CROWD: Crow Search and Deep Learning based Feature Extractor for Classification of Parkinson's Disease

MEHEDI MASUD, Taif University, KSA
PARMINDER SINGH, GURJOT SINGH GABA, and AVINASH KAUR, Lovely Professional University, India
ROOBAEA ALROBAEA ALGHAMDI, Taif University, KSA
MUBARAK ALRASHOUD and SALMAN ALI ALQAHTANI, King Saud University, KSA

Edge Artificial Intelligence (AI) is the latest trend for next-generation computing for data analytics, particularly in predictive edge analytics for high-risk diseases like Parkinson's Disease (PD). Deep learning learning techniques facilitate edge AI applications for enhanced, real-time handling of data. Dopamine is the cause of Parkinson's that happens due to the interference of brain cells that produce the substance to regulate the communication of brain cells. The brain cells responsible for generating the dopamine perform adaptation, control, and movement with fluency. Parkinson's motor symptoms appear on the loss of 60% to 80% of cells, due to the non-production of appropriate dopamine. Recent research found a close connection between the speech impairment and PD. Many researchers have developed a classification algorithm to identify the PD from speech signals. In this article, Adaptive Crow Search Algorithm (ACSA) and Deep Learning (DL)–based optimal feature selection method are introduced. The proposed model is the combination of CROW Search and Deep learning (CROWD) stack sparse autoencoder neural network. Parkinson's dataset is taken for the experiment from the Irvine dataset repository at the University of California (UCI). In the first phase, dataset cleaning is performed to handle the missing values in the dataset. After that, the proposed ACSA algorithm is employed to find the scrunched feature vector. Furthermore, stack spare autoencoder with seven hidden layers is employed to generate the compressed feature vector. The performance of the proposed CROWD autoencoder model is compared with three feature selection approaches for six supervised classification techniques. The experiment result demonstrates that the performance of the proposed CROWD autoencoder feature selection model has outperformed the benchmarked feature selection techniques: (i) Maximum Relevance (mRMR) (ii) Recursive Feature Elimination (RFE), and (iii) Correlation-based Feature Selection (CFS), to classify Parkinson's disease. This research has significance in the healthcare sector for the enhancement of classification accuracy up to 0.96%.

CCS Concepts: • **Computing methodologies** → Machine learning approaches; Logical and relational learning; Search methodologies; Model development and analysis;

Additional Key Words and Phrases: Deep learning, artificial Intelligence, feature extraction, Parkinson disease

Authors' addresses: M. Masud (corresponding author) and R. A. Alghamdi, Taif University, College of Computers and Informationk Technology, Taif-AlHeiwah, KSA, 26513; emails: {mmasud, r.robai}@tu.edu.sa; P. Singh (corresponding author), G. S. Gaba, and A. Kaur, Lovely Professional University, Phagwara, Punjab, India, 144411; emails: {parminder.16479, gurjot.17023, avinash.14557}@lpu.co.in; M. Alrashoud and S. A. AlQahtani, King Saud University, Department of Software Engineering, College of Computer and Information Sciences, Riyadh, KSA, 11543; emails: {malrashoud, salmanq}@ksu.edu.sa.

1533-5399/2021/05-ART77 $15.00
https://doi.org/10.1145/3418500

ACM Reference format:
Mehedi Masud, Parminder Singh, Gurjot Singh Gaba, Avinash Kaur, Roobaea Alrobaea Alghamdi, Mubarak Alrashoud, and Salman Ali AlQahtani. 2021. CROWD: Crow Search and Deep Learning based Feature Extractor for Classification of Parkinson's Disease. *ACM Trans. Internet Technol.* 21, 3, Article 77 (May 2021), 18 pages.
https://doi.org/10.1145/3418500

1 INTRODUCTION

IoT technology in the healthcare domain is advancing fast. It is highly expected that in 2020 this trend will expand further in healthcare, such as privacy of patients, imaging medical data, monitoring of patients, will experience the results of edge IoT and edge AI in greater degree [13]. Edge AI is the latest commuting trend in IoT-enabled healthcare domain for data analytics [25], mainly patient monitoring, real-time care, disease prediction and classification, and high-risk disease analysis, like Parkinson's disease [16, 18]. Parkinson's disease is a moderately expanding neurodegenerative brain sickness that damages brain cells [9]. The cells of the brain create **Dopamine Chemical (DC)** and are concentrated at specific sectors known as **Substantia Nigra (SN)**. DC carries signals from the Substantia Nigra to other brain sectors, which govern the motion of the body. Dopamine helps the human body to make smooth and coordinated movements of the organs. The motor symptoms of Parkinson's disease may happen if more than half of the cells (typically, 60% to 80%) producing dopamine are lost [7].

The premature signs of Parkinson's disease usually emerge at the lower part of the brain stem, olfactory tracts, and enteric nervous system. Subsequently, the disease spreads to the SN and the brain shell. The other symptoms observed by the affected person are loss of the sense of smell, constipation, disturbed sleep, tremor, slow movements, and so on. Therefore, researchers are working on finding out ways to identify these non-motor symptoms that originate earlier than the motor symptoms. According to the study in Reference [29], 90% of Parkinson's disease patients experience speech impairment problems. Hence, there is a need for techniques that can determine the presence of Parkinson's disease using non-motor signs, e.g., speech impairments.

Machine Learning (ML) approach is now widely considered for diagnosis of medical diseases because of its convenient execution and higher precision. Due to the aforementioned characteristics, ML is recommended for the identification of Parkinson's disease. Wan et al. [33] carried out a comprehensive review of the articles covering **Feature Selection (FS)** and ML. ML is being used to detect the actual area to be treated during brain surgery. The authors also focused on the post-diagnosis issues of Parkinson's disease. Salmanpour et al. [30] applied an ML approach for predicting the cognitive results of Parkinson's disease.

Cavallo et al. [6] proposed a prediction model for Parkinson's disease using the motion data collected from the upper limbs of the subject. The research is carried out on two sets of people: Parkinson's affected and healthy ones. The experimental subjects were made to tie a gadget on their upper limbs and instructed to perform various actions. Spatiotemporal and frequency data analysis is utilized to acquire data. Thereafter, numerous learning methods are applied to accomplish the classification process. In another approach [2], several feature extraction and ML methods have been applied to identify Parkinson's disease. Authors observed that phonation is the most convenient parameter to identify Parkinson's disease. The authors applied **Support Vector Machine (SVM)**, Optimum Path Forest, K-NN, and **Multilayer Perceptron (MLP)** classifiers. **Artificial Neural Networks (ANN)** is considered to reduce the verbal impairments, whereas SVM is used for the classification to aid ML-based discovery of Parkinson's disease.

Recently, the authors in Reference [20] introduced an unsupervised method to detect Parkinson's disease. They applied a **Self-Organizing Map (SOM)** as well as incremental **Support Vector Regression (SVR)** for the purpose of clustering and prediction, respectively. In the proposed approach SOM and SVR techniques are executed after the reduction of dimension using **Partial Least Squares (PLS)** and **Unified Parkinson's Disease Rating Scale (UPDRS)** is forecasted. In another approach [23], the authors considered fuzzy-based C-means clustering algorithm to evaluate weights of features and used k-NN classifier technique for detection of Parkinson's disease. The K-NN classifier used a weighted Parkinson's disease dataset with different values of k to identify the best k value.

