification technique, whose objective is PD prediction employing speech signal processing. In the proposed technique, classification approach is used, in which five primary stages are involved. The stages as shown in Fig. 1 included the transformation of speech signals applying DWT, preprocessing, features extraction, feature selection, and classification. DWT are applied to speech signals for extracting vectors a_3 approximations of signal, based on structures shown in Fig. 1. The hamming window and pre-emphasis filter are used in the pre-processing stage. Features such as wavelet Shannon entropies, energies, ZCR, MFCC, and LPC were derived. Following the concatenations of acoustic and compositional characteristics of a_3 approximations during feature extractions, feature vectors with 21 coefficients were obtained. Subsequently, features were

chosen using Modified Brainstorm Optimisations, which aided in reductions of features. Finally, SVMclassifiers categorized PD data.

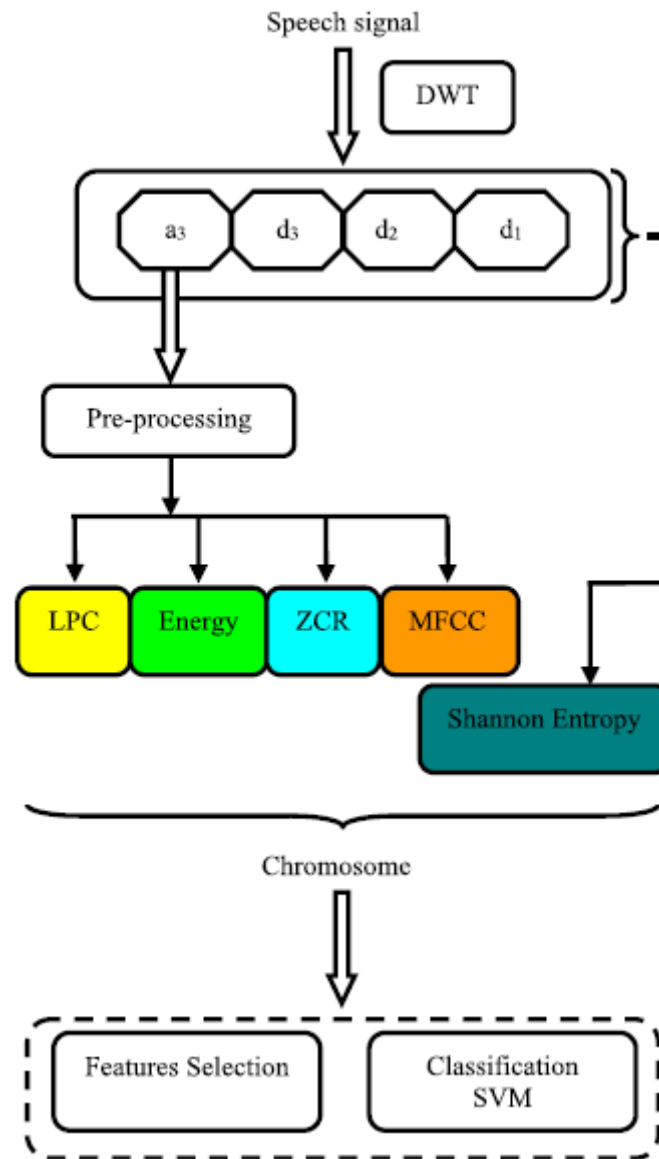


Figure 1. The overall process of the proposed PD classification

3.1. Features extraction using DWT

DWT is acquired through the discretization of the scaling and shifting factors present in continuous wavelet transform [22]. This transform is formulated as given in Eq. (1):

$$Ws_{(a,b)} = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} s(t) \cdot \psi^*\left(\frac{t-b}{a}\right) dt \quad (1)$$

$W(t)$ indicates mother wavelets, while conjugate complexes $W(t)$ are given by $W^*(t)$, $s(t)$ stands for signals whose transformations are to be done. “a” stands for scale factors, and “b” refers to translation parameters.

3.1.1. Formant frequencies

Using LPC, the formant estimate of the speech signal is obtained among the acoustical properties. Formant frequencies are thought of as frequency characteristics. This formant will show up as explicit peaks in

the speech spectrum. These peaks represent the vocal tract's resonance frequencies. Speech sound differentiations are quantified using amplitudes or frequency spectrums.

