Research Article

Predicting Parkinson's Disease Progression: Evaluation of Ensemble Methods in Machine Learning

Mehrbakhsh Nilashi [ID],[1,2] **Rabab Ali Abumalloh** [ID],[3] **Behrouz Minaei-Bidgoli** [ID],[2] **Sarminah Samad** [ID],[4] **Muhammed Yousoof Ismail** [ID],[5] **Ashwaq Alhargan** [ID],[6] **and Waleed Abdu Zogaan** [ID][7]

[1]*Centre for Global Sustainability Studies (CGSS), Universiti Sains Malaysia, USM, Penang 11800, Malaysia*
[2]*School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran*
[3]*Computer Department, Applied College, Imam Abdulrahman Bin Faisal University, Dammam 1982, Saudi Arabia*
[4]*Department of Business Administration, College of Business and Administration,*
 Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
[5]*Department of MIS, Dhofar University, Salalah, Oman*
[6]*Computer Science Department, College of Computing and Informatics, Saudi Electronic University, Riyadh, Saudi Arabia*
[7]*Department of Computer Science, Faculty of Computer Science and Information Technology, Jazan University,*
 Jazan 45142, Saudi Arabia

Correspondence should be addressed to Mehrbakhsh Nilashi; nilashidotnet@hotmail.com

Parkinson's disease (PD) is a complex neurodegenerative disease. Accurate diagnosis of this disease in the early stages is crucial for its initial treatment. This paper aims to present a comparative study on the methods developed by machine learning techniques in PD diagnosis. We rely on clustering and prediction learning approaches to perform the comparative study. Specifically, we use different clustering techniques for PD data clustering and support vector regression ensembles to predict Motor-UPDRS and Total-UPDRS. The results are then compared with the other prediction learning approaches, multiple linear regression, neurofuzzy, and support vector regression techniques. The comparative study is performed on a real-world PD dataset. The prediction results of data analysis on a PD real-world dataset revealed that expectation-maximization with the aid of SVR ensembles can provide better prediction accuracy in relation to decision trees, deep belief network, neurofuzzy, and support vector regression combined with other clustering techniques in the prediction of Motor-UPDRS and Total-UPDRS.

1. Introduction

Parkinson's disease (PD) is the second most common and complex neurodegenerative disorder worldwide [1–4]. Both polygenic and environmental factors can cause PD [5]. It is found that, in about 1%–2% of the PD cases (mainly familial), the disease development occurs through a single gene [5]. The main symptoms of PD are bradykinesia (motor features), muscle stiffness, and tremor, along with other symptoms such as sleep disorders (nonmotor features), cardiac arrhythmia, and constipation. Alteration of voice

and speech is one of the features of PD. Unified Parkinson's Disease Rating Scale or UPDRS, which shows symptoms' presence and severity, is mainly used in tracking PD symptom progression [6–8]. UPDRS is considered as the well-validated test and the most widely used clinical rating scale for patients with PD [6,9–11]. UPDRS includes 4 sections, in which UPDRS I, UPDRS II, UPDRS III, and UPDRS IV are used to evaluate psychiatric symptoms in PD, activities of daily living, reliable motor symptoms measured in PD recognized by physical exam, and complications of treatment [10]. In many studies, this scale is considered

based on Total-UPDRS with the range of 0–176 (176 total disability and 0 representing healthy) and Motor-UPDRS which indicates the UPDRS' motor section with the range of 0–108 (108 indicating severe motor impairment and 0 indicating healthy state) [6].

Machine learning (ML) approaches have demonstrated the capability of handling large volumes of medical datasets and presented perceptive directions [12]. The use of ML-based tools could enhance the safety of individuals [13–15], enhance the quality of medical care [16–18], minimize the costs of medical care [19–21], and support physicians' efforts by manipulating big data of patients' records. ML approaches have been broadly utilized for disorders' classification and prediction [22–30]. Gadekallu et al. [31] investigated the use of machine learning techniques for the prediction of diabetic retinopathy. The authors used the PCA-based Deep Neural Network (DNN) model using the Grey Wolf Optimization (GWO) algorithm for the classification of the extracted features of the diabetic retinopathy dataset. The method was evaluated through the accuracy, recall, sensitivity, and specificity evaluation metrics and compared with the support vector machine (SVM), naïve Bayes classifier, decision tree (DT), and XGBoost. Overall, their method achieved higher accuracy compared with the SVM, DT, and XGBoost techniques. Bhattacharya et al. [32] developed a method for the classification of imbalanced multimodal stroke dataset. The authors implemented the Antlion optimization algorithm on the DNN model to select optimal hyperparameters in minimal time consumption. A positive aspect of their method was that it consumed only 38.13% of the training time on the stroke dataset. An artificial neural network is among the most significant approaches for disease classification and prediction [33–38]. Referring to Berner [39], clinical decision support systems (CDSSs) are special tools that are developed to aid medical specialists in their decision-making, considering particular disorders or diseases. ML approaches can be utilized for designing effective CDSS [36] to aid medical specialists in reaching accurate and timely predictions. CDSSs designed using machine learning approaches have played a significant part in evaluating the existence or the severity of the disease.

In machine learning methods, unsupervised approaches are used to lower the dimensionality of data, which allows the detection of the disease. Besides, these approaches allow manipulating the data, removing the noise from data, calculating the similarity, and segmenting the data [40]. On the other hand, supervised learning approaches are used to enable the final classification, prediction, and diagnosis of the disease [41]. While ML has proven its benefits, the effective deployment of ML needs a great effort from human specialists, considering that no particular approach can present acceptable results in all possible scenarios [12]. Although clinical data are available to researchers to explore, the lack of experience to handle big sources of data might restrict the optimum utilization of these sources. Besides, even though several approaches have been used in disease prediction using various real-world medical datasets, the choice of the deployed approach should consider enhancing

the accuracy of the prediction and minimizing the time of computation [42].

The goal of this paper is to present a comparison of machine learning approaches for remote tracking of Parkinson's disease progression. The comparative study is based on clustering and prediction learning approaches. To further improve the accuracy of UPDRS prediction, this study uses ensemble learning in the final stage of the proposed method. Ensemble learning approaches have proven to be effective in prediction tasks [25]. Few studies have incorporated ensemble learning approaches for the development of the diseases diagnosis systems [43], [44]; [45]. Further investigations are needed for the effectiveness of these approaches in UPDRS prediction. Accordingly, we use ensembles of support vector regression and different clustering techniques for PD data clustering. The results are then compared with other prediction learning approaches, deep belief network (DBN), support vector regression, multiple linear regression, and neurofuzzy techniques.

The rest of this paper is organized as follows. We introduce a summary of related works on Parkinson's disease in Section 2. In this section, the results of previous works are discussed. In Section 3, we introduce a new hybrid method for PD diagnosis. In Section 4, we present the method evaluation through a PD dataset. Finally, we present the conclusion and recommendations for future study in Section 5. To simplify, a list of abbreviations we used in this research is presented in Table 1.

2. Related Work

Previous literature has presented several approaches that allow PD detection, classification, and severity prediction. In Table 2, we present these studies along with the adopted approaches in each study. In the following, we will introduce a summarization of the up-to-date researches in this field.

Prashanth et al. [68] concentrated on utilizing nonmotor signals in the early diagnosis of PD by deploying NB, SVM, Boosted Trees, and RF. The findings indicated that SVM presented the highest accuracy value of 96.40%. Abiyev and Abizade [69] presented a new methodology for PD diagnosis using FNS and NN. The outcomes of the study presented efficient performance of the FNS compared to other approaches. Singh et al. [70] presented a new approach for PD detection using SVM and presented an overall accuracy of 100%. Çimen and Bolat [60] focused on the vocal signals for PD diagnosis using ANN, MLP, and GRNN. The outcomes indicated that the best performance was achieved using GRNN. Shetty and Rao [71] focused on gait signals in PD diagnosis and other neurological disorders using SVM. The presented approach achieved an accuracy of 83.33%. Nilashi, Ibrahim, and Ahani [61] presented a hybrid methodology by using EM, PCA, ANFIS, and SVR techniques. The findings of the study indicated that the presented methodology can detect the severity of the disease accurately. Ozkan [72] concentrated on vocal indicators by using a hybrid methodology based on PCA with K-NN. The result of the study indicated the robustness of the presented approach with an accuracy of 99.1%. Avci and Dogantekin [62] presented a

TABLE 1: List of acronyms in this paper.

Acronyms	Description
AI	Artificial intelligence
ANFIS	Adaptive neurofuzzy inference system
ANN	Artificial neural network
CART	Tree classification and regression
CDSSs	Clinical decision support systems
CGP	Cartesian genetic programming
CSPA	Cluster-based similarity partitioning algorithm
CNN	Convolutional neural network
DBN	Deep belief network
DT	Decision tree
DSS	Decision support systems
DNN	Deep neural network
ebTL	Empirical Bayes transfer learning
ECG	Electrocardiogram
ELM	Extreme learning machines
EM	Expectation-maximization
EMG	Electromyogram
FDR	Fisher discriminant ratio
FNS	Fuzzy neural system
FOG	Freezing of gait
GA	Genetic algorithm
GRNN	Generalized regression neural networks
HGPA	Hypergraph partitioning algorithm
IMU	Inertial measurement unit
ISVR	Incremental support vector regression
K-NN	K-nearest neighbor
LR	Logistic regression
LSTM	Long short-term memory
LSVM	Lagrangian support vector machine
MAE	Mean absolute error
ML	Machine learning
MLP	Multilayer perceptron
MSA	Multiple system atrophy
MLP-LSVM	Multilayer perceptron-Lagrangian support vector machine
MLR	Multiple linear regression
NB	Naïve Bayes
NN	Neural network
NIPALS	Nonlinear iterative partial least squares
OPF	Optimum-path forest
PCA	Principal component analysis
PCG	Phonocardiogram
PD	Parkinson's disease
PSP	Progressive supranuclear palsy
RBM	Restricted Boltzmann machine
RF	Random forest
RBF	Radial basis functions
RNN	Recurrent neural network
RSSD	Sparse signal decomposition
RMSE	Root mean squared error
RBF	Radial basis functions
RNN	Recurrent neural network
RSSD	Sparse signal decomposition
RMSE	Root mean squared error
rTL	Regularized transfer learning
SOM	Self-organizing map
SVD	Singular value decomposition
SPECT	Single-photon emission computerized tomography
SVR	Support vector regression
SVM	Support vector machine

TABLE 1: Continued.