Wang et al. [34] have utilized paramount learning apparatus for Parkinson's disease diagnosis. They used a weighted scheme and a nonlinear mapping of kernel function to structure the imbalanced data. **Artificial Bee Colony (ABC)** algorithm is employed for selecting the features and optimizing the variables. Singh et al. [31] in their proposed scheme have used **Fisher Discriminant Ratio (FDR)**, **Principle Component Analysis (PCA)**, and SVM for feature selection, dimension optimization, and classification, respectively. The proposed hybrid model resulted in a successful diagnosis of Parkinson's disease. Abdulhay et al. [1] took into consideration tremor and gait symptoms for the detection of the PD. They acquired the physical signals using the sensors placed under the feet of the subjects and analyzed the electrical pulses to extract abnormal peaks. The gait features are retrieved from the raw data extracted from the PhysioNet database. Their approach resulted over 92% classification accuracy.

Yaman et al. [38] utilized voice inputs from the common dataset. Feature augmentation has been applied to identify the most prominent features for detecting the illness. They selected approximately 66 features to classify Parkinson's disease. Another interesting study is carried out in Reference [15], where they discussed the possibility of detecting Parkinson's disease through handwriting inspection instead of examining MRI, motion, or voice data.

Through literature study it is realized that machine learning–based determination methods of Parkinson's disease have high classification rates. Though ML is advantageous in determining Parkinson's illness, it faces certain challenges in realization. If a great number of features are considered to detect Parkinson's disease, then the identification rate is high but at the cost of increased resource consumption, i.e., computation cost and execution time, whereas fewer features may not produce reliable results [31, 38]. A tradeoff exists between the reliability and the computation cost.

In the few recent researches, it is proven that the use of limited features, lightweight feature extraction model, and a less complex classifier can extensively decrease the computation time. In addition, it is also summarized that features extraction from the speech signals are comparatively easier than the Magnetic Resonance Imaging–based [31] or motion-based [6] approaches. Few researchers [2, 23] have used a large number of vocal features to enhance the classification rates. In contrast, higher classification accuracy is possible by extracting features from MRI images but the method is more complicated than the one involving speech analysis.

In this article, the operative learning model is used to identify the certain patterns from the voiced features. These features are applied to predict normal or affected cases. A competent pipeline has been designed to generate compressed features in the PD dataset [26] obtained from UCI repository to record the Parkinson's disease vocal dataset. The dataset has been pre-processed, because it consists of missing values and noise. Further, Crow Search Algorithm [3], a metaheuristic global search optimization scheme, has been extended to design the proposed **Adaptive Crow Search Algorithm (ACSA)** to determine the candidate's feature set. To generate the compressed features for the classification of Parkinson's disease, a sparse autoencoder is used on the features of the candidate. The proposed CROWD autoencoder aims to generate compressed features to increase the classification algorithm's speed and accuracy. This is used to extract the latent depiction

from the features, which is then given as an input to the machine learning algorithms. Moreover, K-Fold cross-validation technique is considered in the training and assessment phase. To evaluate the performance of the models, validation metrics are considered. The results of the proposed CROWD autoencoder provides optimal performance on the classification of Parkinson's disease. Furthermore, this study emphasizes the significance of supervised classification techniques with proposed feature selection models in the healthcare sector for the advanced diagnosis of a disease.

The main contribution of the proposed CROWD autoencoder feature extraction model is listed below:

- The swarm intelligence is used at the first phase to select the scrunched feature vector using Adaptive Crow Search Algorithm (ACSA).
- The deep learning–based autoencoder is applied to generate the compressed feature vector for classification techniques.
- The combination of ACSA algorithm and autoencoder is integrated with supervised learning techniques to improve the classification accuracy.
- The proposed CROWD autoencoder model is evaluated with various benchmarked feature selection models.
- The proposed CROWD autoencoder model is evaluated for the performance enhancement with benchmarked classification techniques.
- The proposed approach is better than other approaches with respect to **Root Mean Square Error (RMSE)** and F1-score to generate 36 features after ACSA algorithm in the first phase and seven compressed features in the second phase of the autoencoder.

The article is structured as follows: Section 2 discusses related work. Section 3 presents the dataset that is used in this study, feature selection using ACSA, and a mechanism of generating compressed features using the proposed CROWD autoencoder model. Section 4 shows the outcomes of the CROWD autoencoder and discusses detailed analysis with comparison to the existing models. Section 5 summarizes the article and highlights the importance of the proposed approach.

2 RELATED WORK

This section discusses traditional schemes that are precisely linked with this study.

A research study used a tele-determination and surveillance scheme to determine the vocal impairments to predict Parkinson's disease [28]. The scheme takes into consideration heterogeneous data (vowel, sentence, and word) retrieved from the speech of Parkinson's-affected patient. The authors addressed two major problems: the predictive analysis of the various types of the PD and the competence of different statistical measures (e.g., central tendency) from the recordings. **Neural Networks (NNs)** have also been used simultaneously in conjunction with the feed-forward structure to ensure reliability in the prediction of Parkinson's disease.

Åström and Koker [4] focused on reducing the prediction error of the Parkinson's disease prediction model. A rule-based scheme is used to assess the results obtained from the NN. The training of the machines involves collecting the untaught information individually and feeding it to the subsequent group for training. Besides, this structure executes considerably well on heterogeneous datasets. Data mining schemes are usually recommended on structured data for making accurate predictions. In the proposed approach [37], three methods such as statistical, tree-based, and SVM are applied to segregate the samples of Parkinson's affected and controlled cases. Optimization of the prediction accuracy is another significant consideration of this approach. Apart from this, the aforementioned approach also involves an **Artificial Neural Network (ANN), Support Vector Machine (SVM)** algorithms, and **Decision Support System (DSS)**. ANN, SVM, and DSS are adopted to benefit the medical practitioners in the diagnosis of the disease at less expenditure.

Computer-Aided Diagnosis (CAD) is widely accepted because of its structured processing, compatibility over multiple systems, and accurate predictions on different disease conditions. Moreover, for the diagnosis of Parkinson's disease, a CAD model can be designed considering selection of features and the algorithms for classification. In most approaches, SVMs are utilized as a classifier to find appropriate features. These features are then fed to an RF model, followed by an ensemble classifier to optimize the prediction process [22]. Further, stimulated by intelligent swarm techniques, swarm algorithms are adopted as global search optimization methods. These models are mostly applied for detecting the attributes from the lower dimensional datasets to identify the discriminative patterns in the input data.

An optimized cuttlefish algorithm is proposed to make an optimized decision on the choice of the feature subset [11]. The selection of the features is carried out via an optimized search methodology accompanied by the steps of the traditional cuttlefish algorithm. The performance evaluation of different features has been performed using the k-NN and decision tree algorithms. Furthermore, the proposed scheme provides higher prediction accuracy considering reduced feature size.

In another research by Nilashi et al. [19], **Expectation-Maximization (EM)** and PCA have been used as a hybrid model to predict and keep track of the progression of the PD. Hybridization of the processes increases the efficiency of the learning methodologies, as they possess the characteristics of the two individual methods. The intelligence is governed by eliminating the noise, analyzing the cluster, followed by the utilization of efficient classifier methodologies. The authors demonstrated the use of EM and PCA to resolve the multi-collinearity issues of the datasets. An **Adaptive Neuro-Fuzzy Intelligent System (ANFIS)** and an SVM regression model have been applied for monitoring the advancement and prediction of Parkinson's disease, respectively.

