A Survey of Parkinson's Disease Using Data Mining Alogorithms

Dr. Hariganesh S [1], Gracy Annamary S[2]
Asst.Professor[1] , M.Phil Scholar[2]
Department of Computer Application[1], Department of Computer Science[2]
Bishop Heber College (Autonomous)
Trichirappalli-620 017

Abstract-The data mining techniques is a more popular in many field of medical, business, railway, science...etc; they are most commonly used for medical diagnosis and disease prediction. The data mining is used for medical and health areas of the most important factors in industrial societies. This paper to provide a survey of data mining techniques of using Parkinson's disease.

Key Words-Parkinson's disease, Classification, Neural Network, Speech Disorder

1. INTRODUCTION

Data mining is the technology provides user oriented approach to novel and hidden pattern in the data. The Parkinson's disease (PD) is a type of neurological disease. Parkinson's disease is a degenerative disease of the brain, which affects the nerve cells in brain called by Neurons. The neurons produce the dopamine, they control to brain of movements. Normally, this disease progresses slowly. Some people reside for many years with only minor symptoms. The people start to have symptoms between the ages of 50 and 60. But sometimes symptoms start previous.

1.1 Classification

Classification is perhaps the most familiar and most popular data mining techniques. Estimation and prediction may be viewed as types of classification. The classification is a several types of models, important by random tree, id3, c4.5, classification and regression tree(C-RT), decision tree.

1.2 SPEech Disorders

The speech and voice disorders are common among individual suffering from idiopathic Parkinson's disease. The speech affect by 70-90% people in this disease. The Speech disorders are four subsystems by phonation, respiration, articulation and prosody. The subsystem they are several features measures.

Phonation: jitter, shimmer, noise-to-harmonics (NHR), harmonics-to-noise (HNR).

Respiration: Sound pressure level decline (SPLD).

Articulation: DDK rate, robust formant periodicity correlation (RFPC), spectral distance change variation (SDCV).

Prosody: fundamental frequency variation (F0 SD), intensity of voice variation (intensity SD), number of pauses.

1.3 Neural Network

Neural Networks (NN) offer referred to as Artificial Neural Network (ANN) to distinguish them from biological neural networks, are modeled after the workings of the human brain. They have 3 type of layer by input, output, and hidden layers. The neural networks have been used pattern recognition, speech recognition and synthesis, medical application, fault detection, problem diagnosis, robot control and computer vision. The Artificial Neural Network (ANN) used for analysis the Parkinson's disease.

1.4 Parkinson's Disease

The Parkinson's disease is a neurodegenerative disease. The disease affect by brain cells (neurons) in human brain. Then affected neurons by the brain cells by substantia nigra. The neurons make an important chemical called dopamine. The dopamine sends singles to the part of the brain that controls movements. The little signals can help those parts of the brain work better. The decrease of dopamine of the brain the person unmorally movement.

They are four types of symptoms of Parkinson's disease are tremor, rigidity, bradykinesia and postural instability.

- Tremor: vibrate by hands, arms, legs or jaws.
- Rigidity: limbs and trunk inflexible.
- Bradykinesia: slow movements.
- Postural instability: depression and emotional changes.

The earliest indicators affects by 75-90% people with Parkinson's disease.

2. RELATED WORKS

[1] **R.Geetha Ramani, G.Sivagami et al**., developed by the training set used in tell-monitoring dataset by UCI Irvine machine learning repository. The dataset comprises of two target classes by Motor-UPDRS and Total-UPDRS.

The six features selection algorithm(Fisher Filtering, ReliefF, Runs Filtering, Forward Logistic Regression, Backward Logistic Regression and Stepwise Discriminates) the best of ReliefF feature selection algorithm and thirteen classification algorithm(C4.5, classification tree(C-RT), Iterative Dishomotiser (ID3), K-Nearest Neighbor (K-NN), Cost-Sensitive Classification (CS-CRT), Cost-Sensitive Decision tree Algorithm (CS-MC4), Random Tree (Rnd tree), partial least squares for classification (C-PLS), SVM for classification (C-SVC), Linear Discriminant Analysis (LDA), Partial Least Squares-Discriminant/Linear Discriminant Analysis (PLS-DA/LDA) Support vector machine (SVM)) the best of random tree classification algorithm, they produce by 100% classifier accuracy of test set.

[2] **Peyman Mohammadi et al.,** presented by the data set is 22 attributes and 16 biomedical voice measures by jitter, shimmer, HNR, NHR, RPDE, DFA, PPE. The Parkinson's disease of remote tracking used data mining algorithms

used by eleven methods (simple linear regression (SLR), Multi-layer perceptron (MLP), SMOreg, M5Rules, decision table, M5P, REPTree, decision stump, IBk, LWL, and Regression by discretization).

The 4406 training dataset and 1469 data of test set. They are comparing of algorithms in correlation coefficient, mean absolute error, root mean squared error, relative absolute error and root relative squared error. The higher accuracy of correlation coefficient in dataset is 99.85% of M5Rules algorithm of accuracy is 99.67% of best method. They use weka 3.6.9 tools.