Acronyms	Description
T-F	Time-frequency
UPDRS	Unified Parkinson's disease rating scale
WK	Wavelet kernel

new methodology for PD detection based on GA, wavelet kernel, and ELM. The findings of the study indicated that the hybrid approach presented better prediction accuracy than other state-of-the-art related approaches.

Three ML methods, namely SVM, NB, and RF, were deployed by Rovini et al. [74] and presented an encouraging outcome with a specificity value of 0.967. The vocal signals of PD were assessed by Pahuja and Nagabhushan [73] by using ANN, K-NN, and SVM. The outcomes of the evaluation indicated that ANN presented the most accurate performance with an overall accuracy of 95.89%. In a study by Nilashi et al. [63], SOM, NIPALS, and ISVR approaches were used in PD prediction and presented a robust performance in forecasting UPDRS while minimizing the time of prediction. Parisi et al. [64] presented a hybrid approach based on SVM (MLP-LSVM). The deployed method presented the highest accuracy in comparison with other techniques for PD diagnosis. Prince and De Vos [65] deployed several approaches for PD diagnosis, focusing on severity detection. The deployed approaches entail LR, RF, DNN, and CNN. Among the used approaches, CNN presented a better prediction accuracy of 62.1%. Another study that concentrated on PD's severity prediction was presented by Zou and Huang [66]. In the deployed method, rTL, LASSO, and ebTL approaches were used, among which ebTL presented the best prediction accuracy.

Grover et al. [67] adopted the DNN for severity prediction among PD patients. The deployed method presented encouraging results with an overall prediction accuracy of 81.66% based on the Motor-UPDRS score. Several supervised approaches were presented by Khoury et al. [59] for PD diagnosis, focusing on gait signals, such as K-NN, NB, SVM, RF, and CART. These approaches were combined with other unsupervised approaches to meet the goal of the study. Among the deployed methods, K-NN, RF, and SVM presented the highest accuracy result. Nilashi et al. [46] presented a new approach based on DL and clustering methods. Particularly, they used DBN and SVR for UPDRS prediction. SOM was used as a clustering method to enhance the prediction accuracy. The method was assessed based on a real-world dataset and the proposed approach of clustering, DBN, and SVR presented predictions' accuracy that outperformed other learning approaches. Ghaderyan and Fathi [47] concentrated on analyzing gait signals for PD detection. The basic method is based on separating various parts of the signal, choosing the most related parts that are utilized to measure interlimb divergence in singular value space. The proposed method presented an average accuracy of 95.59% and 97.22%. Nilashi et al. [75] presented a hybrid approach that utilized the clustering method, SVD, and ANFIS. They indicated that the presented methodology outperformed

Table 2: Previous literature on PD diagnosis.

Author(s)	REG	SVM	FMCFW	CART	K-NN	NB	BT	RF	FNS	PCA	GSGP	GSGO	FDR	ANN	MLP	GRNN	EM	ANIFS	SVR	GA	WK	ELM	SOM	NIPALS	ISVR	MLP-LSVM	DNN	HMM	LASSO	CNN	rTL	eBTL	SVD	LSTM	RBM	LR	RNN	DBN
Nilashi et al. [46]																																						✓
Ghaderyan and Fathi [47]																			✓				✓															
Paragliola and Coronato [48]																														✓			✓					
Balaji et al. [49]																																		✓	✓			
de Souza, et al. [50]																																			✓			
Senturk [51]		✓																																				
De Vos et al. [52]								✓																												✓		
Mohammed et al. [53]																														✓								
Goyal et al. [54]																														✓								
Xu and Pan [55]								✓		✓																											✓	
Ribeiro et al. [56]																																					✓	
Parziale et al. [57]																			✓	✓																		
Tsuda et al. [58]		✓						✓						✓																								
Khoury et al. [59]		✓		✓	✓	✓									✓																							
Çimen and Bolat [60]										✓				✓	✓																							
Nilashi et al. [61]		✓															✓		✓																			
Avci and Dogantekin [62]							✓										✓			✓	✓	✓																
Nilashi et al. [63]	✓						✓													✓		✓	✓															
Parisi et al. [64]																								✓	✓	✓												
Prince and De Vos [65]	✓							✓																			✓											
Zou and Huang [66]																											✓		✓		✓							
Grover et al. [67]																													✓		✓	✓						
Prashanth et al. [68]		✓																✓																				
Abiyev and Abizade [69]						✓			✓					✓																								
Singh et al. [70]		✓								✓																												
Shetty and Rao [71]		✓											✓																									
Ozkan [72]					✓					✓																												
Pahuja and Nagabhushan [73]		✓			✓									✓																								
Rovini et al. [74]		✓				✓		✓																									✓					
Nilashi Ibrahim et al. [75]																																		✓				
Ashour et al. [76]		✓												✓																								

other state-of-the-art approaches in terms of detection accuracy and minimizing the time of computation.

Ashour, El-Attar, Dey, Abd El-Kader, and Abd El-Naby (2020) utilized the LSTM, focusing on the FOG signals collected from several sensors worn on parts of the body. Besides, SVM and ANN were used in the classification process. Although they achieved an overall accuracy of 83%, one shortcoming related to the size of the data can impact the generalizability of the outcomes. Paragliola and Coronato [48] investigated the performance of hybrid NN, which entails CNN for reducing the dimensionality and LSTM for the diagnosis. The study concentrated on walking patterns by utilizing gait data and presented an accuracy of 95%. Still, the applied methodology focused on binary classification issues, in which the severity of the disease was not explored.

In a study by Mohammed, He, and Lin [53], a CNN model for discriminating PD patients from healthy controls, based on SPECT modalities, was proposed. The model was assessed based on 10-fold cross-validation and presented an accuracy of 99.34%. In a study by Balaji et al. [49], a DL approach based on LSTM was developed for severity classification of the PD and presented encouraging results with an overall accuracy value of 98.6%. De Souza et al. [50] presented a Fuzzy OPF for PD diagnosis. RBM was used to extract the features and outperformed other baseline models. Senturk [51] developed an ML method for PD diagnosis. In this study, SVM was used for the classification task and presented an overall accuracy of 93.84%. De Vos et al. [52] used a novel method to discriminate between PSP and PD by utilizing two approaches of LR and RF. RF presented higher accuracy results compared to LR.

In a study by Goyal et al. [54], a hybrid method for feature extraction that integrates RSSD and T-F algorithms was presented. The deep learning approach, particularly CNN, was utilized in PD diagnosis based on speech impairments. The hybrid method presented accurate outcomes in classifying PD patients (99.37%). Based on handwriting assessment, a study by Xu and Pan [55] adopted an ensemble learning model based on RF in the PD diagnosis. Dimensionality reduction was performed using PCA. The presented approach has an overall accuracy of 89.4%, which outperformed LR and SVM based on the adopted evaluation method. Another study, which concentrated on handwriting indicators, was presented by Ribeiro et al. [56]. The study adopted a PD diagnosis approach based on RNN by utilizing the bag of samplings to measure several compact representations and presented acceptable performance compared to other methods in the literature. Another study that concentrated on handwriting was proposed by Parziale et al. [57] based on CGP. This approach presented an explicit decision measure that was used in the detection of PD. The authors also compared various AI approaches for PD detection and indicated that CGP outperformed these approaches in terms of accuracy. Tsuda et al. [58] adopted an approach based on NNs to distinguish between PD and MSA. The adopted methodology presented an improved recognition of the patterns compared to other approaches.

3. Methodology

In the prediction of diseases, ML techniques have proven to be effective [77–79]; [25]; [26]; [61]; [27] [42]. This study uses both unsupervised and supervised learning techniques to diagnose PD through UPDRS prediction. Several approaches that entail clustering, reducing dimensionality, and learning of prediction approaches are used to create the PD diagnoses method. Figure 1 depicts the proposed method with its main stages. Data preprocessing, dimensionality reduction using PCA, clustering using ensembles of EM, and prediction using ensembles of SVR are all stages of the method that are utilized to predict UPDRS through a set of real-world PD data.

Step 1 (data preprocessing). As suggested by previous studies, the data is preprocessed [80, 81] to have a more accurate prediction of UPDRS. The goal of data preprocessing in this study is to handle the dataset's null values. In general, we included the preprocessing stage in the proposed method because it is typically completed during the first step of data analysis. The data is then deployed in the data analysis stages, such as clustering and prediction. The datasets are created with null values for method evaluation. Before clustering and classification tasks, these values must be imputed. In this study, SVD is used for missing value prediction.

Step 2 (data clustering). We use an unsupervised learning technique in this stage for clustering the PD data. The objective of this step was to increase patient record readability through the grouping of patients into different groups. We used ensembles of EM to have a better cluster analysis of the data.

Step 3 (dimensionality reduction). To remove the noise of data, the PCA method was used in this phase to lower the dimensionality of the data [82]. Multicollinearity has a considerable impact on the accuracy of predictors and is a major issue in the field of disease diagnosis [46]. The accuracy of SVR predictors has been affected by the multicollinearity of the data. We, therefore, use PCA to solve the multicollinearity problem as the most popular technique for noise removal.

