As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 7(11): 93-97, 2021
Copyright © 2021 International Journal for Modern Trends in Science and Technology
ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0711016
Available online at: http://www.ijmtst.com/vol7issue11.html

Machine Learning Approches for the Classification of Parkinson Disease

VangalaSwaroopa[1] | Dr.S.Jessica Saritha[2]

[1]PG Scholar, Department of Computer science and Engineering,JNTUA college of engineering, Pulivendula, Pulivendula, Andhra Pradesh, India
[1]Assistant Professor, Department of Computer science and Engineering,JNTUA college of engineering, Pulivendula, Pulivendula, Andhra Pradesh, India

To Cite this Article
VangalaSwaroopa and Dr.S.Jessica Saritha. Machine Learning Approches for the Classification of Parkinson Disease. *International Journal for Modern Trends in Science and Technology* 2021, *7*, pp. 93-97. https://doi.org/10.46501/IJMTST0711016.

Article Info
Received: 21 October 2021; Accepted: 03 November 2021; Published: 09 November 2021

ABSTRACT

Parkinson's disease (PD) is a dynamic neurodegenerative disorder. To understand Parkinson's disease (PD), biomarkers obtained from human voice can provide knowledge into the sickness cycle itself. Because Parkinson's disease is difficult to diagnose, doctors have relied on abstract rating scales to focus on certain symptoms while ignoring others. Because Parkinson's disease causes a loss of engine control, the voice can be used to identify and analyse the disease. Considering technological advancements and the widespread use of sound gathering devices in everyday life, solid models that can convert sound data into a diagnostic tool for medical professionals may result in more affordable and precise analysis. This paper provides evidence based on a voice dataset collected from people with and without Parkinson's disease. Our precision on previously concealed information can be determined by fostering all calculations dependent on our precision. Using random woods and xgboost models, we could achieve a precision of nearly 95%.

Keywords: Machinelearning,Classification,Randomforest,Supervised learning algorithm.

INTRODUCTION

In individuals more than 60 years old, Parkinson's sickness (PD) is perhaps the most widely recognized neurodegenerative disease. It is estimated that nearly 30% of the population is affected by this incurable disease. Parkinson's disease is primarily instigated by an accumulation of protein molecules in the neuron, which results in the misfolding of the protein and the development of Parkinson's disease. Parkinson's disease can be detected using an effective methodology, according to the findings of this research work. The results of different classification algorithms in machine learning are compared and contrasted. By using AI draws near, we might have the option to distinguish pertinent components that are not as of now utilized in the clinical conclusion of Parkinson's sickness and depend on these elective measures to recognize Parkinson's illness in its preclinical stages or abnormal indications[1].

This paper investigates the viability of utilising regulated arrangement calculations, for example, AI, to precisely analyse patients. Based on the client's information, we can probably predict the onset of Parkinson's disease[2]. To do this, we will use AI

calculations like KNN, Logistic Regression, Random Forest, and the XGBoost model.

STRUCTURE OF PAPER

Section 1 describes introduction,Section 2 provides Literature survey, Section 3 presentes methodology, Section 4 gives results and discussions,Section 5 concluded paper

LITERATURE SURVEY:

Over the period 1990 to 2016, the assessed worldwide populace influenced by Parkinson's illness dramatically increased (from 2.5 million to 6.1 million), attributable to an increment in the quantity of older individuals and an increment in age-normalized commonness rates (Dorsey et al., 2018). It influences 1–2 individuals for every 1,000 individuals and has a pervasiveness pace of 1% in the populace more than 60 years old (Tysnes and Storstein, 2017). It influences all parts of development, including arranging, commencement, and execution, and influences various parts of development (Contreras-Vidal and Stelmach, 1995).

Engine manifestations have been utilized to make the determination of Parkinson's illness[19,20]. In opposition to the foundation of cardinal indications of Parkinsonism (PD) in clinical appraisals, most of rating scales utilized in the evaluation of infection seriousness have not been completely assessed and approved (Jankovic, 2008).

Parkinson's disease is a common problem in the USA, with 4.5–19 cases per 100,000 people each year[21]. Since this is an ongoing issue with a delayed start, the issue's pervasiveness is more significant than the recurrence of events. Uncommonness varies between 18 and 328 per 100 000 people in Shanghai, China, and 328 per 100 000 people in Mumbai, India, according to two studies[22,23]

PROPOSED METHODOLOGY:

Parkinson's diseases are a neurological disorder that affects the brain. It causes the body to shake, as well as the hands, and causes the body to become stiff. At this advanced stage, there is currently no effective cure or treatment available. Treatment is only possible if it is started as soon as the disease manifests itself or when it is in its early stages. It is possible that one of these will not only reduce the cost of the disease, but will also save a life.

