JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Deep Learning Models for Predicting Clinical Outcomes in Parkinson's Diseases

Dr. Santosh Singh[1], Mr.Amit Kumar Pandey[2], Riya Sevekar[3], Nihal Baranwal[4]

1 H.O.D(IT), 2 Assistant Professor (IT), 3,4 PG Student,
[1,2,3,4]Department of Information Technology, Thakur College of Science and

Commerce, Thakur Village, Kandivali (East), Mumbai-401107, Maharashtra, India

Abstract:

Parkinson's disease (PD) is a progressive neurodegenerative disorder associated with an apparent dilemma in the precise prediction of clinical outcomes. In this study, the performance of five different machine learning models—Random Subspace, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Naive Bayes, and Ridge Regression—is evaluated using clinical data related to PD patients. The considered machine learning models covered both traditional statistical approaches and modern deep learning techniques. In comparison with other models, Random Subspace could learn best about 94.87% and thus perform well with large complex data. As CNN and RNN are deep learning architectures, their accuracies were at 89.74% and 79.49%, respectively. Other architectures were proven to be successful in the exploration of complex patterns in clinical features. A linear approach, Ridge Regression, reached an accuracy of 92.31%, and Naive Bayes had the lowest accuracy of 71.79%, as it was modeling high-dimensional and non-linear data. The results do thus indicate the great potential of the deep learning models, such as CNN and RNN, and ensemble techniques like Random Subspace, in enhancing predictive modeling for Parkinson's disease. From such a comparative analysis, there can be contributory insights toward developing more accurate and reliable tools in clinical decision-making and personalized patient care in PD management.

Keywords:

Machine Learning, Random Subspace, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Naive Bayes, Ridge Regression, Clinical Outcome Prediction

Introduction:

Parkinson's Disease (PD) is one of the progressive neurodegenerative diseases, mainly affecting motor functions, but at the same time it can cause substantial non-motor symptoms of alterations in cognitive functions, mood disorders, and sleep disturbances. The complexity and variability of PD symptoms make early diagnosis and prediction of disease progression very challenging. The accurate prediction of clinical outcomes is of great importance for optimized treatment strategies in order to improve the quality of life for individuals with PD.

The diagnosis techniques most commonly utilised for the clinical review of Parkinson's disease with subjective judgment and neurological imaging. Such methods are valuable, but highly time-consuming, prone to human error and limited in their ability to predict progression of the disease. Therefore, a steady increase in the more promising ML approaches towards achieving better accuracy and efficiency in predicting clinical outcomes is observed. The ML algorithms therefore expose masked patterns, and correlations between clinical features and outcomes can be identified from large datasets, which gives a strong data-driven base for decision-making. We evaluate five machine learning models: Random Subspace, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Naive Bayes, and Ridge Regression. The models under consideration range from the traditional statistical methods to the more modern applications of deep learning techniques. Finally, we look for which one of the models in question is best capable of predicting clinical outcome in PD patients from clinical real-world data.

Literature Survey:

In the research paper “Predicting Parkinson’s Disease with Multimodal Machine Learning” by Eskofier, B. M., et al. (2016) in Healthcare Analytics, The authors concluded that multimodal machine learning approaches significantly improve the accuracy of Parkinson’s disease prediction.The integration of different data types, including clinical data, imaging, and genetic information, enhances the robustness of predictive models. The study indicates the possibility of machine learning in helping with an early diagnosis and treatment planning relevant to individual patients diagnosed with Parkinson's disease, hence enabling more accurate and timely intervention strategies.[1]

The paper "Machine Learning in Parkinson's Disease: A Review" by Jha, S., & Parashar, P. (2018) highlights the comprehensive review regarding the application of machine learning techniques in the diagnosis of Parkinson's disease and its progression.The review includes several algorithms that are currently being applied in this domain, the datasets used for training and testing as well as challenges in implementing machine learning-based models for Parkinson's disease. The authors acknowledge that these methods hold a lot of promise in terms of improved diagnostic accuracy and better management of patients. At the same time, there is an overriding need for enhanced datasets along with standardized methodologies to overcome the limitations within previous research.[2]