In analytical research, an **Electroencephalogram (EEG)** signal is used to identify brain abnormality, which is further inferred for detecting Parkinson's disease [21]. Their approach consists of a 13-layered **Convolution Neural Network (CNN)** architecture for the automatic extraction of the features and classification process. The presented approach delivered favorable outputs and has been considered as a significant method to be utilized at hospitals for assisting the medical staff in the determination of the PD.

Camps et al. [5] discussed the debilitating symptoms of motor of the PD by freezing the gait. Occurrence of this symptom of PD in an affected person may lead to fall or loss of independence. The authors proposed a method to determine the PD in a person using a waist-fused inertial measurement system. This system is used to accumulate the data of a subject's movement, possible falling gestures, and so on. The received information is executed and converted into predictive factors with the help of the CNN model. These predictive factors determine the presence of PD in a person along with intensity.

A deep neural network–based system employs the Tensorflow library for determining the intensity of the PD [10]. PD assessment has been conducted remotely and aided by the DL mechanism for efficient classification. The samples of Alternate Finger Tapping are collected from various remote areas using smartphones. The experimental assessment of machine learning and deep learning is carried out separately. The efficiency of the deep learning models is found more in contrast to conventional ML processes. In the latest research, a new hybrid system is designed using **Synthetic Minority Over-Sampling Technique (SMOTE)** with a blending of random forest. The dataset has a class balancing problem that is solved using SMOTE, and for classification process a random forest model is used [24].

Yanhao et al. [36] studied the vocal features of Parkinson's disease-affected individuals using advanced computerized processes. At an earlier stage, the samples are pre-processed to filter out redundant values. Further, the **Augmented Grey Wolf Optimizer (AGWO)** [17] is used to recognize the predictor data subset from the processed vocal features. The AGWO is a global search

Table 1. Description PD Dataset

Name	Description
Type of Disease	Neurodegenerative (Parkinson's Disease)
Repository	Irvine Machine Learning, University of California
No. of Features	754
No. of Observations	756
Classes	Binary (1-PD, 0-Control)

optimization technique based on a meta-heuristic approach. Eventually, the patient's features are obtained using sparse autoencoders for the proper distinction between Parkinson's disease sick and controlled cases. To classify the features, six supervised ML algorithms were implemented. The presented scheme is processed with a 10-fold cross-validation data split scheme. The validation matrix is further considered to examine the performance. The experimental analysis uses Parkinson's disease dataset from the Irvine Machine Learning Repository, University of California (UCI). The results reveal that the proposed approach [36] performs better than the standard models; it proves its efficiency in determining the affected and healthy samples of Parkinson's disease separately.

J. P. Queralta et al. [25] proposed a system combining Edge computing, IoT, and deep learning algorithms for remotely observing health supervising tasks. Authors mainly show the possibility and usefulness of this system using a use case of fall identification applying recurrent neural networks. The system is implemented using sensor nodes and Edge gateway connected to cloud services to provide end-user support.

Hossain et al. [14] proposed an emotion recognition model based on edge cloud preserving privacy. The model utilized IoT devices to acquire speech signals and facial images from users. The model performs ordinal signal pre-processing mechanism in the edge devices and then transmits to the main cloud. The main cloud applied a convolutional neural network model for the purpose of features extraction from the image and speech signals. Despite the several available approaches for predicting and classifying Parkinson's disease, there is a requirement of the more efficient scheme that can perform the classification at least cost with the use of minimum features while attaining maximum accuracy. The proposed scheme optimizes the classification process and attains a good level of accuracy with minimum computations.

3 MATERIAL AND METHOD

3.1 Dataset

The dataset used for the experimental evaluation of the present study was collected from Irvine Machine Learning repository, University of California [26]. The dataset consists of recording of the voice signals for biomedical voice measurement. The dataset consists of 188 people, 81 women and 107 men candidates. The participant age varies between 33 years to 87 years. There are 64 people aged from 41 to 82 from healthy group, out of which 41 are women and 23 are men. The dataset contains 754 columns and 756 rows, as shown in Figure 1. Table 1 describes the detail of the dataset. The goal is to find the people affected by PD. The "1" value in the status column denotes a person has PD and "0" value denotes a person is healthy.

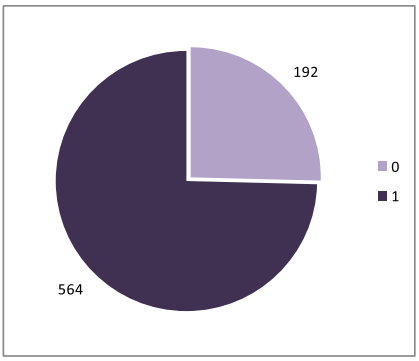

Fig. 1. Parkinson's dataset class distribution.

3.2 Feature Selection

The complexity of a machine learning model can be reduced to an extent with proper feature selection. There are different scenarios needed for the set of techniques, such as wrapper, embedded, and filter methods. A hybrid approach combines these techniques. The relevant features are adopted from the data with the help of a filtering technique [32]. The role of the wrapper technique is used for performance evaluation of the subset at each iteration. This is a greedy approach but still having a lot of advantages instead of cost for computation. To perform the selection, a famous technique is penalization. This helps to improve the selection of prediction parameters. Other popular techniques are Elastic Net, LASSO (L1 Regularization), and so on [27].

The optimal feature selection can be performed with nature-inspired algorithms; these algorithms monitor the action of natural proceedings with a different set of conditions. Some of the nature-inspired algorithms are **Harmony Search Algorithm (HSO), Firefly Optimization Algorithm (FFO), Bee Colony Optimization (BCO), Ant Colony Optimization (ACO), Cuckoo Search Algorithm (CSA), Particle Swarm Optimization (PSO)**, and so on. [39]. All these algorithms apply the nature of insects, animals, or living things. In this article, the parameter selection is performed using the **Adaptive Crow Search Algorithm (ACSA)**. The process of crow search is shown in Figure 2.

CSA is motivated by the crows' hiding and stealing techniques of food. Crows are one of the most intelligent birds in the world. The intelligence factors of the crows are: They hide food throughout the seasons, they use natural language to communicate with each other, tool utilization, remember faces, and mirror test for self precognition. To avoid themselves from being a victim, crows always move their hiding places. The crow is famous for hiding and steals food. A well-known technique is used for this purpose, also called intelligent thieves.

As per the study on crow behavior [3], CSA principles are:

(1) Living in flock.
(2) Memorize hiding locations.
(3) Stealing food by following peers.
(4) Guard caches from thief crows.

3.3 Adaptive Crow Search Optimization

ACSA is presented to extract the usable features from the dataset. There are 754 attributes in the dataset and out of which we are going to find the scrunched feature vector V. Furthermore,

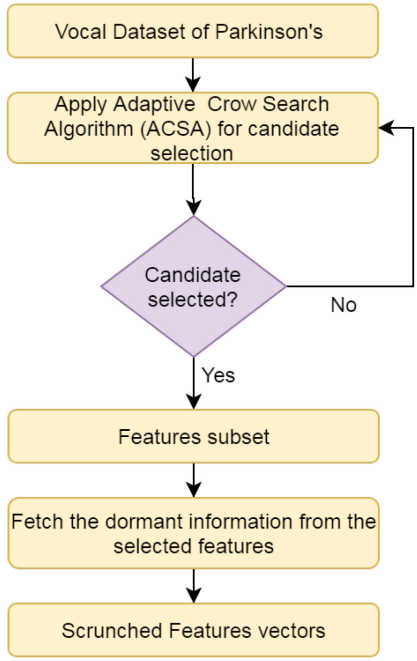

Fig. 2. Feature extraction using Adaptive Crow Search Algorithm (ACSA).

these features are input to the 4-tier classification approach for the prediction of PD. Our ACSA algorithm uses local minima values for the prevention of getting stuck. Algorithm 1 depicts the proposed ACSA.