The Knn Method was utilised in the data classification process to detect the disease. KNN is a classification method that is used when there are insufficient facts for data distribution. This method is divided into two parts: To determine class type, you must first identify K close neighbours, and then determine the classes type using these near neighbor's[3,4,5].

Using a variety of classifiers, et al conducted a comparative analysis to detect Parkinson's disease in patients. Support Vector Machine (SVM), Feedforward Back-Propagation Based Artificial Neural Network (FBANN), and Random Tree (RT) Classifiers Were Used To Differentiate Between Patients With Parkinson's Disease (PD) and Healthy Patients, and A Comparison Between Them Was Made. The UCI Machine Learning Repository has been used in this research[10,11,12].

The following clinical methods are commonly used to diagnose Parkinson's disease. Magnetic resonance imaging (MRI) or computed tomography (CT) scan - Conventional MRI cannot detect early signs of Parkinson's disease. Using a PET scan, it is possible to determine the activity and function of brain regions that are involved in movement. It is possible to detect changes in brain chemistry, such as a decrease in dopamine, using a SPECT scan[8].

RESULTS AND DISCUSSIONS:

Following the previous stage's determination of the best deep learning method and dataset, we use a variety of measurements to evaluate the models delivered by the 10-crease cross approval within the victor analysis. It is critical to grasp the disarray grid before moving on to the measurements[9,18].

A disarray framework is a network of disarray. A disarray framework is a table used to group issues for easier resolution. This table contains four numbers: the number of True up-sides (TP), True negatives (TN), False up-sides (FP), and False negatives (FN) (FN).

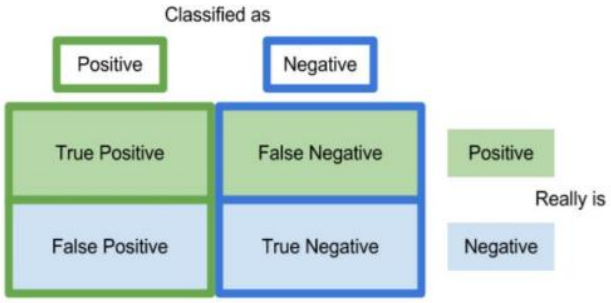

Fig 4: Confusion Matrix block diagram

- In the case of an objective class where there are a significant number of instances of the objective class that are effectively named as being associated with the objective class, this is known as a genuine positive (TP).
- In statistics, the number of occurrences that are effectively labelled as not belonging to the objective class and do not belong to the objective class in question is referred to as a genuine negative (TN).
- If there are more occurrences of something that does not belong to the objective class but is mistakenly assigned to belong to the objective class, this is referred to as a bogus positive, and it is defined as follows: (FP).
- The number of examples that belong to the objective class but are incorrectly delegated to a class that is unrelated to the objective class is referred to as the true negative (TN).

Using the KNN model, we were able to achieve 94% accuracy, but this was not satisfactory. In second part we tried logistic regression in this case we got 93% but training time complexity is less compared to KNN. Next step we build Random forest model with best hyper parameter we got 95.8% accuracy. In final step we tried boosting concept XGBoost model in this model its almost gave same results there is no much differ[13]. We endeavored to choose the most significant and least number of components utilizing distinctive AI methods like KNN, Logistic Regression, and XGboost, with fluctuating levels of achievement. Irregular woodland outflanks any remaining calculations when contrasted with their exhibition[14,15,16].