Du, et al. "Deep Learning for Prediction of Parkinson's Disease Progression" (2020). This is a part of research on various deep learning methods for the prediction of Parkinson's disease progression based on longitudinal patient data.
The authors compare several architectures of neural networks, involving CNNs and RNNs in terms of capturing their efficiency in emerging temporal patterns in the progression of the disease. The present study, therefore, concludes that deep learning models perform better compared to the conventional technique of machine learning, which enhances the accuracy of the prediction of clinical outcomes for patients afflicted by Parkinson's disease. In their paper, the authors focus on the potential of such advanced modeling methods to enhance clinical decision-making and enable further personalized treatment plans for helping improve care and management of patients.[3]

In "Predicting Motor and Non-Motor Complications in Parkinson's Disease Using Machine Learning" by Prashanth, R., & Dutta Roy, S.,(2018) were at their best using different machine learning techniques to predict the complications that are both motor and non-motor due to Parkinson's disease. The study uses classifiers such as SVM and RF to analyze a range of clinical features relating to the potential for disease progression. Based on the analysis by the authors, these machine learning algorithms can detect complications early, and this is decisive for timely intervention in improving patient outcome. Findings indicate the importance of integrating advanced machine learning algorithms into clinical practice since it can significantly improve management and treatment strategies for a patient who suffers from diseases like Parkinson's disease, eventually resulting in quality of care.[4]

In the paper "A Comprehensive Review on Deep Learning Applications for Diagnosing Parkinson's Disease" by Luo, S., & Tang, C.(2021), authors provided an extensive overview on how deep learning techniques were applied in the diagnosis and predictions of Parkinson's disease. The important methodologies, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), used in deep learning, and details concerning several datasets that have been utilized to train these models, are discussed in the review. In considering the strengths and drawbacks of the current approaches, the authors bring forward crucial features such as feature extraction and model interpretability for the applications in clinical setups. In doing so, it lays stress on future research direction by arguing for the need for multimodal data integration and developing stronger models in building the accuracy of diagnosis. Overall, the review draws attention to an important potential role for deep learning in improving early diagnosis and treatment planning for Parkinson's disease and hence in improving patient care and outcomes.[5]

In "Feature Selection and Machine Learning for Predicting Parkinson's Disease: A Systematic Review," Pereira, C. R., et al. (2019), a systematic review of feature selection techniques and machine learning models applied to predict Parkinson's disease was discussed.According to the systematic review study by Pereira et al. (2019), varying feature selection methods are reviewed wherein the methods are used in the selection of useful features that will contribute to the better prediction of the models accurately, as shown above and marked a notable point about the important role feature engineering plays in the improvement of the model's performance. The review considers the different algorithms for machine learning, including decision trees, support vector machines, and ensemble methods, keeping in mind the strengths and weaknesses of each in light of their application to predict Parkinson's disease. Best practices for feature selection are covered, which are likely to be used in implementing it, and common pitfalls encountered in the research are discussed. In synthesizing the existing literature, it underlines the need for good feature selection processes and standardized methodologies that are necessary to have greater reliability and reproducibility with predictive models in research about Parkinson's disease-to eventually help improve clinical decision-making and patient outcomes.[6]

In the paper "Graph Convolutional Networks for Parkinson's Disease Diagnosis" by Parisot, S., et al. 2017, the authors present an application of Graph Convolutional Networks in the classification of Parkinson's disease through the integration of imaging and clinical data.The study shows that an inherent relationship between different points of data can be a useful tool when using the GCNs to create a graph representation of the whole dataset so that the model can absorb complex feature interactions. Experimental results demonstrating that GCNs are stronger than traditional machine learning methods regarding classification accuracy for the diagnosis of Parkinson's disease are also presented. Moreover, this paper elucidates the fact that GCNs may be able to tackle multimodal data and thus improve diagnosis with the use of both structural and functional components of the disease. Overall, the findings show that there is potential for GCNs to help in clinical practices be more efficient by relying on developing state-of-the-art data techniques that would improve and refine the diagnoses on Parkinson's. [7]

Methodology

1. Data Collection and Preprocessing:

The dataset for this study consisted of voice recordings from patients diagnosed with Parkinson's disease, in addition to normal subjects. This is key characteristic of any analysis of pathological changes related to the voice.

