Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 2544 - 2550
Received 15 April 2021; Accepted 05 May 2021.

Detection of Parkinson's disease using Machine Learning Approach

Marimuthu.M[1*] , Vidhya.G[1] , Dhaynithi.J[1], Mohanraj.G[1], Basker.N[1], Theetchenya.S[1] , Vidyabharathi.D[1]
[1]Sona College of Technology, Department of CSE, Salem, Tamil Nadu
*mari.btech@gmail.com, gvidhyacse@gmail.com, dhaya.j@gmail.com, mohanraj134@gmail.combas2k9@gmail.com, theetchenya@gmail.com, dvbharathi77@gmail.com

ABSTRACT

Parkinson's disease has recently become one of the most common chronic global diseases among the elderly. The disease is identified by the motor related symptoms caused due to the lack of production of dopamine from the brain cells .But there are other non-motor related symptoms which occur much earlier which can be identified and predict the various stages of disease. The earlier detection of the disease will help us to halt the progression of the disease; thereby the livelihood of the patients remains easy. Traditional methods fail to detect the earlier symptoms of disease due to its complexity in nature, implementation and accuracy. The short comings of the traditional methods can be overcome by using the more efficient automated Machine learning models. Machine learning plays a key role in the Healthcare area because of its accuracy, less computation time and its adaptability. In this research paper, we propose a machine learning based algorithm for early diagnosis, Machine learning techniques XGBClassifier.

Keywords
Machine Learning; Parkinson's disease; XGBClassifier; amplitude; tonal

Introduction

Parkinson's disease (PD) named after the Scientist James Parkinson's and is initially called as Shaking Palsy later as Neurological Syndrome .It is chronic neurodegenerative disorder that affects the central nervous systems that leads to tremors, shaking, stiffness, and difficulty in walking, balance and coordination's i,e motor related activities which affects the quality of life of the patients . This is mainly caused by the disruption of 60 – 80% of the brain cells or neurons that produce a chemical substance called dopamine which is responsible for smooth coordinating, controlling and fluency of movements in the body by transmitting messages between substantial nigra and other regions of the brain. Apart from Motor Related activities, it leads to depression, memory impairment, behavioral changes, and sleeping disorders. Also Researchers believe that non motor symptoms appear much earlier than motor symptoms which can be detected and treated before the adverse effects of the disease. One of the clear risk factor with this disease is age, people above 60 develop the symptoms or about 5 to 10 percent populations have an early on set before or at the age of 50. So, earlier detection of the disease helps to halt the progression of the disease and the patients can have a normal life as long as possible without any assistance.

Recent studies suggests many methods, tools, model that have been used to detect this disease as early as possible but none of them produce/shown effective or promising results. So in this paper, a machine learning based model (XGBClassifier) for early diagnosis of the disease is proposed. In the era of technology, the evolution of Machine learning and its significance in healthcare industry saves many lives by early diagnosis and treatments. Machine learning becoming an efficient and important tool (or plays a vital Role) in Healthcare Care sectors because of its Accuracy, implementation ease, automation, flexibility. Many Algorithms and models have been proposed and experimented to get the authenticate results. In this research paper work XGBClassifier algorithm is proposed for detection the PD.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 2544 - 2550
Received 15 April 2021; Accepted 05 May 2021.

The remaining part of the paper is organized as follows. The related work is discussed in Section 2. The proposed ML model and algorithm are described in Section 3. The result of the experiment is presented in Section 4 along with a discussion. And finally conclusion of the work is provided in Section 5.

Literature Review

In order to assess the disease symptoms and to identify the features of the Parkinson's, many articles have been studied and the summary of the literature survey has been carried out at end of the section.