3.1.2. The time-domain energies and ZCR

Time-domain energies of signals are additions of squares of samples as given in Eq. (2):

$$E = \sum_{n=1}^N |x_1(n)|^2 \quad (2)$$

In cases of speech signals, to compute ZCR of signal frames are defined as rates at which signal's sign varies during frames. They represent transitions between frequencies of signal change values from positives to negatives and vice versa, and overall lengths of windowed frames ZCR expressed as Eq. (3).

$$ZCR = \frac{1}{N} [\sum_{n=1}^{N-1} x_1(n+1) - x_1(n)] \quad (3)$$

x_1 implies windowed frame signals, and $n = (1, 2, \dots, N)$ stands for lengths of windowed frames.

3.1.3. MFCC

Up until now, the cepstrum domain has made use of the highly probable processing of speech signals. due to the vocal signal's property, which are convolutions of vocal tracts and sources. These convolutions produce results that are difficult to separate into contributions provided by sources and conduits. By passing them via log-spectral domains, cepstral analyses may be used to circumvent this problem [23]. One of the speech signal representations that are thought to be cepstrum along outputs that Mel scale filter banks provide MFCC.

MFCC analysis converts linear frequency scales into Mel scales, utilising properties of human auditory systems. Encoding of final scales are done by sending through sets of 15 to 24 triangular filters with linear spacing. The changes from linear scales to Mel scales are expressed in Eq. (4).

$$Mel(f) = 2595 \times \log\left(1 + \frac{f}{700}\right) \quad (4)$$

Given that signal analysis is the subject of study, the mathematical transformation in the form of auxiliary tools should be taken into account initially. These transformations aid in the frequency and temporal domain representation of this signal. When it comes to identifying both stationary and transient signals, the wavelet transform is the most successful transform among the highly potential transformations. The wavelet transform's ability to characterise signals in frequency domains at multiple localization levels in time lies at the heart of the concept. To implement discrete wavelets, two wavelet filters—one for low pass and one for high pass—are utilised. As filters, only quadrature mirror filters (QMF) are utilised. They are rebuilt in accordance utilising the scaling function and the mother wavelet. The signal will be separated into high-frequency and low-frequency components in a matched fashion. The approximation coefficient "a" represents the low-frequency component, whereas the detail coefficient "d" represents the high-frequency component. This procedure is defined below in Eq. (5):

$$\begin{cases} d_i(n) = \sum_k H(2n-k)a_{i-1}(k) \\ a_i(n) = \sum_k L(2n-k)a_{i-1}(k) \\ a_{i-1}(n) = a_{i-1}(n) + d_i(n) \end{cases} \quad i \geq 1 \quad (5)$$

Where $L(n)$ refers to low-pass filters and $H(n)$ indicate high-pass filters. "i" denotes scales. There are several wavelet families to choose from, in addition to Daubechies, Coiflets, and Symlets. Daubechies are purposefully chosen amongst wavelets in this article. Given that Daubechies level 2 at third scales yield best results. Low-frequency component signals, or approximations a_3 , will be used as pre-processing inputs. There are three phases to pre-processing related to approximations a_3 . Pre-emphasizing, framing, and windowing are some of these stages. The following are expressions of outputs that pre-emphasizing stage produce (refer to Eq (6)):

$$x(n) = a_3(n) - ka_3(n-1) \quad (6)$$

With empirically picked $k = 0.97$ [24]. By reducing spectral tilts in speech spectrums and placing focus on higher frequency formants, pre-emphasis filters make it easier to obtain extremely accurate LPC analysis.

After then, the signal is split into frames with N samples separated at intervals of 10–30 ms, which is when the signal is expected to stabilise, in the framing stage. because of the challenge of the framing stage, which is defined as the appearance of a few anomalies along the edges of the frames. It must be used in conjunction with the sample from each frame to multiply the hamming window in order to decrease such irregularities. In Eq (7), the hammingwindow is stated.

$$w_h(n) = 0,54 - 0,46 \cos\left(\frac{2\pi n}{N-1}\right) \quad (7)$$

Initially, formant frequencies are extracted at steps of extracting acoustical characteristics. LPCs are methods that produce extremely efficient and highly accurate estimations of speech characteristics by estimating F1, F2, and F3 formant frequencies. By using auto-correlation approaches, LP coefficients which are polynomials are discovered. The roots are then used to obtain the poles of the LP filters. The following is an illustration of the LPC algorithm utilising the autocorrelation approach in Algorithm 1.

