IETE Journal of Research

ISSN: 0377-2063 (Print) 0974-780X (Online) Journal homepage: http://www.tandfonline.com/loi/tijr20

A Comparative Study of Existing Machine Learning Approaches for Parkinson's Disease Detection

Gunjan Pahuja & T. N. Nagabhushan

IETE JOURNAL OF RESEARCH
https://doi.org/10.1080/03772063.2018.1531730

Check for updates

A Comparative Study of Existing Machine Learning Approaches for Parkinson's Disease Detection

Gunjan Pahuja[1] and T. N. Nagabhushan[2]

[1]Department of Computer Science & Engineering, JSSATEN affiliated to Dr. A.P.J Abdul Kalam Technical University, Noida, UP, India;
[2]Department of Information Science & Engineering, SJCE, Mysuru, India

ABSTRACT

Parkinson's disease (PD) has affected millions of people worldwide and is more prevalent in people, over the age of 50. Even today, with many technologies and advancements, early detection of this disease remains a challenge. This necessitates a need for the machine learning-based automatic approaches that help clinicians to detect this disease accurately in its early stage. Thus, the focus of this research paper is to provide an insightful survey and compare the existing computational intelligence techniques used for PD detection. To save time and increase treatment efficiency, classification has found its place in PD detection. The existing knowledge review indicates that many classification algorithms have been used to achieve better results, but the problem is to identify the most efficient classifier for PD detection. The challenge in identifying the most appropriate classification algorithm lies in their application on local dataset. Thus, in this paper three types of classifiers, namely, Multilayer Perceptron, Support Vector Machine and K-nearest neighbor have been discussed on the benchmark (voice) dataset to compare and to know which of these classifiers is the most efficient and accurate for PD classification. The Voice input dataset for these classifiers has been obtained from UCI machine learning repository. ANN with Levenberg–Marquardt algorithm was found to be the best classifier, having highest classification accuracy (95.89%). Moreover, we compared our results with those obtained by Resul Das ["A comparison of multiple classification methods for diagnosis of Parkinson Disease," *Expert Systems and applications*, vol. 37, pp 1568–1572, 2010].

KEYWORDS

Artificial neural networks (ANN); K-nearest neighbors (KNN); Parkinson's disease (PD); support vector machine (SVM)

1. INTRODUCTION

Parkinson's disease (PD) is a progressive neurodegenerative disorder of the nervous system that affects our body movements including speech. Dr. James Parkinson in 1817 [1] discovered this disease and described the condition which he called the 'Shaking Palsy'. Neurodegenerative diseases are defined as hereditary and sporadic conditions which are characterized by dysfunction of the progressive nervous system (JPND research, 2015). Out of many neurodegenerative diseases like Alzheimer's disease, Brain Cancer, Degenerative Nerve Diseases and Epilepsy, "Parkinson's Disease" is considered to be the second most common neurodegenerative disease [2].

PD is mainly caused by the progressive loss of dopamine neurons in the area of the midbrain called substantia nigra – the "movement control center" of the brain (Figure 1). Loss of dopamine causes the neurons to fire out-of-control movements called hypokinetic movement disorder [3]. Although this disease can be diagnosed easily in the advanced stage, effective treatment is still very challenging. To date, there exists no cure/medical treatment for PD.

After decades of exhaustive study, the causes of PD are still unknown. Many of the researchers think that a combination of genetic [4] and environmental factors [5], such as exposure to the environmental toxin, head injury, rural living, drinking water, manganese and exposure to pesticides, are responsible for PD. These factors may vary from person to person. Also, there are some specific symptoms that an individual experience and each PD patient experience these symptoms differently. Description of different stages of PD is reported in Table 1. Primary motor symptoms of PD include tremor of the hands, arms, legs, jaw and face, bradykinesia or slowness of movement, rigidity or stiffness of the limbs and trunk and postural instability or impaired balance and coordination [6–8]. In addition to these symptoms, there are some non-motor symptoms like depression and loss of memory which may occur and affect the quality of life [9,10]. At the advanced stage, PD can be easily and accurately diagnosed, but effective treatment is a challenging task. Also, if treatment is started in advanced stages, it might have less effective in controlling PD progression. This situation necessitates the early and accurate diagnosis of PD, thus helps the patients

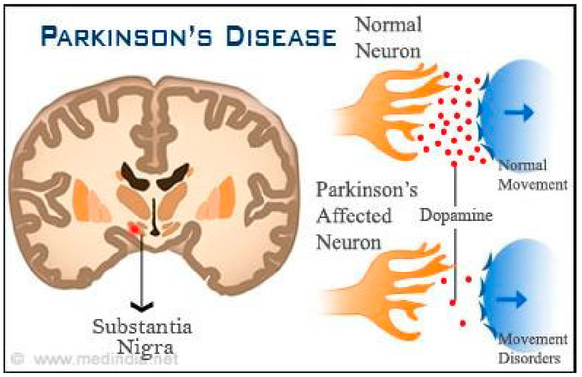

Figure 1: Parkinson's disease (normal movement vs. movement disorders) [www.medindia.net]

in maintaining a good quality of life. To date, no single blood or laboratory test exists that is helpful in the identification of PD and its progression. However, rating methods such as Hoehn and Yahr scale (1967), Unified Parkinson Disease Rating Scale (UPDRS) and its modified version MDS-UPDRS [11] are sometimes used for the early diagnosis of PD. Certain drawbacks associated with these methods are as follows:

(1) Availability of skilled workforce
(2) Time and cooperation required from patients for a longer period [12] etc.

Sometimes, it would be difficult to distinguish between various neurological disorders because they share the same etiology. Approximately 75% of clinical diagnosis of PD is confirmed to be idiopathic PD. Thus, automatic methods based on machine learning are required to improve diagnosis accuracy rate and to help the doctors to make right decisions.

