Machine Learning Approaches for Detection and Diagnosis of Parkinson's Disease- A Review

Iqra Nissar
Department of Computer Engineering
Jamia Millia Islamia
New Delhi, India
iqra.nissar13@gmail.com
\

Izharuddin
Department of Computer Engineering
Aligarh Muslim University
Aligarh, India
wamir@zhcet.ac.in

Waseem Ahmad Mir
Department of Computer Engineering
Aligarh Muslim University
Aligarh, India
wamir@zhcet.ac.in

Tawseef Ayoub Shaikh
Department of Computer Engineering
Baba Ghulam Shah Badshah University
Rajouri, India
tawseef@bgsbu.ac.in

2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) | 978-1-6654-0521-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICACCS51430.2021.9441885

***Abstract*— Parkinson's disease (PD) is disabling disease that affects the quality of life. It belimps due to the death of cells that produce dopamine's in the substantia nigra part of the central nervous system (CNS) which affects the human body. People who have Parkinson's disease feel difficulty in doing activities like speaking, writing, and walking. In the recent past, speech, gait and EEG signals have been investigated for the detection of PD. However, speech analysis is the most considered technique to be used. Researches have shown that 90% of the people who suffer from Parkinson 's disease have speech disorders. With the increase in the severity of the disease, the patient's voice gets more and more deteriorated. The non-invasive treatments for voice analysis are available that helps in ameliorating the life quality of a patient. Thus, for building the telemonitoring and telediagnosis models for prediction, the speech analysis has been tremendously increased. The proper interpretation of speech signals is one of the important classification problems for Parkinson's disease diagnosis. The main purpose of this paper is to contemplate the survey work of the machine learning techniques and deep learning procedures used for Parkinson's disease classification. Deep learning and machine learning techniques have been used as a part of the discovery for the efficient classification of PD. The various classification models like support vector machines, naive Bayes, deep neural networks, decision tree and random forest are effectively employed for classification purposes. The analysis of results of different research works showed that both machine learning and deep learning algorithms have shown promising future and therefore paving a better way for the detection of Parkinson's disease at its earlier stages. The classification accuracy achieved by the machine learning classifier. Among deep learning approaches, the deep neural network has achieved the best accuracy of 99.49%. The results obtained from different works suggest that artificial intelligence is becoming a powerful learning tool that has much to offer to data scientists as well as neurologists. In general the learning methods are adding value to decision-making problems especially in the field of medical diagnosis.**

Keywords—Deep neural networks; Parkinson's disease; speech signals; support vector machines; classification, Machine learning; feature selection.

I. INTRODUCTION

PD is a progressive neuropathological degeneration of the CNS in which the motor functions of the human body are affected [1]. Among the neurological disorders, after Alzheimer's disease, the second most frequently found disease is PD [2], and in North America alone the number of PD patients has crossed more than one million people [3]. PD is also called as shaking palsy as described by Dr. James Parkinson in 1817 [4]. However, this number is expected to increase in an ageing population as the studies show that the PD increases rapidly in people whose age is over 60 [5][6].

The PD is instigated by the loss of certain brain cell clusters that produces the neurotransmitters which include acetylcholine, dopamine, norepinephrine, and serotonin. Particularly the loss of dopamine results in various symptoms that include urinary problems, visual problems, anxiety, depression, panic attacks and weight loss [6]. The other symptoms of PD include voice impairment, poor balance and tremor [7]. Various studies done by researchers have found that over 90% of PD patients suffer from speech and vocal problems [8][9] that includes dysphonia, dysarthria, monotone, hypophonia [10] and thus the initial symptom that can be seen

978-1-6654-0521-8/21/$31.00 ©2021 IEEE

in people with Parkinson (PWP) is the degradation of voice [11]. Also, PWP have disturbances in sleep during the rapid eye movement sleep phase (REMs) [12]. At present, the cure for the disease is yet unknown [13][14] but the various drug therapies are available that offer the significant mitigation of symptoms especially at its earlier stages. Voice measurement is simple to analyse and is non-invasive. Thus, the measurement of voice can also be used to track the progression of PD [15][16]. To assess the PD progression, many vocal tests have been devised. These include sustained phonations and running speech texts [17][18]. The telemonitoring and telediagnosis systems have been widely used for the detection of PD at its earlier stages. Most of these methods rely on motor disorders which are caused by PD.

Generally, the diagnosis of PD consists of three steps: 1) data pre-processing, 2) extraction of features and 3) classification [11]. In the first step, the segmentation of voice signals with time windows is performed. The speech signals are subjected to filtering so as to eliminate the noise present in the signals. From each segment, several features are extracted in the second step. In the last step, the classification is performed. The performance of the classification method has a strong impact on the method used for feature extraction. Hence, selecting the right classification model is an important issue that needs to be maneuvered in the diagnosis of PD. In this paper, there is an attempt to explore the different deep learning and machine learning (ML) techniques that have been used for the diagnosis of PD.

The paper is organized as follows: The previously done studies on PD detection have been substantiated in section 2. The commonly used classification algorithms are described in Section 3. The conclusion is presented in section 4.

II. Literature Survey

For predicting the Parkinson's disease in subjects, various notable attempts were developed by different researchers from time to time. This section presents the literature survey of the work done so far.

A. By using Machine Learning Algorithms

Max A. little et al [14] suggested a novel technique for classifying the subjects into PD and control subjects by detecting dysphonia. In their work, a new robust measure of dysphonia was introduced called as pitch period entropy (PPE) that works properly in a noisy and uncontrollable environment. The data for their study was collected from 31 people (23 were PD patients and the remaining were healthy) and was comprised of 195 sustained vowel phonations.