Senturk et.al., (2020) proposed FS based decision support system along with SVM classification is used for early detection of Disease. Minimum no of features i.e. more relevant features are used to reduce the computational time .The comparison of different FS Algorithms, Classifiers and Optimizers were made and concluded with FS Based SVM Classifiers results attains reasonable accuracy with ease of implementation than others. Wang et.al., (2020) proposed a deep learning model for detecting the PD earlier based on Rapid eye moment, changes in sleeping pattern and behaviors along with olfactory loss. The suggested model results in 96.5% accuracy and seems to be promising when compared with other twelve machine learning models. However the model is tested for small data set comprises of 584 individuals and for the huge data set the performance of the model yet/need to be analyzed.

Challa et.al., (2016) developed a prediction model based on Machine learning Techniques Bayesian Net ,multilayer perceptron, Random Forest and Boosted Logistic regression for earlier detection of PD using Rapid eye moment and olfactory loss. A comparative study and analysis was done over the performance of the techniques and they have concluded that Boosted Logistic regression achieves 97.16% accuracy. Roviniet.al., proposed a preclinical assessment system for PD detection implemented with the wearable inertia device for assessing then lower limb movement which could help the neurologist to evaluate the PD at the starting stage . Additionally the author deals with Idiopathic hypsomia (IH) subjects who have the high risk of developing the disease along with healthy and PD subjects .The Author also have given some future research scope such as upper limb assessment can be considered for detection, address the misclassified data of IH with respect to PD .

Shetty, S., & Rao, Y. S. (2016) suggested a Gaussian radial basis function kernel based SVM Machine learning model using gait analysis . The model can distinguish between Parkinson's disease and other neurological disorders including Amyotrophic lateral sclerosis (ALS) and Huntington's disease. The time series data is considered and feature extraction techniques are applied to acquire more relevant best features. The obtained results achieve overall accuracy of 83.33%. They made an attempt to detect PD and differentiate using Gait analysis, further for future directions, the various stages of the progression of the disease can be experimented with clinical systems like UPDRS and others. Almeida et.al., (2019) proposed a voice based PD detection. The author uses both phonation and speech signals recorded over different microphone channels for diagnosing the disease .The Obtained data is processed through eighteen feature extraction and four different classification methods. Finally, the performance of both the signals is calculated and the comparative study concludes that phonation gives better accuracy than the speech signals.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 2544 - 2550
Received 15 April 2021; Accepted 05 May 2021.

Abdulhayet.al., (2018) demonstrates an innovative approach for PD Detection through gait analysis and tremor investigation. Algorithms like peak detection and pulse duration were used to differentiate normal gait from abnormal gait to predict the stages of disease and also tremor features and its characteristics helps us to realize the severity of the disease. Prashanth, R., & Roy, S. D. (2018) suggests an approach which combines or compromises of MDS – UPDRS scale to measure the PD Features and HY scale to detect the stage of the PD disease along with ML Techniques like SVM, OLR, AdaBoost classifiers to quantify the adverse effect of the Disease in the entire body. The obtained result from the combined approach is effective in prediction of PD.

The study discloses the various methodologies of machine learning techniques like, deep learning models for various datasets that comprises of voice, non-motor symptoms and motor symptoms obtained from repository for earlier detection of the PD disease. Feature based selection classifiers results in reasonable accuracy. In the literature survey, it's observed the minimum features are selected in feature based classifiers that results in reasonable accuracy. Preclinical assessment of PD is instigated with the wearable inertia device for the lower limb assessment of IH subjects, healthy and PD subjects which can be extended further for upper limb assessment. Additionally, Misclassified IH Subjects can be refined with better algorithms. Deep learning models implemented with random forest, classification algorithms and boosting algorithms were tested for small dataset. The performance of the models has to be tested for huge dataset and observations were to be made. Comparative study and performance analysis of the various inputs, i.e. speech signals or phonation based inputs or wearable device based features inputs with respect to the XGBClassifier algorithms can be evaluated.