1.1 Objective

The objectives and contribution of our research paper are as follows:

(a) To present a comprehensive survey including the most recent research papers up to year 2017.
(b) To offer a wide range of comparison in diverse angles and perspectives in terms of data acquisition, feature extraction, feature subset selection, different classifiers and result comparison organization.
(c) To compare the accuracy of existing classifiers on vocal dataset available from UCI repository and also to validate the performance of implemented classifiers on two other benchmark datasets.
  (1) Wisconsin Breast Cancer Database
  (2) Pima Indians Diabetes Dataset
(d) To recommend the potential opportunity for automatic diagnosis of PD.

2. REVIEW OF THE LITERATURE

PD is also termed as idiopathic or primary Parkinsonism /hypokinetic-rigid syndrome. From the literature, it is clear that various machine learning approaches have been used for classification of PD by undertaking the vocal and gait features [13–21]. MA Little et al. [2007] received the speech signals and created a database, in collaboration with National Center for Voice and Speech, Denver, Colorado. The authors used kernel-support vector machine for PD classification [22]. Till now several studies have been conducted using Little's PD voice dataset and different values of accuracy have been achieved using different classification algorithms. Resul Das [13] used the same dataset created by Little et al. and compared four independent classification approaches (neural networks, data mining neural, logistic regression and decision trees) for diagnosis of PD. Among the four approaches, the best performance of 92.9% was yielded by Multi-layer feed-forward neural network with Levenberg–Marquardt algorithm. The authors also compared the results with kernel–SVM results (from the literature) and found that the obtained results are better than kernel SVM approach. Freddie and Rasit [14] used a set of nine parallel feed-forward neural network approach on the same voice dataset for PD prediction. Although complexity has been increased, this approach yielded an improvement of 8.4% on the prediction of PD as compared to the single unique network. On similar voice

Table 1: **Stages of Parkinson's disease**

Stages	Symptoms
Mildest stage (Stage 1)	In this stage, the PD patients have least interference with routine tasks. Tremors and other symptoms are restricted to one side of the body
Moderate stage (Stage 2)	In this stage, symptoms like stiffness, resting tremors and trembling can be sensed on both sides of the body. Also facial expressions of PD patients may get changed
Mid-stage (Stage 3)	During this stage, major changes like balance loss, decreased flexes in addition with stage II symptoms will be observed in PD patients. Occupational therapy combined with medication may help in decreasing the symptoms
Progressive stage (Stage 4)	The condition of PD patient will get worse in this stage and it becomes difficult for the patient to move without some assistive device like a walker
Advanced stage (Stage 5)	Stage V is the most advanced and debilitating stage of PD. Stiffness in legs may cause freezing when standing. Patients are frequently unable to stand without falling. They may experience hallucinations and occasional delusions

dataset, Hui-Ling Chen et al. [23] used fuzzy k-nearest neighbor approach with Principal Component Analysis for predicting PD and constructing the feature subset from the whole feature space. The authors reported that their proposed method outperformed the other methods in the literature.

Omer et al. [16] compared the performance of LS-SVM, SVM, MLPNN and GRNN in the remote tracking of PD progression. It was observed that LS-SVM outperforms the other methods while mapping the vocal features to UPDRS data.

It is clear from the literature that most of the PD patients exhibit gait disorder [20] along with vocal impairment. You-Yin et al. [24] developed a gait regression model for predicting the severity of motor dysfunction from gait image sequence. The studies done so far also indicate that there is a loss of neurons in dopamine region of the brain in the individuals affected by PD. Thus, over the past 2 decades, neuroImaging techniques, such as MRI; SPECT; fMRI and PET, have been used to visually assess and quantify the loss of neurons in different lobes of the brain [25–27]. MRI is preferred over others because of non-invasiveness and high spatial resolution quality [28,29]. In the literature, various machine learning techniques/approaches exist that are found to be effective for diagnosing PD patients using neuroImaging techniques.

The changes in the functional connectivity of motor networks in the resting state in PD, using fMRI and a network model based on graph theory, were demonstrated by Tao et al. [30]. The authors found that functional connectivity in the supplementary motor area, left dorsal lateral prefrontal cortex and left putamen of PD patients at off state had significantly decreased while functional connectivity in the left cerebellum, left primary motor cortex and left parietal cortex had increased as compared to normal subjects in PD. Defeng et al. [31] conducted a real-time case study using deep brain electrode implantation to predict the PD tremor. Similarly, Christian Salvotre et al. [32] used a dataset of MRI scans from 28 controls, 28 PD patients and 28 Progressive Supranuclear Palsy. Supervised machine learning algorithm was used based on PCA as a feature extraction method and SVM as a classification algorithm. The authors have tried to overcome the problem of imbalance dataset by taking the same number of patients of different classes (PD, HC and PSP). Nowadays many classifiers are available for PD detection and their performance is measured with metrics such as accuracy, sensitivity and specificity [15,17,33,34]. In general, the accuracy is a measure of how many cases are correctly identified in total irrespective of positive or negative cases or it measures the overall performance of the method. Table 2 describes some of the studies available in the literature for PD diagnosis and classification using machine learning approaches.

2.1 Feature Subset Selection (FSS) Techniques

The diagnosis of neurodegenerative diseases through machine learning approaches includes the following:

(1) Data acquisition (Brain MRI images, gait movements, vocal data, local field potential etc.).
(2) Feature extraction (extract the features suitable for training and testing a classifier).
(3) Feature subset selection (to reduce the redundant features).
(4) Training and validating the performance of the classifier.

Figure 2 shows the steps involved in medical image processing (MIP) using machine learning techniques.

In the literature, a variety of machine learning algorithms exist such as induction-based (ID3, CART) and instance-based algorithms (IBL) for medical imaging classification. But, these algorithms degrade the prediction accuracy because of the availability of many features that are not necessary for predicting the output. Thus there is a need for FSS methods which optimize the number of features by selecting the relevant subset and thus improve the classification accuracy. A typical FSS consists of 4 basic steps: Subset Generation, Subset Evaluation, Stopping Criterion and Result Validation [39].