The steps of the ACSA are as follows:

(1) Declaration of input to the algorithm defined in the input section, where association of 754 attributes (DA) of the dataset with each crow (C), Flight Length (F), Population ($Flock$), Size (S), Awareness Probability (A). In the output section, the expected output of scrunched feature vector (V) is defined.
(2) First assign the values to C, A, F, T_m, S. The initialization of crow's position and memory is performed (lines 1–2) based on random values.
(3) For loop executes the statements from (lines 3–27) upto T_m.
(4) Choose a random crow for the chasing (lines 5–6). The update of crow position is based on two rules: Crow is aware someone is chasing or crow is not aware of it. If the crow is aware, then the decision is taken as per lines 7–8.
(5) Otherwise, the position p of the crow is updated by Sigmoid function (lines 10–11). Further, the binary conversion is performed. If the new position is smaller, then assign 0. If the random number is greater than the new position, then assign value 1,
(6) For loop, lines 14–19 are used for two purposes. First, calculate the number of crows with value 1 using a $Length$ function. If the number N is greater or equal to $\frac{d}{2}$, then append the index of feature in scrunched feature vector V (lines 15–17).
(7) Another for loop (lines 20–26) is employed to update the memory $M_{i,j}$ with $p_{i,j}$ if $DA_{i,j}$ is equal to $p_{i,j}$.
(8) The new fitness values are calculated with fitness function taking memory M as an argument (line 27).

ALGORITHM 1: Adaptive crow search algorithm for feature selection

Input: Association of 754 attributes (DA) of dataset with each crow (C), Flight Length (F), Population (Flock) Size (S), Awareness Probability (A)
Output: Scrunched features vector (V)

1: Assign the values to C, A, F, T_m, S
2: Set the starting position of crow p
3: **for** i in 1:T_m **do**
4: **for** j in 1:C **do**
5: Randomly select a crow
6: Randomly select the value of R_j
7: **if** $A \leq R_j$ **then**
8: $p_{j,i+1} \leftarrow p_{j,i+1} + R_j \times F \times (M_{j,i} - p_{j,i})$
9: **else**
10: $p_{j,i} \leftarrow Random\ Location$
11: $p_{j,i+1} \leftarrow \frac{1}{1 + exponential((p_{j,i+1} - 0.9))}$ ▷ Binary conversion of position p
12: **end if**
13: **end for**
14: **for** j in 1:C **do**
15: $N \leftarrow Length(p_j, 1)$ ▷ Find the no. of p with value 1
16: **if** $\frac{S}{2} \leq N$ **then**
17: V.add(j) ▷ Append the crow index in Scrunched features vector V
18: **end if**
19: **end for**
20: **for** i in 1:C **do**
21: **for** j in 1:S **do**
22: **if** $DA_{i,j} == p_{i,j}$ **then**
23: $M_{i,j} \leftarrow p_{i,j}$
24: **end if**
25: **end for**
26: **end for**
27: $FT_{new} \leftarrow FT(M)$ ▷ Calculate the new fitness
28: **end for**
29: **if** Size(FT_{new}) < Size(FT) **then**
30: $FT \leftarrow FT_{new}$
31: **end if**
32: $Accuracy \leftarrow \frac{Size(FT)}{C \times 100}$
33: Return V

(9) If the size of FT is greater than the size of FT_{new}, then FT is updated with the FT_{new}.
(10) The accuracy is calculated in line 32 and returns the scrunched feature vector V.

3.4 Autoencoder

There are three important layers of every autoencoder: (i) an input layer, (ii) hidden layer, and (ii) reconstruction. Figure 3 depicts the layers. The units of these layers are n, m, and n, respectively. There is an activation function $f(.)$ between every two layers. In the encoder process, $Y \in R^h$ is

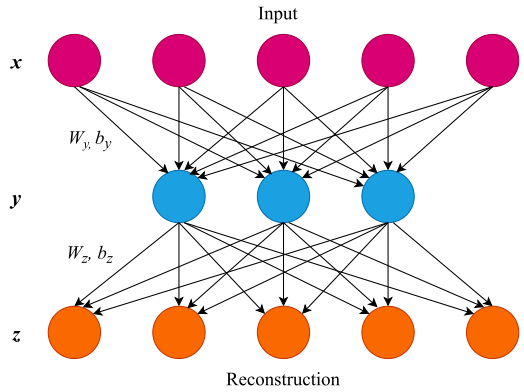

Fig. 3. Single-layer autoencoder.

produced after the mapping of hidden layers and input $X \in R^n$. In the decoder process, the output layer is mapped with Y. Overall, this process is coming under reconstruction.

$Z \in R^n$ refers to the reconstruction. The steps for the process are:

$$Y = f(b_y + W_y X), \tag{1}$$

$$Z = f(b_z + W_z Y). \tag{2}$$

W_y represents the encoder's weight and W_z refers to the decoder's weight. The output layer and hidden layers bias represent with b_z and b_y, respectively. Activation function $f(.)$ in our encoder is sigmoid function.

Once the weight of each layer is decided the b_z ,b_y, W parameters are updated. Further, the error is minimized with the $cost(X, Z)$ function. In this expression, the X is given and Z is based on b_z, b_y defined in Equations (4) and (5), respectively. The learning rate is represented with L that described the counter in each iteration while proceeding towards the minimum loss function. It plays a vital role in optimization techniques and acts as a tuning parameter. It helps to decide the change of new information over the old information and decide the speed of learning. The weight update is defined in Equation (5).

$$b_z = b_z - L\frac{d}{d.b_z}cost(X, Z), \tag{3}$$

$$b_y = b_y - L\frac{d}{d.b_y}cost(X, Z), \tag{4}$$

$$W = W - L\frac{d}{d.W}cost(X, Z). \tag{5}$$

In our approach, the reconstruction layer has been removed along with the parameters. Hidden layers of the autoencoder are used further with the logistic regression for the extraction of more features. The activity Y of the hidden layer has the potential to be the input of the next classification approach. Even if the original input is not recovered from the reconstruction layer, we can minimize the information loss from the stack encoder. The multi-layer classification approach based on the stack encoder is discussed in the next section.

