WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

PARKINSON DISEASE DETECTION COMPARSION USING MACHINE LEARNING TECHNIQUES

Rahul Baranwal[1], Uddeshya Gupta[2], Shivendra Singh[3], Pranjal Mishra[4], Dr. Neelima Agarwal[5]

[1234] RKGIT College. India

[5]UNDER THE GUIDANCE OF Associate Professor, RKGIT College, India.

ABSTRACT

Parkinson`s disease (PD) is an ordinary neurodegenerative illness characterized through manner of approach of revolutionary loss of motor characteristic of the body. Parkinson disease is the most common age-related neurological disorder that is no longer uncommon and causes a variety of motor and cognitive symptoms. A Parkinson`s disease evaluation is hard because of the reality its symptoms and symptoms are quite just like the ones of various disorders, in conjunction with ordinary developing older and critical tremor. For many people, Parkinson`s disease (PD) symptoms and symptoms like walking and communique issues come to be critical throughout the age of 50. While PD is incurable, drugs can offer an assisting hand through manner of approach of alleviating some of the one's symptoms and symptoms. Patients can maintain their lives by taking a proactive approach to managing disease-related consequences. Currently, it is vital to learn more about the disease now and prevent it from progressing. The evaluation of Parkinson disease has been a hassle of masses research. In the context of our paper, we purpose to find out Parkinson disease (PD) using first-rate kinds of Machine Learning (ML), and Deep Learning (DL) models in conjunction with Support Vector Machine (SVM) , Random Forest(RF), Decision Tree(DT), K-Nearest Neighbor(KNN), XG-Boost(XGB).In this paper, device gaining knowledge of-based totally absolutely evaluation of Parkinson`s disease is presented. Python device gaining knowledge of project, using the python libraries scikit-learn, NumPy, pandas, and XG-Boost, we are able to assemble a model using an XGB Classifier, Random Forest (RF), Support Vector Machine(SVM), Decision Tree (DT), K-Nearest Neighbor (KNN).We will load the data, get the abilities and labels, scale the abilities, then break up the dataset, assemble every model, and then calculate the accuracy of our model.

Keywords: Parkinson's disease, XG-Boost, Scalable, Decision tree, Random Forest, K-Nearest algorithm.

INTRODUCTION

Parkinson`s sickness is a not unusual place neurological situation throws a wrench into how your muscular tissues featured influences your capacity to transport freely, communicate clearly, and hold top posture main to tremors , muscle tension and bradykinesia. It happens because of neuronal demise, which lowers dopamine levels inside the brain into a specific format. Signs and symptoms of Parkinson's disease usually start with stiffness or tremors on one side of the body, including the hand or palm. People with Parkinson's disease may develop depression later in life. From 1996 to 2016, the global prevalence of Parkinson's disease increased more than fourfold, from 2.5 million to 6.1 million. (approximately two times the populace of Nevada). It is tough to differentiate among usual cognitive feature losses related to growing old and early PD signs and symptoms. In the US, the general financial effect in 2016 become expected to be $52 billion (approximately $160 in line with character withinside the US) (approximately $160 in line with character in the US) (approximately $160 in line with character withinside the US), together with an oblique value of $14.2 billion (about $44 per person in the US), non-clinical costs of $7.5 billion (about $23 per person in the US), and $4.8 billion (about $15 per person in the US) accruing to incapacity profits for owner`s public The mass of people with Parkinson disease are over 65 years old, and the general financial burden is expected to reach $100 billion (about $310 per person in the US) via way of means of 2050.

PD has five degrees of development and 90% PWP show symptoms and symptoms of vocal twine injuries. A superb 90% of sufferers with PD would possibly show off symptoms and symptoms of vocal impairment even withinside the earliest, frequently symptomless, level (level 0). This opens doorways for far flung analysis of the usage of telemedicine. Imagine skipping the experience to the doctor's office. Patients may want to in reality report their voice on the usage of their telephones and carry out an easy look at home. Tasks like keeping an unmarried vowel sound for so long as viable or analyzing a passage may want to monitor early symptoms and symptoms of impairment. In the case of Preliminary diagnosis, doctor can provide healing solutions and deep mind simulations to reinvigorate the dopamine - generating neurons within the mind, there via