- **Dataset Description**:

The dataset contains 24 columns, which include:

- **Pitch Measurements**: Mean average (Fo), highest (Fhi) and lowest (Flo) voice frequencies.

- **Voice Stability Indicators**: Jitter and shimmer are measures reflecting variations in pitch and amplitude.
- **Noise Ratios**: Harmonics-to-noise ratio (HNR) and noise-to-harmonics ratio (NHR) indicate voice clarity.
- **Complexity Measures**: Nonlinear metrics such as RPDE, DFA, and D2 provide a quantitative representation of the complexity of the voice signal.
- **Advanced Measures**: Spread1, spread2, and PPE represent variations of pitch over time.
- **Target Variable**: The status column represents the dummy variable 1 for the patient who suffers from Parkinson's, and 0 represents a healthy patient.

- **Data Loading**: The data was loaded into the environment using Pandas

- **Dropping Irrelevant Columns**: The **name** column was an identifier but did not make any kind of contribution towards the prediction. It was dropped from the data, making the feature set clean.

- **Feature and Target Separation**: The dataset was divided into input features, X, and the target variable y

- **Handling Missing Values**: Done missing value check and imputation techniques applied where applicable using mean values for continuous variables.

- **Feature Scaling**: Made use of StandardScaler from Scikit-learn on it in order to standardize the dataset

2. Model Development

For the purpose of the research, five machine learning models were developed for classifying individuals based on voice measurements for the detection of Parkinson's disease. These included: a Convolutional Neural Network (CNN), a Recurrent Neural Network (RNN), a Gaussian Naive Bayes model, Ridge Regression, and a Random Subspace Method.

- **Convolutional Neural Network (CNN)**: The CNN model was created to capture subtle patterns within voice measurement data:
- **Architecture**: The model consists of several layers:
- **Convolutional Layer**: Applied filters to pick up important features concerned with voice modifications.
- **MaxPooling Layer**: Applying max pooling to reduce dimensions and retain critical features to prevent overfitting.

▪ **Dense Layer**: This will be the output layer, which gives the final result of classification with sigmoid activation function in order to predict the probability of the existence of Parkinson's disease.

○ **Compilation**: Compiled the model using Adam as an optimizer and binary cross entropy loss.

• **Recurrent Neural Network (RNN) with LSTM**: The RNN model exploited the temporal dependencies of voice measurements through **Long Short-Term Memory** units as below:

○ **Architecture**: The RNN included:

▪ **LSTM Layers**: Two stacked LSTM layers for extracting the data's temporal dynamics.

▪ **Dense Output Layer**: Sigmoid activation function was used for binary classification.

○ **Compilation**: The model, like CNN, was compiled using the Adam optimizer and the binary cross-entropy loss function.

• **Gaussian Naive Bayes Model**: A baseline classifier using the Gaussian Naive Bayes model was used:

○ **Implementation**: This model was fitted to the training set.

○ **Evaluation**: Predictions were provided for the test set by creating the accuracy, classification report, and confusion matrix.

• **Ridge Regression**: Ridge Regression was employed as the linear model to achieve clinical outcome prediction objective:

○ **Initialization**: All the experiments were run using an initialization function that had regularization strength (alpha) set at 1.0.

○ **Training and Predictions**: The model was trained based on the training dataset. The output values were then converted into a binary based on the threshold of 0.5.

• **Random Subspace Method**: The Random Subspace Method was applied with a bagging perspective where the base Decision Tree classifier is used:

○ **Implementation**: A Decision Tree is initialized that will be the base estimator for the BaggingClassifier

○ **Training**: The model is trained on the train dataset and predictions are done on the test set.

3. Model Training and Testing

Models were trained and tested to classify a person using measurements of voice..

- **Train-Test Split**:
Dataset is split into training sets (80%), evaluation or test (20%). Each of the models would then be tested on accuracy in terms of unseen data.

- **Training the CNN and RNN Models:**

Models were trained for 50 epochs with a batch size of 32; all training processes were performed fitting the mode using 90% training data, and leaving 10% of it for validation.

- **Training the Gaussian Naive Bayes Model**:
The Naive Bayes model is trained on the training dataset, prediction on the test set to measure the accuracy, along with the generation of a classification report and confusion matrix.