3.5 Stack Sparse Autoencoder

In the previous section, we have discussed the representation of autoencoders with a single layer. The optimum score is obtained from error minimization. The basic principles of autoencoders are

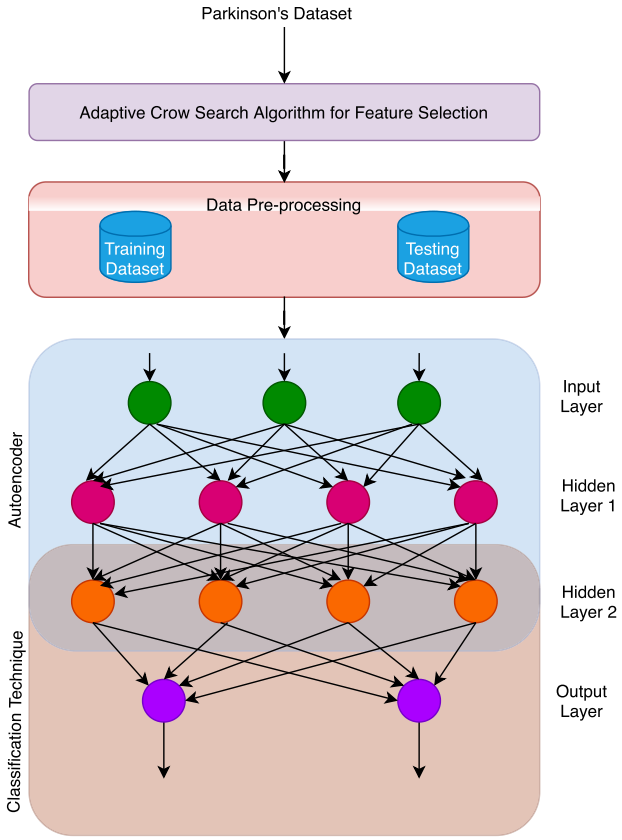

Fig. 4. Proposed CROWD autoencoder feature selection model.

used to derive the optimum scores. The four-tier architecture is used for increasing the classification accuracy of Parkinson's disease. The input data is used for optimal feature extraction and prediction. This network is embedded with a sigmoid (σ) function. The activation function is described in Equation (6).

$$A = \sigma(b + Wx) \tag{6}$$

Here, b refers to the deviation vector, and W represents the weight matrix. Further, the hidden layers of the neuron's activation function are defined in Equation (7).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

The proposed CROWD autoencoder PD feature selection model is shown in Figure 4. The training of the proposed PD classification technique is used with b_z, b_y, W parameters. The error rule must be defined before the weight update, because sigmoid function is considered as an activation function. The cost function is defined as in Equation (8).

$$Cost = -\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n}[x_{ij}\ log(z_{ij}) + (1 - x_{ij})\ log(1 - z_{ij})] \tag{8}$$

The subset size is represented with n and input size is defined as m. The jth element in input ith is represented with $x_{ij}(Z_{ij})$. Gradient method is used to optimize the cost function defined in

Equation (8). The partial differential is applied to b_y, b_z, and W. Process of reconstruction is defined in Equations (9), (10), and (11).

$$input_{ih}^{Y} = b_{yh} + \sum_{l=1}^{d} x_{il} W_{lh}, \tag{9}$$

$$input_{ib}^{Z} = b_{zb} + \sum_{q=1}^{p} W_{bq}\, f(input_{y}^{iq}), \tag{10}$$

$$f(input_{ib}^{z}) = f_{z_{ib}} = b_{zb} + \left(\sum_{h=1}^{p} W_{bh} f \left(\sum_{l=1}^{d} x_{il} W_{lq} + b_{yh} \right) \right). \tag{11}$$

The hidden layer is represented with hth and all the input is represented with $input_{ih}^{Y}$.

Softmax regression is required for neural network integration with learned features. The cost function can be modified for softmax regression. The proposed design can be combined with unsupervised learning through softmax regression. The binary classification is possible with a simple proposed CROWD autoencoder process and multi-class classification softmax is required. The scope of this study is to integrate the CROWD autoencoder with supervised classification models only. To extend this work for unsupervised classification, the output of the hidden layer is fed as input to softmax regression.

4 EXPERIMENTAL EVALUATION

In the training process, the supervised machine learning algorithms consider label data as input. However, the evaluation of the model is performed on the unlabeled data. There are many machine learning algorithms available for the classification. The performance of these algorithms varies with different datasets and circumstances. We used six algorithms for the performance evaluation of PD similar to the Xiong et al. [36] Parkinson's classification study. The algorithms are gone through the step mentioned in Figure 4. The experiment study is performed on **Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Gradient Boosting Model (GBM), Random Forest (RF), Naive Bayes (NB)**, and **Logistic Regression (LR)**. The implementation of all the algorithms is performed in Python 3.7. The implementation of sparse autoencoder is performed using tensorflow library.

4.1 Results and Analysis

The initial input to the ACSA algorithm is a total of 754 features, out of which 36 features are generated as a candidate. Stack sparse autoencoder is employed to increase the prediction performance. The input feature is compressed to seven features with the help of hidden layers. The supervised machine algorithms are further employed to check the performance of the proposed CROWD autoencoder feature selection model. The K-fold mechanism is used to evaluate the proposed model with proper training. The performance of the proposed and existing models are evaluated considering standard metrics listed below, like **Root Mean Square Error (RMSE)**, as per Equation (12), F1-Score as per Equation (13), specificity as per Equation (14), sensitivity as per Equation (15) and accuracy as per Equation (16).

(1) Root Mean Square Error (RMSE): The prediction errors or residuals in standard deviation are known as RMSE. It is used to gauge the difference of actual and predicted output. The common usage of RMSE is in regression analysis, forecasting, and climatology to evaluate

the results. RMSE can be calculated as per Equation (12).

$$RMSE = \sqrt{\sum_{j=1}^{m} \frac{(\hat{y}_j - y_j)^2}{m}} \tag{12}$$

(2) F1-Score: It is widely used to measure the accuracy of binary classification. The score is calculated using recall and precision. All similar samples that must be *true* are considered in denominator, and total correct *true* outputs are put in numerator to calculate the recall. Furthermore, the precision is the total of *true* results divided with total *true* results given by the classifier. The best value of F1-score is 1 and it can be calculated as per Equation (13).

$$F1 - Score = \frac{Recall \times Precision}{Precision + Recall} \times 2 \tag{13}$$

(3) Specificity: It is another vital metric to test binary classification. The proportion of actual $false$ cases identified is calculated as per Equation (14). The correct prediction of the healthy people's percentage is represented with specificity, also known as true negative rate.

$$Specificity = \frac{TN}{TN + FP} \tag{14}$$

(4) Sensitivity: It is also used in binary classification to test the actual positive cases identified. The percentage of Parkinson's disease-infected people is represented with sensitivity, also known as true positive rate. Sensitivity can be calculated as per Equation (15).

$$Sensitivity = \frac{TP}{TP + FN} \tag{15}$$

(5) Accuracy:The percentage of accurate classification of all the observations is known a accuracy. Equation (16) is used to calculate the accuracy.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{16}$$

The above equations refer to symbols as True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). The performance of the proposed CROWD autoencoder is described in Figure 5. The comparison of various benchmarked feature selection algorithms with proposed CROWD autoencoder for RMSE is depicted in Figure 6 and F1-score is shown in Figure 7.

Linear Discriminant Analysis (LDA) model's performance is best out of all benchmark models. The benchmark models are compared based on performance metrics. Although, the performance of other benchmark algorithms is also quite good. However, LDA outperforms in the comparison. We are able to enhance the accuracy of LDA 0.96%. This happens due to appropriate feature selection in the CROWD autoencoder process.

The line graph for RMSE and F1-score is shown in Figure 6 and Figure 7. Different feature selection techniques along with the proposed CROWD autoencoder model are compared for different classification techniques. The results demonstrate that the proposed CROWD autoencoder model has minimum RMSE and maximum F1-score. Thus, the proposed CROWD autoencoder technique claims better performance from the existing feature selection mechanisms. The supervised classification technique is applied to check the performance of each benchmark feature selection algorithm.