Step 4 (UPDRS prediction). This stage was performed to predict UPDRS according to the input features. In contrast to the previous prediction methods for PD diagnosis, we used ensembles of SVR to perform this task. SVR is trained to build prediction models with training datasets. It is a common practice to seek the advice of several doctors who are experts in the field in various clinical settings. The ultimate decision for a specific therapy is thus normally made through consultation and a combination of opinions of a committee of specialists. Ensemble learning systems serve a similar function in the machine learning context [83]. In general, ensemble learning systems can be utilized effectively
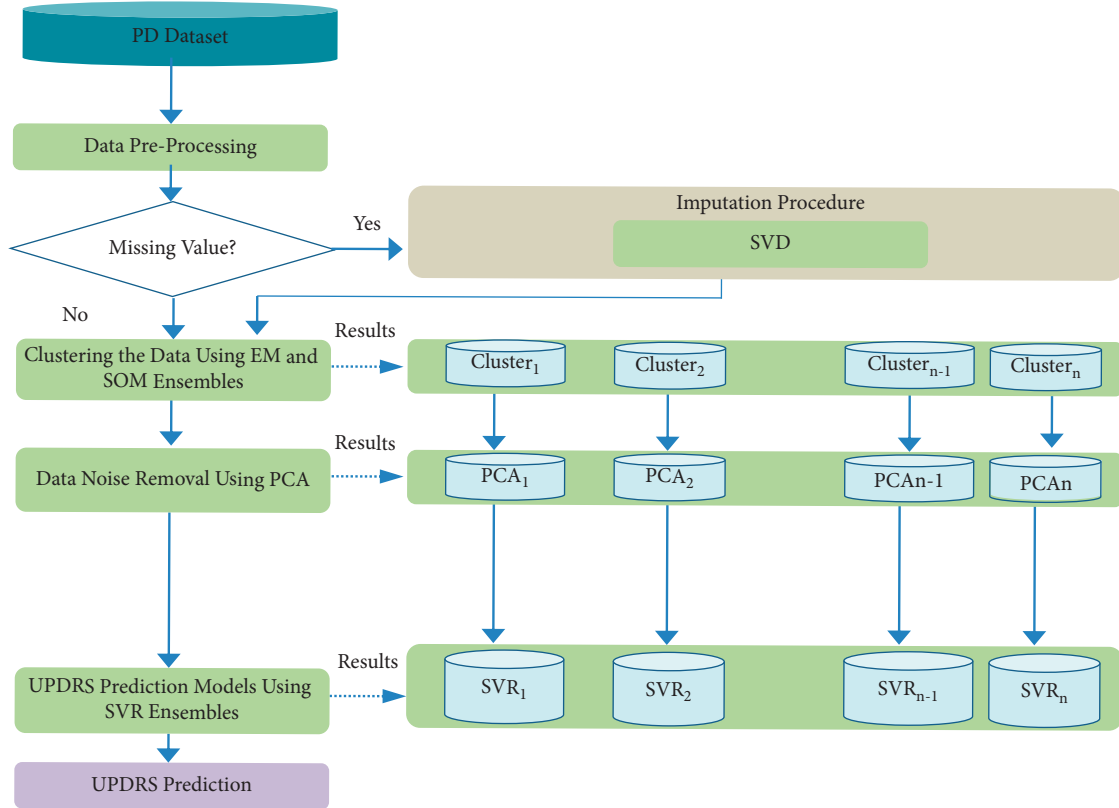
FIGURE 1: The proposed method for PD diagnosis.

in classification and regression problems and provide more reliable predictions than any individual learning model [84]. In fact, several weak hypotheses are combined in ensemble learning systems to form a stronger theory. Note that the success and effectiveness of ensemble learning approaches are heavily dependent on the diversity of the individual predictors that construct the ensemble. The total error can be reduced by combining the output of different prediction models through an algebraic expression (e.g., mean value of the predictions), as the various errors of the prediction models are averaged out.

3.1. SVR Ensembles.
Cortes and Vapnik [85] developed SVM as a machine learning technique for forecasting problems with the potential to be extensively used as a benchmark. Support vector classification (SVC) and SVR are the two main branches of SVM. SVR performs the prediction of a new sample by training the data with target values. This is done by finding $\Phi(x)$ function to map data to a flat space. The SVR can effectively solve complex prediction problems through linear and nonlinear regression. The kernel functions are used to transform the data into a high-dimensional feature space. Radial basis functions (RBF) and polynomial functions are the most widely utilized kernel functions in SVR.

Let us have a training dataset of length N: $T = (x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$, where $y \in \mathfrak{R}$ and $x_k \in \mathfrak{R}^n, k = 1, 2, \ldots, N$, to model a single output (y) in the original SVR; a linear model formulation is presented as

$$\widehat{y}(x) = \langle w, x \rangle + b, \tag{1}$$

where $\widehat{y}$ indicates the predicted output, b is a bias term, w is a weight vector, and $\langle .,. \rangle$ indicates vector inner product. To solve nonlinear problems, in SVR, it is possible to use nonlinear kernel functions, $\Phi(x)$, and we have

$$\widehat{y}(x) = \langle w, \Phi(x) \rangle + b. \tag{2}$$

One of the most widely employed nonlinear SVR kernels is radial basis functions. In SVR, to have good generalization performance, weight vector w is required to be as flat as possible. Thus, for every data point in the dataset, we need to minimize the norm (.) of w as

$$\text{minimize } \frac{1}{2}\|w\|^2, \tag{3}$$

subject to $y_i \langle w, \Phi(x_i) \rangle - b \leq \varepsilon \langle w, \Phi(x_i) \rangle + b - y_i \leq \varepsilon.$

To guarantee feasible constraints, slack variables (ξ_i, ξ_i^*) are introduced as well as ε-insensitive loss function $L(\varepsilon, y, \widehat{y})$. Thus, the following cost function needs to be minimized by estimating the parameters w and b.

$$\text{minimize } \frac{1}{2}\|w\|^2 C \sum_{il}^{N} (\xi_i + \xi_i^*) y_i - \langle w, \Phi(x_i) \rangle - b \le \varepsilon + \xi_i, \tag{4}$$

$$\text{subject to } y_i \langle w, \Phi(x_i) \rangle - b \le \varepsilon \le \varepsilon + \xi \partial_i^* \xi_i, \quad \xi_i^* \ge 0,$$

where C is the regularization parameter. The above optimization problem can be solved in its dual through handling the constraints by employing the Lagrange multipliers η, η^*, β, and β^*. Thus, Lagrangian (L) is presented as

$$L = \frac{1}{2}\|w\|^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) - \sum_{i=1}^{N} \beta_i (\varepsilon + \xi_i - y_i + \langle w, \Phi(x_i) \rangle + b)$$
$$- \sum_{il}^{N} \beta_i^* (\varepsilon + \xi_i^* + y_i - \langle w, \Phi(x_i) \rangle - b) - \sum_{il}^{N} \eta_i \xi_i - \eta_i^* \xi_i^*. \tag{5}$$

In order to find the minimum, the following final problem of quadratic programming results in Karush–Kuhn–Tucker conditions as

$$\text{minimize} \frac{1}{2} \sum_{i-1}^{N} \sum_{i-1}^{N} K_{ij} (\beta_i - \beta_i^*)(\beta_j - \beta_j^*) + \varepsilon \sum_{i-1}^{N} (\beta_i + \beta_i^*) - \sum_{i-1}^{N} y_i (\beta_i + \beta_i^*), \tag{6}$$

$$\text{subject to } \sum_{i-1}^{N} (\beta_i + \beta_i^*) = 0 \, 0 \le \beta_i, \quad \beta_i^* \le C.$$

Here, K_{ij} indicates the kernel function. Finally, in SVR, the model form in the dual space can be written as

$$\hat{y}(x) = \sum_{i-1}^{N} (\beta_i - \beta_i^*) K(x, x_i) + b. \tag{7}$$

3.2. PCA.
Pearson [86] introduced PCA as a statistical technique to simplify the complexity of high-dimensional data. This can be accomplished by the orthogonal projection of the correlated variables into uncorrelated variables. The uncorrelated variables are then known as principal components (PCs).

We tested Pearson Correlation Coefficient (PCC) for the interdependencies in the data. If y is the output and x_i is the ith observation in the dataset, the PCC is defined as

$$R_{(i)} = \frac{\text{cov}(x_i, y)}{\sqrt{\text{var}(x_i)\text{var}(y)}}$$
$$= \frac{\sum_{k=1}^{m}(x_{k,i} - \overline{x}_i)(y_k - \overline{y})}{\sqrt{\sum_{k=1}^{m}(x_{k,i} - \overline{x}_i)^2 \sum_{k=1}^{m}(y_k - \overline{y})^2}}. \tag{8}$$

Here, cov is the covariance and var indicates a variable variance.

3.3. EM and SOM Ensembles

3.3.1. EM Clustering.
EM is a probabilistic and iterative algorithm that switches between the maximization (M) and expectation (E) phases in a sequential fashion. In the E-phase, EM calculates the expected value of the likelihood function. In M-phase, however, EM obtains an estimation of the parameters to maximize the likelihood function. The parameters obtained in the M-phase are then used in a subsequent E-phase. This process is repeated until convergence occurs (i.e., it convergences to the final values of the parameters). To perform E-phase for the probabilities calculation of which input in the dataset belongs to which mixture model or cluster, EM performs the following formula:

$$p(m|n)^t = \frac{w_c^t h(X_n, \mu_m^t, \sigma_m^t)}{\sum_{k=1}^{M} w_k^t h(X_n, \mu_k^t, \sigma_k^t)}, \tag{9}$$

where h indicates the probability density function of input X_n in the dataset for the cluster m with standard deviation (SD) σ_m^t and mean μ_m^t at iteration t.

Here, normal univariate, bivariate, or multivariate probability density function may be employed and should be according to the dimensionality of data. In addition, under the constraint $\sum_{k=1}^{M} w_k^t = 1$, the allocation of data to mixture models is influenced by a weighting factor w^t. Accordingly, performing M-phase, for the maximization of the likelihood

(1) Computation of the mean feature vector, $\mu = (1/n) \sum_{t=1}^{n} x_t$ for n patterns x_t ($t = 1$ to n).
(2) Computation of the covariance matrix C using $C = (1/n) \sum_{t=1}^{n} x \{x_k - \mu\}\{x_t - \mu\}^T$ where T indicates matrix transposition.
(3) Computation of the δ_i (eigenvectors) of the covariance matrix and λ (eigenvalues) of q features using $v_i = \lambda_i \delta_i$ ($i = 1, 2, 3, \ldots q$).
(4) High-valued eigenvectors estimation.
 (i) λ_i in descending order are arranged.
(ii) Define a threshold value σ.
(iii) The number of high-valued eigenvalues (s) can be selected in order to meet the relationship $(\sum_{i=1}^{s} \lambda_i)(\sum_{i=1}^{q} \lambda_i)^{-1} \geq \sigma$.
(iv) Choose v_i that correspond to high-valued λ_i.
(5) The PCs are extracted x, $PC = H^T x$, where H presents the matrix of PCs.

ALGORITHM 1: PCA.

to the data, EM needs the following computations for each cluster:

$$\mu_m^{t+1} = \frac{\sum_{n=1}^{N} p(m|n)^t x_n}{\sum_{n=1}^{N} p(m|n)^t},$$

$$\sigma_m^{t+1} = \sqrt{\frac{\sum_{n=1}^{N} p(m|n)^t |x_n - \mu_m^{t+1}|^2}{\sum_{n=1}^{N} p(m|n)^t}}, \quad (10)$$

$$w_k^{t+1} = \frac{\sum_{n=1}^{N} p(m|n)^t}{N}.$$

The above procedure is repeated by executing E and M phases until convergence occurs.