way of means of slowing the development of PD. There may be no therapy for Parkinson`s until date. However, early detection and appropriate treatment can reduce the patient's tremors and imbalance symptoms, and lead to a normal life. The focus of this research paper is on early detection using a machine learning algorithm. Compared to K-nearest Neighbor, Support Vector Machine, and Decision trees, our early results show that the Random Forest version, trained on attributes of Kaggle data, has an accuracy of 91.83%. There can be an opportunity that Patient with Parkinson Disease (PD) be afflicted by mobility troubles and are not able to journey for fitness check-unsour studies compares and contrasts diverse ML fashions for sickness class that aren't best reminiscence efficient, however additionally quicker as compared to Machine getting to know community getting to know fashions.

LITERATURE SURVEY

A lot of researchers have categorized Parkinson`s ailment the use of numerous methods. Studied records to predict the onset of Parkinson's disease in sufferers using the Support Vector Machine version. Raundale, Thosar, and Rane used a random forest classifier to predict the severity of Parkinson's disease in older patients using keyboard records from the UCI telemonitoring dataset. Our studies make use of open-supply fashions educated in Python, which are quicker and reminiscence efficient. Applied a linear class version with 95%curacy to signify shuffling motion of PD sufferers.al applied selection trees, random woodland and K-Nearest neighbor to discover Mild Cognitive Impairment (MCI) in Patient with Parkinson ailment.

Based on our studied review, we've applied a PD class version on program getting to know modules. Our studies objectives are to discover K-Nearest Neighbor, Decision tree, Random Forest regression and support vector machine models to categories Parkinson`s affected person records.

SOFTWARE REQUIREMENTS SPECIFICATION

A. Python

1. Python 3

B. Libraries

2. Numpy
3. pandas
4. sklearn
5. matplotlib, etc.

C. Operating System

6. Windows or Ubuntu

D. Hardware Requirements Specification

7. Laptop with basic hardware

RESEARCH METHODOLOGY

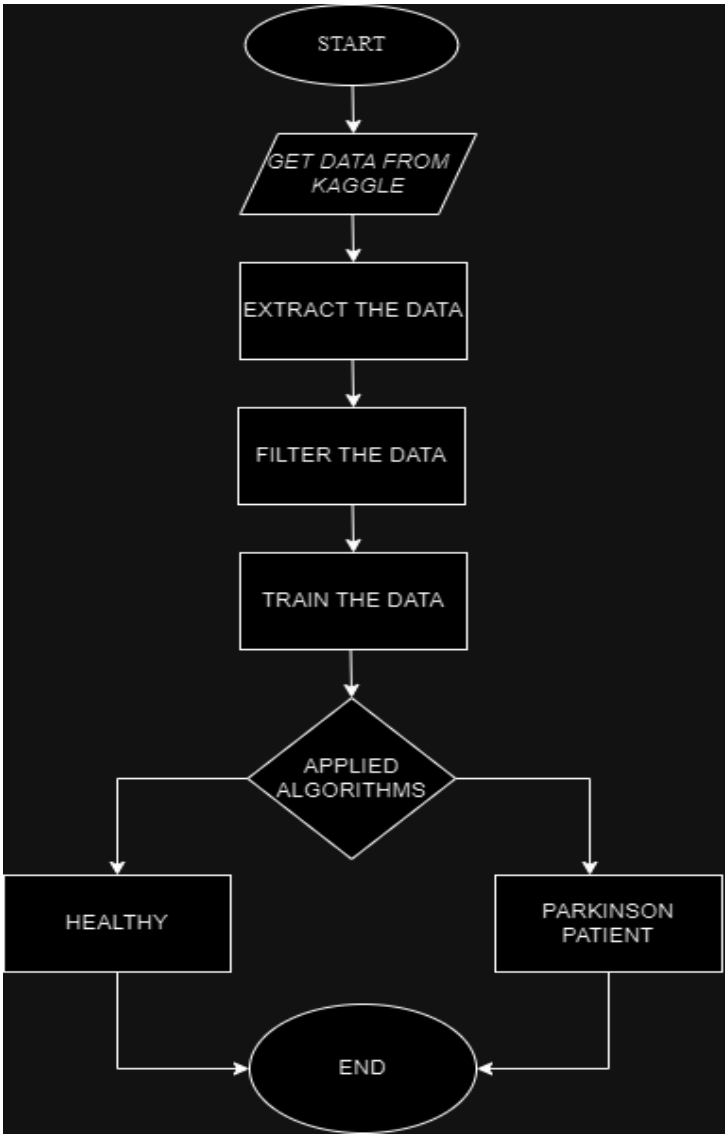
The suggested approach gathers information from PPMI and Kaggle. For a comprehensive grasp of the attribute's algorithms, including K-nearest neighbors, XG-boost, Random Forest, Support Vector Machine, and Decision tree, data is preprocessed, analyzed, cleaned, trained, and visualized.