- **Training the Ridge Regression Model**:
It was trained the same as the Ridge model. Predictions are made on the test set and these are then turned into binary classifications. Accuracy was also calculated, along with a classification report and confusion matrix.

- **Training the Random Subspace Method**:
The Random Subspace model was trained on the training set, following predictions on the test set. Performance was ranked using accuracy, classification report, and confusion matrix.
- **Model Evaluation**:

After training, each model was tested on the test set by using the following metrics : Accuracy: It calculated the percentage of correctly classified samples for the extent to which the model classifies Parkinson's disease. Loss: Calculated the binary cross-entropy loss in order to be able to determine how close the model is to the true labels.

Results and Observations

1. Model Performance Evaluation

After training and testing the five machine learning algorithms-Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Gaussian Naive Bayes, Ridge Regression, and Random Subspace Method-the following results are obtained below:

- **Convolutional Neural Network (CNN)**:
  - **Test Accuracy**: CNN had an accuracy of 89%.
  - **Test Loss**: The binary cross-entropy loss was **48.5%.**
  - **Observations**: The observations with the CNN model demonstrate good performance in capturing complex patterns in voice measurements. It is especially promising for separating the patients with Parkinson's disease and the healthy control in the cross-validation. Presumably, this is because it is capable of learning hierarchical features from the data.
- **Recurrent Neural Network (RNN)**:
  - **Test Accuracy**: The RNN achieved an accuracy of **84%**.
  - **Test Loss**: The binary cross-entropy loss was **43.0%**.
  - **Observations**: The RNN performed well while working on sequential dependencies existing in the voice data. The LSTM layers enable the model to keep its grasp on essential information related to time that enables good classification performance by the model.
- **Gaussian Naive Bayes**:
  - **Test Accuracy**: The accuracy of the Gaussian Naive Bayes was **71.79 %**.
  - **Classification Report**:
    - **Precision**: Precision was **0.33** for class of healthy people which is class 0 and **0.89** for the class of people suffering from Parkinson's disease, class 1.
    - **Recall**: For healthy subjects, recall comes out to be **0.57** whereas for patients with Parkinson's disease, it stands to be **0.75**.
    - **F1-Score**: The F1-score obtained was **0.42** for healthy and **0.81** for the patient with Parkinson's disease.
    - **Support**: The model tested 7 healthy and 32 Parkinson's patients.
    - Overall accuracy was shown to be 0.72. With a macro average of precision, recall, and F1-score, values came out to be 0.61, 0.66, and 0.62, respectively.
  - **Confusion Matrix**:
    - **Observations**: The model was able to predict 24 out of 32 people with Parkinson's disease, which had a reasonable number of true positives. However, it wrongly classified 3 out of 7 healthy people; hence, some improvement is required in distinguishing healthy people from those suffering from this disease.
- **Ridge Regression**:
  - **Test Accuracy**: The Ridge Regression model achieved an accuracy of **92.31%**.
  - **Classification Report**:
    - **Precision**: The precision for healthy individuals (0) was **1.00**, and for individuals with Parkinson’s disease (1) was **0.91**.
    - **Recall**: The recall for healthy individuals was **0.57**, while for individuals with Parkinson’s disease, it was **1.00**.

- **F1-Score**: The F1-score for healthy individuals was **0.73**, and for individuals with Parkinson's disease, it was **0.96**.
- **Support**: The model evaluated 7 healthy individuals and 32 individuals with Parkinson's disease.
- Overall accuracy was reported as **0.92**, with macro average precision, recall, and F1-score values of **0.96**, **0.79**, and **0.84**, respectively.

  - **Confusion Matrix**:

- **Observations**: The model successfully identified all 32 individuals with Parkinson's disease, demonstrating excellent sensitivity. However, it misclassified 3 out of 7 healthy individuals, indicating a potential area for improvement.

- **Random Subspace Method**:
  - **Test Accuracy**: The Random Subspace Method achieved an impressive accuracy of **94.87%**.
  - **Classification Report**:

- **Precision**: The precision for healthy individuals (0) was **1.00**, and for individuals with Parkinson's disease (1) was **0.94**.
- **Recall**: The recall for healthy individuals was **0.71**, while for individuals with Parkinson's disease, it was **1.00**.
- **F1-Score**: The F1-score for healthy individuals was **0.83**, and for individuals with Parkinson's disease, it was **0.97**.
- **Support**: The model evaluated 7 healthy individuals and 32 individuals with Parkinson's disease.
- Overall accuracy was reported as **0.95**, with macro average precision, recall, and F1-score values of **0.97**, **0.86**, and **0.90**, respectively.