Moreover, the state-of-the-art techniques of feature selection are evaluated with benchmark classification techniques. We compared the proposed CROWD autoencoder model for candidate selection with **minimum Redundancy Maximum Relevance (mRMR)** [8], **Recursive Feature Elimination (RFE)** [35], and **Correlation-based Feature Selection (CFS)** [12]. mRMR model finds the 68 prominent features, RFE model extracts the 52 features, and CFS models are able to

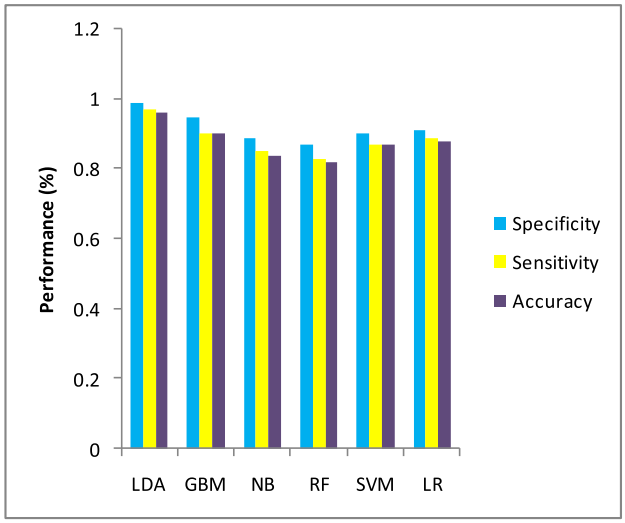

Fig. 5. Performance enhancement of benchmark models with proposed CROWD autoencoder.

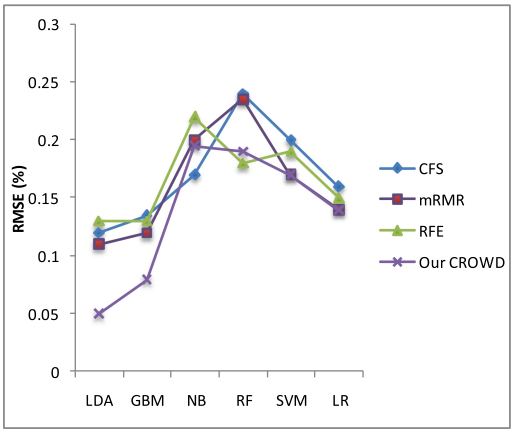

Fig. 6. Comparison of RMSE for proposed CROWD autoencoder and benchmarked feature selection models.

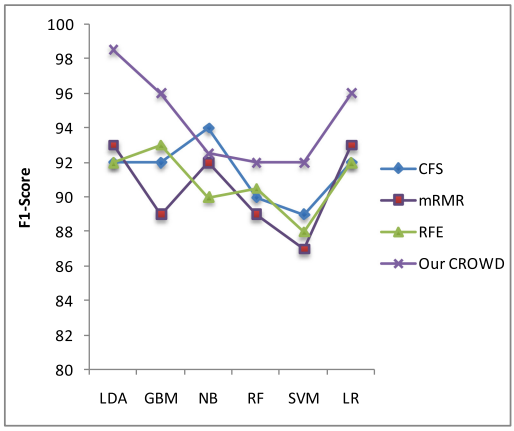

Fig. 7. Comparison of F1-score for proposed CROWD autoencoder and benchmarked feature selection models.

fetch the 47 important features from the PD dataset. We employed seven hidden layers stacked sparse encoders to fetch the scrunched feature vector from the given set. The performance comparison of mRMR, RFE, CFS, and our CROWD autoencoder model is shown in Figure 8 for the generated features from stacked sparse autoencoder. The performance enhancement of benchmarks models using mRMR-sparse autoencoder, RFE-sparse autoencoder, and CFS-sparse autoencoder are shown in Figures 9, 10, and 11, respectively.

The proposed CROWD autoencoder performs better among the existing feature selection techniques. The performance of the mRMR technique is also good from the RFE and CFS candidate feature selection technique. Figures 9, 10, and 11 depicts the performance metrics (specificity, sensitivity, and accuracy). Out of existing techniques, mRMR-sparse autoencoder is better, but still,

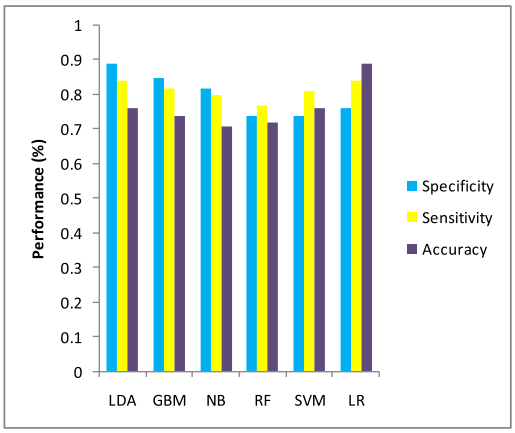

Fig. 8. Performance enhancement of benchmarks models with sparse autoencoder.

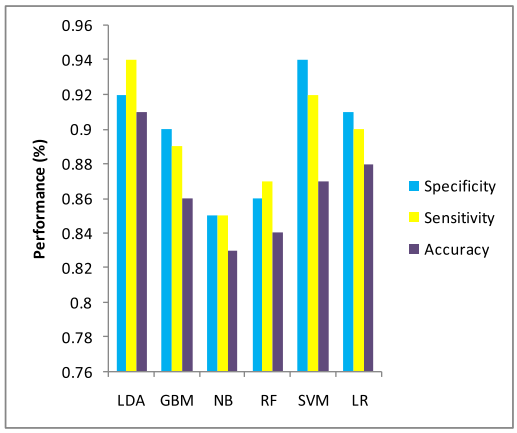

Fig. 9. Performance enhancement of benchmark models with mRMR-sparse autoencoder.

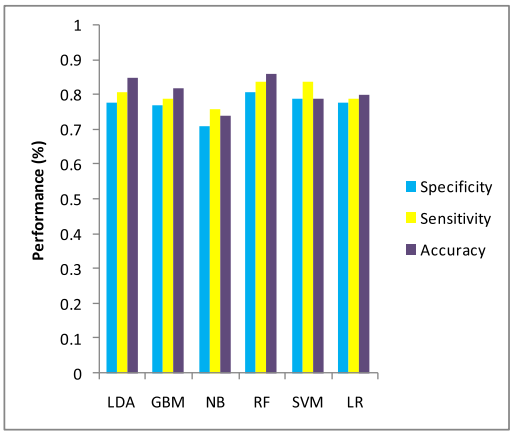

Fig. 10. Performance enhancement of benchmark models with RFE-sparse autoencoder.

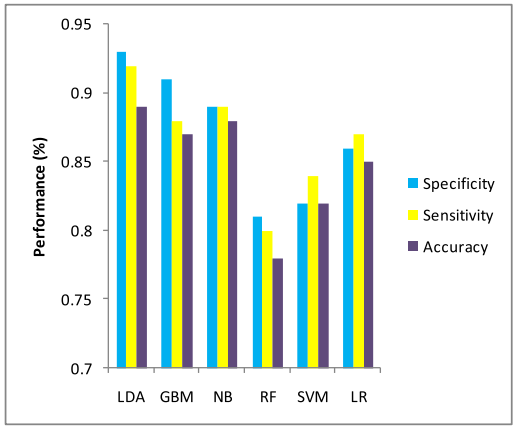

Fig. 11. Performance enhancement of benchmark models with CFS-sparse autoencoder.

our CROWD autoencoder is better from all the benchmark feature selection techniques in case of supervised learning techniques.

