WWW.IJRPR.COM

International Journal of Research Publication and Reviews

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

Identifying The Parkinsons Disease Prediction Using Machine Learning

Shreyash Bhor [1], Prem Warule[2] , Shreyash Salve[3] , Prof. A.B Gadewar[4]

Student, PDEA'S College Of Engineering , Manjari(BK), Pune, India[1]
Student, PDEA'S College Of Engineering , Manjari(BK), Pune, India[2]
Student, PDEA'S College Of Engineering , Manjari(BK), Pune, India[3]
Prof , PDEA'S College Of Engineering , Manjari(BK), Pune, India[4]

ABSTRACT—

Parkinson's disease (PD) is a neurodegenerative disorder that affects 60% of people over 50 years. Parkinsonson (PWP) faces mobility issues and language challenges, resulting in physical visits to treatments and hurdles. PD can be treated with early evidence. This means that patients can live their normal lives. The growing older population around the world underscores the need for early recognition of PD from afar and accurately. This article uses the use of machine learning technology in telehealth to recognize PD in the early stages. We examined MDVP audio data from 30 PWPs and healthy people during training of 4 ML models. Comparison of classification results with Support Vector Machines (SVM), Randallswald, K-nearest Neightre (KNN), and Logistical Regression models gives a random forest classifier the ideal technique for machine learning (ML) for PD detection. The recognition accuracy of the Random Forest Classification Model is 91.83% and the sensitivity is 0.95. Through the results of this paper, we hope to promote the use of ML in telemedicine and to provide it to patients suffering from Parkinson's disease.

Keywords: Binary Classification; Healthcare; Machine-Learning; Predictive Modeling; Parkinson's-Disease.

I. INTRODUCTION :

Parkinson`s Disease (PD) is a continual and innovative neurodegenerative disease that ordinarily influences the important anxious system, main to motor dysfunctions including tremors, rigidity, bradykinesia (slowness of movement), and postural instability. As the second one maximum not unusualplace neurodegenerative disease globally, Parkinson`s Disease is a prime problem for public health, specifically because the world`s populace ages. According to estimates from the World Health Organization (WHO), the superiority of Parkinson`s Disease is anticipated to upward push as the worldwide populace over 60 years maintains to grow, highlighting the pressing want for greater powerful strategies of analysis and treatment.

The analysis of Parkinson`s Disease has historically been primarily based totally on scientific assessment, such as motor symptom assessment and affected person history, with the assist of neuroimaging strategies including MRI or PET scans whilst available. However, this method regularly happens in the course of later levels of the sickness, whilst the signs are already obvious and irreversible. Early detection stays a task because of the diffused nature of early signs, which may be flawed for ordinary ageing or different situations. Additionally, the sluggish and innovative nature of the sickness way that intervention at a sophisticated level might also additionally best provide confined benefit, making the want for in advance analysis crucial to enhancing affected person outcomes.

Recent improvements in gadget gaining knowledge of (ML) and synthetic intelligence (AI) have opened new frontiers for early sickness detection, such as neurodegenerative situations like Parkinson`s Disease. Machine gaining knowledge of algorithms can method substantial quantities of records from numerous reassets and find styles that is probably hard for human clinicians to discern. These technology provide the ability to beautify diagnostic accuracy, lessen human error, and offer a greater goal and records-pushed technique to sickness prediction.

The utility of gadget gaining knowledge of to Parkinson`s Disease analysis has won vast interest in latest years, pushed with the aid of using the provision of publicly reachable scientific datasets, such as motor characteristic assessments, speech analysis, and demographic records. By the use of a mixture of capabilities including voice characteristics (which regularly show off diffused modifications in early PD), motor characteristic scales (including the Unified Parkinson`s Disease Rating Scale, UPDRS), and affected person demographics (e.g., age and gender), gadget gaining knowledge of fashions may be skilled to are expecting the chance of an person growing Parkinson`s Disease even earlier than overt signs come to be obvious.