Subset generation procedure is a search procedure that produces feature subsets for evaluation based on predefined criterion [40]. An evaluation function is used to evaluate the subset under examination, the stopping criterion is used to decide when to stop and a validation procedure is used to check whether the subset is valid. Based on different evaluation criteria, FSS algorithms are categorized into three categories (1) the filter model, (2) the wrapper model [39] and (3) hybrid model. In all the categories, algorithms can be further differentiated by how the space of feature subsets is explored and the exact nature of their evaluation function.

The filter model relies on general characteristics of the data to evaluate and select the feature subsets without involving any learning algorithm. But, sometimes the filter method fails to select the right subset of features if the applied criterion deviates from the one that is used for training purpose. Also, the filter approach may fail to

Table 2: Literature survey for diagnosis of Parkinson's disease using machine learning approaches

Study	Dataset	Method	Results
Song Pan et al. [15]	Local field potential signals	Radial Basis Function+ Support Vector Machine + Multilayer Perceptron	Accuracy SVM: 81.14% RBF:80.13% MLP:79.25%
Sang-Hong Lee and Joon S. Lim [17]	Gait characteristics	Wavelet-based feature extraction, +Neural Network with weighted fuzzy membership functions	Accuracy:77.33%
G. Sateesh Babu and S. Suresh [18]	Gene expressions	ICA+ Meta-cognitive neural classifier	Accuracy:95.55%
R. Armananzas et al. [35]	Movement disorder	Wrapper feature selection + 5 classifiers: Naïve Bayes (NB), k-nearest neighbors LDA, C4.5 decision trees, ANN	Accuracy 1. NB:82.08% 2. KNN:80.06% 3. LDA:83.24% 4. C 4.5:81.50% 5. ANN:64.74%
G.S. Babu et.al [33]	Brain MRI images	Voxel-Based Morphometry + PBL-McRBFN+ RFE	Accuracy:87.21%
F.J. Martinez-Murcia et al. [36]	DaTSCAN Images	Independent Component Analysis (ICA) + Support Vector Machines(SVM)	Accuracy on 1. PPMI dataset = 91.3% and 2. Virgen dela Victoria" Hospital in Málaga (VV), Spain-94.7%
G. Singh and L. Samavedham [37]	T1-weighted MRI Images	Kohonen Self Organizing Map+ Least Square Support Vector Machine	Accuracy: 99.9% (For classifying PD, HC and SWEDD subjects)
A. Benba et al. [38]	Voice Assessment	Principal Component Analysis+ Support Vector Machine	Accuracy: 87.50% (On 3 vowel samples /a/,/o/,/u/)
L. Naranjo [21]	Acoustic features y extracted from replicated voice recordings	Gibb's Sampling Algorithm +Bayesian Approach	Accuracy: 86.2% Sensitivity:82.5% Specificity:90.0%

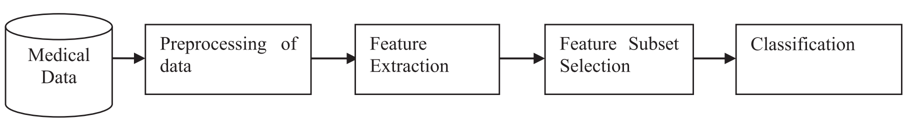

Figure 2: Steps involved in medical image processing (MIP) using machine learning techniques

find a feature subset that would jointly maximize the criterion, thus degrading the performance of the learning model [41,42]. On the other hand, the wrapper method requires a learning algorithm and uses its performance as the evaluation criterion. Wrappers can show even better results than others by considering prediction accuracy. But, wrapper models are less general and are computationally expensive than filter models because they need more computational resources and use specific learning algorithm [43,44]. Since filters execute many times faster than wrappers, there is a much better chance of scaling to databases with a large number of features in filter approach than wrappers. Also, filters do not require re-execution for different learning algorithms. Thus filters can provide the same benefits for learning as wrappers do.

The hybrid model combines the advantage of the filter and wrapper model by utilizing different evaluation criteria in different search stages [45]. Although the hybrid methods are more efficient than wrapper and filter approaches, they are much complex and limited to a specific learning machine [46,47]. Much work has been done in this field as well [48]; different researchers have mentioned advantages and disadvantages of filter and wrapper approaches. Table 3 indicates FSS/dimensionality reduction methods currently available in the literature for reducing the dimensionality or removing the irrelevant/redundant features in the case of PD detection and classification using machine learning methods.

2.2 Classification

After feature extraction and subset selection, the next phase is the classification. It is an instance of supervised learning and can be defined as a problem of identifying the category, to which new observation will belong. Various methods used for classification are categorized as (a) Statistical Algorithms, (b) Pattern Recognition and

Table 3: Feature subset selection techniques

Study	Dataset	Feature subset selection technique
Y.-Y. Chen et al. [24]	Monocular image sequences	Linear discriminant analysis
R. Armananzas et al. [35]	Movement Disorders	Wrapper feature selection scheme
G.S. Babu et al. [33]	Brain MRI images	Wrapper Method based on Recursive feature elimination
M. Hariharan et al. [19]	Vocal dataset	Principal Component Analysis +Linear Discriminant Analysis
F.J. Martinez-Murcia et al. [36]	DaTSCAN Images	Independent Component Analysis (ICA)
B. Rana et al. [34]	T1-weighted MRI images	Filter feature selection approach (based on mutual information)
P. Shrivastava et al. [49]	Gait and Voice dataset	Evolutionary approaches (like Bat Algorithm, Cuckoo Search algorithm, PSO and Genetic Algorithms)

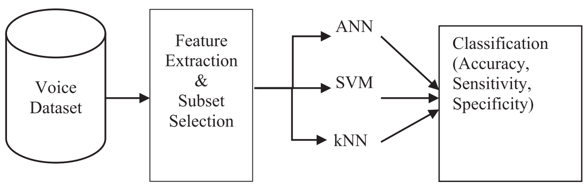

Figure 3: Methods applied for PD classification

learning-based algorithms, (c) Search heuristics and a combination of algorithms.