The developed procedure is depicted in Figure. It shows the steps involved in gathering data from the PPMI and Kaggle databases, separating the dataset into clean and training variable, training five models with the data's, and validating the outcomes using test data.

Our goal is to determine which characteristics are most important for classifying Parkinson's disease (PD) and how imbalanced medical data affects classification. All methods will be put into practice while keeping in mind the requirements. Training will be done on the entire dataset, which acts as a baseline test for PD classification. Below, various algorithms are employed.

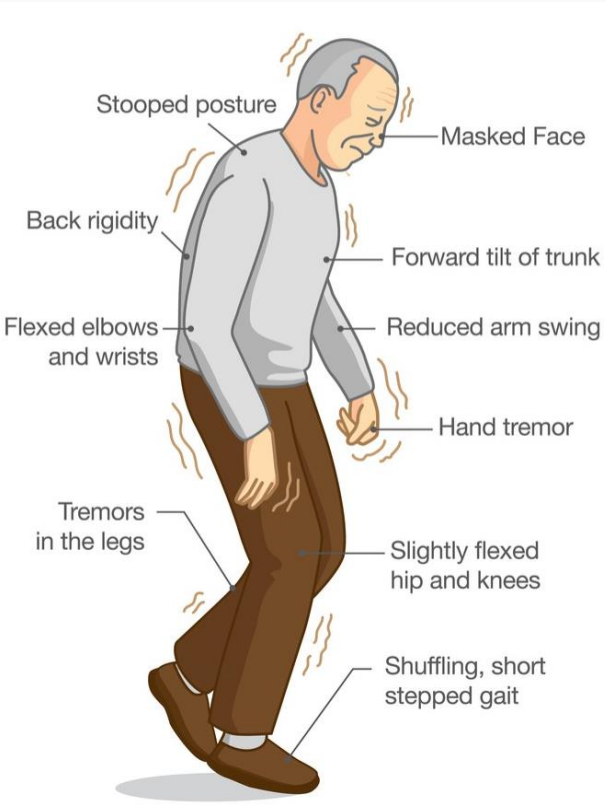
1. SUPPORT VECTOR MACHINE
2. DECISION TREE
3. RANDOM FOREST
4. XG-BOOST
5. K-NEAREST NEIGHBOR

SYSTEM DESIGN


START
GET DATA FROM KAGGLE
EXTRACT THE DATA
FILTER THE DATA
TRAIN THE DATA
APPLIED ALGORITHMS
HEALTHY
PARKINSON PATIENT
END