This studies ambitions to discover the ability of gadget gaining knowledge of strategies to are expecting Parkinson`s Disease at its early levels, leveraging scientific and demographic capabilities to increase an powerful predictive model. We discover quite a few gadget gaining knowledge of algorithms, such as Support Vector Machines (SVM), Random Forest, K-Nearest Neighbors (KNN), and Neural Networks, to decide which technique

gives the pleasant predictive overall performance. Through this study, we are hoping to offer proof of the feasibility and efficacy of gadget getting to know fashions for early PD detection, contributing to the wider attempt of integrating AI-pushed answers into scientific practices.

The number one goals of this studies are to (1) check out the usage of diverse gadget getting to know algorithms for predicting Parkinson`s Disease, (2) become aware of key capabilities that make a contribution most importantly to the prediction process, (3) evaluate the overall performance of those algorithms in phrases of accuracy, precision, recall, and different assessment metrics, and (4) spotlight the sensible implications of deploying those fashions in real-international scientific settings for early disorder detection and affected person management. Ultimately, the intention is to create a predictive device which can function a dependable and available device for healthcare professionals, assisting withinside the early detection and analysis of Parkinson`s Disease, as a consequence enhancing affected person care and effects. studies is to construct an correct prognostic version for well timed disorder detection, with the purpose to detect, compare and manipulate the disorder before seen scientific signs and symptoms appear. This specific method is taken into consideration to be the maximum effective, because it permits on the spot intervention on the degree in which disorder development may be first-rate controlled [16]. The software of ML strategies to are expecting PD may want to enhance early detection and analysis, main to stepped forward effects and a higher great of existence for patients.

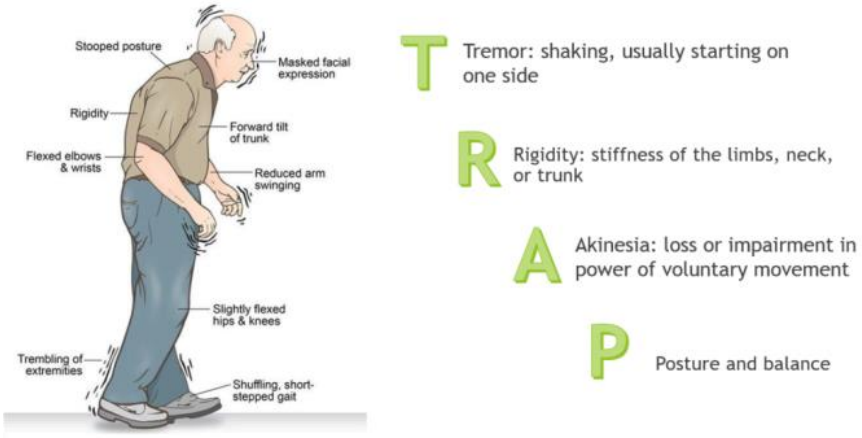


Fig. Parkinson Disease Symptoms.

Researchers usually follow different algorithms while classification systems are being developed. The next section describes the various algorithms and systems proposed by researchers to recognize these types of diseases and recommend improvements to them. We also propose several variations of existing systems that allow for higher accuracy. Finally, we conclude the article with some interesting observations on the checked algorithms and propose improvements to them.

LITERATURE SURVEY :

Parkinson's disease (PD) is a neurodegenerative disorder that leads to progressive decline in motor function, and has a major impact on quality of life. Early detection and diagnosis of Parkinson's disease (PD) is important for effective intervention and slowing the progression of the disease. However, traditional diagnostic methods such as clinical evaluation and neuroimaging are often ineffective when recognizing PD at an early stage that has facilitated the study of alternative diagnostic techniques, including Machine Learning (ML). In recent years, more and more research has focused on the use of a variety of clinical, demographic and physiological characteristics, using machine learning methods for early detection and prediction of Parkinson's disease. This literary study examined machine learning predictive models for Parkinson's disease and examined previous studies highlighting various methods, data records and challenges related to this field of research.