In statistical approaches, the computation of mean, standard deviation of the features in the template is done. Distance techniques such as Euclidean distance, weighted Euclidean distance and Manhattan distance are used for comparing the training data with the testing data.

Pattern recognition is defined as an act of taking raw data and classifying them into different categories based on machine learning algorithms such as K-NN rule, Bayes classifier, SVM, artificial neural networks (ANN) [13] and clustering techniques like K-means [13,16,19,35,50].

Various evolutionary algorithms such as Ant colony optimization and Particle swarm optimization can also be used for classification purpose [18,49,51,52]. The advantage of using these evolutionary algorithms is that they can handle large databases. Figure 3 depicts the methods applied for PD classification in this study.

3. MATERIALS AND METHODS

This section describes the methods and materials used in this study for classifying the PD patients from healthy subjects.

3.1 Dataset

In this paper, dataset of PD patients regarding general voice disorders has been used. MA Little [53] of the University of Oxford, in collaboration with the National Centre for Voice and Speech, Denver, Colorado, recorded the speech signals and created the database. This database is freely available and can be easily downloaded from UCI repository. There are six recordings per patient. The first column of the dataset specifies the name of the patient and the last column specifies the status which is set to 1 for PD and 0 for healthy subjects.

In addition to classifying the PD patients using voice dataset, we had also evaluated the classifiers performance on two other benchmark datasets available from UCI repository. Table 4 summarizes the benchmark datasets used in this study.

3.2 Applied Methods

3.2.1 Artificial Neural Network (ANN)

ANN symbolizes a parallel architecture that is motivated by the way how biological neural processing takes place. Although many types of ANN architectures exist, MLP (multi-layer feed-forward neural network) is the most commonly used architecture (Figure 4). Backpropagation algorithm proposed by Rumelhart in 1986 is a generalized delta rule that is utilized by MLP Network for the adjustment of weights [13,16,54]. Levenberg–Marquardt, Gradient descent scaled conjugate gradient, and Resilient back propagation are some of the variants of the Backpropagation algorithm. According to M.T. Hagan and M. Menhaj [55], for small- and medium-sized networks, the Levenberg–Marquardt algorithm is efficient and strongly recommended for neural network training; therefore, the same algorithm has been implemented here.

3.2.2 Support Vector Machine (SVM)

SVM is considered to be a supervised classification approach. Vapnik [56] first proposed SVM for binary classification. Binary classification is based on the concept of dividing the data into classes using a hyperplane.

For the linear classification problems, SVM is considered as an extension of the perceptron. From Figure 5, it is clear that the distance between the 2 hyperplanes is $2/||w||$. So, the optimization problem is to reduce/minimize $||w||$ or to maximize the margin

Table 4: Summary of Benchmark datasets

Title	Features	Instances	Classes
Parkinson's disease – voice dataset (https://archive.ics.uci.edu/ml/datasets/Parkinsons)	23	197	2 (Binary)
Wisconsin Breast Cancer database (http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original))	10	699	2 (Binary)
Pima Indians Diabetes Dataset (http://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes)	8	768	2 (Binary)

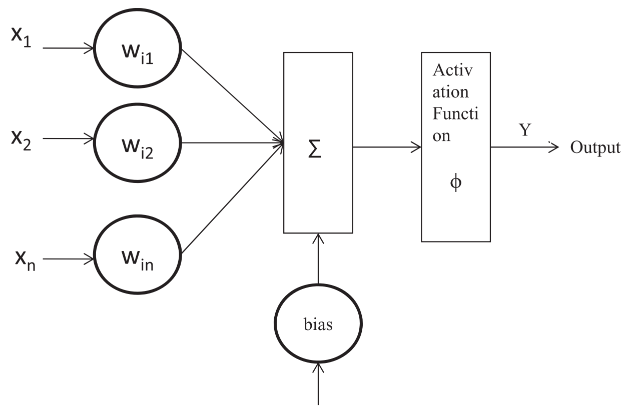

Figure 4: Artificial neural network architecture

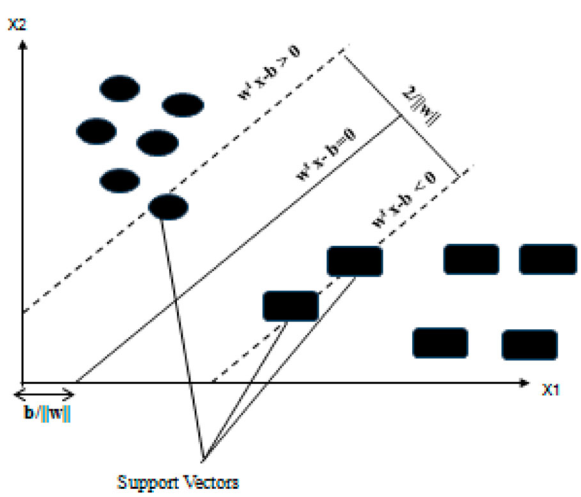

Figure 5: SVM trained with data/samples from 2 classes

between the support vectors. Thus the binary optimization problem can be stated as

$$\arg\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{n}\xi$$

where C is a tuning parameter and $C > 0$ and ξ is required to tolerate misclassification.

If $w^T x + b \geq 0$, then SVM predicts "1", otherwise it will predict "-1". The decision boundary is given by $w^T x + b = 0$. In this paper, SVM has been chosen as a classifier for classifying PD patients from normal subjects because of its computational capability in dealing with overfitting and dimensionality problems that generally occurs during classification [15,36–38].