The ages of the people vary from 46 to 85 years and from each subject 6 phonation were recorded. The filtering of features was done and 10 highly uncorrelated measures were selected and all the possible combinations of the features were searched. It was found that 4 in that combination produced the best classification performance. Their proposed model achieved an accuracy of 91.4%. In their work, they found that when combining the traditional harmonic to noise ratios with non-standard methods are best able to perform the classification.

Resul Das [19] compared the performance of different methods of classification in order to diagnose the PD very effectively. The different classifiers relied on Sas based software for PD detection. The classifiers used were DMneural, decision tree, neural network, and regression respectively. The accuracy obtained was 84.3% for DMneural, 88.6% for regression, 84.3% for decision tree and the best accuracy of 92.9% was reported for the neural network. The dataset used was divided into 65% and 35% for training and testing respectively. The hyperparameter tuning was done separately for each classifier. The Levenberg Marquardt (LM) algorithms were used in their study.

In [20] proposed the normal subjects from PD subjects. In their work, the dimension of the feature vector was reduced and optimized features were obtained by using a genetic algorithm (GA) and k nearest neighbor (k-NN) was used for classification. In their study, Parkinson's dataset [21] was taken from the UCI repository which includes 197 voice samples that were recorded from 31 subjects (23 PD patients and 8 normal subjects).

In another significant work, B.E Sakar et al [10] suggested a model for classifying the subjects into control and PD subjects. In the context of their study, the data was collected from 40 subjects of which 20 were PWP and 20 were healthy. Each subject had to record 26 voice samples that include words, short sentences, numbers, and sustained vowels. For classification, they used SVM and k-NN with Summarized Leave-One-Out (s-LOO) and Leave-One-Subject-Out (LOSO) as cross-validation schemes. For evaluating the model performance accuracy, sensitivity, specificity and Matthews's correlation coefficient (MCC) scores were calculated. The Praat acoustic analysis software was used for feature extraction. For the k-NN classifier, the value of 1, 3, 5 and 7 was chosen for k and for SVM they used linear and RBF kernels. The accuracy achieved was 82.50% and 85% for k-NN and SVM respectively.

Dr. R. Geetha Ramani et al [22] used data mining algorithms for the classification of control and PD subjects. The dataset used for their study comprises of 197 voice samples from which 22 features were extracted [21]. They used various classification models like a random tree (Rnd tree), k-NN, SVM, C4.5, ID3, binary logistic regression, LDA and PLS. An accuracy of 100% was achieved for the random tree and above 90% for k-NN, C4.5 and LDA respectively. The least accuracy of 69.74% was reported for the C-PLS algorithm.

In another work, David Gil A. et al [23] suggested methods that were based on SVM and artificial neural networks (ANN). The dataset used was retrieved from the UCI repository [21]. An ANN-based multilayer perceptron (MLP) network with two layers was used. It was found that SVM produced better significant results than that of MLP. SVM with the linear and puk kernel was used which produced an accuracy of 91.79% and 93.33% respectively. The MLP achieved an accuracy of 92.31%.

Ipsita Bhattacharya et al [24] used SVM-a supervised machine learning algorithm to separate the PD subjects from healthy subjects. They used weka -a tool for data mining. Pre-processing of data was performed on the dataset [21] before applying the classification algorithm. The random split was done repeatedly and the best possible accuracy was found on the different kernel values by applying libsvm. It was found that the RBF kernel and polykernel SVM produced the highest accuracy of 60. 8696% whereas, the best accuracy of 65.2174% was reported on using the linear kernel SVM.

To perform the classification, K.Uma Rani et al [25] used a multilayer perceptron and RBF network. The data they used consists of 136 sustained vowel phonations in which 83 phonations were recorded from the people suffering from PD and 53 phonations were recorded from the healthy individuals. For training, the network 112 phonations were used and for testing 24 phonations were used. It was found that the RBF network yielded better results as compared with the multilayer perceptron. An accuracy of 86.66% and 83.33% was achieved for training and test set on using MLP while the RBF network produced an accuracy of 90.12% and 87.5% for training and test set respectively.

Achraf Benba et al [26] used a dataset comprised of 34 sustained vowels which were collected from 34 individuals in which 17 were suffering from PD. From each subject 1 to 20 Mel-frequency cepstral coefficients (MFCCs) were obtained. Different kernel types of SVM were used along with leave one subject out validation technique. The linear kernel SVM showed the best accuracy of 91.17% by taking only the top 12 coefficients of MFCCs. In another of their work [27], they aimed to separate the people with Parkinson from the people who suffer from other neurological diseases. They collected voice samples from 50 subjects in which 30 patients were PD patients and 20 were other neurological disease patients. From each voice sample, cepstral coefficients were extracted by using the three cepstral techniques- MFCC, ReAlitiveSpecTrAl PLP (RASTA-PLP) and Perceptual Linear Prediction (PLP). Five supervised classifiers were used with different kernel types. These include SVM, discriminant analysis, K nearest neighbor, naive Bayes and classification tree (CT). The first 11 PLP coefficients showed 90% accuracy with linear kernel SVM.