Proposed System

Machine learning algorithms are trained to find pattern and features in a raw data in order to make decision or prediction based on a given data. In this proposed system, XGBClassifier algorithm is used to implement UCI ML Parkinson's dataset to predict Parkinson's disease. Figure 1 depicts the proposed ML model's systematic flow diagram. The dataset has been uploaded in UCI repository for further reference. The dataset consist of 197 records and 24 attributes or features including maximum vocal fundamental frequency, average vocal fundamental frequency, several measures of variation in fundamental frequency, minimum vocal fundamental frequency, several measures of variation in amplitude, two measures of ratio of noise to tonal components in the voice and status. The status represents as follows: 1 is denotes Parkinson's disease and 0 denotes the person is healthy. The entire model is implemented in Spyder (Python3.8) version.

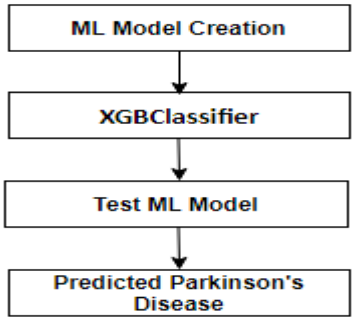

Figure 1. Proposed ML Model

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 2544 - 2550
Received 15 April 2021; Accepted 05 May 2021.

The ML approach implementation steps as follows:

1. A Proposed machine learning model has been developed using XGBClassifier algorithm to test with accuracy (Algorithm1).
2. In order to prove better performance, the proposed model cross validate with few ML algorithms like LogisticRegression, DecisionTreeClassifier and Support Vector Classification (SVC).

Algorithm 1: Proposed ML Model

Model Creation
Input : Features and Labels of Parkinson's Disease
Output : MODEL
ACCURACY: Accuracy of MODEL
foreach features & labels in Input do
data ← data in a file path
data ← empty values replaced as 0
x ← data [All_rows, column1_to_columnN]
y ← data [All_rows, 0th column]
x ← value x is label Encoded
x_tra,x_tes,y_tra,y_tes ← tra_tes_split(x, y, test_size=0.2,random_state=0)
mod=XGBClassifier()
mod.fit(x_tra,y_tra)
y_pred← mod.predict(x_test)
accuracy_score(y_test, y_pred)*100) to Accuracy
end
return**MODEL, ACCURACY**

The proposed model is implemented using various Python libraries such as scikit-learn, NumPy, pandas, and xgboost to build a model using XGBClassifier. The step by step procedure to implement the proposed ML model is given below.

1. Obtaining the data
2. Obtaining the features and labels
3. Feature scalability
4. Partitioning the data
5. Create an XGBClassifier
6. Calculating the model's accuracy

Results Discussion

The entire algorithm works on the probability to build a classification tree for the given dataset. The simulation result for the model is shown in fig.2. The accuracy of the model is 94.87%, the same dataset has been implemented other ML algorithms like LogisticRegression, DecisionTreeClassifier and Support Vector Classification (SVC). It shows that the accuracy of

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 2544 - 2550
Received 15 April 2021; Accepted 05 May 2021.

the proposed model is best compare with other models is shown in fig.3. The comparison results of the accuracy are shown in Table. 1 and it depicts graphically in fig.4.