Algorithm 1: LPC

1. Inputs: speech frames x_1 (speech weighed by windows). $m = \{0, 1, 2, \dots, e\}$ with e indicating LPC model orders.

2. Do while ($m \leq e$). Frames' autocorrelation functions (ACF) are computed applying equations $r(i) = \sum_{l=0}^{L-1-m} x_1(l)x_1(l+i)$ With L indicating lengths of x_1 .

1. Specify Toeplitz matrices, diagonal-constant matrices:

$$\begin{bmatrix} r(1) & r(2) & r(e) \\ r(2) & \ddots & r(2) \\ r(e) & r(2) & r(1) \end{bmatrix}$$

2. Finding solutions to Yule-Walker $\sum_{k=0}^p a(k)R(m-k) = -r(m)$ by using Levinson-Durbin recursive algorithms.

3. Output: $a = R^{-1}$ the LP coefficients to be the polynomial. Once LP polynomial is extracted, the formants are got. The time-domain energy and the ZCR are computed using the Eqs. (2) and (3).

The twelve initial coefficients of MFCC are computed using energies from a triangle filter bank along the Mel frequency scale. The discrete cosine transform (DCT) is used to convert the logarithm of this energy across time. The DCT output has a problem due to the small higher order of the cepstral coefficients. As a result, modifying the coefficients is required to increase the amplitudes. As illustrated in Fig. 2, the following is an example of how to compute the MFCC coefficients.

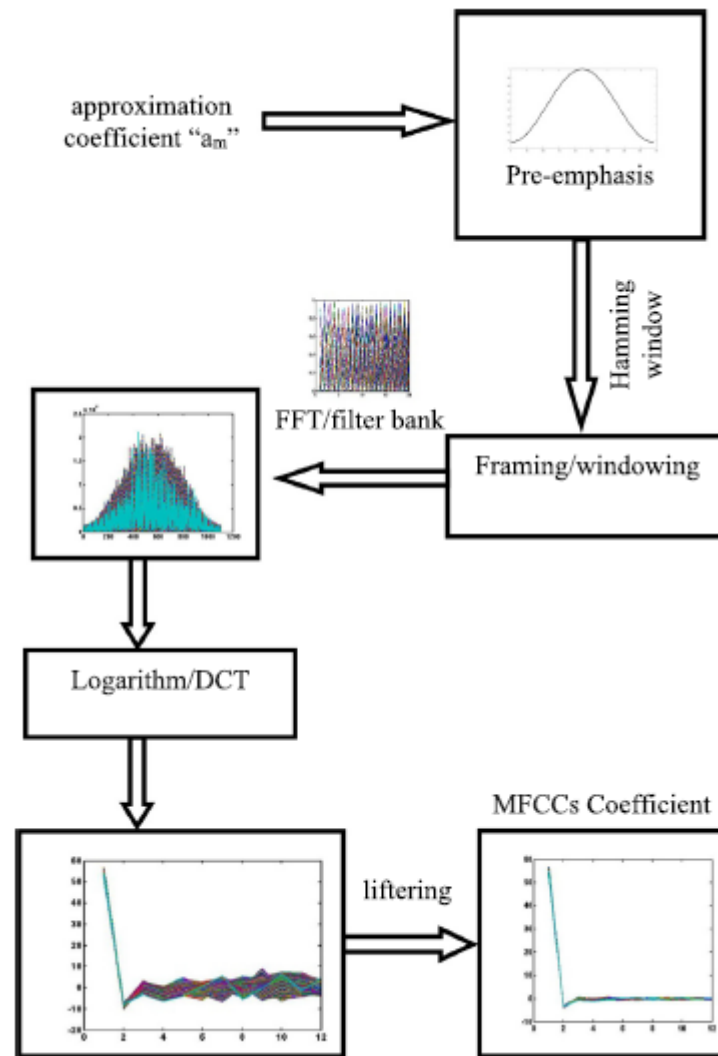


Figure 2. Process of MFCC

3.1.4. Shannon entropy decompositional features

Vectors ($a_m, d_m, d_{m-1}, \dots, d_1$) are extracted once the Shannon entropy is applied to all of them. Shannon entropies constitute vital roles in information theories which define degrees of confusions of systems. Highly ordered systems have lesser entropies. Shannon entropies “ H ” are computed as in the Eq. (8):