C.O. Sakar et al [28] gave the comparison of different speech signal processing algorithms for PD detection. To extract the features from voice signals, the tunable Q-factor wavelet transform (TQWT) was first time introduced. The TQWT effectiveness was compared and it was found that it outperformed the state of art methods used for feature extraction in PD diagnosis. In the context of their study, the voice samples from 252 subjects were recorded and multiple features were extracted. Different classifiers were used on different feature subsets and using the ensemble learning approaches the prediction of these classifiers were combined. It was shown that MFCCs and TQWT produced the highest accuracies and thus are important in the problem of PD classification. The feature selection technique used was the minimum redundancy maximum relevance (mRMR) as the step for pre-processing to filter the most relevant features. The mRMR filter selected the top 50 relevant features and the algorithms were applied on different feature subsets. The RBF kernel SVM performed better on all feature subsets with 86% accuracy.

Mohammad Shahbakhi et al [29] proposed an algorithm for the diagnosis of PD. From all the extracted features, the genetic algorithm was used to select the optimal features on a given dataset [21]. The different number of optimized features were obtained and then fed into SVM for classification. For 4 optimized features, 7 optimized features and 9 optimized features, 94.5%, 93.66%, and 94.22% of accuracy were reported.

Richa Mathur et al [30] suggested a method for predicting PD. They used a weka tool for implementing the algorithms to perform pre-processing of data, classification and the result analysis on the given dataset [21]. They used k-NN along with Adaboost.M1, bagging, and MLP. Results showed that k-NN + Adaboost.M1 and k-NN + MLP yielded the same accuracy of 91.28 % while the time taken for building the model with Adaboost.M1 was 0.01 and with MLP it was 0.43. Therefore, the k-NN + Adaboost.M1 is the best classification algorithms which produced significant results and was also time-efficient.

Amin ul Haq et al [31] proposed a classifier model that classifies the subjects into control and people with Parkinson subjects. In their study, they used the L1-norm SVM algorithm to get the appropriate features selected and then classification was performed. Their proposed system involved 1) data pre-processing, 2) selection of appropriate features, 3) cross-validation (CV) and 4) evaluation of classifiers. 10-fold CV was used for SVM (with both linear and RBF kernel). The highest accuracy of 95% and 97% were reported on full features with RBF and linear kernel SVM respectively. However, it was observed that by selecting the best features, an accuracy of 99% was reported which proved that by the selection of most relevant features, the classifier performance can be increased considerably.

Diogo Braga et al [32] presented a new technique to detect the early signs of PD through the analysis of speech. They used algorithms like SVM, random forest, NB and neural network. In their framework, three speech datasets were used. The first one was collected from 22 PD subjects that contain 1002 speech lines. The second dataset consists of 785 speech lines that were collected from 30 healthy individuals. The third dataset was used for validating the performance of machine learning algorithms. This dataset consists of sustained vowels that were collected from 28 PD patients. An accuracy of 94.77% was obtained by using random forest along with the LOSO validation scheme, but after optimizing the algorithms, the accuracy of 99.94% was reported with random forest. The SVM with RBF kernel and neural network produced an accuracy of 92.38% and 91.10% respectively.

In another significant work, Salama A. Mostafa et al [33] proposed a method based on multiple feature evaluation approach (MFEA) which selected the 11 top features. The classification algorithms like NB, SVM, neural network, RF and DT were implemented with a 10-fold CV method. For result analysis, the original dataset was used first and then the filtered features. It was observed that the highest accuracy of 87.755% was produced by the neural network on the original dataset but on filtered features, the random forest produced 99.492% accuracy. The NB, NN, decision tree and SVM achieved an accuracy of 89.340%, 96.950%, 96.954%, and 95.431% respectively.

In this work, Ozcift [34] presented a rotation forest (RF) ensemble method which was based on linear SVM feature selection for improving the PD diagnosis. The dataset [21] consisted of 195 voice samples that were collected from 31 individuals of which 23 were PD patients. The dataset has 22 speech attributes. Their method worked as follows. Firstly, the 22 attributes were reduced to 10 relevant features by using linear SVM. In the second step, on a feature subset, six different classification algorithms were used. Finally, an ensemble method RF of IBk which is a variant of k-NN was used to improve the model accuracies. Different metrics were calculated to evaluate the classification performance. These include Kappa Error, classification accuracy. The accuracy of 96.93% was achieved by using IBk while the least accuracy of 88.71% was yielded by RBF kernel SVM.

I. Nissar et al [36] employed the ensemble machine learning approach for the classification of PD. In their framework, the data pre-processing was done on a dataset [29] which involved data normalization followed by feature relevance analysis. The data was normalized by using min-max normalization. RFE and mRMR were used to select the relevant features. The selected features were fed to the classifiers to accomplish the task. They used nine different machine learning algorithms for performance analysis. Among all the classifiers, the XGBoost along with the mRMR feature selection technique produced the highest accuracy of 95.39% with a precision of 0.95, recall of 0.95 and F1-score of 0.95 respectively when taking both MFCC and TQWT features into consideration.

C.D. Anisha et al [37] employed the use of dataset [29] on which they applied LDA and PCA to select the highly correlated features. Gradient boosting machine (GBM), extreme gradient boosting (XGBoost), bagging and adaptive boosting (AdaBoost) were used as ensemble classifiers for performing the classification. The grid search and random search were used to select the optimal parameters, and the model performance was evaluated based on precision, support, F1-score, accuracy and recall. The result analysis showed that PCA performed better than LDA. The hyperparameter tuning done by grid search was better than the random search. Also, the boosting method produced good results than bagging methods. The highest accuracy of 94% was achieved by the AdaBoost classifier when using the grid search and 10-fold cross-validation.