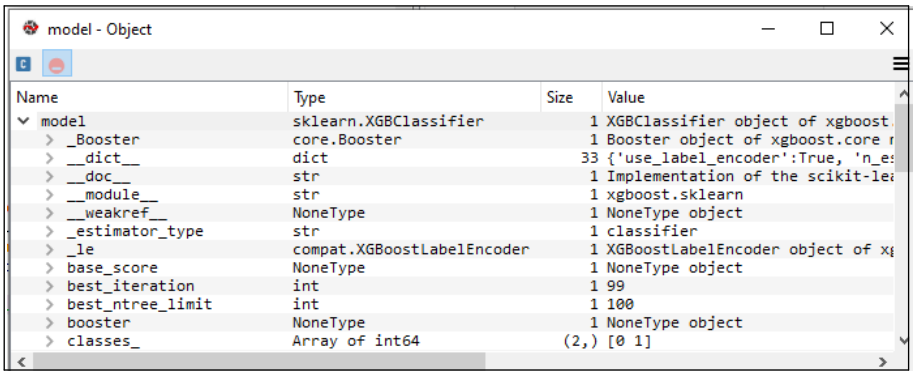

Figure 2. Simulation Result of ML Model

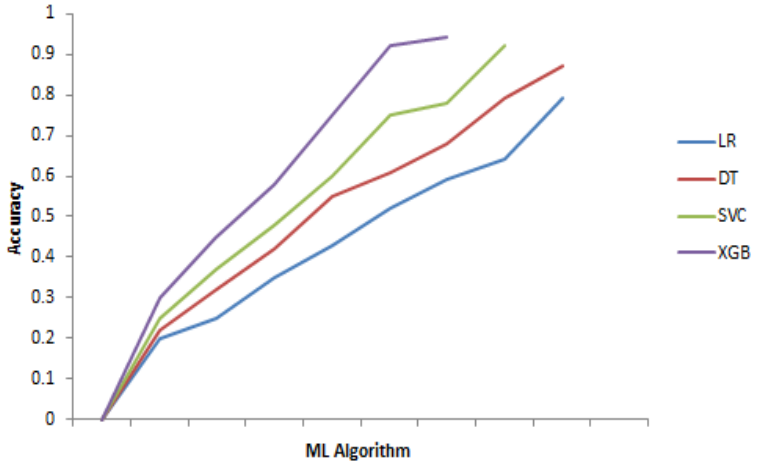

Figure 3. Accuracy of proposed ML Model

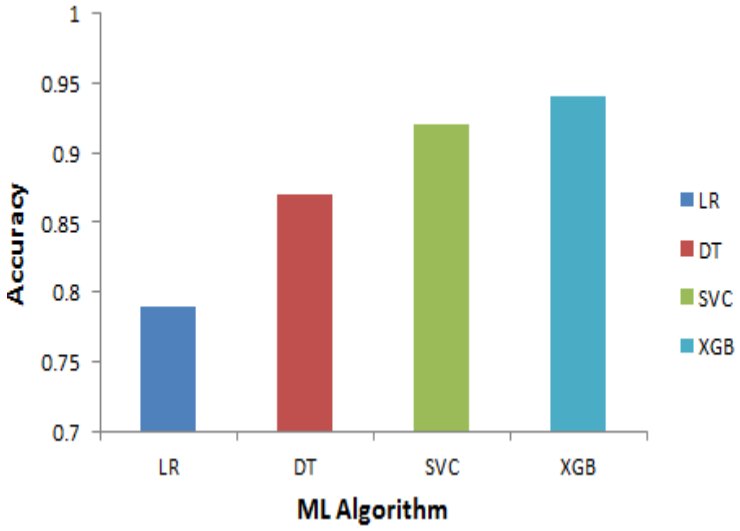

Figure 4. Graphical representation of Accuracy comparisons with other models

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 5, 2021, Pages. 2544 - 2550
Received 15 April 2021; Accepted 05 May 2021.

Table 1. Accuracy comparisons with other models

S.No	Algorithm	Accuracy
1	LogisticRegression	0.79
2	DecisionTreeClassifier	0.87
3	SVC	0.92
4	XGBClassifier	0.94

Conclusion

Machine learning models processhuge amounts of data. In this proposedML model, early diagnosis/predictions of Parkinson's disease in the people are experimented using various ML algorithms. The XGBClassifierimplemented in the proposed the model achieves an accuracy of 94.87%. Furthermore, comparison and review about the performance of different machine learning algorithms from given dataset along with the assessment report is done. The obtained results shows the proposed XGB classifier achieves better accuracy with 0.94% and provides better efficiency /overall performance. The XGB classifier based system provides earlier and more reliable results in predicting the dangerous risks but it can consume more time and cost for a complete and proper training. In Future, we can combine machine learning with AI and computational technology to improve the accuracy in early disease prediction and diagnosis.

References

[1] Senturk, Z. K. (2020). Early diagnosis of Parkinson's disease using machine learning algorithms. Medical hypotheses, 138, 109603.