$$H = \sum_{j=1}^J p_j \log_2(p_j) \quad (8)$$

where p_j gives probabilities of energy distributions for every wavelet coefficients [25]. In final features extraction stages, vector sized 1×21 are obtained concatenating vector sized 1×17 leading to phases of acoustical features extractions and vector sized 1×4 including decompositional features.

3.2. Feature selection

Choose features The method of feature selection looks for a linear subspace with smaller dimension than the current feature space, where the new PD features have the largest variance. For the purpose of choosing the optimal features, the modified brain storm optimisation technique is presented in this work.

3.2.1. Brain storm optimization

Three crucial processes are involved in BSO [26]: grouping people into clusters, uprooting cluster centres, and creating new solutions. Initially, BSO generates random solutions, which are then evaluated using a fitness function. Subsequently, BSO offers the k-mean clustering algorithm for grouping n solutions into m

groups. After that, a new solution is created with the intention of replacing a cluster centre that is selected at random. A cluster centre is upset in this stage. At last, an individual is randomly chosen on the basis of one or more fusions of two cluster center(s), as given:

$$X_{selected} = \begin{cases} X_i, & \text{one cluster} \\ rand \times X_{1i} + (1 - rand) \times X_{2i}, & \text{two cluster} \end{cases} \quad (9)$$

where *rand* stands for random values ranging between 0 and 1, X_{1i} and X_{2i} indicate i -th dimensions of chosen clusters and chosen ideas are revised as below:

$$X_{new} = X_{selected} + \xi * random(0,1) \quad (10)$$

where *random(0,1)* stands for Gaussian random values having 0 and 1 to be means and variances, correspondingly; and ξ indicates adjusting factors, which are specified as:

$$\xi = \logsin\left(\frac{0.5 * m_i - c_i}{k}\right) \times rand \quad (11)$$

where *logsin()* stands for logarithmic sigmoid functions, k indicates rates of changes for slopes of *logsin()* functions, *rand()* indicates random values between 0 and 1, m and c stands for maximum and current iteration counts, correspondingly.

3.2.2. Modified Brain Storm Optimization using Easy Grouping Process (EGP)

The grouping operator used a k-mean clustering approach called BSO. Nevertheless, the BSO algorithm's implementation was made more difficult by the use of the k-mean clustering approach, and its computing complexity was also significant. A highly accurate method, such as the k-mean clustering technique, must be included for classifying the concepts into different groups because the evolutionary process involves stochastic parameters during executions and BSO executes grouping operators in generations. Using simple grouping procedures known as EGP adopting the following steps, MBSO may construct grouping operators.

Step 1: Randomly chosen M diverse concepts from current generations are seeds of M groups where M seeds are represented as S_j ($1 \leq j \leq M$).

Step 2: For every idea X_i ($1 \leq i \leq N$) in present generations, distances to groups j are computed as:

$$d_{ij} = |X_i, S_j| = \sqrt{(\sum_{d=1}^D (x_{id} - s_{jd})^2) / D} \quad (12)$$

Step 3: M distance values are compared and idea X_i are assigned into closest groups. It is to be observed that the seed S_j of this group remains unmodified.

Step 4: Move to Step 2 for the upcoming idea. Else the EGP stops once the assignment of all the ideas is done.

3.3. Classification using SVM

The primary concept of the SVM [27] is the estimation of a model that can help finding the optimal hyper-plane capable of separating the data. The hyperplane is mathematically defined as:

$$H_p(x) = w^T x + bv = 0 \quad (13)$$

where x stands for inputs, w denotes weight vectors, and bv denotes the bias value. A Hard Margin optimality may be applied if the training dataset classes can be precisely divided. The goal of choosing the hyperplane decision boundary in this case, to maximise separations between hyperplanes and closest training data points. When nonlinear classification is taken into account, the boundary is still the hyperplane, and the only difference is that the input is replaced by the hyperplane that is located in the feature space. The form of Eq. (13) $H_p(x)$ is:

$$w^T \phi(x) + bv = 0 \quad (14)$$

where $\phi(x)$ form non-linear transformations of input vectors x . Optimal weight vectors are attained by executions of Lagrange multipliers and computed using Eq. (15):

$$w = \sum_{i=1}^N \alpha_i y_i \phi(x) \quad (15)$$

with α_i stands for Lagrange multiplier coefficients where optimal decisions can be formulated as Eq. (16):

$$\sum_{i=1}^N \alpha_i y_i \phi(x) + bv = 0 \quad (16)$$

By giving $u_i = \alpha_i y_i$ and $k(x, x_i) = \phi^T x \phi(x_i)$ are consequent y functions of decision expressed as follows (see Eq. (17)):

$$y = \sum_{i=1}^N u_i k(x, x_i) + b \quad (17)$$

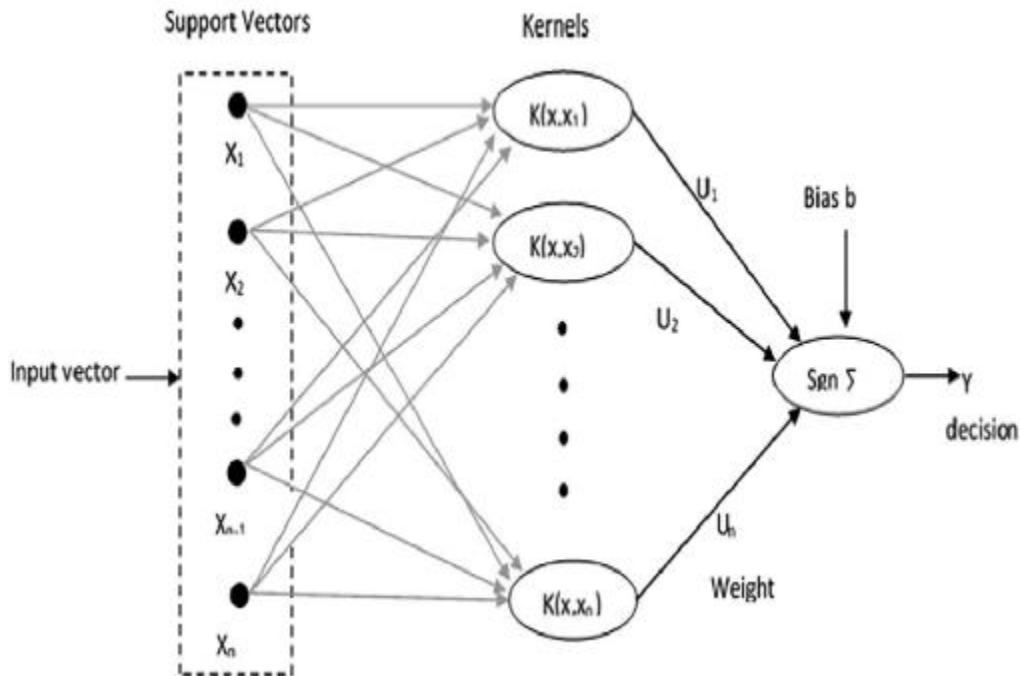


Figure 3. Process of SVM

Fig. 3 gives earlier Eq. (17) of decision functions and final words forms signs of output values y . if $\text{sign}(y) = -1$ labels of inputs give classes -1, else form classes 1.

4. Results and Discussion

Performance evaluation of MBSO-SVM technique were carried out on standard datasets. These datasets are also classified as training, and test sets for evaluating various classifiers. The classification accuracies attained with the proposed MBSO-SVM technique and other classifiers considered for PD classification are compared. Each one of the available classifiers and the proposed SVM based learning model are executed in static environment whereas the entire dataset is considered and 10-fold cross validations compute accuracies. For testing the classifiers, except the classification accuracy, few statistical measurements, provided in equation (18)–(21) are also carried out along with the average outcomes for the classifiers. Precision is specified to be the proportion between the rightly got positive observations and all of the expected positive observations.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (18)$$

Sensitivity is given by the proportion of rightly detected positive observations and the over-all number of observations.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (19)$$

F1 score is specified as the weighted average of Precision and Recall. Consequently, it uses false positives and false negatives.