The performance of LDA is better among all the benchmark classification techniques. To evaluate the combined performance of CROWD autoencoder and LDA, the models are trained for 754, 37, and 7 features, respectively. The performance for all three subsets is described in Table 2. Our approach reduces the 37 features from 754 in the first phase of the ACSA algorithm; here, the accuracy achieved up to 75.4%. Further, these features are input to the stack sparse autoencoder, where the candidate features are reduced up to 7, and accuracy increased up to 96.4%. This shows that the performance of classification increased with the CROWD autoencoder feature selection model for supervised learning techniques.

Table 2. Performance of CROWD Autoencoder with Linear Discriminant Analysis

Performance Metric	No. of Candidate Feature Selected 754	36	7
Specificity	72.3	77.2	98.5
Sensitivity	74.8	86.6	97.1
Accuracy	72.2	75.4	96.4

5 CONCLUSION

This article studied the pattern in Parkinson's disease dataset to identify the scrunched feature vector for classification techniques. A superior feature selection technique is designed that runs in a pipeline.The dataset with 754 features is input to the Adaptive Crow Search Algorithm (ACSA), out of which 36 prominent features are selected. The latent features are extracted from these 36 features using deep learning techniques with a stacked sparse autoencoder. The combined model with CROW search and deep learning is proposed. The CROWD autoencoder is able to transform the inputted 36 features to 7 compressed feature vectors. The existing benchmarked feature selection techniques are compared with the proposed CROWD algorithm for the PD dataset. The performance of the proposed CROWD algorithm shows better results than the existing techniques with respect to specificity, sensitivity, and accuracy metrics. The proposed model also improves the performance of the classification technique up to 0.96%. This will aid in the diagnosis of Parkinson's disease. In the future, the same technique can be applied to identify the scrunched feature vectors in other datasets to enhance the performance of classification techniques. We have tested the proposed model for supervised learning techniques. Further, CROWD autoencoder can be extended to work with unsupervised classification techniques. To perform the likelihood prediction of Parkinson's disease, the CROWD autoencoder can be integrated with fuzzy models.

REFERENCES

[1] Enas Abdulhay, N. Arunkumar, Kumaravelu Narasimhan, Elamaran Vellaiappan, and V. Venkatraman. 2018. Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease. *Fut. Gen. Comput. Syst.* 83 (Mar. 2018), 366–373. DOI: https://doi.org/10.1016/j.future.2018.02.009

[2] Jefferson S. Almeida, Pedro P. Rebouças Filho, TiagoCarneiro, Wei Wei, Robertas Damaševičius, Rytis Maskeliūnas, and Victor Hugo C. de Albuquerque. 2019. Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. *Pattern Recog. Lett.* 125 (July 2019), 55–62. DOI: https://doi.org/10.1016/j.patrec.2019.04.005

[3] Alireza Askarzadeh. 2016. A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm. *Comput. Struct.* 169 (June 2016), 1–12. DOI: https://doi.org/10.1016/j.compstruc.2016.03.001

[4] Freddie Åström and Rasit Koker. 2011. A parallel neural network approach to prediction of Parkinson's Disease. *Exp. Syst. Applic.* 38, 10 (Sept. 2011), 12470–12474. DOI: https://doi.org/10.1016/j.eswa.2011.04.028

[5] Julia Camps, Albert Sama, Mario Martin, Daniel Rodriguez-Martin, Carlos Perez-Lopez, Joan M. Moreno Arostegui, Joan Cabestany, Andreu Catala, Sheila Alcaine, and Berta Mestre. 2018. Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit. *Knowl.-based Syst.* 139, 1 (Jan. 2018), 119–131. DOI: https://doi.org/10.1016/j.knosys.2017.10.017

[6] Filippo Cavallo, Alessandra Moschetti, Dario Esposito, Carlo Maremmani, and Erika Rovini. 2019. Upper limb motor pre-clinical assessment in Parkinson's disease using machine learning. *Parkinson. Relat. Disord.* 63 (June 2019), 111–116. DOI: https://doi.org/10.1016/j.parkreldis.2019.02.028

[7] Hsiao-Chun Cheng, Christina M. Ulane, and Robert E. Burke. 2010. Clinical progression in Parkinson disease and the neurobiology of axons. *Ann. Neurol.* 67, 6 (May 2010), 715–725. DOI: https://doi.org/10.1002/ana.21995

[8] Chris Ding and Hanchuan Peng. 2005. Minimum redundancy feature selection from microarray gene expression data. *J. Bioinf. Comput. Biol.* 3, 2 (Apr. 2005), 185–205. DOI: https://doi.org/10.1142/S0219720005001004

[9] Lysia S. Forno. 1988. The neuropathology of Parkinson's disease. In *Progress in Parkinson Research*. Springer, Boston, MA, 11–21. DOI: https://doi.org/10.1007/978-1-4613-0759-4_2

[10] Srishti Grover, Saloni Bhartia, Abhilasha Yadav, and K. R. Seeja. 2018. Predicting severity of Parkinson's disease using deep learning. *Procedia Comput. Sci.* 132 (June 2018), 1788–1794. DOI: https://doi.org/10.1016/j.procs.2018.05.154

[11] Deepak Gupta, Arnav Julka, Sanchit Jain, Tushar Aggarwal, Ashish Khanna, N. Arunkumar, and Victor Hugo C. de Albuquerque. 2018. Optimized cuttlefish algorithm for diagnosis of Parkinson's disease. *Cog. Syst. Res.* 52 (Dec. 2018), 36–48. DOI: https://doi.org/10.1016/j.cogsys.2018.06.006

[12] Mark A. Hall. 2000. Correlation-based feature selection of discrete and numeric class machine learning. In *Proceedings of the 17th International Conference on Machine Learning (ICML'00)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, 359–366.

[13] Yixue Hao, Yiming Miao, Long Hu, M. Shamim Hossain, Ghulam Muhammad, and Syed Umar Amin. 2019. Smart-Edge-CoCaCo: AI-enabled smart edge with joint computation, caching, and communication in heterogeneous IoT. *IEEE Netw.* 33, 2 (Mar. 2019), 58–64. DOI: https://doi.org/10.1109/MNET.2019.1800235

[14] M. S. Hossain and G. Muhammad. 2019. Emotion recognition using secure edge and cloud computing. *Inf. Sci.* 504 (July 2019), 589–601. DOI: https://doi.org/10.1016/j.ins.2019.07.040

[15] C. Kotsavasiloglou, N. Kostikis, Dimitrios Hristu-Varsakelis, and M. Arnaoutoglou. 2017. Machine learning-based classification of simple drawing movements in Parkinson's disease. *Biomed. Sig. Proc. Contr.* 31 (Jan. 2017), 174–180. DOI: https://doi.org/10.1016/j.bspc.2016.08.003