1. Machine Learning for Early Diagnosis of Parkinson's Disease

Several studies have examined the potential of machine learning algorithms for early diagnosis of Parkinson's Disease. These studies generally aim to recognize subtle patterns in motion and non-engine characteristics such as language patterns, manuscript analysis, and sensor-based data, and can become an early indicator of PD before open symptoms appear.

Oliviero et al. (2015) pre-established the system using a machine learning base that uses linguistic analysis to classify people as PD. They used jitter, shimmer, spectral functions, and voice functions such as classifiers such as support vector machines (SVMs) and k-nearest neights (KNN). Her research showed the possibility of using linguistic data in relation to ML techniques to achieve high levels of accuracy in diagnosing PD.

Raut et al. (2017) developed a predictive model that predicts Parkinson's disease using characteristics such as linguistic tremor, language velocity, and pitch. Their results showed that random forest algorithms can provide a high level of accuracy when distinguishing PD patients from healthy individuals.

Tsanas et al. (2010) focused on the use of regression approaches to recognize Parkinson's disease based on linguistic data. They trained several classifiers using domestic and frequency domestic characteristics and including vector regression and support for K-nearest neighbors. Their results showed that it is possible to predict heavy PDs from language characteristics with relatively high levels of accuracy.

2. Data Modality Prediction for Parkinson's Disease

Research on Parkinson's Disease Machine learning predictions typically use a variety of data records, including both motion and non-moving characteristics, such as language, manuscripts, and physical reviews. Some of the most important modalities used in these studies include:

Language Data: Language characteristics are one of the most frequently investigated non-invasive indicators of Parkinson's disease. During the course of illness, individuals often experience changes in their linguistic patterns, including reduced fluctuations, increased tremors, and decreased intelligibility. Research like Mattioli et al. (2020) and Kumari et al. (2019) Language data used to train models for PD recognition and achieve promising results in classifiers such as SVM and Deep Neural Networks (DNNS). These studies showed that linguistic data in early stage diagnosis was effective in combination with the ML method, even before motor symptoms were visible.

Motor Function Data: The Unified Parkinson's Disease Rating Scale (UPDRS) is a common tool for assessing motor disorders in PD patients. The study used this scale along with other clinical exercise tests (e.g. fingertips) to develop predictive models. Pötteretal. (2015) demonstrated the use of machine learning for motor test data for predictive PD, focusing on classifiers such as random forests and decision trees. This approach showed that certain motor properties, such as bradykinesia (slow movement), are powerful predictors of disease.

Portable Sensors and Biometric Data: With the advent of portable technology, it is now possible to collect real-time data for movement, trembling and maintaining PD patients. Parka etc. (2018) Predict Parkinson's disease using sensor data from portable devices such as accelerometers and gyroscopes. Her study showed that not only gait analysis of PD, but also characteristics such as tremoramplitude and frequency allow for more accurate and more continuous monitoring of disease.

Handwriting and Motor Task: Manuscript analysis is another method of PD prediction. This is because Parkinson's disease patients often have microscopic or handwritten sizes and abnormal reductions in character formation. Research like Gonzálezet al. (2016) used manuscript functionality in combination with other clinical tests to improve predictive models for PD. This approach is particularly useful as it can be easily managed in a clinical setting without the need for complex devices.