O. Asmae et al [38] used ANN and k-NN for distinguishing the PD and healthy ones. The dataset [22] consisted of a total of 31 subjects in which 23 were of the PD category. They used MATLAB tool for the purpose of classification. For ANN, 70% of data was used for training the model, 5% for validation purpose and 25% for testing respectively. An accuracy of 96.7% was reported by ANN. For k-NN, the data consisted of 70% of training data and 30% of testing data. They used a 10-fold CV technique to avoid the overfitting of data. An accuracy of 79.31% was achieved when taking k=1 and by using cosine as a distance metric.

Z.K. Senturk [39] presented a framework for the early diagnosis of PD. They employed the RFE method as a feature selection (FS) task on a dataset [22]. They used ANN, CART and support vector machines as classification methods. It was found that the model performed better after the feature selection. Without feature selection, an accuracy of 85.23%, 79.98% and 80.25% was achieved for CART, SVM and ANN respectively. After employing the FS, the accuracy was improved to 90.76%, 91.54% and 93.84% for CART, ANN and SVM respectively.

Tuncer et al [40] presented a novel method to detect the PD. The minimum average maximum tree was used along with a combination of singular valued decomposition (SVD). On the original dataset [29], the top 50 features were selected by the relief feature selection method. The k-NN was used as a classifier with a 10-fold CV. The highest classification accuracy of 92.46% was achieved. Post-processing was also done to obtain the individual results which improved the accuracy to 96.83%. The literature review summary of ML algorithms is tabulated in Table 1.

TABLE 1. MACHINE LEARNING ALGORITHMS

Author	***Technique Applied***	***Highest Accuracy Achieved (%)***
Little et al (2009)	SVM	91.4
David Gil A. et al (2009)	SVM	93.33
Resul Das (2010)	Neural network	92.9
Ipsita Bhattacharya et al (2010)	Linear SVM	65.21
R. Arefi Shirvan et al (2011)	k-NN	98.2
R. Geetha Ramani et al (2011)	Random tree	100
Uma Rani et al (2012)	SVM (RBF)	87.5
A. Ozcift et al (2012)	IBk	96.93
B.E Sakar et al (2013)	Linear SVM	85.0
Mohammad Shahbakhi et al (2014)	SVM	94.50
Achraf Benba et al (2015)	Linear SVM	91.17
Achraf Benba et al (2016)	Linear SVM	90.0
Richa Mathur et al (2018)	k-NN + Adaboost.M1	91.28
Salama A. Mostafa et al (2019)	Random forest	99.49
Diogo Braga et al (2019)	Random forest	99.94
Amin ul Haq et al (2019)	SVM	99.0
C.O. Sakar et al (2019)	SVM (RBF)	86.0
I.Nissar et al (2020)	XGBoost	95.39
C.D. Anisha (2020)	AdaBoost	94.0
O. Asmae et al (2020)	ANN	96.7
Z.K. Senturk (2020)	SVM	93.84
Tuncer et al (2020)	k-NN	96.83

B. By using Deep Learning Algorithms

Various symbolic approaches of deep learning methods can be seen in the medical health field [41]. Alex Frid et al [42] presented a novel deep learning architecture for signal processing using one dimension. They eliminated the use of signal processing field expertise for the feature extraction stage, and thus they worked on raw speech signals. They adapted convolutional neural networks (CNN) for their work and signals of raw speech were fed into the deep network. During the training of the network, some changes were applied to a dataset. The entire recording was processed without the selection of phonemes and windowing was applied in which the speech signal was divided into processing windows each of 20 milliseconds with an overlap and each window was used separately as data input. A majority voting scheme was used to perform the final classification. 85% of the windows were used for training the data and 15% were used for validation. Between the subsequent stages at the window level, the binary classification of the disease progression achieved the classification accuracy of 83.63%. The binary classification was done by using leave one out scheme between the severity stages and an accuracy of 80.5% was reported.

Abdullah Caliskan et al [43] suggested a deep neural network (DNN) classifier for the PD diagnosis. The datasets used were OPD [21] and PSD [10] in their work. Their proposed classifier had two main components. The features were learned and reduced by the use of autoencoders (AEs) and the classification was done by the supervised layer known as the softmax layer whose working was based on the softmax function. The features of the speech data were taken as the input for the DNN classifier and the output was labeled with either 0 or 1, corresponding to the healthy and PD subject respectively. The tuning of the network was done to get efficient results. 10-fold cross-validation was used for the comparison with the state of art methods and also for validating the performance. 70% and 30% were used for training and testing respectively. The mean accuracy of 65.549% was reported on the PSD dataset whereas the mean accuracy of 86.0955% was achieved by the OPD dataset.

Ali H.et al [44] suggested a deep belief network (DBN) as an efficient model for the diagnosis of PD. The dataset used was retrieved from the UCI repository [21]. In the context of their work, DBN was composed of two Stacked Restricted Boltzmann machines (RBMs) with an output layer. For optimizing the network parameters, the learning involved two stages. Unsupervised learning was applied in the first stage that used RBMs to tackle the problem which was caused by the random initialization of weights. The second stage was needed to fine-tune the network hyper-parameters which used a supervised back-propagation learning algorithm. Mean square error (MSE) was used as a performance measure of the network during the training phase. It was found that supervised learning involved 8700 iterations to obtain a steady state of MSE. Their proposed methodology achieved a testing accuracy of 94%.