$$\text{F1 Score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}) \quad (20)$$

Accuracy is computed in terms of positives and negatives as given:

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (21)$$

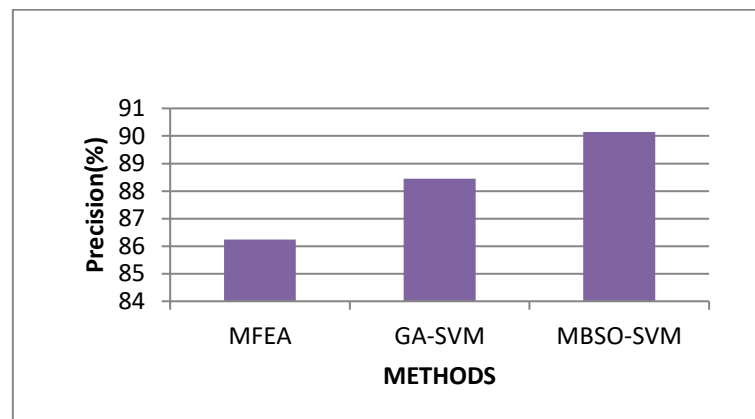


Fig.2. comparison of Precision in the proposed MBSO-SVM

Fig.2 illustrates comparative Precision values of MBSO-SVM and available techniques. With the proposed MBSO-SVM, improved Precision values of 90.15% is achieved, while the accuracy of GA-SVM technique is 88.45% and the accuracy of MFEA technique is obtained at 86.24%. It can be proven that a better precision is achieved with the proposed technique.

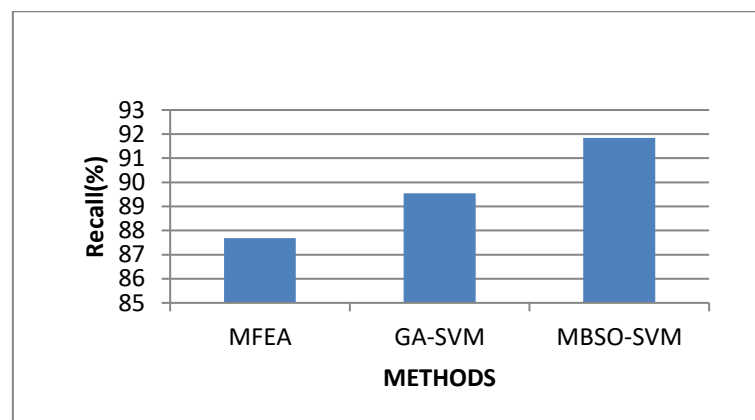


Fig.3. comparison of Recall of the proposed MBSO-SVM

Fig.3 demonstrates the comparison analysis in terms of recall between the proposed MBSO-SVM and the available techniques. When using the proposed MBSO-SVM, improved recall of 91.84% is achieved, while the accuracy of GA-SVM technique is obtained at 89.54% and that of MFEA technique is 87.68%. It is proven that an improved recall value is achieved with the proposed technique.

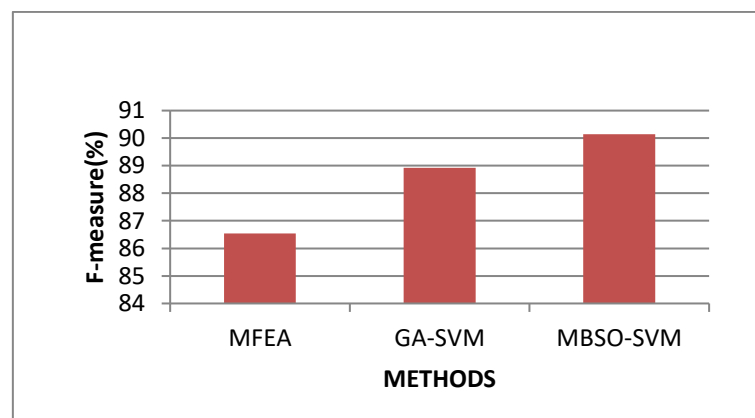


Fig.4. comparison of F-measure with the proposed MBSO-SVM

Fig.4 comparative F-Measure values of MBSO-SVM and available techniques. While using the proposed MBSO-SVM technique, better F-measure results of 90.14% is attained, while the accuracy of GA-SVM technique is 88.92% and that of MFEA technique is 86.54%. It can be proven that the proposed technique yields increased F-measure.

