"DETECTION OF PARKINSON'S DISEASE BY USING MATLAB"

[1]AKASH P. MANWATKAR
M. Tech. Scholar, (VLSI), Bapurao Deshmukh College of Engineering, Sewagram, India
akashmanwatkar@ymail.com

[2]PROF. S. R. SALWE
Assistant Professor, Bapurao Deshmukh College of Engineering, Sewagram, India

[3]PROF. A. P. BAGADE
Assistant Professor, Bapurao Deshmukh College of Engineering Sewagram, India

[4]DR. R. D. RAUT
Associate Professor, Department of Applied Electronics, SGB, Amaravati, India

ABSTRACT: This paper present different method of diagnosis of Parkinson's disease through voice in early stages. The aim of this study is to provide a simple, fast, cheaper method of detection of Parkinson's disease for the patients as there is no cure for this disease and the available therapies which provide some relief to them are too much costlier for this disease. We used Linear and Non-linear kernel function together for improving the accuracy of the Support Vector Machines (SVMs) classification. We have used kernel function with a number of parameters associated with the use of the SVM algorithm that can impact the results. A comparative analysis of Linear SVM versus Non-linear SVMs for data classifications is also presented to verify the effectiveness of the kernel function. We seek an answer to the question: ''which kernel can achieve a highest accuracy classification among all kernel functions''. The support vector machines are evaluated in comparisons with different kernel functions by application to a variety of non-separable data sets with several attributes.

Keywords: False positive, False negative, Multi- Layer perception, Parkinson's disease, Radial basis function, Speech Analysis, Support Vector Machine, True positive, True negative.

1. INTRODUCTION

PARKINSON'S disease (PD, also known as idiopathic or primary parkinsonism, hypokinetic rigid syndrome (HRS), or paralysisagitans) is a degenerative disorder of the central nervous system mainly affecting the motor system. The motor symptoms of Parkinson's disease result from the death of dopamine-generating cells in the substantianigra, a region of the midbrain. The causes of this cell death are poorly understood. Early in the course of the disease, the most obvious symptoms are movement-related. These include shaking, rigidity, slowness of movement and difficulty with walking and gait[1]. Parkinson's disease is more common in older people, with most cases occurring after the age of 50; when it is seen in young adults, it is called young onset PD (YOPD) [4].

The main motor symptoms are collectively called Parkinsonism, or a "parkinsonian syndrome". The disease can be either primary or secondary. Primary Parkinson's disease is referred to as idiopathic (having no known cause), although some atypical cases have a genetic origin, while secondary parkinsonism is due to known causes like toxins. Many risks and protective factors have been investigated: the clearest evidence is for an increased risk of PD in people exposed to certain pesticides and a reduced risk in tobacco smokers. The pathology of the disease is characterized by the accumulate ion of a protein into Lewy bodies in neurons, and from insufficient format ion and activity of dopamine in certain parts of the midbrain. Where the Lewy bodies are located is often related to the expression and degree of the symptoms of an individual. Diagnosis of typical cases is mainly based on symptoms, with tests such as neuro imaging being used for confirmation.

Treatments are effective at improving the early motor symptoms of the disease. This is typically with the medications L-DOPA and dopamine agonists. As the disease progresses and dopaminergic neurons continue to be lost, these drugs eventually become ineffective at treating the symptoms and at the same time produce a complication marked by involuntary writhing movements. Diet and some forms of rehabilitation have shown some effectiveness at improving symptoms. Surgery and deep brain stimulation have been used to reduce motor symptoms as a last resort in severe cases where drugs are ineffective. Research directions include investigations into new animal models of the disease and of the potential usefulness of gene therapy, stem cell transplants and neuro protective agents. Medications to treat non-movement-related symptoms of PD, such as sleep disturbances and emotional problems, also exist.

In 2013 PD resulted in 103,000 deaths up from 44,000 deaths in 1990. The disease is named after the English doctor James Parkinson, who published the first detailed description in An Essay on the Shaking Palsy in 1817. Several major organizations promote research and improvement of quality of life of those with the disease and their families. Public awareness campaigns include Parkinson's disease day (on the birthday of James Parkinson, 11 April) and the use of a red tulip as the symbol of the disease. People with parkinsonism who have increased the public's awareness of the condition include actor Michael J. Fox, Olympic cyclist Davis Phinney, and professional boxer Muhammad Ali. Parkinson's not only affects humans, but other primates as well, which have often been used in researching the disease and testing approaches to its treatment.