Srishti Grover et al [45] proposed a model based on deep learning for predicting the severity of PD. The dataset was taken from the UCI repository [7]. The data was collected from 42 patients and from each patient 200 recordings were taken which comprised a total of 5875 recordings of voice. The dataset had sixteen speech biomedical attributes. The voice data was normalized by using min-max normalization. 80% and 20% of normalized data were used for training and testing respectively. The number of input neurons was equal to the number of speech attributes present in voice data. The output layer had two neurons corresponding to severe and non-severe classes respectively. The motor UPDRS and total UPDRS scores were used as the evaluation metrics for the output variable. The accuracy of 94.4422% and 62.7335% was obtained for training and testing on using the total UPDRS as the output variable. When using the motor UPDRS score as an output variable, the classification accuracy of 83.367% and 81.6667% was obtained for training and testing respectively.

Savitha S. Upadhya et al [46] used a neural network for classifying the subjects into Parkinson diseased and control people. In their work, they extracted the cepstral features and phonation of speech. The phonation included 14 features and the cepstral features were 12 MFCCs. These features were fed in a neural network and classification was done. They used a feed-forward network as a neural network (NN) classifier that had two layers and one hidden layer. During the training of the network, a scaled conjugate back-propagation algorithm was used. The output layer used a softmax function to perform the classification. Sensitivity, recognition accuracy, specificity, and equal error rate were used for performance evaluation. The experiments were done on phonation, cepstral features and hybrid of these features. The cepstral and hybrid features achieved an accuracy of 81.1% and 96.7% respectively. On phonation, the best accuracy of 98% was obtained.

In this paper, P. Chitra Rajagopal et al [47] used a neural network to diagnose the PD. The dataset used [21] consists of 23 speech attributes which were collected from 31 individuals of which 8 were healthy subjects and the remaining were suffering from PD. Their proposed neural network used 23 neurons in the input layer and the output layer had two neurons one for the control subject and one for the PD subject. The classification was done by using the sigmoid function at the output layer. The accuracy was obtained by varying the number of epochs and batch size. It was shown that on choosing the 500 epochs and batch size of 6, an accuracy of 99.49% was reported.

D.R. Rizvi et al [48] employed deep learning for classifying the subjects into PD and control subjects. In their framework, they used DNN and LSTM models. Different experiments were done on a dataset [11] to find the optimal parameters of the network which eventually improved the network performance. The results of their work were compared with conventional machine learning algorithms. An accuracy of 97.12% and 99.03% was obtained while using the DNN and LSTM models respectively.

S. Kaur et al [49] proposed a methodology in which they used grid search optimization to develop a deep learning model to predict Parkinson's disease at its earliest stages. The grid search optimization had three stages: optimizing the topology of deep learning model, hyperparameters and performance. The deep learning model was tuned for various parameters like the number of layers, a number of neurons in different layers, epochs, batch size, activation functions, optimizers, learning rate, performance metrics and cross-validation. It was found that after successful fine-tuning of deep learning model, the test accuracy of 89.23% was achieved and the average accuracy of 91.69% was reported.

K. Akyol [50] compared the performance of DNN and extreme learning machine (ELM) which were based on a growing and pruning approach. Various experiments were performed on a dataset [29] and a different set of optimal parameters was selected. It was observed that the ELM model achieved the best accuracy of 83.70% when using 6 hidden layers, each of which has 2, 4, 8, 16, 32 and 64 nodes and sigmoid as an activation function. When using the 7-layer DNN model having a number of nodes as 2, 4, 8, 16, 32, 64 and 64 respectively with SGD optimization and Relu activation function, the best prediction accuracy of 95.15% was achieved which suggested that the DNN model performed better than the ELM model. The deep learning classification techniques are presented in Table 2.

TABLE 2. DEEP LEARNING ALGORITHMS

Author	*Technique Applied*	*Highest Accuracy Achieved (%)*
Ali H. Al-Fatlawi et al (2016)	DBN	94
Alex Frid et al (2017)	CNN	83.63
Abdullah Caliskan et al (2017)	DNN classifier	86.09
Savitha S. Upadhya et al (2018)	NN classifier	98.0
Srishti Grover et al (2018)	DNN	81.66
Chitra Rajagopal et al (2019)	NN classifier	99.49
D. R . Rizvi et al (2020)	LSTM	99.03
S. Kaur et al (2020)	DNN	91.69
K. Akyol (2020)	DNN	95.15

III. ALGORITHMS

A. *k-Nearest Neighbor (k-NN)*

k-NN is a supervised ML algorithm that is used for classification as well as regression problems. It is instance-based learning which describes the problem-solving process that is based on solutions for already known similar problems. Instance-based learning methods are a part of lazy learners, which means that there is no computation done on the data until a query is given to the system. The classification of data is done based on different distance metrics [51]. The value of k is considered as the hyper-parameter that varies from problem to problem. The distance metrics used are Manhattan distance, Euclidean distance, and Minkowski distance. However, the distance function is used for continuous variables and hamming distance is used for categorical variables. A new data point is classified by doing the voting of the neighbors, which means that the data point is assigned to the class which has the largest common points among the k nearest neighbors that are measured by some distance function. The commonly used distance functions in k-NN are given below:

$$Euclidean = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \quad (1)$$

$$Manhattan = \sum_{i=1}^{k}|x_i - y_i| \quad (2)$$

$$Minkowski = \left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q} \quad (3)$$

B. *Support Vector Machines*

Support Vector Machines belong to the supervised learning algorithms, which requires the labeled data for the classification of unseen data. It works on the concept of hyperplanes or decision planes that define the decision boundaries. Hyperplane separates the set of data objects that belong to different classes. The working of SVM is to classify the data by creating a function which split the data points into their corresponding labels with

- the number of errors as least as possible.
- the largest (maximum) possible margin.