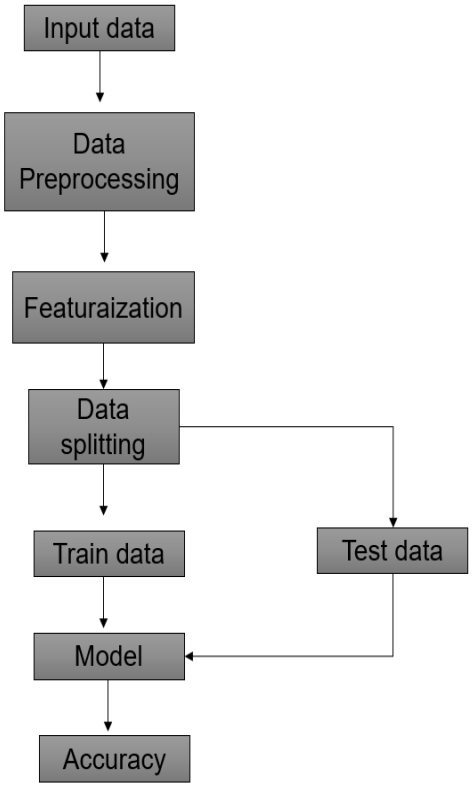

Fig 5: Proposed system architecture

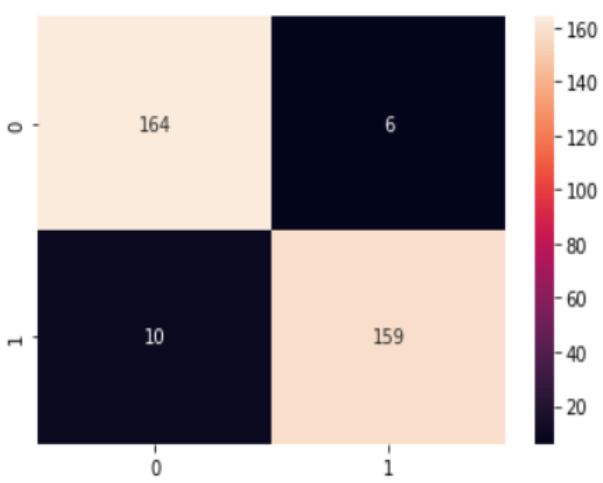

Fig 6: KNN model confusion matrix

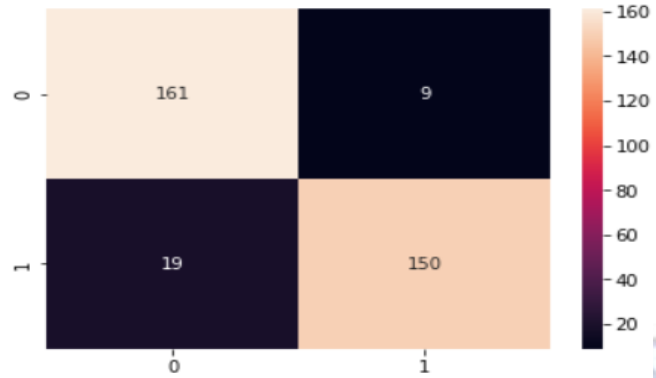

Fig 6: Logistic regression confusion matrix

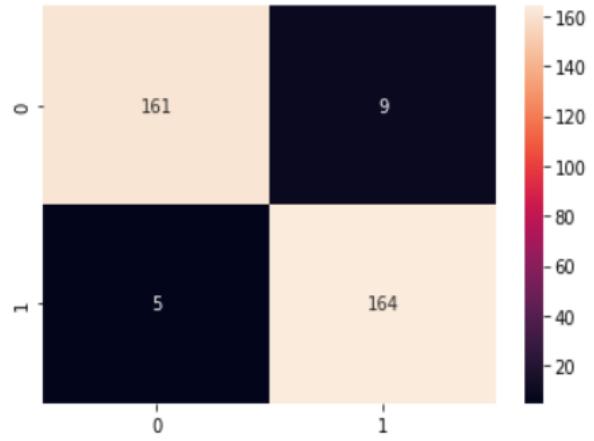

Fig 7: Random forest model confusion matrix

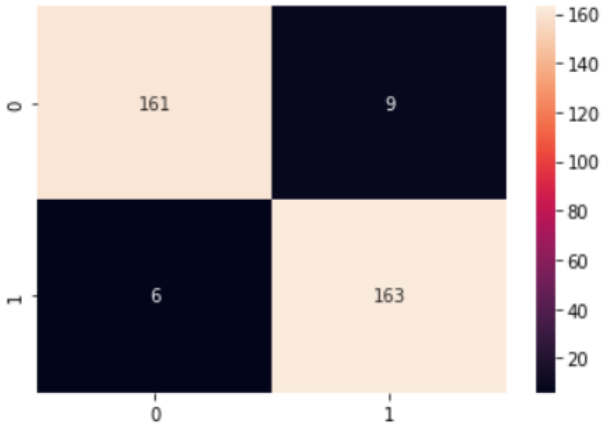

Fig 8: XGBOOST model confusion matrix

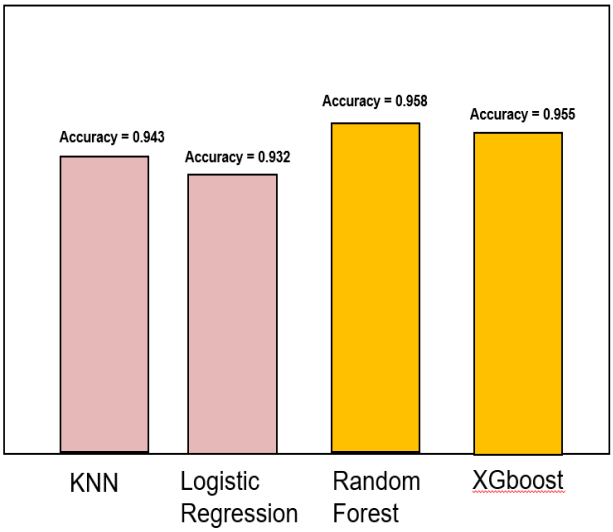

Fig 9: Results comparison table

CONCLUSION AND FUTURE WORK:

It is possible that the scope of the work will be expanded in the future to include the use of auto encoders to reduce the number of elements and concentrate the most important ones from the remaining components]. Furthermore, because the dataset used in this study is not particularly complex, the auto encoder did not perform particularly well on it; however, a more complex dataset would almost certainly produce better results.

REFERENCES

[1] Kamal Nayan Reddy, Challa, VenkataSasankPagolu and Ganapati Panda, "An Improved Approach for Prediction of Parkinson's Disease using Machine Learning Techniques", in Procedings of the International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)-2016, pp. 1446-145, 2016.

[2] eetaYadav, Yugal Kumar and G. Sahoo, "Predication of Parkinson's disease using Data Mining Methods: a comparative analysis of tree, statistical and support vector machine classifiers", in Procedings of the National Conference on Computing and Communication Systems (NCCCS), pp. 1-4, 2012.

[3] Paolo Bonato, Delsey M. Sherrill, David G. Standaert, Sara S. Salles and MetinAkay, "Data Mining Techniques to Detect Motor Fluctuations in Parkinson's Disease", in Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4766-4769, 2004.

