ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 5, May 2015**

# A Review on Parkinson Disease Classifier Using Patient Voice Features

Priyanka Holkar[1], Payal Gatti[2], Sonali Meher[2], Pooja Sable[2]

UG Student, Dept. of CSE, VP Engineering College, Baramati, Maharashtra, India [1]

UG Student, Dept. of CSE, VP Engineering College, Baramati, Maharashtra, India[2]

UG Student, Dept. of CSE, VP Engineering College, Baramati, Maharashtra, India[3]

UG Student, Dept. of CSE, VP Engineering College, Baramati, Maharashtra, India[4]

**ABSTRACT**: Parkinson's disease (PD) known as chronic and progressive movement disorder, means that symptoms continues and becomes worst over time.PD affects on person's moves, also affects how they speak and write. After Alzheimer's disease, around whole world 6.3 million people live with Parkinson's disease which makes it the second most common neurological disorder. The cause of disease is unknown, and also there is presently no cure and no treatment options such as medication and surgery to manage its syndromes. Approximately 90% of PD patients have suffered speech difficultiesi.e., dysphonia which is impaired speech production and dysarthria is referred as speech articulation difficulties. These mobility deficits are difficult to treat with drugs or neurosurgery. Parkinson disease people must visit clinician to track their progressions regularly. It will become simple process to anticipate harshness of disease with the help of voice recording of patients. This can be achieved by using Hoehn Yahr Score and Parkinson disease Rating Scale (PDRS) Score. Combination of machine learning algorithms are used for classification of voice features according to severities of disease.

**KEYWORDS**: Parkinson's Disease, disease severity, machine learning algorithms;

## I.INTRODUCTION

Parkinson's disease (PD) was first described by James Parkinson in England. PD is nothing but a progressive neurological disorder of the brain. Basically it is characterized by shaking and difficulty with walking, movement, postural instability or impaired balance and coordination. It affects muscle control and movements of both men and women. Sometimes it is referred to as unknown cause disease and there are treatment options such as medication and surgery to manage its symptoms because there is no cure. Currently tracking PD symptom progression requires the patient to go through time-consuming examinations by trained medical staff by visiting clinic physically. All current methods of symptom monitoring are inconvenient. One of method is Brain Imaging and some others which are very costly. To reduce cost for finding whether or not person has PD, Max little the head of Parkinson's disease Initiative, has introduced an automated medical diagnosis system. Lots of efforts has been taken to rapidly screen PD with voice recordings of patients [1].

There are more than six million people worldwide affected by PD according to WHO (World Health Organization). The group of symptoms typically appears around the age of 60. PD affects more in men than in women. Recently 2% people out of the population affected by PD, if this situation remains same then it will double around 2020 year. In coming years, because of those reasons, PD poses a significant public health burden, which is likely to increase [2]. Our aim is to create a classifier that classify extracted features from voice recordings of PD patients and also detect the severity level of the disease. Harshness of disease can be found with the help of machine learning algorithm such as support vector machine.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 5, May 2015**

## II.RELATED WORK

Han Lee, Yonghyun Ro and Changwhan Yea's goal was to predict proper severities - PDRS and Hoehn Yahr - of Parkinson's disease patients using their voice recordings [1]. About 62.2% accuracy they achieved in predicting Hoehn Yahr score. Also achieved RMS error of 9.57 while predicting PDRS score. They minimized the noise ratio from the original recordings by using three different filters such as band filter, cepstral mean normalization, and spectral subtraction. They extracted additional features to make the number of features to be 1,582 from top of 38 features. Finally, they got best result with the help of decision tree, linear regression, and k-means clustering. They proposed new method using support vector machine in decision tree fashion from many machine learning algorithms.

Nicolas Genain, Madeline Huberthy and Roshan Vidyashankary worked alot on predicting the accuracy of Parkinson's disease severity from patient's voice features [2]. At the start they used UCI dataset for successful accuracy and it resulted with help of smoothed voice features by a moving average in a Random Forest model. Prediction of PD severity from voice features will be efficient way of monitoring PD progressions by simply recording their voice; this reduces trips to the clinician.

Bocklet, Steidl, Noth and Skodda firstly performed classification experiments which focuses on the automatic detection of PD as a Two-class problem i.e. PD patients versus healthy speakers [3]. They described detection of PD severityasa3-classtaskbasedontheUnifiedParkinson'sDisease Rating Scale (UPDRS) standards. They utilized phonic, prosodic and grating features on various kinds of speech tests such as various read sentences, syllable repetition tasks and, monologues and texts. That classification performed by SVM. They reported recognition results of 81.9% while they were differentiating between speakers with PD and normally speaking persons. They achieved a recognition result of 59.1% with system fusion on the duty classification by using UPDRS. Actually they didn't have sufficient data for classification and feature selection. Due to high number of false alarms, results are still not good for application in a clinical monitoring task.

Bocklet, Noth, and Stemmer, Ruzickova, and Rusz used phonic features, prosodic features and also features which are obtained from a 2-mass model of the vocal folds on various kinds of speech tests such as sustained phonations, syllable repetitions, monologues and read texts [4]. Classification performed by SVM. Feature selection was performed on the basis of correlation, in order to find the most important features. Finally they reported recognition results of 91% while trying to distinguish between speakers with PD and normal speaking persons in early stages with prosodic modeling. They achieved a recognition rate of 88% with acoustic modeling and they achieved 79% with vocalmodeling.Theyachievedeffectiveresultsusingfeature selection. According to them, results are still not optimistic. Read texts and monologues are the most meaningful texts when it comes to the automatic detection of PD based on articulation, voice, and prosodic evaluations according to them.