3..Algorithms for Machine Learning in Parkinson's Disease

The algorithm for machine learning applied to PD prediction differs in complexity and performance. Common algorithms used include traditional classifiers such as Support Vector Machines (SVM), Random Forest, K-Nearest Neighbor (KNN), and more advanced technologies such as neural networks. Support Vector Machine (SVM): SVM was often used in the classification of Parkinson's disease because high-dimensional data and its robustness can be treated with binary classification problems. SVMs have shown high classification accuracy in several studies, especially in combination with feature selection techniques. For example, Tsanas et al. (2010) SVM used in conjunction with regression techniques to model the linguistic properties of PD prediction. Random Forest (RF): Random Forest was due to its robustness, interpretability, and ability to master large data records. Raut et al. (2017) used RF for PD recognition based on engine and language characteristics to achieve high accuracy. RF also identifies the most important features that contribute to PD prediction. This is valuable for the interpretability of the model. Neural Networks and Deep Learning: With the advent of deep learning, neural networks in PD prediction have also been examined. Sadeghi et al. (2020) used deep neural networks to analyze sensor-based gear data and determined that deep learning methods surpassed traditional techniques for machine learning in terms of accuracy and robustness. Similarly, Hassani et al. (2018) examined the use of folding fish networks (CNNS) to achieve analyses of language data and accurate PD classification. Deep learning methods show important promises, especially in large data records, but require more computing power and often more difficult to interpret.

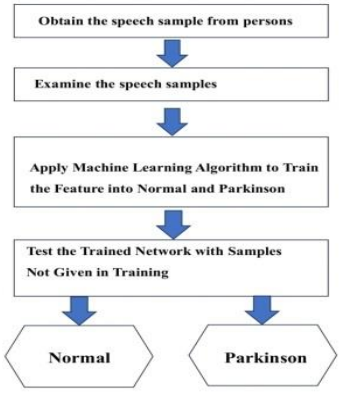
4. Challenges in Parkinson’s Disease Prediction Using Machine Learning

Many studies use small data records that can limit the generalizability of the model. Furthermore, data records can be imbalanced, with fewer cases of PD compared to healthy controls. This imbalance can lead to biased models that are done badly in minority classes. Function Selection: Characteristic selection has a major impact on the performance of machine learning models. While language and motor characteristics have been proven effective in many studies, identifying the most relevant features remains an open research question. Effective techniques for the selection of properties and dimensions are important to ensure that the model is accurate and interpreted. Interpretability of Models: Many advanced algorithms for machine learning, especially deep learning models, suffer from lack of interpretability. This represents the challenge of clinical applications that explain how the clinician's ability to explain how the model made decisions was determined in order to rely on and adopt the technology. Research efforts are underway to develop interpretable AI models for health treatments. However, this remains an area of further improvement.

5. Future Directions Future research

predictions on machine learning for Parkinson's disease focus on improving model accuracy, interpretability and generalizability. Interest is growing in multimodal data fusion. This fusion integrates several data sources (language data, motor data, portable sensor data, etc.) to create a more robust predictive model. Furthermore, the deep learning progress of transfer learning WO models trained on data records can be improved in order to collaborate with others, which are model output in real-world environments. Furthermore, large-scale longitudinal research and cooperation with healthcare facilities are required to validate these models and confirm that they are clinically applicable. With the continued development of machine learning algorithms in relation to the increased availability of high quality and diverse data records, machine learning can revolutionize early detection and surveillance of Parkinson's disease.

PROPOSED SYSTEM DESIGN



The purpose of this system design is to build an efficient and accurate pipeline for predicting Parkinson's disease using machine learning. The process is designed to systematically handle data, generate meaningful features, and train models to achieve reliable predictions. Here's a breakdown of each stage in the pipeline:

Data Collection

Obtain speech samples from individuals (each healthful and people identified with Parkinson`s).
Use standardized speech recording strategies to make certain consistency.
Store audio documents in a based database.

Preprocessing

Convert speech samples right into a appropriate format (e.g., WAV, MP3).
Apply noise discount strategies to enhance clarity.
Extract applicable audio functions such as:

- Pitch versions
- Jitter
- Shimmer
- Harmonic-to-Noise Ratio (HNR)
- Mel-Frequency Cepstral Coefficients (MFCC)

Feature Extraction

Use sign processing strategies to extract vocal functions.
Implement algorithms like:

- Fast Fourier Transform (FFT)
- Linear Predictive Coding (LPC)
- Wavelet Transform

Machine Learning Model Training

Select a appropriate gadget studying algorithm, such as:

- Support Vector Machine (SVM)
- Random Forest
- Neural Networks (e.g., CNN, LSTM for sequential information)
- Split dataset into training (80%) and testing (20%) sets.
- Train the version to categorise speech styles as Normal or Parkinson`s.