[3] R. Geetha Ramani, G. Sivagami et al., developed by the voice measurements of disease mainly focuses the speech signals. The Parkinson dataset is range of biomedical voice measurement from 31 people 23 characteristic features in Parkinson's disease. The error rate of confusion matrix of 2*2 matrix of output. The main goal of minimum error rate of minimum characteristic of Parkinson dataset. The random tree is the 100% accuracy and zero error rates. Tanagra data mining tool are used.

[4] Shianghan Wu, Jjiannjong Geo et al., presented by three models to analysis the Parkinson's disease for error probability calculated by, logistic regression analysis, decision tree analysis and neural net analysis.

The logistic regression analysis is 5.15% found error probability. The classification analysis error probability by 8.47% and neural net analysis calculated by 23.73% of error probability. Finally, the neural net analysis is the highest error probability is best analysis in Parkinson's disease.

[5] J. Rusz, R. Cmejla et al., presented by speech of vocal sound test for the Parkinson's disease patients to compare the health control (HC) people. The speech test for four features by NHR, SPLD, RFPC and F0 SD.

The disease affected person speaks to microphone. The voice test for vocal task performed by speech subsystem measures calculated by correct evaluation rate. The classification accuracy is the best performances by 85%.

[6] David Gil A, Magnus Johnson B et al., developed by the evaluate in Artificial Neural Networks (ANN) and Support Vector Machines (SVM) of performances classifier. The ANN and SVM is calculated by parameters in accuracy, sensitivity, specificity, positive predictive value and negative predictive value.

[7] Farhad Soleimanian Gharehehopogh, Peyman Mohammadi has presented by the ANN is a two type of Multi Layer Percepton (MLP) with back- propagation learning algorithm and Radia Basis Function (RBF) classifier the accuracy of data set. They use of Matlab neural network toolbox. The MLP network has classifier instances accuracy output in 93.22% and RBF has classifier instances accuracy output in 86.44%.

[8] Tripti Kapoor, R.K.Sharma is presented by speech recognized by Mel-frequency cepstral coefficients (MFCC) and Vector Quantization (VQ). The MFCC has to speech analysis the frames in signal to frequent domain and VQ is the codebook of lowest distortion is calculated.

The 20 phonations used for normal speech and patient with Parkinson's disease. The finally, VQ result with codebook in normal speech of rate in classifier 90% and patient with PD in 95%.

[9] Mehmet can present the boosting committee a machine has developed. Boosting is the filtering. The filtering techniques used for the neural networks with back propagation, they majority voting scheme.

They are three type experts training set. The majority of voting increase the performance rate an imbalanced data set. The highest positive rate in 74% by Neural Networks.

S.No	Classification Techniques	Accuracy (%)
1	Random Tree	100
2	M5Rules algorithm	99.67
3	MLP Network	93.22
4	Neural Network	74

CONCLUSION

The survey paper is contain by Parkinson's disease by using data mining techniques. They are much type of data mining techniques used in Parkinson's disease. They are analysis different techniques of data mining. The different algorithm to comparative performance classifier the high accuracy.

REFERENCES

[1] Dr. R.Geetha Ramani, G.Sivagami, Shomona Gracia jacob " Feature Relevance Analysis and Classification of Parkinson's Disease Tele-Monitoring data Through Data Mining" , International Journal of Advanced Research in Computer Science and Software Engineering,vol-2,Issue 3, March 2012.

[2] Peyman Mohammadi, Abdolreza Hatamlou and Mohammed Msdaris "A Comparative Study on Remote Tracking of Parkinson's Disease Progression Using Data Mining Methods" , International Journal in Foundations of Computer Science and Technology(IJFCST),vol-3,No.6, Nov 2013.

[3] Dr. R.Geetha Ramani and G.Sivagami "Parkinson Disease Classification using Data Mining Algorithms", International Journal of Computer Applications (IJCA),Vol-32,No.9, October 2011

[4] Shanghais Wu, Jiannjong Guo "A Data Mining Analysis of the Parkinson's Disease", Scientific Research, iBusiness 2011, 3, 71-75.

[5] J.Rusz, R. Cmejla, H. Ruzickova, J.Klempir et al., "Acoustic Analysis of Voice and Speech Characteristics in Early Untreated Parkinson's Disease",
http://sami.fel.cvut.cz/Articles/Rusz_et_al_MAVEBA2011.pdf

[6] David Gil A, Maguns Johnson B, "Diagnosing Parkinson by Using Artificial
Neural Networks and Support Vector Machines", Global Journal of Computer Science and Technology, page 63-71.

[7] Farhad Soleimanian Gharehehopogh, Peymen Mohammadi, "A Case Study of Parkinson's Disease Diagnosis Using Artifical Neural Networks" , International Journal of Computer Applications, Vol-73,No.19, July 2013.

[8] Tripti Kapoor, R.K.Sharma, "Parkinson's Disease Diagnosis Using Mel-Frequency Cepstral Coefficients and Vector Quantization", International Journal of Computer Applications, Vol-4, No.3, Jan 2011.

[9] Mehmet can "Boosting committee Machines to Detect the Parkinson's disease by Neural Networks", http://www.ieice-europe.org/ictf/COMM6-3.pdf