2. DATABASE COLLECTION OF VOICE SIGNALS AND SPEECH FEATURE ANALYSIS

A. Source

The dataset was created by Max Little of the University of Oxford, in collaboration with the National Centre for Voice and Speech, Denver, Colorado, who recorded the speech signals [11]. The original study published the feature extraction methods for general voice disorders.

B. Data Set Information

This dataset is composed of a range of biomedical voice measurements from 31 people, 23 with Parkinson's disease (PD). Each column in the table is a particular voice measure, and each row corresponds one of 195 voice recording from these individuals ("name" column). The main aim of the data is to discriminate healthy people from those with PD, according to "status" column which is set to 0 for healthy and 1 for PD. The time since diagnoses ranged from 0 to 28 years, and the ages of the subjects ranged from 46 to 85 years (mean 65.8, standard deviation 9.8). Averages of six phonations were recorded from each subject, ranging from one to 36 seconds in length. The phonations were recorded in an IAC sound-treated booth using a head-mounted microphone (AKG C420) positioned at 8 cm from the lips. The voice signals were recorded directly to computer using CSL 4300B hardware (Kay Elemetrics), sampled at 44.1 kHz, with 16 bit resolution. Although amplitude normalization affects the calibration of the samples, the study is focused on measures insensitive to changes in absolute speech pressure level. Thus, to ensure robustness of the algorithms, all samples were digitally normalized in amplitude prior to calculation of the measures. The data is in ASCII CSV format. The rows of the CSV file contain an instance corresponding to one voice recording [5] [6].

C. Feature Extraction

In this dataset, 22 linear and non-linear features were extracted from the data. 14 features are based on four factors: F0 (fundamental frequency or pitch), jitter, shimmer and noise to harmonics ratio, which are the most important factors of the voice signal. It was concluded that the change in these factors is remarkable in people with Parkinson’s disease compared to healthy people, therefore, optimized features are selected among them. Each feature is described below [4] [6]:

Fo (Hz): Average vocal fundamental frequency.
Fhi (Hz): Maximum vocal fundamental frequency.
Flo (HZ): Minimum vocal fundamental frequency.

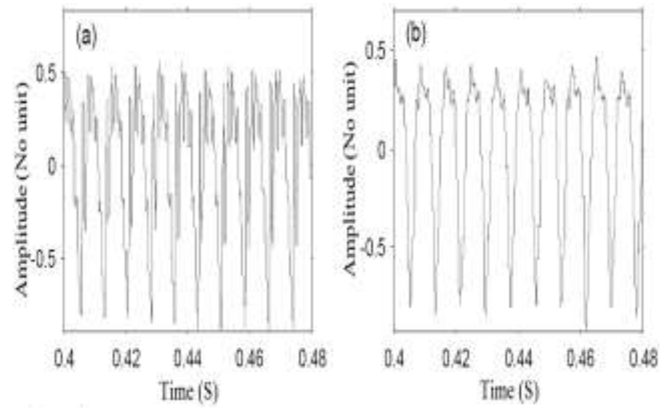

Figure 1: (a) healthy; (b) subject with Parkinson’s disease.

This are the 22 linear and non-linear features extracted from the data

MDVP: Fo (Hz), MDVP: Fhi (Hz), MDVP: Flo (Hz), Jitter (%), Jitter (Abs), MDVP: RAP, MDVP: PPQ, Jitter: DDP, Shimmer, Shimmer (dB), Shimmer: APQ3, Shimmer: APQ5, MDVP: APQ, Shimmer: DDA, NHR, HNR, RPDE, D 2, DFA, spread 1, spread 2, PPE.

3. DISEASE DETECTION USING SUPPORT VECTOR MACHINE (SVM)

The most representative example of local neural network is the Support Vector Machine (SVM) of the Gaussian kernel function. It is a two layer neural network employing hidden layer of radial units and one output neuron. The procedure of creating this network and learning its parameters is organized in the way in which it deal only with kernel functions instead of direct processing of hidden unit signals. Basic SVM is linear but it can be used for non-linear data by using kernel function to first indirectly map non-linear data into linear feature space. Basic SVM is also a two-class classifier however; with some modification, multiclass classifier can be obtained.