It is the powerful learning model which applies for both classification and regression problem and it handles the continuous as well as categorical attributes. In this classification, the function maps the training set into a higher dimensional space. A linear separating hyperplane is found by using a maximum margin. It identifies the best hyperplane that divides the dataset into different classes. Fig.1, shows the example of SVM, in which a group of instances is separated by the optimal hyperplane and the maximum margin which is on either side of the hyperplane.

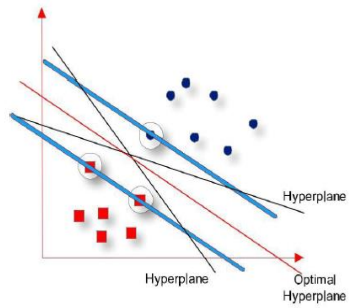

Fig. 1. *An example of SVM.*

C. *Deep Neural Network*

It is based on an artificial neural network which is now known to have excellent model building capabilities for a wide variety of problems. This network consists of an input layer, a number of hidden layers and finally an output layer of processing units called neurons [52]. The number of neurons present in the input layer is the number of attributes of input data. The number of nodes in the output layer depends upon the number of output classes and it varies from problem to problem. For example, in a binary classification, there is one output node corresponding to the predicted class and for the multiclass classification, there are n output nodes where n corresponds to the number of classes in the dataset. The number of nodes to be chosen for the hidden layer depends on the classification problem. The neurons of one layer communicate with the neurons of another layer by using signals that are real numbers. By applying the nonlinearity on the inputs, the neurons calculate the output by finding the correct mathematical manipulation. Initially, the weights are assigned to the neurons which get updated by the process of learning. The input is processed at each layer and the output is then passed to the next layer. Their training is harder than the shallow nets but they are very powerful and are typically used in natural language processing, speech recognition, image recognition, bioinformatics, computer vision problems, etc. The architecture of DNN is shown in Fig.2.

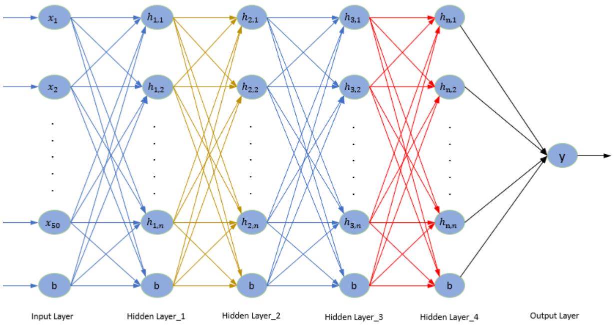

Fig. 2. *DNN architecture.*

IV. Conclusion

Currently, the neurodegenerative disease research area is of much significance and its diagnosis at its earlier stage can make the patient's life better. The recent developments in the methodologies of speech analysis have yielded promising results. In this paper, different types of deep learning and ML algorithms have been explored for the detection of PD. Our main aim is to show the PD diagnosis by analysing the voice signals. Since the voice measurements are non-invasive, therefore speech processing has been widely used in many diverse applications and has incredible potential in the classification and diagnosis of PD for many years. This paper is intended to analyse and ascertain the performance of many classification models. These models can be effectively used to monitor and diagnose the PD remotely which reduces the need for physical visits of the patients to the clinics. The different classifiers were applied on different voice datasets and it was found that the random forest outperforms the other classifiers in machine learning algorithms which achieved an accuracy of above 99%. Among the deep learning techniques, the neural network classifier also produced a reasonable accuracy of 99.49%. Thus, for the early diagnosis and detection of PD, the techniques of artificial intelligence have been widely employed to produce effective and efficient results.

References

[1] C. W. Olanow, M. B. Stern, and K. Sethi , "The scientific and clinical basis for the treatment of Parkinson disease," Neurology, vol. 72(21 Supplement 4), pp. S1-S136, 2009.

[2] M. C. de Rijk, L. J. Launer, K. Berger, M. M. Breteler, J. F. Dartigues, M. Baldereschi, L. Fratiglioni, A. Lobo, J. M. Lage , C. Trenkwalder, and A. Hofman, "Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group," Neurology, vol. 54(11 Suppl 5), pp. S21-3, 2000.

[3] A. E. Lang, and A. M. Lozano, "Parkinson's disease—First of two parts," New. Engl. J. Med., vol. 339, pp. 1044–1053, 1998.

[4] J. W. Langston, "Parkinson's disease: current and future challenges," Neurotoxicology, vol. 23(4-5), pp. 443-450, 2002.

[5] S. K. Van Den Eeden, C. M. Tanner, A. L. Bernstein, R. D. Fross, A. Leimpeter, D. A. Bloch, and L. M. Nelson, "Incidence of Parkinson's disease: variation by age, gender, and race/ ethnicity," American journal of epidemiology, vol. 157(11), pp. 1015-1022, 2003.

[6] D. M. Huse, K. Schulman, L. Orsini, J. Castelli-Haley, S. Kennedy, and Lenhart, "Burden of illness in Parkinson's disease. Movement disorders: official journal of the Movement Disorder Society," vol. 20(11), pp. 1449-1454, 2005.