3.3.2. SOM Clustering.
Kohonen's SOM system is an unsupervised machine learning method. Thus learning method is widely used in the visualization of complex data, image processing, speech recognition, data mining, process control, and diagnostics. Based on the characteristics of features, SOM's algorithm tries to map m-dimensional input vectors x_j to two-dimensional maps. SOM aims to reduce the dimensions of the data, which aids in the understanding of high-dimensional data. SOM by this way can present the data in similar groups. Two layers make up a basic SOM. Input space' nodes are included in the first layer, and the output space' nodes are included in the second layer. SOM's idea is to adjust the nodes to represent the distribution of the data. The nodes represent clusters that reflect the distribution of the data. The SOM algorithm starts by assigning random weights to the variables. SOM algorithm in three main stages is shown in Figure 2.

3.3.3. Cluster Ensembles.
A number of various clustering approaches form cluster ensembles to partition the initial dataset and concentrate on the enhancement of the clustering outcomes resulting from a mixture of the results of different clustering. This is performed to overcome the instability of the methods of clustering. The fusion approach begins with the clusters formed during the combination phase and determines the optimal number of clusters in the dataset based on certain predetermined criteria. Next, we describe the cluster ensemble approaches, hypergraph
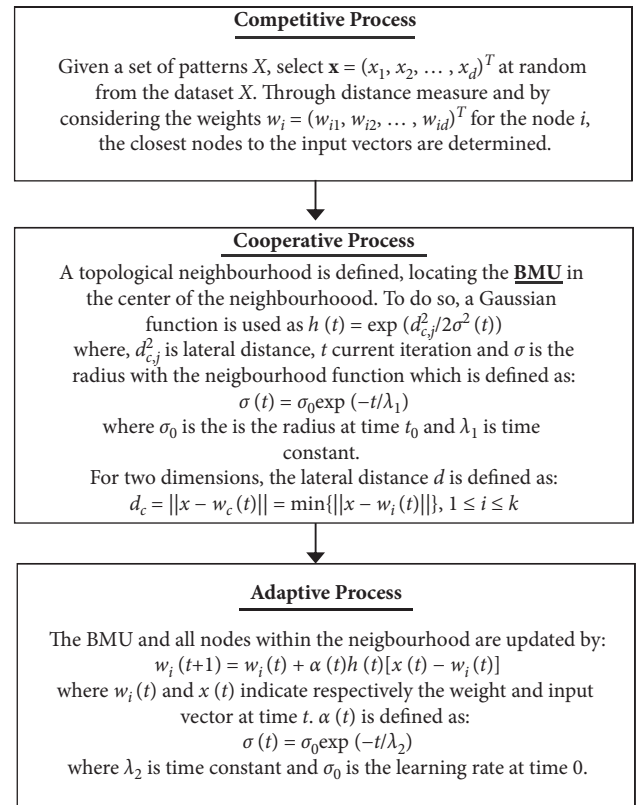

Competitive Process

Given a set of patterns X, select $\mathbf{x} = (x_1, x_2, \ldots, x_d)^T$ at random from the dataset X. Through distance measure and by considering the weights $w_i = (w_{i1}, w_{i2}, \ldots, w_{id})^T$ for the node i, the closest nodes to the input vectors are determined.

Cooperative Process

A topological neighbourhood is defined, locating the **BMU** in the center of the neighbourhoood. To do so, a Gaussian function is used as $h(t) = \exp(d_{c,j}^2/2\sigma^2(t))$ where, $d_{c,j}^2$ is lateral distance, t current iteration and σ is the radius with the neigbourhood function which is defined as:
$\sigma(t) = \sigma_0 \exp(-t/\lambda_1)$
where σ_0 is the is the radius at time t_0 and λ_1 is time constant.
For two dimensions, the lateral distance d is defined as:
$d_c = ||x - w_c(t)|| = \min\{||x - w_i(t)||\}, 1 \leq i \leq k$

Adaptive Process

The BMU and all nodes within the neigbourhood are updated by:
$w_i(t+1) = w_i(t) + \alpha(t)h(t)[x(t) - w_i(t)]$
where $w_i(t)$ and $x(t)$ indicate respectively the weight and input vector at time t. $\alpha(t)$ is defined as:
$\sigma(t) = \sigma_0 \exp(-t/\lambda_2)$
where λ_2 is time constant and σ_0 is the learning rate at time 0.

FIGURE 2: SOM algorithm.

partitioning, and majority voting. They are also called consensus functions.

3.3.4. Hypergraph Partitioning.
The cluster label vectors for hypergraph partitioning are transformed into a hypergraph image. In particular, there are vertices and hyperedges in a hypergraph. The clusters are represented in a graph as hyperlinks, whose vertices match the clustered objects. A set of objects that belong to the same group are described on every hyperlink. There are three common functions for transforming the cluster set into the representation of the hypergraph. They are CSPA and HGPA. The CSPA uses a Metis algorithm to induce an association matrix graph and cluster it [25]. A clustering in CSPA refers to a link between objects in the same segment of data and can be employed for measuring the similarity in pairs. Then, the similarity will be

deployed to recluster the items in order to generate an integrated clustering. In the HGPA algorithm, the problem of cluster ensembles is viewed as a partitioning issue of a suitably identified hypergraph, which represents the clusters. Minimal cut algorithms are mainly used to control the partition size and to find good hypergraph partitions when they are combined with the proper objective functions.

3.3.5. Majority Voting. In the majority voting mechanism, the cluster that is the one with the most votes is selected. Assume a dataset contains four Parkinson's disease patients (PDP1, PDP2, PDP3, and PDP4), and there are three clustering algorithms (Algorithms 1–3). Assume that Algorithms 1 and 2 have both assigned PDP1 to cluster A, while Algorithm 3 has assigned it to cluster B. Cluster A would then be chosen as the best cluster for PDP1 based on the majority vote.

3.4. Imputing Missing Value. In this research, we use SVD for missing values imputation. The procedure for missing value prediction by SVD is provided in Algorithm 2. Five steps are used in the SVD algorithm to predict the missing value in the PD dataset. In the first stage of the algorithm, the data is converted to a dense matrix $\mathbf{B}^{m,n}$. In the next stage, we perform a normalization procedure on the $\mathbf{B}^{m,n}$. In the third stage of the algorithm, we apply the SVD technique to the matrix produced in the second step. Then, we use matrix Z to approximate $\mathbf{Z}_d$. In the last step, the missing value is predicted.

4. Methods Evaluations and Results

In the first step of our data analysis, EM and SOM algorithms were performed on the PD dataset. We implemented EM and SOM to generate a different number of clusters to have their ensembles for final clustering results. Specifically, we run EM for $k = 8$, 10, 12, and 13 and SOM for different SOM maps, SOM2×3, SOM2×4, SOM3×3, and SOM3×4. The results of EM clustering for 13 ($k = 13$) clusters and SOM clustering for 9 clusters (SOM3×3) are provided in Figure 3. These clusters are then used for ensemble learning to be used in UPDRS prediction by the SVR technique.

SVR was trained through the RBF kernel. The SVR parameters are penalty factor C and loss function ε, and the parameter for RBF kernel is γ. These parameters can have a significant impact on the prediction quality of SVR. The radial basis function is defined as

$$
\begin{aligned}
K_{ij} &= K\left(x_i, y_j\right) \\
&= \exp\left(-\frac{\left\|x_i - y_j\right\|^2}{\gamma^2}\right),
\end{aligned}
\tag{11}
$$

where a single parameter γ in (11) indicates the spread of the function.

We used 5-fold cross-validation with a grid search mechanism to select the best parameters of each SVR. In the grid search, γ was explored on $\gamma \in [0.01$ to $0.1]$ at an interval of 0.005, loss function ε was explored on $\varepsilon \in [0.0001, 0.002]$ at an interval of 0.0001, and the penalty factor C was explored on $C \in [3, 4.1]$ at an interval of 0.005. We developed the ensembles on bootstrap samples drawn from the selected data points. The training was repeated 8 times to get 8 SVRs. To measure the performance of the presented methodology, we use several metrics such as adjusted coefficient of determination (R^2), prediction accuracy (PA), Index of Agreement (IA), MAE, and RMSE. Their formulas for n observations are shown as follows:

$$
\mathbf{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\text{Predicted}_i - \text{Observed}_i\right)^2}{n}},
$$

$$
\mathbf{MAE} = \frac{\sum_{i=1}^{N}\left|\text{Predicted}_i - \text{Observed}_i\right|^2}{n},
$$

$$
\mathbf{Index\ of\ Agreement\ (IA)} = 1 - \left[\frac{\sum_{i=1}^{N}\left(\text{Predicted}_i - \text{Observed}_i\right)^2}{\sum_{i=1}^{N}\left(\text{Predicted}_i - \overline{\text{Observed}_i}\right|^2| + \left\|\text{Observed}_i - \overline{\text{Observed}_i}\right)^2}\right]
$$

$$
\mathbf{prediction\ accuracy} = \frac{\sum_{i=1}^{N}\left(\text{Predicted}_i - \overline{\text{Observed}_i}\right)^2}{\sum_{i=1}^{N}\left(\text{Observed}_i - \overline{\text{Observed}_i}\right)^2},
$$

$$
\mathbf{R}^2_{\mathbf{adjusted}} = 1 - \left(1 - \frac{\sum_{i=1}^{N}\left(\text{Observed}_i - \overline{\text{Observed}}\right)\left(\text{Predicted}_i - \overline{\text{Predicted}}\right)}{\sqrt{\sum_{i=1}^{N}\left(\text{Observed}_i - \text{Observed}\right)} \times \sqrt{\left(\text{Predicted}_i - \overline{\text{Predicted}}\right)}}\right)\left(\frac{n-1}{n-m-1}\right)
$$

$$
= 1 - \left(1 - R^2\right)\left(\frac{n-1}{n-m-1}\right),
$$

$$
\tag{12}
$$

Step 1: The data in each cluster is converted to the dense matrix $\mathbf{B}^{m,n}$.
Step 2: The normalization procedure is performed on $\mathbf{B}^{m,n}$ through
$\mathbf{Z}_{ij} = B_{ij} - \overline{B^j}/\sigma_j$,
where $\overline{B^j}$ is the average value and σ_j indicate the SD for B^j which are calculated by
$\overline{B^j} = 1/m \sum_{i=1}^{m} B_{ij}$, $\sigma_j^2 = 1/m - 1 \sum_{i=1}^{m} (B_{ij} - \overline{B^j})^2$
Step 3. The SVD technique is applied to $\mathbf{Z}$.
Step 4. $\mathbf{Z}_d$ is computed for the approximation of $\mathbf{Z}$.
Step 5. The missing value is provided using $\overline{B^j} + \sigma_j (\mathbf{Z}_d)_{ij}$.