## IIIEXISTING APPROACHES

### A. Random Tree Classification Algorithm

The Random tree algorithm [7] can be applied to both regression problems and classification. Random trees are a collection or assembly of tree predictors which is called as forest. Classification works in following way : firstly random trees classifier takes the input feature vector then classifies it with every tree in the forest, which outputs the class label that received the majority of "votes". In regression the average of the responses over all the trees in the forest is classifier response. The Random tree algorithm is given below.

Algorithm 1: Random Tree Classification

```
Start {
    FF = {collection of all predictor features (forest)}
    // the feature selection algorithm is used to obtain
      Forest RF = input data - feature vector
     Repeat {
```

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 5, May 2015**

//Compare the Attribute Values (av) of RF with FF.

If (RF.av == FF.av) then

take the positive branch

else

take the negative branch

} for all RF until leaf node is reached.

} End

• Disadvantages of Random Tree Classification:

1) Random Tree Classification is slightly accurate but it is expensive in terms of both time and space.
2) Random Tree Classification generally needs larger number of instances to work its randomization concept.

B.Probabilistic neural networks

Probabilistic Neural Networks (PNN) are radial basis neural networks consisting four layers. PNN have been introduced by Donald Specht. Bayes probabilistic rule based approach used for this type of neural network which gives a solution for classification problems. Network architecture involves Parzen estimators for estimating the classes probability density function fi (p.d.f.).

Finding p.d.f. of the pattern layers with the help of PNN which uses a supervised training set is possible. Through training data we have to pass only once which is advantage of PNN.

• Disadvantages of Probabilistic neural networks :

1) Probabilistic neural networks.

2) Slow execution of the network.

3) Requires a representative training set.

C.Decision Tree

DT is a popular classification tool because it can represent rules for classifying data. A decision tree can be formed by using a tree structure where each node is either a decision node or a leaf node. A decision tree can be implemented to classify data by starting at the root of the tree and then moving down through a leaf node. The leaf node is able to provide a classification outcome.

• Disadvantages of Decision Tree :

1)For data including categorical variables with different number of levels, information gain

in decision trees are biased in favor of those attributes with more levels.

2) Calculations can get very complex particularly if many values are uncertain and/or if

Many outcomes are linked.

## IV.PROPOSED SYSTEM

Our aim is to create a classifier that classifies extracted features from voice recordings of Parkinson's disease patients and detect the severity level of the disease by running various data mining models with the feature set to predict the final severity score.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 5, May 2015**

A. Classificationusing Support Vector Machine:

SVM is technique provides solutions to classification problems. This technique is based on the notation of margin. The margin means the minimal distance from the closest data points to a hyperplane that separates two data classes. SVM aims to find the maximal margin that can create the largest distance between the separating hyperplane and the data points of different classes. These data points are known as support vectors.

SVM can be explained in four basic concepts: the separating hyperplane, the maximal margin hyperplane, the soft margin, and the kernel function.

• Advantage of Support Vector Machine:

1)SVM is that it can provide a good generalisation when the parameters of SVM are selected appropriately.

2)It is also a robust model, even if the training instances have some bias.

3) Requires a representative training set.

## VI.CONCLUSION

Our goal is to predict proper severities of Parkinson's Disease patients using their voice recordings. Parkinson's disease is the second most common Neuro degenerative action only surpassed by Alzheimer's disease. Diagnosis of PD at an early stage can result in significant life saving. An automated medical diagnosis system would reduce the cost effects and also enhance the accuracy of the diagnosis. The present study compares the accuracy of machine learning method such as Support Vector Machines (SVM) for diagnosis of parkinson's disease.

## REFERENCES

[1] Han Lee, Yonghyun Ro, Changwhan Yea "Parkinson Disease Classifier Using Patient Voice Recording Data," Department of Computer Science, Stanford University, Stanford, California

[2] Genain, Madeline Huberthy, Roshan Vidyashankary"Predicting Parkinson's Disease Severity from Patient Voice Features," Department of Management Science and Engineering, Stanford University.

[3] Bocklet, T., Steidl, S., Noth, E., and Skodda, S., "Automatic Evaluation of Parkinson's Speech - Acoustic, Prosodic and Voice Related Cues," INTERSPEECH 2013, 2013.

[4] "About Parkinson's Disease," Facts about Parkinson's Disease Neurological Movement Disorder. Web. 17 May 2014.

[5] M. M. Hoehn and M. D. Yahr, " Parkinsonism: Onset, progression and mortality," Neurology, vol. 17, no. 5, pp. 427- 442, May 1967.

[6] Noyes K, Liu H, Li Y, Holloway R, Dick AW 2006. Economic burden associated with Parkinson's disease on elderly Medicare beneficiaries. Movement disorders official journal of the Movement Disorder Society.

[7] Dr.R.Geetha Ramani, G. Sivagami, Shomona Gracia Jacob" Feature Relevance Analysis and Classification of Parkinson Disease Tele-Monitoring Data through Data Mining," Computer Science and Software Engineering Research Paper.