[7] J. A. Logemann, H. B. Fisher, B. Boshes, and E. R. Blonsky, "Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients," Journal of Speech and hearing Disorders, vol. 43(1), pp. 47-57, 1978.

[8] D. J. Gelb, E. Oliver, and S. Gilman, "Diagnostic criteria for Parkinson disease," Arch Neurol., vol. 56, pp. 33–39, 1999.

[9] A. Tsanas, M. A. Little, P. E. McSharry, and L. O Ramig, "Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests," IEEE transactions on Biomedical Engineering, vol. 57(4), pp. 884-893, 2009.

[10] K. S. Perez, L. O. Ramig, M. E. Smith, and C. Dromey, "The Parkinson larynx: tremor and videostroboscopic findings," Journal of Voice, vol. 10(4), pp. 354- 361, 1996.

[11] P. Bugalho, and M. Viana-Baptista, "REM sleep behavior disorder and motor dysfunction in Parkinson's disease - A longitudinal study," Park Relat Disord, vol. 19(12), pp.1084–1087, 2013.

[12] N. Singh, V. Pillay, and Y. E. Choonara, "Advances in the treatment of Parkinson's disease," Progress in neurobiology, vol. 81(1), pp. 29-44, 2007.

[13] B. E.Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil, and O. Kursun, "Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings," IEEE Journal of Biomedical and Health Informatics, vol. 17(4), pp. 828-834, 2013.

[14] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," IEEE Trans. Biomed. Eng., vol. 56(4), pp. 1010 –1022, 2009.

[15] S. Sapir, J. L. Spielman, L. O. Ramig, B. H. Story, and C. Fox, "Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and perceptual findings," Journal of Speech, Language, and Hearing Research, vol. 50 (4), pp. 899-912, 2007.

[16] D. A. Rahn, M. Chou, J. J. Jiang, and Y. Zhang, "Phonatory impairment in Parkinson's disease: evidence from nonlinear dynamic analysis and perturbation analysis," Journal of Voice, vol. 21(1), pp. 64-71, 2007.

[17] R. J. Baken, and R. F. Orlikoff, "Clinical measurement of speech and voice," Cengage Learning, 2000.

[18] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, and V. Woisard, "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques," European Archives of Oto-rhino-laryngology, vol. 258(2), pp. 77-82, 2001.

[19] R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," Expert Syst with Appl., vol. 37 (2), pp. 1568–1572, 2010.

[20] R.A. Shirvan and E. Tahami, "Voice analysis for detecting Parkinson's disease using genetic algorithm and KNN classification method," 18th Iran Conf Biomed Eng ICBME, pp. 278–283, 14-16 Dec 2011, Iran, IEEE, 2012,

[21] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," Behavioural neurology, vol. 11(3), pp. 131-137, 1999.

[22] M. A. Little, P. E. McSharry, S.J. Roberts, D.A.E. Costello, and I. M. Moroz, “Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection,” Biomed Eng Online, vol. 6, pp. 1–19, 2007.

[23] R. G. Ramani, and G. Svigami, “Parkinson disease classification using Data Mining algorithms,” Int J Comput Appl,.vol. 32(9), pp. 17-22, 2011.

[24] D. Gil, and M. Johnson, “Diagnosing parkinson by using artificial neural networks and support vector machines,” Glob J Comput Sci Technol., pp. 63–71, 2009.

[25] I. Bhattacharya, and M. P. S. Bhatia, “SVM classification to distinguish Parkinson disease patients.” Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India (p. 14). ACM, September 2010.

[26] K. U. Rani, and M.S. Holi, “Analysis of speech characteristics of neurological diseases and their classification,” 3rd Int Conf Comput Commun Netw Technol ICCCNT 2012,Coimbatore India, DOI: 10.1109/ICCCNT.2012.6395886, 26-28 July 2012.

[27] A. Benba, A. Jilbab, A. Hammouch, and S. Sandabad, “Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease,” International conference on electrical and information technologies (ICEIT), pp. 300-304, March 2015.

[28] A. Benba, A. Jilbab, A. Hammouch, “Discriminating between patients with parkinson’s and neurological diseases using cepstral analysis,” IEEE Trans Neural Syst Rehabil Eng, vol. 24(10), pp. 1100–1108, 2016.

[29] C. O. Sakar, G. Serbes, A. Gunduz, H. C. Tunc, H. Nizam, B. E. Sakar, and H. Apaydin, “A comparative analysis of speech signal processing algorithms for Parkinson’s disease classification and the use of the tunable Q-factor wavelet transform,” Applied Soft Computing, vol. 74, pp. 255-263, 2019.

[30] M. Shahbakhi, D.T. Far, E. Tahami, “Speech analysis for diagnosis of Parkinson’s Disease using Genetic Algorithm and Support Vector Machine,” J Biomed Sci Eng, vol. 7, pp. 147–156, 2014.

[31] R. Mathur, V. Pathak, & D. Bandil, “Parkinson Disease Prediction Using Machine Learning Algorithm,” In Emerging Trends in Expert Applications and Security, pp. 357-363 Springer, Singapore, 2019.

[32] A. Haq, J.P. li , M.H. Memon , J. Khan , A. Malik , T. Ahmad , A. Ali , S. Nazir , I. Ahad, and M. Shahid, “Feature selection based on L1-Norm Support Vector Machine and Effective Recognition System for Parkinson’s disease using voice recordings,” IEEE Access, vol. 7, pp. 37718 – 37734, 2019.