3.2.3 K-nearest Neighbor (K-NN)

K-NN is one of the nonparametric classification approaches used in machine learning [57]. The result of K-NN approach depends on the type of output required for the particular applications. The class is assigned to the object that is most common among the K-nearest neighbors. If $K = 1$, then the class of that single nearest neighbor is assigned to that object. Hui-Leng Chen et al. [23] presented an efficient diagnosis system for PD detection using fuzzy k-nearest neighbor approach.

4. RESULTS AND DISCUSSIONS

This section discusses the results obtained using ANN, K-NN and SVM for classifying the PD patients from healthy subjects using voice dataset. The same dataset has been used by various researchers to prove their studies for PD detection, Resul Das [13], Hui-Leng Chen et al. [23], G.S. Babu [18] to name a few. All the implementations in this study are carried out using Matlab R2013a.

4.1 ANN with Levenberg–Marquardt Algorithm and Scaled Conjugate Gradient Algorithms

The neural network architecture used for classification is a feed-forward back-propagation network. Backpropagation has been used in this study based on Levenberg–Marquardt optimization and scaled conjugate gradient method with 10 neurons in the hidden layer. The input dataset is randomly partitioned into training, testing and validation dataset for ANN classification. The initial weights were chosen randomly. The tuning of all the parameters for ANN classification has been done as in [13].

4.2 K-Nearest Neighbor (K-NN)

K-NN (lazy learning) is considered to be the simplest algorithm among the various machine learning algorithms available for classification. The predictions made by this method are based on the outcome of 'K' neighbors that are closest to that point. Various distance metrics such as Euclidean, Euclidean squared, Cityblock distances can be used for calculating the distance between the sample cases (q) and query point (y). Here, the statistical features of K-NN have been evaluated using Euclidean and cityblock distance metrics.

$$D(y,q) = \sqrt{(y-q)^2} \qquad \text{Euclidean Distance}$$
$$= \text{abs}(y-q) \qquad \text{Cityblock}$$

4.3 Support Vector Machine (SVM)

SVM classifier with RBF, polynomial and linear kernel functions has been implemented in Matlab. To validate classifier accuracy, 10-fold cross-validation [43] was used. The advantage of 10-fold CV approach is that all the test sets are independent and thus the reliability of results could be improved. Since the voice dataset is an imbalanced dataset, we are using "stratified sampling" to split the data. Stratified 10-fold CV ensures the same class distribution in the subset, thus sample proportion in each data subset is the same as that in population. 10-fold CV means that cross-validation process is repeated 10 independent times and then results are averaged to produce a single estimation.

Table 5 shows the performance comparison of ANN, SVM and K-NN with different variants on PD voice dataset. It is clear from Table 5 that ANN with Levenberg–Marquardt, K-NN with Euclidean distance and SVM with RBF kernel outperformed the other variants that are investigated in this study. Since it is an imbalanced dataset along with performance parameters such as accuracy, sensitivity and specificity, geometric mean of the true rates is also calculated. The significance of this metric is that geometric mean tries to maximize the accuracy on each of the 2 classes with a good balance. Another approach to produce evaluation criteria is to make use of ROC curve. Since the aim of this study is to compare the existing machine learning approaches for PD classification, ROC (Receiver Operating Characteristic) curve only for this case is shown in Figure 6. ROC curve is used for validating the classifier performance. This curve specifies the true positive rate vs. false positive rate for different thresholds of the classifier output. From Figure 6, true positive rate vs false positive rate can be easily identified. Figure 7 demonstrates the validation performance graph using Levenberg-Marquardt Algorithm. There is no problem of overfitting the data, because test curve and validation curve are similar. Some overfitting could have occurred if the test curve had increased significantly before validation curve increased. For comparison purpose, classification accuracies of the previous methods which were investigated on for PD diagnosis using voice data are listed in Table 6. In order to further validate the effectiveness of these methods, we implemented the same algorithms on two other benchmark datasets (Table 4) and results are compiled in Table 7.

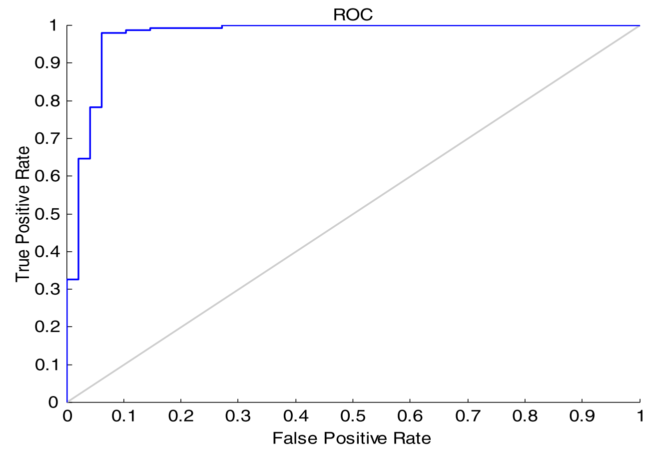

Figure 6: ROC curve showing true positive rate vs. false positive rate (Levenberg–Marquardt algorithm)

5. CHALLENGES AND ISSUES

Ingrid Scholl et al. [58] discussed Kilo-to-Tera byte challenges in MIP. These challenges are related to management and mining of medical images, bioimaging, neuroImaging and virtual reality in medical visualizations. Technological advancement has enabled Peta-byte availability for medical imaging and hence

Table 5: Performance comparison of ANN, KNN and SVM on PD voice dataset

	ANN		KNN		SVM		
Variants → Performance parameters ↓	Levenberg– Marquardt algorithm	Scaled conjugate gradient	Euclidean distance	Cityblock distance	RBF kernel	Polynomial kernel	Linear kernel
Classification accuracy	95.89	85.12	72.31	69.74	88.21	81.03	82.9
Sensitivity	93.75	70	68.75	66.67	91.67	79.17	87.33
Specificity	96.59	96.59	73.47	70.75	77.55	87.76	78.56
Geometric mean	95.16	82.23	71.07	68.68	84.31	83.35	82.83