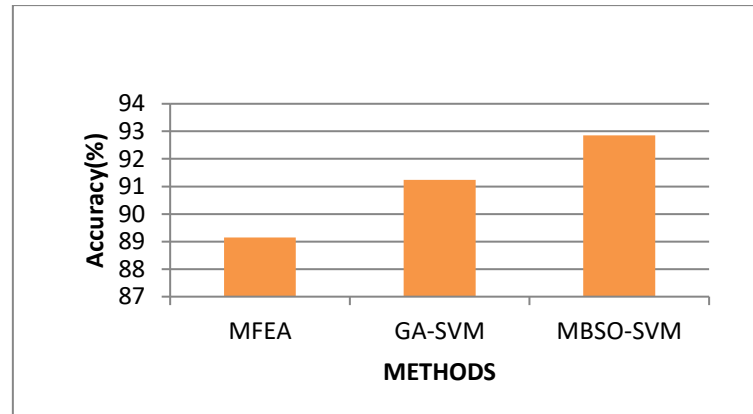


Fig.5. comparison of Accuracy with the proposed MBSO-SVM

Fig.5 depicts that accuracy achieved with the proposed MBSO-SVM is extremely superior compared to the available techniques. With the proposed MBSO-SVM, improved accuracy results of 92.85%, is achieved, while GA-SVM method yields an accuracy of 91.24% and MFEA technique provides 89.15%. It can be concluded that the proposed technique is highly accurate.

5. Conclusion

Early detection of any type of disease forms a significant aspect. In this research work, introduced an Modified brain storm optimizations based feature selections and classification model employing DWT for signal transformations. Moreover, BSO performs a random disruption in the cluster center for generating a fresh solution, and it is diverse from the available ones. This updating strategy aids BSO in avoiding the local optima and yields improved outcomes in comparison with that of GA. However, BSO needs extended execution times while the optimal feature subset is found. This is primarily because of the application of distance-based k -mean clustering during every iteration. The resolution to this issue can be achieved with the substitution of k -mean clustering and easy grouping procedure. The outcomes of simulations show that the proposed framework provides superior predictions for given datasets than the existing models. This research work is helpful in providing proactive treatment to the patient. In this research work, SVM based classifier yields the 93.4% accuracy.

References

- [1] Srivastava, A. K., Jeberson, K., & Jeberson, W. (2022). A systematic review on data mining application in Parkinson's disease. *Neuroscience Informatics*, 2(4), 100064.
- [2] Mathur, R., Pathak, V., & Bandil, D. (2019). Parkinson disease prediction using machine learning algorithm. In *Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018* (pp. 357-363). Springer Singapore.
- [3] Miljkovic, D., Aleksovski, D., Podpečan, V., Lavrač, N., Malle, B., & Holzinger, A. (2016). Machine learning and data mining methods for managing Parkinson's disease. *Machine learning for health informatics: state-of-the-art and future challenges*, 209-220.
- [4] Sonu, S. R., Prakash, V., Ranjan, R., & Saritha, K. (2017, August). Prediction of Parkinson's disease using data mining. In *2017 international conference on energy, communication, data analytics and soft computing (ICECDS)* (pp. 1082-1085). IEEE.
- [5] Sriram, T. V., Rao, M. V., Narayana, G. S., & Kaladhar, D. S. V. G. K. (2015). Diagnosis of Parkinson disease using machine learning and data mining systems from voice dataset. In *Proceedings of the 3rd*