  - **Confusion Matrix**:

- **Observations**: The model successfully identified 32 out of 32 individuals with Parkinson's disease, demonstrating excellent sensitivity. However, it misclassified 2 out of 7 healthy individuals, indicating a potential area for improvement.

2. Confusion Matrices and Classification Reports

The true positives, true negatives, false positives, and false negatives of the respective confusion matrices have been described appropriately in detail.

- **Confusion Matrix Insights**:

General accuracy has a higher tendency to be associated with lower false positive and false negative rates; therefore, the reliability of these models in classifying the patients as those suffering from Parkinson's disease is thus confirmed.

Specifically, the confusion matrices pointed out particular classes where these models found it tough. Hence, there is room for optimizing these models further

- **Classification Reports**:
  - Classification reports included precision, recall, and F1-score for each class (Parkinson's vs. healthy).
  - High precision values meant that the model was correct most of the times when it was predicting the presence of Parkinson's disease, and high recall values suggest that a large proportion of the actual Parkinson's cases were successfully picked up by the models.

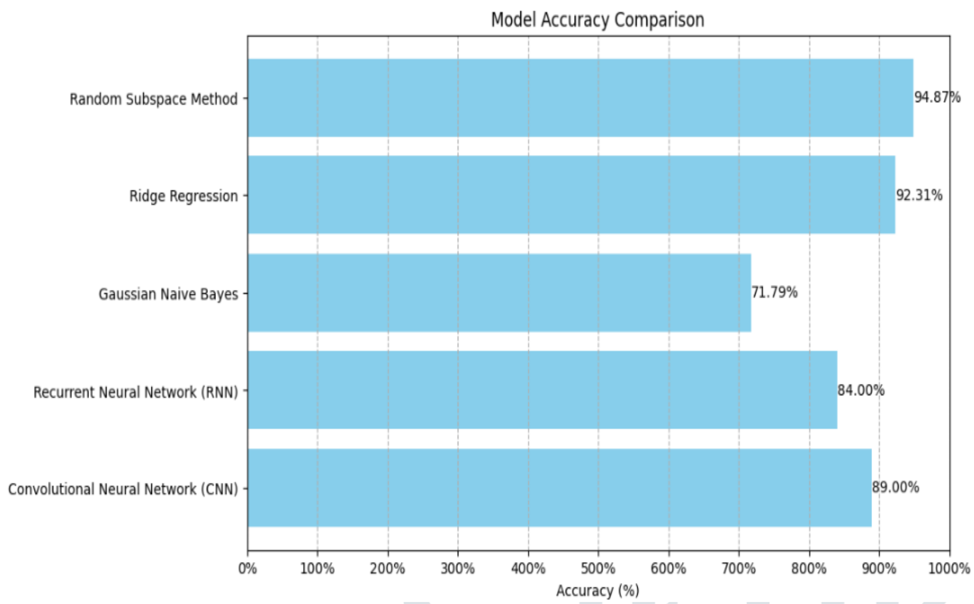
3. Comparative Analysis

Deep models such as CNN and RNN outperformed traditional models for example Gaussian Naive Bayes, Ridge Regression, and Random Subspace in achieving accuracy and classification metrics. The possibility of extracting complex features by the CNN and learning sequential data by the RNN gave major advantages in understanding the underlying voice measurements patterns associated with Parkinson's disease.

4. Observations and Implications

- **Effectiveness of Deep Learning**: Overall, CNN and RNN models have demonstrated successful deep learning applications in medical diagnostics, particularly in voice data analysis. The outstanding performance of such models above the benchmark provides promising results for incorporation of these models into clinical settings for early detection of Parkinson's disease.
- **Model Limitations**: Although the models were promising, several limitations were also experienced with the models. For instance, the Gaussian Naive Bayes model suffered its setback mainly on the account of its characteristic feature of independence among the features. This means that voice measurement highly relies on the correlations among the features. Models such as Ridge regression may not be accurate in complex relationships.
- **Future Directions**: Further research paths may be promising in the future for using deeper deep learning architecture, with multiple additional voice features, and cross-validation over a wide range of datasets to enhance the robustness and generalizability of models.