[4] Sonu S. R., Vivek Prakash and Ravi Ranjan, "Prediction of Parkinson's Disease using Data Mining", in Proceedings of the International Conference on Energy, Communication, Data

Analytics and Soft Computing (ICECDS), pp. 1082-1085, 2017.

[5] Aarushi Agarwal, SprihaChandrayan and Sitanshu S Sahu, "Prediction of Parkinson's Disease using Speech Signal with Extreme Learning Machine", in Proceedings of the International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 1-4, 2016.

[6] Akshaya Dinesh and Jennifer He, "Using Machine Learning to Diagnose Parkinson's Disease from Voice Recording", in Proceedings of the IEEE MIT Undergraduate Research Technology Conference (URTC), pp. 1-4, 2017.

[7] Giulia Fiscon, Emanuel Weitschek, Giovanni Felici and Paola Bertolazzi, "Alzheimer's disease patients classification through EEG signals processing", in Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining (CIDM). pp 1-4, 2014.

[8] Pedro Miguel Rodrigues, DiamantinoFreitas and Joao Paulo Teixeirab, "Alzheimer electroencephalogram temporal events detection by K-means", in Proceedings of the International Conference on Health and Social Care Information Systems and Technologies HCIST. pp. 859 – 864, 2012.

[9] Elva Maria Novoa-del-Toro, Juan Fernandez-Ruiz, Hector Gabriel Acosta-Mesa and Nicandro Cruz-Ramirez, "Applied Macine Learning to Identify Alzheimer's Disease through the Analysis of Magnetic Resonance Imaging", in Proceedings of the International Conference on Computational Science and Computational Intelligence, pp. 577-582, 2015.

[10] Daniel Johnstone1, Elizabeth A. Milward1, Regina Berretta1 and Pablo Moscato1, "Multivariate Protein Signatures of Pre-Clinical Alzheimer's Disease in the Alzheimer's Disease Neuroimaging Initiative (ADNI) Plasma Proteome Dataset", in Proceedings of the Disease Neuroimaging Initiative, vol-7, pp. 1-17, 2017.