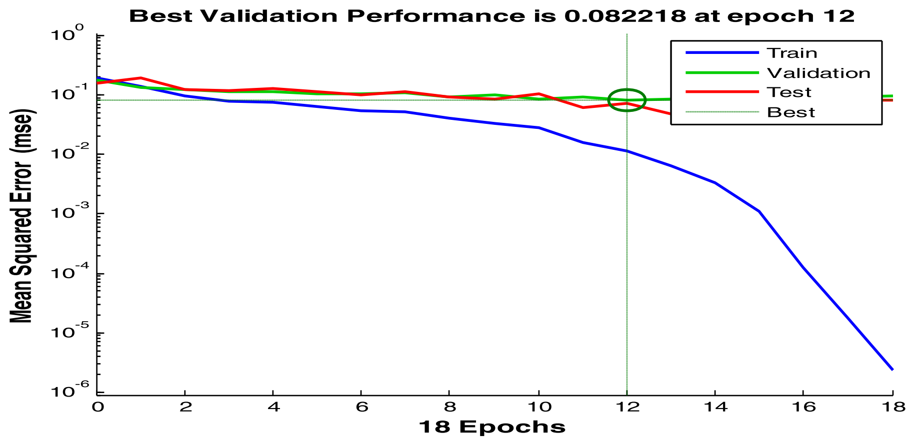

Figure 7: Graph showing the validation performance (Levenberg–Marquardt algorithm)

Table 6: Classifier performance comparison with studies available in the literature on vocal dataset

Study	Method	Accuracy (%)
R. Das [13]	ANN	92.9
F. Astrom and R. Kokar [14]	9 parallel neural networks	91.2
A. Khemphila and V. Boonjing [54]	Information Gain+ ANN	83.33
H.-L. Chen et al. [23]	PCA+FKNN	96.07
A. Benba et al. [38]	PCA+SVM	87.21

addresses byte challenge. The other two biggest challenges that still exist in MIP are the dataset and computational power, e.g. G.S. Babu et al. [33] developed the meta-cognitive algorithm for the identification of the brain regions responsible for PD using RFE approach. 87.21% accuracy was achieved but the computational cost was high. From machine learning perspective, dataset must be clean and of significant size to solve the problem. However, availability of clean dataset is limited due to the nature of complexity.

Dataset collection has some inherent challenges like "class imbalance problem" [23] and presence of noise and outliers in the dataset. Class imbalance problem means that the total number of samples from one class of data (+ve) are not equal to the total number of samples from other class of data (−ve). This problem exists not only in medical diagnosis but also approximately in all fields where "Machine Learning" is used such as face recognition and biometrics. This problem may be overcome by using balanced dataset, so that decision model can learn without bias. The presence of noise and outliers during data collection can lead to poor diagnosis. Thus, preprocessing of medical data is a necessary step and must be handled automatically. Post removal of noise and outliers, medical images can be processed and analyzed to extract meaningful information such as volume, shape, motion of organs which are helpful in the diagnosis of the disease and abnormalities.

6. CONCLUSION

Research highlights that 90% of people with PD exhibit vocal impairment. Vocal impairment or disorders of voice means that voice will sound hoarse, strained or effortful. Several studies have been done to automate the PD diagnosis using voice dataset. In this paper, the performance of ANN, KNN and SVM classifiers has been

Table 7: Performance Comparison of ANN, KNN and SVM on Wisconsin breast cancer dataset and Pima Indians diabetes dataset

	Variants →	ANN		KNN		SVM		
Datasets	Performance parameters↓	Levenberg–Marquardt algorithm	Scaled conjugate gradient	Euclidean distance	Cityblock distance	RBF kernel	Polynomial kernel	Linear kernel
Wisconsin Breast Cancer Database	Classification accuracy	98	97	73.33	72.31	96.71	90.1	95.02
	Sensitivity	97.8	97.16	68.75	66.67	96.29	92.16	96.72
	Specificity	95.85	98.3	74.83	74.15	97.51	88.8	94.51
	Geometric mean	96.82	97.73	71.73	70.31	96.90	90.46	95.61
Pima Indians Diabetes Dataset	Classification accuracy	81.11	78.51	72.82	72.31	75.01	73.16	74.61
	Sensitivity	90	80.62	68.75	68.75	73.4	77.4	78.3
	Specificity	68.33	73.3	74.15	73.47	72.76	69.4	71.04
	Geometric mean	78.42	76.87	71.40	71.07	73.08	73.29	74.58

evaluated using sensitivity, specificity, total classification accuracy and geometric mean on voice database. Similar discussion is also carried out for Wisconsin Breast Cancer database and Pima Indians Diabetes Dataset. It is observed that Artificial Neural Networks with Levenberg–Marquardt algorithm gives the highest classification accuracy of 95.89% for voice dataset. We believe that the use of machine learning techniques as discussed here will be a great support to the doctors. Although a large number of techniques are available for PD diagnosis their performance is still imperfect. Hence, to improve the accuracy of CAD algorithms, there is a need for further enhancements. In future, we will attempt to use other evolutionary algorithms like Genetic algorithm and Extreme Learning Machine for PD detection and classification.