[16] Kai Lin, Chensi Li, Daxin Tian, Ahmed Ghoneim, M. Shamim Hossain, and Amin Syed Umar. 2019. Artificial-intelligence-based data analytics for cognitive communication in heterogeneous wireless networks. *Wirel. Commun.* 26, 3 (June 2019), 83–89. DOI: https://doi.org/10.1109/MWC.2019.1800351

[17] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Andrew Lewis. 2014. Grey wolf optimizer. *Adv. Eng. Softw.* 69 (May 2014), 46–61. DOI: https://doi.org/10.1016/j.advengsoft.2013.12.007

[18] Ghulam Muhammad, M. S. Hossain, and Muhammad A. Yasmine. 2020. Tree-based deep networks for edge devices. *IEEE Trans. Industr. Inf.* 16, 3 (Mar. 2020), 2022–2028. DOI: https://doi.org/10.1109/TII.2019.2950326

[19] Mehrbakhsh Nilashi, Othman Ibrahim, and Ali Ahani. 2016. Accuracy improvement for predicting Parkinson's disease progression. *Sci. Rep.* 6, 1 (Sept. 2016), 1–18. DOI: https://doi.org/10.1038/srep34181

[20] Mehrbakhsh Nilashi, Othman Ibrahim, Hossein Ahmadi, Leila Shahmoradi, and Mohammadreza Farahmand. 2018. A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. *Biocyber. Biomed. Eng.* 38, 1 (Oct. 2018), 1–15. DOI: https://doi.org/10.1016/j.bbe.2017.09.002

[21] Shu Lih Oh, Yuki Hagiwara, U. Raghavendra, Rajamanickam Yuvaraj, N. Arunkumar, M. Murugappan, and Rajendra U. Acharya. 2018. A deep learning approach for Parkinson's disease diagnosis from EEG signals. *Neural Comput. Applic.* 32 (Aug. 2018), 1–7. DOI: https://doi.org/10.1007/s00521-018-3689-5

[22] Akin Ozcift. 2012. SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease. *J. Med. Syst.* 36, 4 (Mar. 2012), 2141–2147. DOI: https://doi.org/10.1007/s10916-011-9678-1

[23] Kemal Polat. 2012. Classification of Parkinson's disease using feature weighting method on the basis of fuzzy C-means clustering. *Int. J. Syst. Sci.* 43, 4 (May 2012), 597–609. DOI: https://doi.org/10.1080/00207721.2011.581395

[24] Kemal Polat. 2019. A hybrid approach to Parkinson disease classification using speech signal: The combination of smote and random forests. In *Proceedings of the Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT'19)*. IEEE, 1–3. DOI: https://doi.org/10.1109/EBBT.2019.8741725

[25] Jorge Peña Queralta, Tuan Nguyen Gia, Hannu Tenhunen, and Tomi Westerlund. 2019. Edge-AI in LoRa-based health monitoring: Fall detection system with fog computing and LSTM recurrent neural networks. In *Proceedings of the 42nd International Conference on Telecommunications and Signal Processing (TSP'19)*. IEEE, 601–604. DOI: https://doi.org/10.1109/TSP.2019.8768883

[26] UCI ML Repository. 2008. Parkinson's Data Set. Retrieved from https://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classification.

[27] Yvan Saeys, Thomas Abeel, and Yves Van de Peer. 2008. Robust feature selection using ensemble feature selection techniques. In *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, Berlin, 313–325. DOI: https://doi.org/10.1007/978-3-540-87481-2_21

[28] Betul Erdogdu Sakar, M. Erdem Isenkul, C. Okan Sakar, Ahmet Sertbas, Fikret Gurgen, Sakir Delil, Hulya Apaydin, and Olcay Kursun. 2013. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE J. Biomed. Health Inf.* 17, 4 (July 2013), 828–834. DOI: https://doi.org/10.1109/JBHI.2013.2245674

[29] C. Okan Sakar and Olcay Kursun. 2010. Telediagnosis of Parkinson's disease using measurements of dysphonia. *J. Med. Syst.* 34, 4 (Aug. 2010), 591–599. DOI: https://doi.org/10.1007/s10916-009-9272-y

[30] Mohammad R. Salmanpour, Mojtaba Shamsaei, Abdollah Saberi, Saeed Setayeshi, Ivan S. Klyuzhin, Vesna Sossi, and Arman Rahmim. 2019. Optimized machine learning methods for prediction of cognitive outcome in Parkinson's disease. *Comput. Biolo. Med.* 111 (Aug. 2019), 1033–1047. DOI: https://doi.org/10.1016/j.compbiomed.2019.103347

[31] Gurpreet Singh, Meet Vadera, Lakshminarayanan Samavedham, and Erle Chuen-Hian Lim. 2016. Machine learning-based framework for multi-class diagnosis of neurodegenerative diseases: A study on Parkinson's disease. *IFAC-PapersOnLine* 49, 7 (Aug. 2016), 990–995. DOI: https://doi.org/10.1016/j.ifacol.2016.07.331

[32] Noelia Sánchez-Maroño, Amparo Alonso-Betanzos, and María Tombilla-Sanromán. 2007. Filter methods for feature selection—A comparative study. In *Proceedings of the 18th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL'07)*. Springer, 178–187. DOI: https://doi.org/10.1007/978-3-540-77226-2_19

[33] Kai Rui Wan, Tomasz Maszczyk, Angela An Qi, Justin Dauwels, and Nicolas Kon Kam King. 2019. A review on microelectrode recording selection of features for machine learning in deep brain stimulation surgery for Parkinson's disease. *Clin. Neurophys.* 130, 1 (Jan. 2019), 145–154. DOI: https://doi.org/10.1016/j.clinph.2018.09.018

[34] Yang Wang, An-Na Wang, Qing Ai, and Hai-Jing Sun. 2017. An adaptive kernel-based weighted extreme learning machine approach for effective detection of Parkinson's disease. *Biomed. Sig. Proc. Contr.* 38 (Aug. 2017), 400–410. DOI: https://doi.org/10.1016/j.bspc.2017.06.015

[35] Xue wen Chen and Jong Cheol Jeong. 2007. Enhanced recursive feature elimination. In *Proceedings of the 6th International Conference on Machine Learning and Applications (ICMLA'07)*. IEEE, 429–435. DOI: https://doi.org/10.1109/ICMLA.2007.35

[36] Yanhao Xiong and Yaohua Lu. 2020. Deep feature extraction from the vocal vectors using sparse autoencoders for Parkinson's classification. *IEEE Access* 8 (Jan. 2020), 27821–27830. DOI: https://doi.org/10.1109/ACCESS.2020.2968177

[37] Geeta Yadav, Yugal Kumar, and Gadadhar Sahoo. 2012. Predication of Parkinson's disease using data mining methods: A comparative analysis of tree, statistical and support vector machine classifiers. In *Proceedings of the National Conference on Computing and Communication Systems*. IEEE, 1–8. DOI: https://doi.org/10.1109/NCCCS.2012.6413034

[38] Orhan Yaman, Fatih Ertam, and Turker Tuncer. 2020. Automated Parkinson's disease recognition based on statistical pooling method using acoustic features. *Med. Hypoth.* 135 (Feb. 2020), 109483. DOI: https://doi.org/10.1016/j.mehy.2019.109483

[39] Xin-She Yang. 2014. *Nature-inspired Optimization Algorithms*. Elsevier, Oxford, UK. DOI: https://doi.org/10.1016/C2013-0-01368-0

Received January 2020; revised July 2020; accepted August 2020