SVM training involves solving a convex quadratic programming (QP) problem with equality and inequality constraints obtained by the objective of margin maximization. In multiclass SVM, many two class SVMs are trained and in classification, voting schemes are used for selecting the correct class. SVM is preferred here because it can directly measure the extent to which people with Parkinson can be discriminated from healthy controls on the basis of measures of dysphonia, Addressing the problem of classifying subjects as healthy or PD [5].

In this paper we are using Linear SVM and Non-Linear SVM. In Both method 50% data is used for training purpose. In Non-linear SVM method 4 kernel tricks are used. Which are Quadratic, Polynomial, RBF and MLP. In both method 50% voice samples train first by using MATLAB program. Then for Test purpose testing program is used. Which test 195 samples among which it detect how many people are affected with Parkinson's Disease. In Non-linear SVM method 4 test program are used for 4 Kernel function.

Following sequence gives idea about methods:

1. Linear SVM
    1. Train data
    2. Test data
2. Non-Linear SVM (Kernel tricks)
    3. Train data
    4. Test data
    5.

Test the data by using four kernel function.

1. Quadratic
2. Polynomial
3. RBF (Radial basis function)
4. MLP (Multi-Layer Perception)

4. CONFUSION MATRIX

- True Positive
- True Negative
- False Positive
- False Negative

Confusion Matrix gives the values of affected people and the healthy people. True Positive gives the value of affected people. True Negative gives the value of healthy people. False Positive gives the values of affected people, which is wrongly calculated because calculated people are healthy and False Negative gives the values of healthy people but they are affected.

Following formulae are used to calculate Sensitivity, Specificity, Precision and Accuracy. This formulae are taken from the internet. [https://en.m.Wikipedia.org/wiki/Confusion_matrix]. By using TP, TF, FP and FN values we calculated Sensitivity, Specificity, Precision and Accuracy. In following table we find out which kernel function have highest accuracy and according to values of accuracy in different kernel function we plot an accuracy graph.

Sensitivity = TP / (TP + FN)
Specificity = TN / (TN + FP)
Precision = TP / (TP + FP)
Accuracy = (TP + TN) / (TP + TN + FP +FN)

5. Table :1 Result comparison of all kernel function

Method	TP	TN	FP	FN	Sensitivity	Specificity	Precision	Accuracy
Linear	125	37	11	22	0.85034	0.77083	0.91912	0.83077
Quadratic	147	9	39	0	1	0.1875	0.79032	0.8
Polynomial	131	47	1	16	0.89116	0.97917	0.99242	0.91282
RBF	141	46	2	6	0.95918	0.95833	0.98601	0.95897
MLP	96	30	18	51	0.65306	0.625	0.84211	0.64615

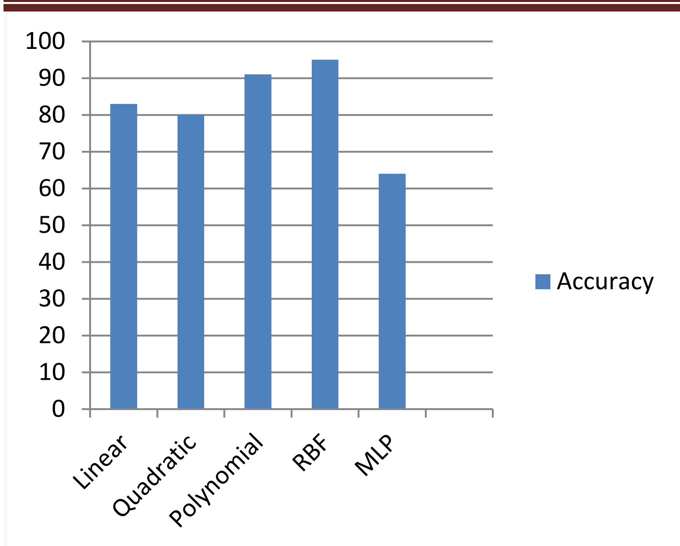

Figure 2: The relation between mean accuracy of SVM kernel function.

We have compared our result with the result of [10] and we find out that the kernel functions we have used have more accuracy. So by using our method of detecting Parkinson's disease we can conclude that our method is cheaper, accurate and less time consuming [10].