ALGORITHM 2: The procedure for missing value prediction by SVD.
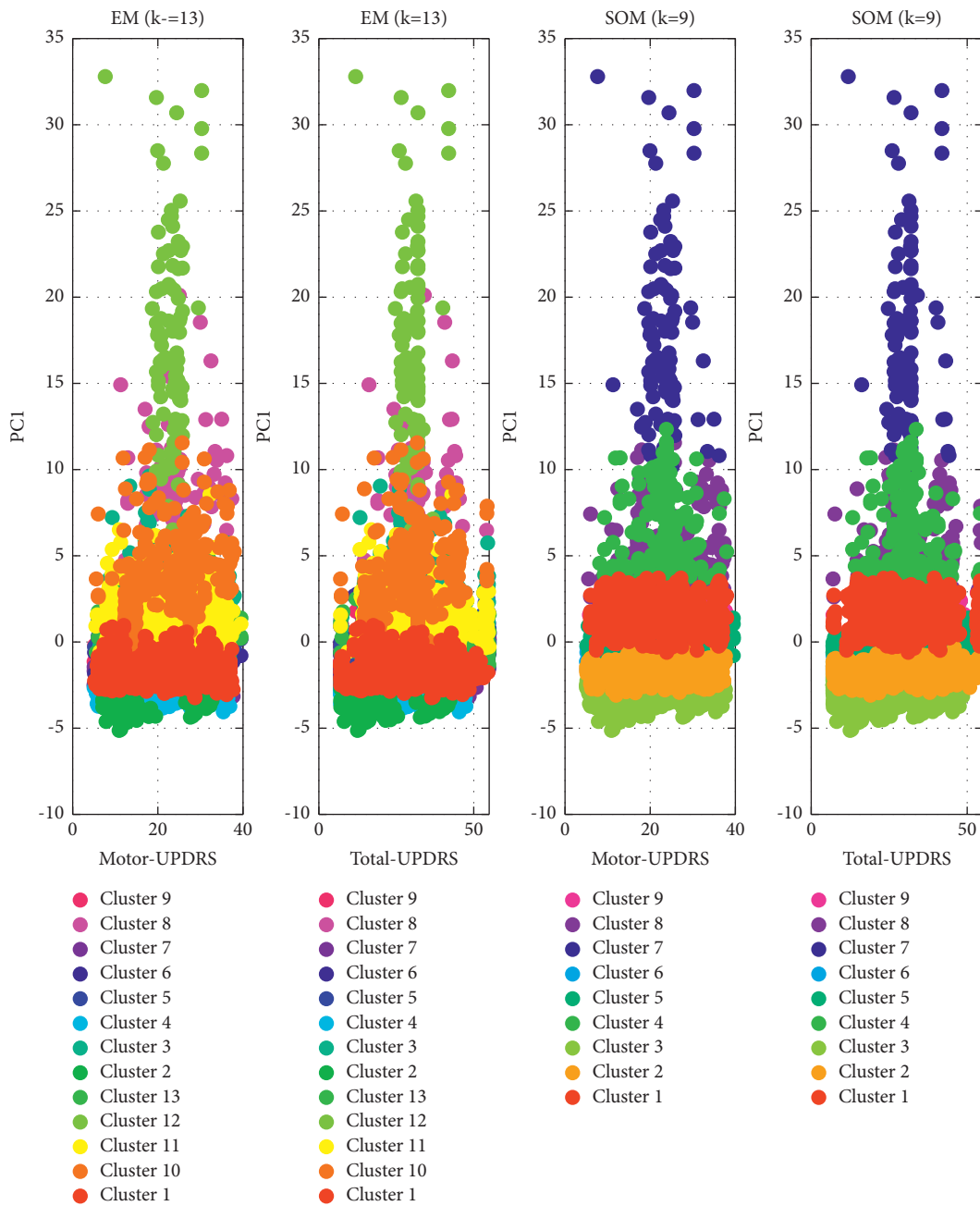
FIGURE 3: EM and SOM clusters.

TABLE 3: SOM and EM ensembles by CSPA, HGPA, and majority voting for UPDRS prediction.

Ensemble size (SOM)	Ensemble technique	RMSE	MAE	IA	PA	$R^2_{adjusted}$
	Motor-UPDRS					
$2(SOM2 \times 3 + SOM2 \times 4)$	CSPA	0.5980	0.4437	0.9208	0.9155	0.8999
	HGPA	0.5960	0.4416	0.9238	0.9200	0.9019
	Majority voting	0.5720	0.4152	0.9306	0.9265	0.9078
$3(SOM2 \times 4 + SOM3 \times 3 + SOM3 \times 4)$	CSPA	0.5620	0.4150	0.9309	0.9266	0.9085
	HGPA	0.5540	0.4116	0.9335	0.9277	0.9139
$4(SOM2 \times 3 + SOM2 \times 4 + SOM3 \times 3 + SOM3 \times 4)$	CSPA	0.5772	0.4287	0.9283	0.9261	0.9067
	HGPA	0.5756	0.4286	0.9287	0.9263	0.9070
	Total-UPDRS					
$2(SOM2 \times 3 + SOM2 \times 4)$	CSPA	0.6053	0.4463	0.9141	0.9101	0.8872
	HGPA	0.6016	0.4432	0.9166	0.9104	0.8928
	Majority voting	0.5756	0.4269	0.9207	0.9173	0.9043
$3(SOM2 \times 4 + SOM3 \times 3 + SOM3 \times 4)$	CSPA	0.5700	0.4250	0.9230	0.9203	0.9043
	HGPA	0.5565	0.4179	0.9289	0.9240	0.9058
$4(SOM2 \times 3 + SOM2 \times 4 + SOM3 \times 3 + SOM3 \times 4)$	CSPA	0.5853	0.4395	0.9182	0.9138	0.8983
	HGPA	0.5799	0.4376	0.9193	0.9140	0.9026
Ensemble size (EM)	Ensemble technique	RMSE	MAE	IA	PA	$R^2_{adjusted}$
	Motor-UPDRS					
$2(k = 8,10)$	CSPA	0.6122	0.4557	0.9103	0.9080	0.8904
	HGPA	0.5974	0.4521	0.9118	0.9092	0.8963
	Majority voting	0.5897	0.4414	0.9149	0.9112	0.8998
$3(k = 8,10,12)$	CSPA	0.5795	0.4397	0.9156	0.9117	0.9010
	HGPA	0.5789	0.4287	0.9173	0.9140	0.9028
$4(k = 8,10,12,13)$	CSPA	0.5623	0.4171	0.9257	0.9198	0.9049
	HGPA	0.5594	0.4165	0.9319	0.9254	0.9130
	Total-UPDRS					
$2(k = 8,10)$	CSPA	0.6137	0.4574	0.9085	0.9049	0.8837
	HGPA	0.6086	0.4531	0.9092	0.9071	0.8848
	Majority voting	0.5921	0.4480	0.9120	0.9081	0.8872
$3(k = 8,10,12)$	CSPA	0.5912	0.4466	0.9135	0.9100	0.8904
	HGPA	0.5846	0.4398	0.9173	0.9101	0.8933
$4(k = 8,10,12,13)$	CSPA	0.5788	0.4370	0.9216	0.9180	0.8957
	HGPA	0.5665	0.4186	0.9281	0.9209	0.9018

where $\overline{\text{Predicted}}$ is the mean of the predicted values, $\overline{\text{Observed}}$ is the mean of the observed values, $S_{\text{Predicted}}$ is the standard deviation of the predicted values, S_{Observed} is the standard deviation of the observed values, and m is the number of independent variables.

In Table 3, we present the performances of EM and SOM ensembles by majority voting, CSPA, and HGPA. The results are provided for different numbers of clusters and their ensembles. From the results, it is found that, on average, SOM and EM ensembles by the use of the HGPA approach perform best, providing the highest rates of R^2, PA, IA, MAE, and RMSE. The best accuracies are provided by ensemble size 3 (SOM2 × 4 + SOM3 × 3 + SOM3 × 4) for SOM and ensemble size 4 ($k = 8$, 10, 12, 13) for EM. In addition, majority voting has provided the lowest accuracy for EM and SOM ensembles in UPDRS prediction.

To evaluate the deployed methodology compared with previous methods in the literature, we performed several experiments on the PD dataset and presented the results in Table 4. The proposed method was compared with the SVR, ANFIS [87], MLR, HSLSSVR [88], neural network (NN), and DBN [89]. These perdition learning methods were compared with the HGPA + SOM + SVR ensemble and

HGPA + EM + SVR ensemble as they provided better accuracy compared with other ensemble learning approaches. The results are provided for RMSE, MAE, and R^2. The results reveal that ensemble learning methods, HGPA + SOM + SVR ensemble (Motor-UPDRS: MAE = 0.5540; RMSE = 0.4116; $R^2 = 0.9139$; Motor-UPDRS: MAE = 0.5565; RMSE = 0.4179; $R^2 = 0.9058$) and HGPA + SOM + SVR ensemble (Motor-UPDRS: MAE = 0.5594; RMSE = 0.4165; $R^2 = 0.9130$; Motor-UPDRS: MAE = 0.5665; RMSE = 0.4186; $R^2 = 0.9018$), can significantly impact the accuracy of both Total-UPDRS and Motor-UPDRS prediction. In addition, as shown in Table 4, it is found that clustering and dimensionality reduction techniques have significantly reduced the computation time of the proposed method for UPDRS prediction.

To assess the performance of the deployed method on the PD dataset with null values, we randomly considered 10% of patients' data as null values and predicted them with the SVD algorithm. The dense matrices are then used in EM and SOM for clustering and ensemble learning. We finally predict the UPDRS using the dataset with predicted values. The results are provided for HGPA + SOM + SVR ensemble and HGPA + EM + SVR ensemble in Figure 4. The results

TABLE 4: Methods' comparisons.