DATASETS

```
name,MDVP:Fo(Hz),MDVP:Fhi(Hz),MDVP:Flo(Hz),MDVP:Jitter(%),MDVP:Jitter(Abs),MDVP:RAP,MDVP:PPQ,Jitter:DDP,MDVP:Shimmer,MDVP:Shimmer(dB),Shimmer:APQ3,Shimmer:
,status,RPDE,DFA,spread1,spread2,D2,PPE
phon_R01_S01_
1,119.99200,157.30200,74.99700,0.00784,0.00007,0.00370,0.00554,0.01109,0.04374,0.42600,0.02182,0.03130,0.02971,0.06545,0.02211,21.03300,1,0.414783,0.81528
4654
phon_R01_S01_
2,122.40000,148.65000,113.81900,0.00968,0.00008,0.00465,0.00696,0.01394,0.06134,0.62600,0.03134,0.04518,0.04368,0.09403,0.01929,19.08500,1,0.458359,0.8195
58674
phon_R01_S01_
3,116.68200,131.11100,111.55500,0.01050,0.00009,0.00544,0.00781,0.01633,0.05233,0.48200,0.02757,0.03858,0.03590,0.08270,0.01309,20.65100,1,0.429895,0.8252
32634
phon_R01_S01_
4,116.67600,137.87100,111.36600,0.00997,0.00009,0.00502,0.00698,0.01505,0.05492,0.51700,0.02924,0.04005,0.03772,0.08771,0.01353,20.64400,1,0.434969,0.8192
58975
phon_R01_S01_
5,116.01400,141.78100,110.65500,0.01284,0.00011,0.00655,0.00908,0.01966,0.06425,0.58400,0.03490,0.04825,0.04465,0.10470,0.01767,19.64900,1,0.417356,0.8234
10335
phon_R01_S01_
6,120.55200,131.16200,113.78700,0.00968,0.00008,0.00463,0.00750,0.01388,0.04701,0.45600,0.02328,0.03526,0.03243,0.06985,0.01222,21.37800,1,0.415564,0.8250
57775
phon_R01_S02_
1,120.26700,137.24400,11                                                                                                                    .02337,0.00607,24.88600,1,0.596040,0.7641
11756
phon_R01_S02_
2,107.33200,113.84000,10                                                                                                                    .02487,0.00344,26.89200,1,0.637420,0.7632
53755
phon_R01_S02_
3,95.73000,132.06800,91.                                                                                                                    03218,0.01070,21.81200,1,0.615551,0.773587,
571
phon_R01_S02_
4,95.05600,120.10300,91.                                                                                                                    04324,0.01022,21.86200,1,0.547037,0.798463,
362
phon_R01_S02_
5,88.33300,112.24000,84.                                                                                                                    03237,0.01166,21.11800,1,0.611137,0.776156,
740
phon_R01_S02_
6,91.90400,115.87100,86.                                                                                                                    04272,0.01141,21.41400,1,0.583390,0.792520,
```

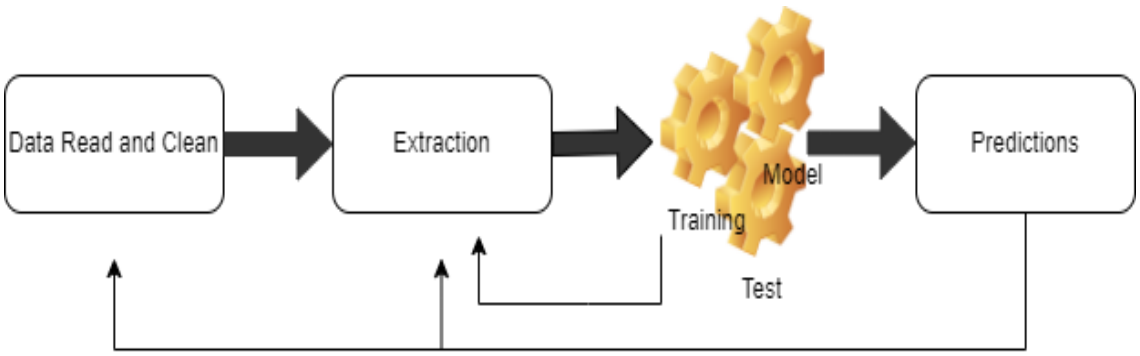


We will find the datase
The age of patients is 2

IMPLEMENTAT

Parkinson's Disease Pr

Running Models



MODEL TRAINING

Decision Tree

Decision trees, a type of machine learning algorithm, are emerging as a valuable tool to assist in this process. Decision trees are showing promise in aiding Parkinson's disease diagnosis. They can analyze data like speech patterns, motor function tests, or even brain scans to identify patterns that separate Parkinson's patients from healthy individuals. This can be helpful for early detection, though it's important to remember these are still under development and shouldn't replace a doctor's evaluation.

Support Vector Machine

SVMs are a powerful machine learning technique well-suited for diagnosing Parkinson's disease (PD) due to their ability to handle complex data and achieve good accuracy. SVMs excel at classifying Parkinson's disease due to their ability to handle high-dimensional medical data and find optimal separation between healthy and Parkinson's patients in that space. This is achieved by maximizing the margin between data points using a support vector formulation, though it requires careful hyperparameter tuning and might not provide the most interpretable results.