REFERENCES

1. J. Parkinson, *An Essay on Shaking Palsy*. London: Whittingham and Rowland Printing, 1817.
2. D. B. Calne, "Is idiopathic parkinsonism the consequence of an event or a process," *Neurology*, Vol. 44, no. 15, pp. 5–5, 1994.
3. A. E. Lang and A. M. Lozano, "Parkinson's disease first of Two parts," *New England J. Med.*, Vol. 339, pp. 1044–53, 1998.
4. A Samii, J. G. Nutt, and B. R. Ransom, "Parkinson's disease," *Lancet*, Vol. 363, no. 9423, pp. 1783–93, 2004.
5. L. M. de Lau and M. M. Breteler, "Epidemiology of Parkinson's disease," *Lancet Neurol.*, Vol. 5, pp. 525–35, 2006.
6. E. M. Morris, "Movement disorder in people with Parkinson disease: A model for physical therapy," *Phys. Ther.*, Vol. 80, pp. 578–97, 2000.
7. A. Schrag, C. D. Good, K. Miszkiel, H. R. Morris, C. J. Mathias, A. J. Lees, and N. P. Quinn, "Differentiation of atypical parkinsonian syndromes with routine MRI," *Neurology*, Vol. 54, pp. 697–702, 2000.
8. R. Angel, W. Alston, and J. R. Higgins, "Control of movement in Parkinson's disease," *Brain*, Vol. 93, no. 1, pp. 1–14, 1970.
9. S. L. Wu, R. M. Liscic, S. Kim, S. Sorbi, and Y. H. Yang, "Nonmotor symptoms of Parkinson's disease," *Parkinson's Dis.*, 2017. DOI:10.1155/2017/4382518.
10. T. Yousaf, H. Wilson, and M. Politis, "Imaging the nonmotor symptoms in Parkinson's disease," *Int. Rev. Neurobiol.*, Vol. 133, pp. 179–257, 2017.
11. C. G. Goetz, et al., "Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Mov. Disord.*, Vol. 23, no. 15, pp. 2129–70, 2008.
12. B. Post, M. P. Merkus, R. M. de Bie, R. J. de Haan, and J. D. Speelman, "Unified Parkinson's disease rating scale motor examination: Are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable?," *Movement Dis.*, Vol. 20, pp. 1577–84, 2005.
13. R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Syst. Appl.*, Vol. 37, pp. 1568–72, 2010.
14. F. Astrom and R. Koker, "A parallel neural network approach to prediction of Parkinson's disease," *Expert Syst. Appl.*, Vol. 38, pp. 12470–4, 2011.
15. S. Pan, S. Iplikci, K. Warwick, and T. Z. Aziz, "Parkinson's disease tremor classification – a comparison between support vector machines and neural networks," *Expert Syst. Appl.*, Vol. 39, pp. 10764–71, 2012.
16. O. Eskidere, F. Ertas, and C. Hanilci, "A comparison of regression methods for remote tracking of Parkinson's disease progression," *Expert Syst. Appl.*, Vol. 39, pp. 5523–8, 2012.
17. S.-H. Lee and J. S. Lim, "Parkinson's disease classification using gait characteristics and wavelet-based feature extraction," *Expert Syst. Appl.*, Vol. 39, pp. 7388–44, 2012.
18. G. Sateesh Babu, and S. Suresh, "Parkinson's disease prediction using gene expression – a projection based learning meta-cognitive neural classifier approach," *Expert Syst. Appl.*, Vol. 40, pp. 1519–29, 2013.
19. M. Hariharan, K. Polat, and R. Sindhu, "A new hybrid intelligent systems for accurate detection of Parkinson's disease," *Comp. Methods Prog. Biomed.*, Vol. 113, pp. 904–13, 2014.
20. W. Zeng and C. Wang, "Classification of neurodegenerative diseases using gait dynamics via deterministic learning," *Inform. Sci.*, Vol. 317, pp. 246–58, 2015.
21. L. Naranjo, C. J. Perez, J. Martín, and Y. Campos-Roca, "A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications," *Comp. Methods Prog. Biomed.*, Vol. 142, pp. 147–56, 2017.
22. 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, pp. 1015–22, 2009.
23. H.-L. Chen, C-C. Huang, X-G. Yu, X. Xu, X. Sun, G. Wang, and S-J. Wang, "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach," *Expert Syst. Appl.*, Vol. 40, pp. 263–71, 2013.

24. Y.-Y. Chen, et al., "A vision-based regression model to evaluate parkinsonian gait from monocular image sequences," *Expert Syst. Appl.*, Vol. 39, pp. 520–6, 2012.

25. P. Piccini, and A. Whone, "Functional brain imaging in the differential diagnosis of Parkinson's disease," *Lancet Neurol.*, Vol. 3, pp. 284–90, 2004.

26. A. G. Filler, "The history, development and impact of computed imaging in neurological diagnosis and neurosurgery: CT, MRI, and DTI," *Nature*, Vol. 7, pp. 1–69, 2009.

27. B. S. Mahanand, S. Suresh, N. Sundararajan, and M. Aswatha Kumar, "Identification of brain regions responsible for Alzheimer's disease using a self-adaptive resource allocation network," *Neural Netw.*, Vol. 32, pp. 313–22, 2012.

28. A. Schrage, M. Jahanshahi, and N. Quinn, "How does Parkinson's disease affect quality of life? A comparison with quality of life in the general population," *Movement Disord.*, Vol. 15, pp. 1112–8, 2000.

29. B. Ravina, et al., "The role of radiotracer imaging in Parkinson disease," *Neurology*, Vol. 64, pp. 208–15, 2005.

30. T. Wus, L. Wang, Y. Chen, C. Zhao, K. Li, and P. Chan, "Changes of functional connectivity of the motor network in the resting state in Parkinson's disease," *Neurosci. Lett.*, Vol. 460, pp. 6–10, 2009.

31. D. Wu, K. Warwick, Z. Ma, J. G. Burgess, S. Pan, and T. Z. Aziz, "Prediction of Parkinson's disease tremor onset using radial basis function neural networks," *Expert Syst. Appl.*, Vol. 37, pp. 2923-2928, 2010.

32. C. Salvatore, et al., "Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and progressive supranuclear palsy," *J. Neurosci. Methods*, Vol. 222, pp. 230–37, 2014.