Method	Measure	MAE	RMSE	R^2	Computation time (ms)
NN	Motor-UPDRS	0.977	2.3836	0.7191	1072250
	Total-UPDRS	0.951	2.3135	0.7343	1043529
MLR	Motor-UPDRS	0.997	2.4142	0.6972	8953573
	Total-UPDRS	0.987	2.3911	0.7094	8845565
SVR	Motor-UPDRS	0.721	1.4942	0.8143	6743563
	Total-UPDRS	0.689	1.4526	0.8192	6633586
ANFIS	Motor-UPDRS	0.771	1.7047	0.7854	1534643
	Total-UPDRS	0.743	1.6062	0.7984	1525675
HSLSSVR	Motor-UPDRS	—	0.8158	—	—
	Total-UPDRS	—	0.8004	—	—
SOM + SVR	Motor-UPDRS	0.6340	0.5921	0.8518	538643
	Total-UPDRS	0.6421	0.6039	0.8421	535623
DBN	Motor-UPDRS	0.7645	1.6112	0.7914	974246
	Total-UPDRS	0.7321	1.5744	0.7996	964633
HGPA + SOM + SVR ensemble	Motor-UPDRS	0.5540	0.4116	0.9139	417435
	Total-UPDRS	0.5565	0.4179	0.9058	394352
HGPA + EM + SVR ensemble	Motor-UPDRS	0.5594	0.4165	0.9130	372223
	Total-UPDRS	0.5665	0.4186	0.9018	363422
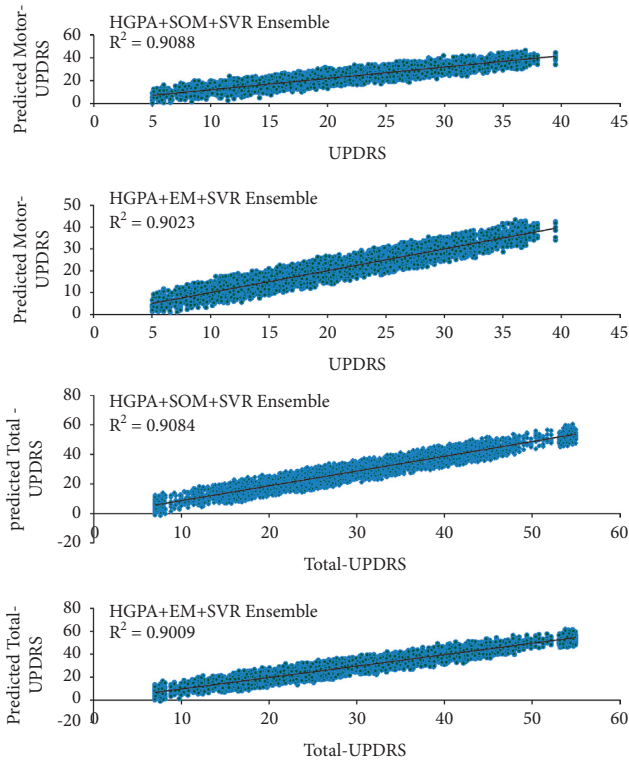
FIGURE 4: Results of the method on predicted null values for UPDRS prediction.

demonstrated that the SVD algorithm has been effective in null values prediction.

5. Discussion

Efficient detection of PD is crucial, as timely diagnosis and appropriate medication can delay the development of symptoms and difficulties resulting from this disorder [90]. Despite the significance of fast detection of PD, it is not an easy task because current detection measures are usually based on subjective indicators [91]. Besides, in the initial phases of PD, nonmotor signs, like depression, sleep disorder, and rapid eye movement, are more recognized than motor signs, which impacts the fast detection of the PD [92,93].

ML has been used for medical disease detection lately and particularly PD treatment [51]. This can be explained by the convenient performance and accurate results of ML techniques [94]. Classification of diseases is a significant type of predictive modeling. It is considered an important data mining approach because it clusters the population referring to a predetermined criterion. It is vital to compare the outcomes of various classification methods to decide which approach presents the best performance [95]. Hence, the main goal of this research is to assess several approaches that are utilized for PD prediction and classification. Even though ML methods have been assessed in several studies separately, the evaluation of these methods based on various datasets makes it complex to perform an accurate comparison among the deployed methodologies. Hence, it is vital to evaluate these methods in one comparative study based on a chosen dataset.

In this study, in the first step of data analysis, EM and SOM algorithms were implemented to produce various numbers of segments to have their ensembles for final clustering outcomes. Following that, the resulting clusters were used for ensemble learning UPDRS prediction by the SVR technique. Following that, the results are provided for different numbers of clusters and their ensembles. Referring to the outcomes, we can conclude that the performance of SOM and EM ensembles by the use of HGPA is the best among the deployed approaches based on R^2, PA, IA, MAE, and RMSE measures. Besides, the evaluation of the proposed methodology in relation to previous methods in the literature was conducted based on various experiments on the PD dataset. The result of the deployed approach was compared with other approaches (SVR, ANFIS, MLR,

HSLSSVR, and NN). Based on the findings, the HGPA + EM + SVR ensemble provided better accuracy compared with the other ensemble learning approaches.

6. Conclusion

Most of the presented methods for PD prediction depend strongly on human proficiency [96]. The benefits of deploying the ML in the medical sector are that they provide objective, context-independent, and data-driven analysis [97]. ML approaches have been utilized effectively in disease diagnosis and severity prediction [27,42,54,61,77–79,98–100]. Particularly, ML has also been utilized in analyzing the data collected from wearable IMU sensors for automated evaluation of motor disorders like PD [101–103]. Hence, the practical aim of this study entails providing supplementary, quick, and accurate methods that can aid experts in reaching more objective medical decisions considering the PD diagnosis. By deploying these methods in the appropriate systems, several gains can be acquired that entail reducing the expenses of manual diagnosis and minimizing diagnosis time.

Continuing this line of research and supporting previous literature, this study uses both unsupervised and supervised learning techniques to diagnose PD through UPDRS prediction. Besides, clustering, dimensionality reduction, and prediction learning techniques are used to create the PD diagnosis method. The basic aim of this paper is to conduct comparative research of the ML approaches for PD diagnosis. We concentrated on clustering and prediction learning methods to conduct the comparative study. Particularly, several clustering approaches for PD data segmentation and SVR ensembles to predict Motor-UPDRS and Total-UPDRS were used. The findings are then evaluated based on other prediction learning methods, MLR, neurofuzzy, and SVR techniques based on a real-world PD dataset. The finding of the study indicated the superiority of deploying EM with SVR ensembles in relation to decision trees, neurofuzzy and SVR combined with other clustering approaches in the prediction of Motor-UPDRS and Total-UPDRS.

Many previous works have been conducted focusing on patients' classifications, severity prediction, and remote monitoring. Still, there are future routes in each field to be investigated. Besides, several sensors such as magnetometer, accelerometer, and gyroscope have been utilized and assessed. Additionally, MRI, EEG signals, f-MRI, and DATSCAN images were utilized to present accurate predictions of the disease. Other research directions can be followed by utilizing other brain signal images such as ECG, EMG, and PCG. Other sensing modalities can be explored and combined to present a more accurate classification of the disease.

Even though ML methods in previous literature have presented high classification accuracy for PD detection, still, there are some obstacles related to feature extraction and selection which need to be addressed [104]. The utilization of several features can increase the computation time [105,106]. On the other hand, if fewer features were utilized, this will increase the complexity of extracting the features, which will accordingly impact the computation time. This paper has some shortcomings which should be considered in future research. The study is based on a real-world dataset to assess the proposed approaches, which has one limitation considering the number of features used in the prediction process. Other PD datasets with a larger number of features can be utilized in the evaluation of the deployed approaches. Large datasets can present more generalized outcomes. Emerging technologies can be used to collect data from patients using particular applications, as suggested by Bot et al. [107], in which the authors developed an application to collect the data from PD patients using their iPhones. This approach can ease the data collection from the public because of the availability of smartphones and help to present more generalizable outcomes. Furthermore, this study can be extended by incremental machine learning approaches to improve the computation time of previous PD diagnosis methods in processing large datasets [44,89].

Data Availability

The data are freely available at https://archive.ics.uci.edu/ml/datasets.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] S. Bharath, M. Hsu, D. Kaur, S. Rajagopalan, and J. K. Andersen, "Glutathione, iron and Parkinson's disease," *Biochemical Pharmacology*, vol. 64, pp. 1037–1048, 2002.

[2] R. Cacabelos, "Parkinson's disease: from pathogenesis to pharmacogenomics," *International Journal of Molecular Sciences*, vol. 18, p. 551, 2017.

[3] S. Engelender and O. Isacson, "The threshold theory for Parkinson's disease," *Trends in Neurosciences*, vol. 40, pp. 4–14, 2017.

[4] G. Natale, L. Pasquali, S. Ruggieri, A. Paparelli, and F. Fornai, "Parkinson's disease and the gut: a well known clinical association in need of an effective cure and explanation," *Neuro-Gastroenterology and Motility*, vol. 20, pp. 741–749, 2008.

[5] L. Krohn, F. P. Grenn, M. B. Makarious et al., "Comprehensive assessment of PINK1 variants in Parkinson's disease," *Neurobiology of Aging*, vol. 91, pp. 168–161, 2020.

[6] Ö. Eskidere, F. Ertaş, and C. Hanilçi, "A comparison of regression methods for remote tracking of Parkinson's disease progression," *Expert Systems with Applications*, vol. 39, pp. 5523–5528, 2012.

[7] M. Nilashi, O. Bin Ibrahim, A. Mardani, A. Ahani, and A. Jusoh, "A soft computing approach for diabetes disease

classification," *Health Informatics Journal*, vol. 24, pp. 379–393, 2018a.

[8] A. Tsanas, M. Little, P. McSharry, and L. Ramig, "Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests," *Nature Precedings*, vol. 57, no. 4, pp. 884–893, 2009.

[9] C. G. Goetz, W. Poewe, and O. Rascol, "Movement disorder society task force on rating scales for Parkinson's disease: the unified Parkinson's disease rating scale (UPDRS): status and recommendations," *Movement Disorders*, vol. 18, pp. 738–750, 2003.

[10] G. Moguel-Cobos, C. Saldivar, P. W. Goslar, and H. A. Shill, "The relationship between social anxiety disorder and motor symptoms of Parkinson disease: a pilot study," *Psychosomatics*, vol. 61, pp. 321–326, 2020.