XG-Boost

XG-Boost (Extreme Gradient Boosting) is a machine learning algorithm gaining traction in Parkinson's disease (PD) diagnosis. Unlike Support Vector Machines, it doesn't have a single, definitive formula but rather builds a series of decision trees in a sequential manner. Each tree in the XG-Boost model corrects the errors of the previous one, focusing on data points that the earlier models struggled with. While it doesn't have a single formula, XG-Boost optimizes an objective function to balance training accuracy and model complexity, making it a powerful tool for PD analysis. This sequential refinement leads to a robust and accurate overall model.

Random Forest

Random Forests (RF) are a machine learning ensemble method well-suited for Parkinson's disease (PD) diagnosis. Unlike Support Vector Machines with a single hyperplane or XG-Boost's sequential trees, Random Forests leverage a multitude of decision trees. Random Forests build a collection of decision trees, each using a random subset of features and data points, votes on a patient's health status (healthy or Parkinson's). The majority vote wins, achieving high accuracy and robustness against overfitting. While lacking a single formula, this forest-like approach offers some interpretability through feature importance analysis, making it a valuable tool for PD detection.

K-NEAREST NEIGHBOR Nearest Algo

K-NEAREST NEIGHBOR, or K-Nearest Neighbors, is a straightforward machine learning algorithm that can be applied to Parkinson's disease (PD) diagnosis. While it doesn't have a complex formula, it excels in its ease of use and interpretability, a user-friendly approach for Parkinson's diagnosis, classifies new patients based on the majority vote of their closest neighbors in the training data (healthy or Parkinson's). This method shines in its simplicity and interpretability but might not be the most accurate compared to other algorithms. While it lacks a complex formula, K-NEAREST NEIGHBOR offers a valuable starting point for analyzing Parkinson's disease data.

PROGRAM LINKS

https://colab.research.google.com/drive/1hxoIYn3tXTqIM_IC5_e12knk2u5iRWjp?authuser=0#scrollTo=9T-e8ldoyPGg

CONCLUSION

In the battle against Parkinson's disease diagnosis, several machine learning algorithms have emerged as contenders. Support Vector Machine, with its ability to find optimal separation in complex data, has proven effective. However, interpretability limitations can be a hurdle. Random Forests and XG-Boost, on the other hand, excel in accuracy, often exceeding 90%. Random Forests offer some interpretability through feature importance, while XG-Boost excels in handling various data types. K-NEAREST NEIGHBOR, the user-friendly option, provides clear interpretability but might fall short in raw accuracy compared to the others.

XG-Boost often emerges as the top performer due to its ability to combine high accuracy with some interpretability, making it a powerful tool for both diagnosis and feature analysis. With the Score of 94.871%.

REFERENCES

1. Prabhavathi, K., Patil, S. (2022). “Tremors and Bradykinesia. In: Arjunan, S.P., Kumar, D.K. (eds) Techniques for Assessment of
2. Parkinsonism for Diagnosis and Rehabilitation”. Series in Bioengineering. Springer. 135–149 https://doi.org/10.1007/978-981-16-3056-9_9
3. Fothergill-Misbah N., Maroo H., Hooker J., Kwasa J., Walker R. (2020). Parkinson's disease medication in kenya–situation analysis. Pharmaceutica l J. Kenya. 24, 38–41
4. Yu T., Zhu H. (2020). Hyper-parameter optimization: a review of algorithms and applications. arXiv. 10.48550/arXiv.2003.05689
5. International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), 706-
6. 711, doi: 10.1109/PerComWorkshops53856.2022.9767372.
7. Neighbors C, Song SA. “Dysphonia” (2022) StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing.
8. Alatas Bilal, Moradi Shadi, Tapak Leili, Afshar Saeid (2022), “Identification of Novel Noninvasive Diagnostics Biomarkers in the Parkinson’s Diseases and Improving the Disease Classification Using Support Vector Machine”, BioMed Research International, Hindawi
9. P. Raundale, C. Thosar and S. Rane (2021), "Prediction of Parkinson’s disease and severity of the disease using Machine Learning and Deep.