- International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014: Volume 1* (pp. 151-157). Cham: Springer International Publishing.
- [6] Ramani, R. G., & Sivagami, G. (2011). Parkinson disease classification using data mining algorithms. *International journal of computer applications*, 32(9), 17-22.
 - [7] Suganya, P., & Sumathi, C. P. (2015). A novel metaheuristic data mining algorithm for the detection and classification of Parkinson disease. *Indian Journal of Science and Technology*, 8(14), 1.
 - [8] Suganya, P., & Sumathi, C. P. (2015). A novel metaheuristic data mining algorithm for the detection and classification of Parkinson disease. *Indian Journal of Science and Technology*, 8(14), 1.
 - [9] Bind, S., Tiwari, A. K., Sahani, A. K., Koulibaly, P., Nobili, F., Pagani, M., ... & Tatsch, K. (2015). A survey of machine learning based approaches for Parkinson disease prediction. *Int. J. Comput. Sci. Inf. Technol*, 6(2), 1648-1655.
 - [10] Sharma, A., & Giri, R. N. (2014). Automatic recognition of Parkinson's Disease via artificial neural network and support vector machine. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 4(3), 2278-3075.
 - [11] Sriram, T. V., Rao, M. V., Narayana, G. S., Kaladhar, D. S. V. G. K., & Vital, T. P. R. (2013). Intelligent Parkinson disease prediction using machine learning algorithms. *Int. J. Eng. Innov. Technol*, 3, 212-215.
 - [12] Khemphila, A., & Boonjing, V. (2012). Parkinsons disease classification using neural network and feature selection. *International Journal of Mathematical and Computational Sciences*, 6(4), 377-380.
 - [13] Mostafa, S. A., Mustapha, A., Mohammed, M. A., Hamed, R. I., Arunkumar, N., Abd Ghani, M. K., ... & Khaleefah, S. H. (2019). Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease. *Cognitive Systems Research*, 54, 90-99.
 - [14] Joshi, D., Khajuria, A., & Joshi, P. (2017). An automatic non-invasive method for Parkinson's disease classification. *Computer methods and programs in biomedicine*, 145, 135-145.
 - [15] Bhattacharya, I., & Bhatia, M. P. S. (2010). SVM classification to distinguish Parkinson disease patients. In *Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India* (pp. 1-6).
 - [16] Oung, Q. W., Basah, S. N., Muthusamy, H., Vijejan, V., & Lee, H. (2018, March). Evaluation of short-term cepstral based features for detection of Parkinson's Disease severity levels through speech signals. In *IOP Conference Series: Materials Science and Engineering* (Vol. 318, No. 1, p. 012039). IOP Publishing.
 - [17] Sharma, A., & Giri, R. N. (2014). Automatic recognition of Parkinson's Disease via artificial neural network and support vector machine. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 4(3), 2278-3075.
 - [18] Polat, K. (2019, April). A hybrid approach to Parkinson disease classification using speech signal: the combination of smote and random forests. In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)* (pp. 1-3). Ieee.
 - [19] Bhardwaj, S., Arora, D., Devi, B., Shankar, V. G., & Srivastava, S. (2022). Machine Learning Assisted Binary and Multiclass Parkinson's Disease Detection. In *Intelligent Sustainable Systems: Proceedings of ICISS 2022* (pp. 191-206). Singapore: Springer Nature Singapore.
 - [20] Sahu, L., Sharma, R., Sahu, I., Das, M., Sahu, B., & Kumar, R. (2022). Efficient detection of Parkinson's disease using deep learning techniques over medical data. *Expert Systems*, 39(3), e12787.
 - [21] Naseer, A., Rani, M., Naz, S., Razzak, M. I., Imran, M., & Xu, G. (2020). Refining Parkinson's neurological disorder identification through deep transfer learning. *Neural Computing and Applications*, 32, 839-854.
 - [22] Soumaya, Z., Taoufiq, B. D., Benayad, N., Yunus, K., & Abdelkrim, A. (2021). The detection of Parkinson disease using the genetic algorithm and SVM classifier. *Applied Acoustics*, 171, 107528.
 - [23] Rabiner, L. R., & Schafer, R. W. (2007). Introduction to digital speech processing. *Foundations and Trends® in Signal Processing*, 1(1-2), 1-194.
 - [24] Benba, A., Jilbab, A., Hammouch, A., & Sandabad, S. (2015, March). Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease. In *2015 International conference on electrical and information technologies (ICEIT)* (pp. 300-304). IEEE.

- [25] Bafroui, H. H., & Ohadi, A. (2014). Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions. *Neurocomputing*, 133, 437-445.
- [26] Cheng, S., Qin, Q., Chen, J., & Shi, Y. (2016). Brain storm optimization algorithm: a review. *Artificial Intelligence Review*, 46, 445-458.
- [27] Suthaharan, S., & Suthaharan, S. (2016). Support vector machine. *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, 207-235.