[11] Jason Orlosky, YutaItoh, Maud Ranchet, Kiyoshi Kiyokawa, John Morgan, and Hannes Devos, "Emulation of Physician Tasks in Eye-tracked Virtual Reality for Remote Diagnosis of Neurodegenerative Disease", in Proceedings of the IEEE Transactions on Visualization and Computer Graphics, vol. 23, pp. 1302 – 1311, 2017.

[12] Mathew J. Summers, Vienna, Austria, Alessandro E. Vercelli, Georg Aumayr, Doris M. Bleier and LudovicoCiferri,"Deep Machine Learning Application to the Detection of Preclinical Neurodegenerative Diseases of Aging", in Proceedings of the Scientific Journal on Digital Cultures, vol. 2, pp. 9-24, 2017.

[13] Bianca Torres, Raquel Luiza Santos, Maria Fernanda Barroso de Sousa, Jose Pedro SimoesNeto, Marcela Moreira Lima Nogueira, Tatiana T. Belfort1, Rachel Dias1, Marcia and Cristina NascimentoDourado, "Facial expression recognition in Alzheimer's disease: a longitudinal study", pp. 383-389, 2014.

[14] Smitha Sunil and Kumaran Nair, "An exploratory study on Big data processing: a case study from a biomedical informatics", 3rd MEC International Conference on Big Data and Smart City, pp. 1-4, 2016.

[15] C. Kotsavasilogloua, N. Kostikis, D. Hristu-Varsakelis and M. Arnaoutoglouc, "Machine learning-based classification of simple drawing movements in Parkinson's disease", in Proceedings of the Biomedical Signal Processing and Control, pp. 174–180, 2017.

[16] Santosh S. Rathore and Sandeep Kumar, "An empirical study of some software fault prediction techniques for the number of faults prediction", in Proceedings of the Soft Computing, vol. 21, pp 7417–7434, 2017.

[17] Arvind Kumar Tiwari, "Machine Learning Based Approaches for Prediction of Parkinson's Disease", in Proceedings of the Machine Learning and Applications: An International Journal (MLAIJ), vol.3, pp. 33-39, 2016.

[18] PolinaMamoshina, Armando Vieira, Evgeny Putin and Alex Zhavoronkov, "Applications of Deep Learning in Biomedicine", in Proceedings of the American Chemical Sociecty Mol. Pharmaceutics, pp. 1445−1454, 2016.

[19] Alexis Elbaz,James H. Bower, Brett J. Peterson, Demetrius M. Maraganore, Shannon K. McDonnell, J. Eric Ahlskog, Daniel J. Schaid, Walter A. Rocca, "Survival Study of Parkinson Disease in Olmsted County, Minnesota ", Arch Neurol. Vol. 60 pp. 91-96, 2003.

[20] Tanner CM, Ross GW, Jewell SA ," Occupation and risk of Parkinsonism: a multicenter case- control study" Arch Neurol,66(9):1106–1113,2009.

[21] V. A. Sukhanov, I. D. Ionov, and L. A. Piruzyan, "Neurodegenerative Disorders: The Role of Genetic Factors in Their Origin and the Efficiency of Treatment" in Proceedings of the Human Physiology US National Library of Medicine National Institutes of Health, vol. 31, pp. 472–482, 2005.

[22] Marras C, Tanner C."Epidemiology of Parkinson's Disease", Movement Disorders: Neurologic Principles and Practice, 2nd ed.2004, Watts, RL, Koller, WC (Eds). The McGrawHillCompanies:New York, pp. 177.

[23] http://www.orionpharma.co.uk/Products-and-Services-Orion/Parkinsons-disease/10-factsabout-Parkinsonsdisease/

[24] Cnockaert, L., Schoentgen, J., Auzou, P., Ozsancak, C.,Defebvre, L., &Grenez, F., "Low frequency vocal modulations in vowels produced by Parkinsonian subjects", Speech Communications, vol 50, pp. 288-300, 2008.

[25] Kenneth Revett, Florin Gorunescu and Abdel-Badeeh Mohamed Salem, "Feature Selection in Parkinson's disease: A Rough Sets Approach", Proceedings of the International Multi onference on Computer Science and Information Technology, pp. 425 – 428,2004, ISBN 978-83- 60810- 22-4.