[33] D. Braga, A. M. Madureira, L. Coelho, and R. Ajith, “Automatic detection of Parkinson’s disease based on acoustic analysis of speech,” Eng Appl Artif Intell., vol. 77, pp. 148–158, 2019.

[34] S. Mostafa, A. Mustapha, M. A. Mohammed, R.I. Hamed, N. Arunkumar, M. K. A. Ghani, M. M. Jaber, and, S. H. Al-Dulaimi, “Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson’s disease,” Cogn Syst Res., vol. 54, pp. 90–99, 2019.

[35] A. Ozcift, “SVM feature selection based rotation forest ensemble classifiers to improve computer-Aided diagnosis of Parkinson disease,” J Med Syst., vol. 36(4), pp. 2141–2147, 2012.

[36] I. Nissar, D.R. Rizvi, S. Masood, and A. Mir, “Voice-based detection of Parkinson’s disease through ensemble machine learning approach: A Performance study,” Phat. 19(19): e2, vol. 5, pp. 1-8, 2020.

[37] C.D. Anisha, N. Arunlanand, “Early prediction of Parkinson’s Disease (PD)usingEnsembleClassifiers”,Doi:10.1109/ICITIIT49094.2020.9071562, 2020.

[38] O. Asmae, R. Abdelhadi, C. Bouchaib, S. Sara, and K. Tajeddine, “Parkinson’s disease identification using KNN and ANN Algorithms based on Voice Disorder,” **DOI:** 10.1109/IRASET48871.2020.9092228, 2020.

[39] Z.K. Senturk, “Early diagnosis of Parkinson’s disease using machine learningalgorithms”.MedicalHypotheses,138,doi.org/10.1016/j.mehy.2020.109603, 2020.

[40] Tuncer, Turker, Dogan, Sengul, Acharya, and U. Rajendra, “Automated detection of Parkinson's disease using minimum average maximum tree and singular value decomposition method with vowels,” Biocybernetics and Biomedical Engineering, vol. 40(1), pp. 211-220, 2020.

[41] W.A. Mir, Izharuddin and I. Nissar, “Contribution of application of deep learning approaches on biomedical data in the diagnosis of neurological disorders: A review on recent findings,” pp. 87-97, ICCISIoT, CCIS, 13-14 Dec 2019, Agartala India, springer,2020.

[42] A. Frid, A. Kantor, D. Svechin and L. M. Manevitz, “Diagnosis of Parkinson’s disease from continuous speech using deep convolutional networks without manual selection of features,” IEEE Int Conf Sci Electr Eng ICSEE, 16-18 Nov 2016, Eilat Israel, IEEE, 2017.

[43] A. Caliskan, H. Badem, A. Basturk, M .E. Yuksel, “Diagnosis of the Parkinson disease by using deep neural network classifier,” Istanbul Univ - J Electr Electron Eng., vol. 17 (2), pp. 3311–3318, 2017.

[44] A.H. Al-Fatlawi, M.H. Jabardi, and S.H. Ling, “Efficient diagnosis system for Parkinson’s disease using deep belief network,” IEEE Congr Evol Comput CEC, pp. 1324-1330, 24-29 July 2016, Vancouver, BC, Canada. IEEE, 2016.

[45] S. Grover, S. Bhartia, Akshama, A. Yadav, and K. R. Seeja, “Predicting severity of Parkinson’s disease using deep learning,” Procedia Comput Sci., vol. 132, pp. 1788–1794, 2018.

[46] S.S.Upadhya, and A.N. Cheeran, “Discriminating Parkinson and healthy people using phonation and cepstral features of speech,” Procedia Comput Sci., vol. 143, pp. 197–202, 2018.

[47] P. C. Rajagopal, T. Choudhury, A. Sharma, and P.Kumar, “Diagnosis of Parkinson’s diseases using classification based on voice recordings,” Emerging Trends in Expert Applications and Security,Springer Singapore, vol. 841, pp. 575-581, 2019.

[48] D. Rizvi, I. Nissar, S. Masood, M. Ahmed, and F. Ahmad, “An LSTM based Deep Learning model for voice-based detection of Parkinson’s disease,” IJAST, vo. 29, pp. 337-343, 2020.

[49] S. Kaur, H. Aggarwal, and R. Rani, “Hyper-parameter optimization of deep learning model for prediction of Parkinson’s disease,” Machine Vision and Applications, doi.org/10.1007/s00138-020-01078-1, 2020.

[50] K. Akyol, “Comparing of deep neural networks and extreme learning machines based on growing and pruning approach,” Expert Systems With Applications, vol. 140, doi.org/10.1016/j.eswa.2019.112875 , 2020

[51] Y. Song, J. Huang, D. Zhou, H. Zha, C.L. Giles. “IKNN: Informative K-Nearest Neighbor pattern classification,’ Knowl Discov Databases PKDD, vol. 4702, pp. 248–264, 2007.

[52] Y.Bengio, “Practical recommendations for gradient-based training of Deep Architectures. Neural Networks: Tricks of the Trade,” 2nd Edition, pp.437–478,2012.

[53] Ali Reshi, Junaid, and Satwinder Singh. "Investigating the role of code smells in preventive maintenance." *Journal of Information Technology Management* 10.4 (2018): 41-63.

[54] Shaikh, Tawseef Ayoub, and Rashid Ali. "Applying machine learning algorithms for early diagnosis and prediction of breast cancer risk." *Proceedings of 2nd International Conference on Communication, Computing and Networking*. Springer, Singapore, 2019.