33. G. Sateesh Babu, S. Suresh, and B. S. Mahanand, "A novel PBL-McRBFN-RFE approach for identification of critical brain regions responsible for Parkinson's disease," *Expert Syst. Appl.*, Vol. 41, pp. 478–88, 2014.

34. B. Rana, A. Juneja, M. Saxena, S. Gudwani, S. Senthil Kumaran, R.K. Agrawal, and M. Behari, "Regions-of-interest based automated diagnosis of Parkinson's disease using T1-weighted MRI," *Expert Syst. Appl.*, Vol. 42, pp. 4506–16, 2015.

35. R. Armananzas, C. Bielza, K. R. Chaudhuri, P. Martinez-Martin, and P. Larrañaga, " Unveiling relevant non-motor Parkinson's disease severity symptoms using a machine learning approach," *Artif. Intell. Med.*, Vol. 58, pp. 195–202, 2013.

36. F. J. Martinez-Murcia, J. M. Górriz, J. Ramírez, I. A. Illán, A. Ortiz, and the PPMI, "Automatic detection of parkinsonism using significance measures and component analysis in DaTSCAN imaging," *Neurocomputing*, Vol. 126, pp. 58–70, 2014.

37. G. Singh, and L. Samavedham, "Unsupervised learning based feature extraction for differential diagnosis of neurodegenerative diseases: a case study on early-stage diagnosis of Parkinson disease," *J. Neurosci.Methods*, Vol. 256, pp. 30–40, 2015.

38. A. Benba, A. Jilbab, and A. Hammouch, "Voice assessments for detecting patients with Parkinson's diseases using PCA and NPCA," *Int J. Speech Technol.*, Vol. 19, no. 4, pp. 743–54, 2016.

39. M. Dash, and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, Vol. 1, pp. 131–56, 1997.

40. L. Huan and H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*. Boston, MA: Kluwer Academic, 1998.

41. M. E. ElAlami, "A filter model for feature subset selection based on genetic algorithm," *Knowledge Based Syst.*, Vol. 22, pp. 356–62, 2009.

42. H. Yoon, C-S. Park, J. S. Kim, and J-G. Baek, "Algorithm learning based neural network integrating feature selection and classification," *Expert Syst. Appl.*, Vol. 40, pp. 231–41, 2013.

43. R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, Vol. 97, pp. 273–324, 1997.

44. M. Hall, "Correlation based feature selection for machine learning," Ph.D. dissertation, Dept. of Computer Science, University of Waikato, 1999.

45. A. H. Hadjahmadi and T. J. Askari, "A decision support system for Parkinson's disease diagnosis using classification and regression tree," *J. Math. Comp. Sci.*, Vol. 4, pp. 257–63, 2012.

46. O. Uncu and I. B. Turksen, "A novel feature selection approach: combining feature wrappers and filters," *Inf. Sci.*, Vol. 177, pp. 449–66, 2007.

47. J. Huang, Y. Cai, and X. Xu, "A hybrid genetic algorithm for feature selection wrapper based on mutual information," *Pattern Recog. Lett.*, Vol. 28, pp. 1825–1844, 2007.

48. D. Guan, W. Yuan, Y-K. Lee, K. Najeebullah, and M. K. Rasel, "A review of ensemble learning based feature selection," *IETE Tech. Rev.*, Vol. 31, pp. 190–198, 2014.

49. P. Shrivastava, A. Shukla, P. Vepakomma, N. Bhansali, and K. Verma, "A survey of nature-inspired algorithms for feature selection to identify Parkinson's disease," *Comp. Methods Prog. Biomed.*, Vol. 139, pp. 171–9, 2017.

50. M. Poletti, M. Emre, and U. Bonuccelli, "Mild cognitive impairment and cognitive reserve in Parkinson's disease," *Parkinsonism Related Disord.*, Vol. 17, pp. 579–586, 2011.

51. K Chandrasekaran, S. P. Simon, and N. P. Padhy, "Cuckoo search algorithm for emission reliable economic multi-objective dispatch problem," *IETE J. Res.*, Vol. 60, pp. 128–38, 2014.

52. V. Mangat and R. Vig, "Dynamic PSO-based associative classifier for medical datasets," *IETE Tech. Rev.*, Vol. 31, pp. 258–65, 2014.

53. M. A. Little, P. E. McSharry, S. J. Roberts, D. A. Costello, and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *BioMed. Eng. Online*, Vol. 6, p. 23, 2007.

54. A. Khemphila and V. Boonjing, "Parkinsons disease classification using neural network and feature selection," *World Acad. Sci. Tech*, Vol. 64, pp. 1–15, 2012.

55. M. T. Hagan and M. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, Vol. 5, pp. 989–93, 1994.

56. V. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.

57. P. Hall, U. B. Park, and R. J. Samworth, "Choice of neighbor order in nearest-neighbor classification," *Annals Stat.*, Vol. 36, pp. 2135–52, 2008.

58. I. Scholl, T. Aach, M. T. Deserno, and T. Kuhlen, "Challenges of medical image processing," *Comput Sci Res. Develop.*, Vol. 26, pp. 5–13, 2011.

Authors

Gunjan Pahuja received M.Tech. from Guru Jambheshwar University of Science and Technology in 2006. Currently, she is pursuing her PhD in the department of Computer Science and Engineering from Dr. A.P.J. Abdul Kalam University, Lucknow. Her research focuses on Machine Learning and Medical Image Processing. She is a member of ISTE.

Corresponding author. Email: gunjanpahuja.04@gmail.com

T.N. Nagabhushan received his Master's degree and PhD in Electrical Engineering from Indian Institute of Science, Bangalore in the year 1989 and 1996, respectively. He has over 32 years of experience in teaching, research and industry besides holding the position of principal. His main area of focus is machine learning, development of new tools for supervised learning and applications to image processing. He is a member of ISTE and CSI.

Email: tnnagabhushan@gmail.com