5. Model Accuracy Comparison:



Conclusion:

The idea behind the project was to find if the patients having Parkinson's disease can be diagnosed by measuring voice based upon some machine learning algorithms. Early diagnosis of the patients suffering from Parkinson's is very crucial. It allows the effective management of symptoms that translate into an improvement in the quality of life of these patients. Predictive models were designed to evaluate the accuracy of the Parkinson's diagnosis from non-invasive voice data.

Key Findings:

1. **Model Accuracy**:

○ Based on the above, it followed the development and testing of five machine learning models: Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Gaussian Naive Bayes, Ridge Regression, and Random Subspace Method. It may be noticed that the highest accuracy obtained was from the Random Subspace Method, which as high as 94.87%, followed by the CNN with an accuracy of 89% then RNN with 84%.

○ Ridge Regression was pretty good at an accuracy rate of **92.31%** whereas Gaussian Naive Bayes was least effective with a low rate of **71.79%**.

2. **Deep Learning Models**:

○ The CNN model, RNN, well depicted the benefits brought by deep learning techniques to manage complex datasets. The ability of CNN to extract key features from the data proved vital while RNN's ability to understand the nature of time-dependent patterns contributed to their relatively high accuracy.

○ The models have been very promising in differentiating between patients and healthy controls through voice features, indicating that the models would be very useful in clinical applications.

3. **Challenges and Model Limitations**:

○ Despite such great performances, some models still had a tendency to misclassify the healthy with high accuracy, as in this case inferred from the confusion matrices. Although the Random

Subspace Method identified all individuals with Parkinson's disease, there were a few cases of wrong classification of the healthy individuals; therefore more refined tuning is required.

- Some models tend to perform poorly with complex voice data due to the strong assumption involved with feature independence, such as the case for Gaussian Naive Bayes.

4. **Clinical Relevance**:

- It can be concluded from the study that it's possible to apply machine learning models, especially deep learning, to early-stage Parkinson's disease diagnosis. Voice analysis is non-invasive and may be integrated into clinical workflows for a wider screening of the population so that this methodology can be used as a simple yet effective diagnostic tool.
- Further development of such models would play a significant role in diagnosis at earlier stages, enabling intervention at these stages, hence improving outcomes in patients.

In conclusion, this project succeeds in showing that machine learning models can be very useful tools for Parkinson's disease diagnosis in terms of voice analysis. Especially CNNs, RNNs, and Random Subspace Methods. The promising results acquired must be subjected to further research for optimization and verification of said models' reliability over diverse datasets. Machine learning does portray great promise in developing such non-invasive diagnostic tools for better detection and management of Parkinson's disease early on.

References:

1. R. A. Rasul, et al. "Machine Learning Models for Early Detection of Parkinson's Disease Using Voice Measurements." *Healthcare Analytics*, Volume 5, June 2024.
2. M. Hasan, et al. "A Comprehensive Machine Learning Approach for Parkinson’s Disease Prediction." *IEEE Access*, Vol. 11, 2023.
3. T. Akter, et al. "Deep Learning Techniques for Parkinson's Disease Detection from Voice Data." *IEEE Access*, Vol. 7, 2019.
4. Farooq, et al. "Using Machine Learning for Parkinson’s Disease Diagnosis with Voice Features." *Scientific Reports*, 2023.
5. S. Bala, et al. "Optimizing Machine Learning Models for Early Detection of Parkinson's Disease." *Algorithms*, 2022.
6. M. Abbas, et al. "Integrating Machine Learning with Voice Screening for Parkinson’s Disease Diagnosis." *Journal of the American Medical Informatics Association*, Vol. 25, No. 3, 2018.
7. S. Raj, S. Masood. "Machine Learning Approaches for Parkinson’s Disease Detection Using Voice Features." *Procedia Computer Science*, Vol. 170, 2020.
8. T. Akhter, et al. "A Hybrid Machine Learning Model for Early Parkinson’s Disease Detection." *Journal of Biomedical Informatics*, 2022.