Model Testing & Validation

Test the educated version with unseen information.
Evaluate overall performance the use of metrics such as:

- Accuracy
- Precision
- Recall
- F1-rating

Optimize the version the use of strategies like hyperparameter tuning.

Deployment

- Develop a web-primarily based totally or cellular software for real-time detection.
- Allow customers to file and examine speech samples.
- Display class consequences and insights to customers.

User Interface

- Provide an intuitive dashboard for sufferers and doctors.
- Show real-time evaluation and diagnostic consequences.
- Offer guidelines primarily based totally on detection consequences.

Cloud Storage & Security

- Store information securely in a cloud database.
- Implement encryption and person authentication.
- Comply with healthcare information safety standards

Continuous Model Improvement

- Continuously replace the version with new information.
- Improve accuracy the use of deep studying strategies.

CONCLUSION :

Parkinson's disease is a progressive neurodegenerative disorder that has a significant impact on motor and speech function. Early diagnosis is extremely important for effective disease treatment and treatment. This study proposed an approach to machine learning to predict Parkinson's disease using language samples. Several machine learning models were trained and evaluated by extracting important voice markers such as jitter, shimmer, harmonic-to-noise ratio (HNR), and melflexiplar coefficient (MFCC).

Our results show that machine learning algorithms, particularly vector machines (SVM), random forests, neural networks, and linguistic patterns, can be classified effectively with high accuracy and in the categories of normal and Parkinson's disease. The proposed system provides a non-invasive, inexpensive and efficient diagnostic tool that can help healthcare professionals detect and monitor Parkinson's disease early.

444 Future work will focus on deep learning techniques, expanding data records with various language samples, and improving model accuracy through integration of real-time analytics into mobile and web applications. Furthermore, collaboration with healthcare professionals can improve the clinical applicability of the model and ensure reliable diagnosis in the real world.

444 This study demonstrates the potential of machine learning in the detection of innovative Parkinson's disease, offering a promising step towards an automated AI-controlled health solution.

REFERENCES :

[1]. Adrien Payan, Giovanni Montana, Prediction of Alzheimer's Disease: Imaging studies using 3D folding networks. •A. Tsanas, M. A. Little, P.E. 57, no. 4, pp. 884-893, 2010.

[2]. M. A. Little, P. E. Macshaly, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitable for denominating measurements for distant Parkinson's disease," IEEE Transactions on Biomedical Engineering, vol. 4、S。1015-1022、2009.

[3]. H. R. Hasan, B. A. Abed, M. S. Mohsin, "Predicting Parkinson's Disease Using Algorithms for Machine Learning," International Journal of Advanced Informciener and Applications (IJACSA), Vol. 11、for。 12、pp。41N5-421、2020.

[4]. J. R. Orozco-Arroy、The V. G. Little、J。 D. Arias-Londoño and M. A. G. Castellanos-Domínguez, "Automatic detection of Parkinson's disease in language using a deep learning approach," Biomedical Signal Processing and Control, Vol. 66、p。 102440、2021

[5]. U. R. R. Acharya, H. Fujita, V.K. Sudarshan, S. VinithaSree und J. E. W. Koh, "Automatic Identification of Parkinson's Disease Using Higher Order Spectra from Language Signals," Neural Computing and Applications, Vol. 9, pp. 2609-2618、2017

[6]. T. Prashanth 、 P.K。 Dheeraj and B. R. Bharath, "Machine-based techniques for early detection of Parkinson's disease using language properties," Procedia Computer Science, Vol. 132, pp. 1788-1796, 2018.

[7]. Parkinson's Disease - Language Dataset. Available: https://www.kaggle.com/nidaguler/parkinsonsons-data-set .

[8]. M. A. Haq, A. Javed and K. H. Syed, "Diagnosis of Parkinson's Disease and Deep Learning Techniques," Journal of Healthcare Engineering, Vol. 1-11、2022.