Table 2: Comparison of reference paper result with our result

Kernel Function	SVM	MLP	Kernel Function	SVM	MLP
Polynomial	0.8438	84.90	**Polynomial**	0.91282	0.64615
RBF	0.8577		**RBF**	0.95897	
PRBF	0.9238		**PRBF**	0.80	
GRPF	0.9579		**GRPF**	0.83077	

6. CONCLUSION

In this project, we have used UCI dataset of voice samples. We have used here both type of SVM. i.e. Linear and Non-linear with kernel function. We have constructed SVMs and computed its accuracy. We held a comparison between the Linear SVM and Non-linear SVMs. We have used different sizes of data sets with different attributes. Then, we have compared the results of the SVM algorithm. Comparing the classification accuracy of the support vector machine, we have concluded that by using Linear and Non-linear SVM techniques we can find out the PD affected patient with more accuracy. Result Table shows that RBF has more accuracy among all. The great advantage of SVM approach is the formulation of its learning problem, leading to the quadratic optimization task. It greatly reduces the number of operations in the learning mode. It is well seen for large data sets, where SVM algorithm is usually much quicker.

7. REFERENCES

[1] Vikas and R.K. Sharma 'Early Detection of Parkinson's disease through Voice', IEEE – International Conference on Advances in Engineering and Technology (ICAET 2014).

[2] Wen-Shao Wu and Wen-Yen Lin and Ming-Yih Lee 'Forward-flexed Posture Detection for the Early Parkinson's Disease Symptom', 2014 IEEE International Conference on Systems, Man, and Cybernetics October 5-8, 2014, San Diego, CA, USA.

[3] Gertrud Laura Sorensena,b,c, Jacob Kempfnera,b,c, PoulJennumb,c and Helge B. D. Sorensena 'Detection of arousals in Parkinson's disease patients', 33rd Annual International Conference of the IEEE EMBS Boston, Massachusetts USA, August 30 - September 3, 2011.

[4] S.Bouchikhi1, A.Boublenza2, A.Benosman3, M. A. Chikh4 'Parkinson's disease Detection With SVM classifier and Relief-F Features Selection Algorithm', Biomedical Engineering Laboratory Tlemcen University, Algeria.

[5] Mohammad Shahbakhi1*, DanialTaheri Far1, Ehsan Tahami2 'Speech Analysis for Diagnosis of Parkinson's Disease Using Genetic Algorithm and Support Vector Machine', J. Biomedical Science and Engineering, 2014, 7, 147-156 Published Online March 2014 in SciRes.

[6] Mohammad S Islam, ImtiazParvez, Hai Deng 'Performance Comparison of Heterogeneous Classifiers for Detection of Parkinson's Disease Using Voice Disorder (Dysphonia)', 3rd International Conference On Informatics, Electronics & Vision 2014.

[7] Braveena K. Santhiranayagam_, Daniel T. H. Laiy,Cancan Jiang z, Alistair Shiltonx and Rezaul Begg'Automatic Detection of Different Walking Conditions Using Inertial Sensor Data', WCCI 2012 IEEE World Congress on Computational Intelligence June, 10-15, 2012 - Brisbane, Australia.

[8] Han Lee, Yonghyun Ro, ChangwhanYea "Parkinson Disease Classifier Using Patient Voice Recording Data," Department of Computer Science, Stanford University, Stanford, California.

[9] Jonathan A. Stamford, Peter N. Schmidt, and Karl E. Friedl,' What Engineering Technology Could Do for Quality of Life in Parkinson's Disease: A Review of Current Needs and Opportunities', IEEE Journal Of Biomedical And Health Informatics, Vol. 19, No. 6, November 2015.

[10] 'Support Vector Machines versus Multilaye Perception in data classification(Egyptian Informatics Journal 2012)'..

[11] 'Predicting Parkinson's Disease Severity from Patient Voice Features'(Department of Management Science and Engineering,Standford University.

[12] Shyamal Patel, KonradLorincz, Richard Hughes, Nancy Huggins, John Growdon, David Standaert, MetinAkay,' Monitoring Motor Fluctuations in Patients With Parkinson's Disease Using Wearable Sensors' IEEE Transactions On Information Technology In Biomedicine, Vol. 13, No. 6, November 2009.