[11] M. Nilashi, O. Ibrahim, S. Samad, H. Ahmadi, L. Shahmoradi, and E. Akbari, "An analytical method for measuring the Parkinson's disease progression: a case on a Parkinson's telemonitoring dataset," *Measurement*, vol. 136, pp. 545–557, 2019b.

[12] J. Waring, C. Lindvall, and R. Umeton, "Automated machine learning: review of the state-of-the-art and opportunities for healthcare," *Artificial Intelligence in Medicine*, vol. 104, Article ID 101822, 2020.

[13] S. M. Lundberg, B. Nair, M. S. Vavilala et al., "Explainable machine-learning predictions for the prevention of hypoxaemia during surgery," *Nature biomedical engineering*, vol. 2, pp. 749–760, 2018.

[14] W. M. Marella, E. Sparnon, and E. Finley, "Screening electronic health record–related patient safety reports using machine learning," *Journal of Patient Safety*, vol. 13, pp. 31–36, 2017.

[15] S. Saria, D. Koller, and A. Penn, "Learning individual and population level traits from clinical temporal data," in *Proceedings of Neural Information Processing Systems*, pp. 1–9, Citeseer, Princeton, NY, USA, December 2010.

[16] C.-C. Kuo, C.-M. Chang, K.-T. Liu et al., "Automation of the kidney function prediction and classification through ultrasound-based kidney imaging using deep learning," *NPJ digital medicine*, vol. 2, pp. 1–9, 2019.

[17] H. Liang, B. Y. Tsui, H. Ni et al., "Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence," *Nature Medicine*, vol. 25, pp. 433–438, 2019.

[18] J. S. Rumsfeld, K. E. Joynt, and T. M. Maddox, "Big data analytics to improve cardiovascular care: promise and challenges," *Nature Reviews Cardiology*, vol. 13, p. 350, 2016.

[19] D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, "Big data in health care: using analytics to identify and manage high-risk and high-cost patients," *Health Affairs*, vol. 33, pp. 1123–1131, 2014.

[20] W.-H. Lo-Ciganic, J. M. Donohue, J. M. Thorpe et al., "Using machine learning to examine medication adherence thresholds and risk of hospitalization," *Medical Care*, vol. 53, p. 720, 2015.

[21] A. T. Özdemir and B. Barshan, "Detecting falls with wearable sensors using machine learning techniques," *Sensors*, vol. 14, pp. 10691–10708, 2014.

[22] G. Arji, H. Ahmadi, M. Nilashi et al., "Fuzzy logic approach for infectious disease diagnosis: a methodical evaluation, literature and classification," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 4, pp. 937–955, 2019.

[23] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big data from healthcare communities," *Ieee Access*, vol. 5, pp. 8869–8879, 2017.

[24] A. Mardani, R. E. Hooker, S. Ozkul et al., "Application of decision making and fuzzy sets theory to evaluate the healthcare and medical problems: a review of three decades of research with recent developments," *Expert Systems with Applications*, vol. 137, pp. 202–231, 2019.

[25] M. Nilashi, H. Ahmadi, L. Shahmoradi, O. Ibrahim, and E. Akbari, "A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique," *Journal of infection and public health*, vol. 12, pp. 13–20, 2019c.

[26] M. Nilashi, H. Ahmadi, L. Shahmoradi, A. Mardani, O. Ibrahim, and E. Yadegaridehkordi, "Knowledge discovery and diseases prediction: a comparative study of machine learning techniques," *Journal of Soft Computing and Decision Support Systems*, vol. 4, pp. 8–16, 2017.

[27] M. Nilashi, H. Ahmadi, L. Shahmoradi, M. Salahshour, and O. Ibrahim, "A soft computing method for mesothelioma disease classification," *Journal of Soft Computing and Decision Support Systems*, vol. 4, pp. 16–18, 2017.

[28] M. Nilashi, O. bin Ibrahim, H. Ahmadi, and L. Shahmoradi, "An analytical method for diseases prediction using machine learning techniques," *Computers & Chemical Engineering*, vol. 106, pp. 212–223, 2017.

[29] M. Nilashi, O. Ibrahim, H. Ahmadi, and L. Shahmoradi, "A knowledge-based system for breast cancer classification using fuzzy logic method," *Telematics and Informatics*, vol. 34, pp. 133–144, 2017.

[30] B. Sabeena, S. Sivakumari, and P. Amudha, "A technical survey on various machine learning approaches for Parkinson's disease classification," *Materials Today Proceedings*, 2020.

[31] T. R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. K. R. Maddikunta, and G. Srivastava, "Deep neural networks to predict diabetic retinopathy," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–14, 2020.

[32] S. Bhattacharya, P. K. R. Maddikunta, S. Hakak et al., "Antlion re-sampling based deep neural network model for classification of imbalanced multimodal stroke dataset," *Multimedia Tools and Applications*, 2020.

[33] H. Abe, K. Ashizawa, F. Li et al., "Artificial neural networks (ANNs) for differential diagnosis of interstitial lung disease: results of a simulation test with actual clinical cases1," *Academic Radiology*, vol. 11, pp. 29–37, 2004.

[34] J. Jeyaranjani, T. D. Rajkumar, and T. A. Kumar, "Coronary heart disease diagnosis using the efficient ANN model," *Materials Today Proceedings*, 2021.

[35] G. Z. Papadakis, A. H. Karantanas, M. Tsikankis, A. Tsatsakis, D. A. Spandidos, and K. Marias, "Deep learning opens new horizons in personalized medicine," *Biomedical Reports*, vol. 10, pp. 215–217, 2019.

[36] K. Paydar, S. R. N. Kalhori, M. Akbarian, and A. Sheikhtaheri, "A clinical decision support system for prediction of pregnancy outcome in pregnant women with systemic lupus erythematosus," *International Journal of Medical Informatics*, vol. 97, pp. 239–246, 2017.

[37] S. Punitha, F. Al-Turjman, and T. Stephan, "An automated breast cancer diagnosis using feature selection and parameter optimization in ANN," *Computers & Electrical Engineering*, vol. 90, Article ID 106958, 2021.

[38] R. K. Yadav, "PSO-GA based hybrid with Adam Optimization for ANN training with application in Medical Diagnosis," *Cognitive Systems Research*, vol. 64, pp. 191–199, 2020.

[39] E. S. Berner, *Clinical Decision Support Systems*, Springer, New York, NY, USA, 2007.

[40] N. Mantel, "The detection of disease clustering and a generalized regression approach," *Cancer Research*, vol. 27, pp. 209–220, 1967.

[41] K. B. Nahato, K. H. Nehemiah, and A. Kannan, "Hybrid approach using fuzzy sets and extreme learning machine for classifying clinical datasets," *Informatics in Medicine Unlocked*, vol. 2, pp. 1–11, 2016.

[42] M. Nilashi, O. Ibrahim, M. Dalvi, H. Ahmadi, and L. Shahmoradi, "Accuracy improvement for diabetes disease classification: a case on a public medical dataset," *Fuzzy Information and Engineering*, vol. 9, pp. 345–357, 2017.

[43] R. Das, I. Turkoglu, and A. Sengur, "Diagnosis of valvular heart disease through neural networks ensembles," *Computer Methods and Programs in Biomedicine*, vol. 93, pp. 185–191, 2009a.

[44] R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles," *Expert Systems with Applications*, vol. 36, pp. 7675–7680, 2009b.

[45] A. Ortiz, J. Munilla, J. M. Gorriz, and J. Ramirez, "Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease," *International Journal of Neural Systems*, vol. 26, Article ID 1650025, 2016.

[46] M. Nilashi, H. Ahmadi, A. A. Manaf et al., "Coronary heart disease diagnosis through self-organizing map and fuzzy support vector machine with incremental updates," *International Journal of Fuzzy Systems*, vol. 22, pp. 1376–1388, 2020.

[47] P. Ghaderyan and G. Fathi, "Inter-limb time-varying singular value: a new gait feature for Parkinson's disease detection and stage classification," *Measurement*, vol. 177, Article ID 109249, 2021.

[48] G. Paragliola and A. Coronato, "Gait anomaly detection of subjects with Parkinson's disease using a deep time series-based approach," *IEEE Access*, vol. 6, pp. 73280–73292, 2018.

[49] E. Balaji, D. Brindha, V. K. Elumalai, and R. Vikrama, "Automatic and non-invasive Parkinson's disease diagnosis and severity rating using LSTM network," *Applied Soft Computing*, vol. 108, Article ID 107463, 2021.

[50] R. W. de Souza, D. S. Silva, L. A. Passos et al., "Computer-assisted Parkinson's disease diagnosis using fuzzy optimum-path forest and Restricted Boltzmann Machines," *Computers in Biology and Medicine*, vol. 131, Article ID 104260, 2021.

[51] Z. K. Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Medical Hypotheses*, vol. 138, Article ID 109603, 2020.

[52] M. De Vos, J. Prince, T. Buchanan, J. J. FitzGerald, and C. A. Antoniades, "Discriminating progressive supranuclear palsy from Parkinson's disease using wearable technology and machine learning," *Gait & Posture*, vol. 77, pp. 257–263, 2020.

[53] F. Mohammed, X. He, and Y. Lin, "An easy-to-use deep-learning model for highly accurate diagnosis of Parkinson's disease using SPECT images," *Computerized Medical Imaging and Graphics*, vol. 87, 2021.

[54] J. Goyal, P. Khandnor, and T. C. Aseri, "A hybrid approach for Parkinson's disease diagnosis with resonance and time-frequency based features from speech signals," *Expert Systems with Applications*, Article ID 115283, 2021.

[55] S. Xu and Z. Pan, "A novel ensemble of random forest for assisting diagnosis of Parkinson's disease on small handwritten dynamics dataset," *International Journal of Medical Informatics*, vol. 144, Article ID 104283, 2020.

[56] L. C. Ribeiro, L. C. Afonso, and J. P. Papa, "Bag of samplings for computer-assisted Parkinson''s disease diagnosis based on recurrent neural networks," *Computers in Biology and Medicine*, vol. 115, Article ID 103477, 2019.

[57] A. Parziale, R. Senatore, A. Della Cioppa, and A. Marcelli, "Cartesian genetic programming for diagnosis of Parkinson disease through handwriting analysis: performance vs. interpretability issues," *Artificial Intelligence in Medicine*, vol. 111, Article ID 101984, 2021.