[2] Wang, W., Lee, J., Harrou, F., & Sun, Y. (2020). Early detection of Parkinson's disease using deep learning and machine learning. IEEE Access, 8, 147635-147646.

[3] Challa, K. N. R., Pagolu, V. S., Panda, G., &Majhi, B. (2016, October). An improved approach for prediction of Parkinson's disease using machine learning techniques. In 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES) (pp. 1446-1451). IEEE.

[4] Rovini, E., Maremmani, C., Moschetti, A., Esposito, D., &Cavallo, F. (2018). Comparative motor pre-clinical assessment in Parkinson's disease using supervised machine learning approaches. Annals of biomedical engineering, 46(12), 2057-2068.

[5] Shetty, S., & Rao, Y. S. (2016, August). SVM based machine learning approach to identify Parkinson's disease using gait analysis. In 2016 International Conference on Inventive Computation Technologies (ICICT) (Vol. 2, pp. 1-5). IEEE.

[6] Almeida, J. S., RebouçasFilho, P. P., Carneiro, T., Wei, W., Damaševičius, R., Maskeliūnas, R., & de Albuquerque, V. H. C. (2019). Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. Pattern Recognition Letters, 125, 55-62.

[7] Abdulhay, E., Arunkumar, N., Narasimhan, K., Vellaiappan, E., &Venkatraman, V. (2018). Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease. Future Generation Computer Systems, 83, 366-373.

[8] Prashanth, R., & Roy, S. D. (2018). Novel and improved stage estimation in Parkinson's disease using clinical scales and machine learning. Neurocomputing, 305, 78-103.

[9] Wroge, T. J., Özkanca, Y., Demiroglu, C., Si, D., Atkins, D. C., &Ghomi, R. H. (2018, December). Parkinson's disease diagnosis using machine learning and voice. In 2018 IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (pp. 1-7). IEEE.

[10] Nilashi, M., Ibrahim, O., Ahmadi, H., Shahmoradi, L., &Farahmand, M. (2018). A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. Biocybernetics and Biomedical Engineering, 38(1), 1-15.

[11] Prashanth, R., Roy, S. D., Mandal, P. K., & Ghosh, S. (2016). High-accuracy detection of early Parkinson's disease through multimodal features and machine learning. International journal of medical informatics, 90, 13-21.

[12] Grover, S., Bhartia, S., Yadav, A., &Seeja, K. R. (2018). Predicting severity of Parkinson's disease using deep learning. Procedia computer science, 132, 1788-1794.

[13] Prince, J., & De Vos, M. (2018, July). A deep learning framework for the remote detection of Parkinson's disease using smart-phone sensor data. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 3144-3147). IEEE.

[14] Miljkovic, D., Aleksovski, D., Podpečan, V., Lavrač, N., Malle, B., &Holzinger, A. (2016). Machine learning and data mining methods for managing Parkinson's disease. In Machine Learning for Health Informatics (pp. 209-220). Springer, Cham.

[15] Cavallo, F., Moschetti, A., Esposito, D., Maremmani, C., &Rovini, E. (2019). Upper limb motor pre-clinical assessment in Parkinson's disease using machine learning. Parkinsonism & related disorders, 63, 111-116.

[16] Parisi, L., RaviChandran, N., &Manaog, M. L. (2018). Feature-driven machine learning to improve early diagnosis of Parkinson's disease. Expert Systems with Applications, 110, 182-190.

[17] Asraf Yasmin, B., Latha, R., & Manikandan, R. (2019). Implementation of Affective Knowledge for any Geo Location Based on Emotional Intelligence using GPS. International Journal of Innovative Technology and Exploring Engineering, 8(11S), 764–769. https://doi.org/10.35940/ijitee.k1134.09811s19

[18] Muruganantham Ponnusamy, Dr. A. Senthilkumar, & Dr.R.Manikandan. (2021). Detection of Selfish Nodes Through Reputation Model In Mobile Adhoc Network - MANET. Turkish Journal of Computer and Mathematics Education, 12(9), 2404–2410. https://turcomat.org/index.php/turkbilmat/article/view/3720