[58] M. Tsuda, S. Asano, Y. Kato, K. Murai, and M. Miyazaki, "Differential diagnosis of multiple system atrophy with predominant parkinsonism and Parkinson's disease using neural networks," *Journal of the Neurological Sciences*, vol. 401, pp. 19–26, 2019.

[59] N. Khoury, F. Attal, Y. Amirat, L. Oukhellou, and S. Mohammed, "Data-driven based approach to aid Parkinson's disease diagnosis," *Sensors*, vol. 19, p. 242, 2019.

[60] S. Çimen and B. Bolat, "Diagnosis of Parkinson's disease by using ANN," in *Proceedings of the Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), 2016 International Conference on*, pp. 119–121, IEEE, Jalgaon, India, December 2016.

[61] M. Nilashi, O. Ibrahim, and A. Ahani, "Accuracy improvement for predicting Parkinson's disease progression," *Scientific Reports*, vol. 6, Article ID 34181, 2016.

[62] D. Avci and A. Dogantekin, "An Expert diagnosis system for Parkinson disease based on genetic algorithm-wavelet kernel-extreme learning machine," *Parkinson's disease*, vol. 2016, Article ID 5264743, 9 pages, 2016.

[63] M. Nilashi, O. Ibrahim, H. Ahmadi, L. Shahmoradi, and M. Farahmand, "A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques," *Biocybernetics and Biomedical Engineering*, vol. 38, pp. 1–15, 2018.

[64] L. Parisi, N. RaviChandran, and M. L. Manaog, "Feature-driven machine learning to improve early diagnosis of Parkinson's disease," *Expert Systems with Applications*, vol. 110, pp. 182–190, 2018.

[65] J. Prince and M. De Vos, "A deep learning framework for the remote detection of Parkinson's disease using smart-phone sensor data," in *Proceeedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3144–3147, IEEE, Honolulu, HI, USA, July 2018.

[66] N. Zou and X. Huang, "Empirical Bayes transfer learning for uncertainty characterization in predicting Parkinson's disease severity," *IISE Transactions on Healthcare Systems Engineering*, vol. 8, pp. 209–219, 2018.

[67] S. Grover, S. Bhartia, A. Yadav, and K. Seeja, "Predicting severity of Parkinson's disease using deep learning," *Procedia Computer Science*, vol. 132, pp. 1788–1794, 2018.

[68] R. Prashanth, S. D. Roy, P. K. Mandal, and S. Ghosh, "High-accuracy detection of early Parkinson's disease through multimodal features and machine learning," *International Journal of Medical Informatics*, vol. 90, 2016.

[69] R. H. Abiyev and S. Abizade, "Diagnosing Parkinson's diseases using fuzzy neural system," *Computational and mathematical methods in medicine*, vol. 2016, Article ID 1267919, 9 pages, 2016.

[70] G. Singh, M. Vadera, L. Samavedham, and E. C.-H. Lim, "Machine learning-based framework for multi-class

diagnosis of neurodegenerative diseases: a study on Parkinson's disease," *Age*, vol. 65, pp. 64–94, 2016.

[71] S. Shetty and Y. Rao, "SVM based machine learning approach to identify Parkinson's disease using gait analysis,"vol. 2, pp. 1–5, in *Proceedings of the Inventive Computation Technologies (ICICT)*, vol. 2, IEEE, Coimbatore, India, August 2016.

[72] H. Ozkan, "A comparison of classification methods for telediagnosis of Parkinson's disease," *Entropy*, vol. 18, p. 115, 2016.

[73] G. Pahuja and T. Nagabhushan, "A comparative study of existing machine learning approaches for Parkinson's disease detection," *IETE Journal of Research*, pp. 1–11, 2018.

[74] E. Rovini, C. Maremmani, A. Moschetti, D. Esposito, and F. Cavallo, "Comparative motor pre-clinical assessment in Parkinson's disease using supervised machine learning approaches," *Annals of Biomedical Engineering*, vol. 46, pp. 2057–2068, 2018.

[75] M. Nilashi, O. Ibrahim, S. Samad, H. Ahmadi, L. Shahmoradi, and E. Akbari, *An Analytical Method for Measuring the Parkinson's Disease Progression*, Vol. 136, A Case on a Parkinson's Telemonitoring Dataset, , Irvine, CA, USA, 2019a.

[76] A. S. Ashour, A. El-Attar, N. Dey, H. Abd El-Kader, and M. M. Abd El-Naby, "Long short term memory based patient-dependent model for FOG detection in Parkinson's disease," *Pattern Recognition Letters*, vol. 131, pp. 23–29, 2020.

[77] N. Ahmadi, M. Nilashi, S. Samad, T. A. Rashid, and H. Ahmadi, "An intelligent method for iris recognition using supervised machine learning techniques," *Optics & Laser Technology*, vol. 120, Article ID 105701, 2019.

[78] H. Kahramanli and N. Allahverdi, "Design of a hybrid system for the diabetes and heart diseases," *Expert Systems with Applications*, vol. 35, pp. 82–89, 2008.

[79] K. B. Nahato, K. N. Harichandran, and K. Arputharaj, "Knowledge mining from clinical datasets using rough sets and backpropagation neural network," *Computational and mathematical methods in medicine*, vol. 2015, 13 pages, 2015.

[80] H. Ahmed, E. M. Younis, A. Hendawi, and A. A. Ali, "Heart disease identification from patients' social posts, machine learning solution on Spark," *Future Generation Computer Systems*, 2019.

[81] D. Pumo, G. Carlino, S. Blenkinsop, E. Arnone, H. Fowler, and L. V. Noto, "Sensitivity of extreme rainfall to temperature in semi-arid Mediterranean regions," *Atmospheric Research*, vol. 225, pp. 30–44, 2019.

[82] M. Nilashi, D. Jannach, O. bin Ibrahim, and N. Ithnin, "Clustering- and regression-based multi-criteria collaborative filtering with incremental updates," *Information Sciences*, vol. 293, pp. 235–250, 2015.

[83] H. M. Gomes, J. P. Barddal, F. Enembreck, and A. Bifet, "A survey on ensemble learning for data stream classification," *ACM Computing Surveys*, vol. 50, pp. 1–36, 2017.

[84] Y. Peng, "A novel ensemble machine learning for robust microarray data classification," *Computers in Biology and Medicine*, vol. 36, pp. 553–573, 2006.

[85] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[86] K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, pp. 559–572, 1901.

[87] J.-S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE transactions on systems, man, and cybernetics*, vol. 23, pp. 665–685, 1993.

[88] Y.-P. Zhao, B. Li, Y.-B. Li, and K.-K. Wang, "Householder transformation based sparse least squares support vector regression," *Neurocomputing*, vol. 161, pp. 243–253, 2015.

[89] M. Nilashi, H. Ahmadi, A. Sheikhtaheri et al., "Remote tracking of Parkinson's disease progression using ensembles of deep Belief network and self-organizing map," *Expert Systems with Applications*, Article ID 113562, 2020b.

[90] D. L. Murman, "Early treatment of Parkinson's disease: opportunities for managed care," *American Journal of Managed Care*, vol. 18, p. S183, 2012.

[91] J. Jankovic, "Parkinson's disease: clinical features and diagnosis," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 79, pp. 368–376, 2008.

[92] K. R. Chaudhuri, D. G. Healy, and A. H. Schapira, "Non-motor symptoms of Parkinson's disease: diagnosis and management," *The Lancet Neurology*, vol. 5, pp. 235–245, 2006.

[93] K. R. Chaudhuri and A. H. Schapira, "Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment," *The Lancet Neurology*, vol. 8, pp. 464–474, 2009.

[94] Z. Rayan, M. Alfonse, and A.-B. M. Salem, "Machine learning approaches in smart health," *Procedia Computer Science*, vol. 154, pp. 361–368, 2019.

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

[96] C.-W. Cho, W.-H. Chao, S.-H. Lin, and Y.-Y. Chen, "A vision-based analysis system for gait recognition in patients with Parkinson's disease," *Expert Systems with Applications*, vol. 36, pp. 7033–7039, 2009.

[97] B. M. Eskofier, S. I. Lee, J. F. Daneault et al., "Recent machine learning advancements in sensor-based mobility analysis: deep learning for Parkinson's disease assessment," *Institute of Electrical and Electronics Engineers Inc*, vol. 2016, pp. 655–658, 2016.

[98] J. Goyal, P. Khandnor, and T. C. Aseri, "Classification, prediction, and monitoring of Parkinson's disease using computer assisted technologies: a comparative analysis," *Engineering Applications of Artificial Intelligence*, vol. 96, 2020.

[99] C. Taleb, L. Likforman-Sulem, C. Mokbel, and M. Khachab, *Detection of Parkinson's Disease from Handwriting Using Deep Learning: A Comparative Study*, Evolutionary Intelligence, Germany, 2020.

[100] Y. Xiong and Y. Lu, "Deep feature extraction from the vocal vectors using sparse autoencoders for Parkinson's classification," *IEEE Access*, vol. 8, pp. 27821–27830, 2020.

[101] C. Ahlrichs and M. Lawo, "Parkinson's disease motor symptoms in machine learning: a review," 2013, https://arxiv.org/abs/1312.3825.

[102] M. Bachlin, M. Plotnik, D. Roggen et al., "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, pp. 436–446, 2009.

[103] A. Salarian, H. Russmann, F. J. Vingerhoets et al., "Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 1434–1443, 2004.

[104] H. Khachnaoui, R. Mabrouk, and N. Khlifa, "Machine learning and deep learning for clinical data and PET/SPECT imaging in Parkinson's disease: a review," *IET Image Processing*, 2020.

[105] E. Abdulhay, N. Arunkumar, K. Narasimhan, E. Vellaiappan, and V. Venkatraman, "Gait and tremor investigation using

machine learning techniques for the diagnosis of Parkinson disease," *Future Generation Computer Systems*, vol. 83, pp. 366–373, 2018.

[106] J. S. Almeida, P. P. Rebouças Filho, T. Carneiro et al., "Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques," *Pattern Recognition Letters*, vol. 125, pp. 55–62, 2019.

[107] B. M. Bot, C. Suver, E. C. Neto et al., "The mPower study, Parkinson disease mobile data collected using ResearchKit," *Scientific Data*, vol. 3, pp. 1–9, 2016.