International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(12s) 2026 pp. 583-595
https://doi.org/10.70917/ijcisim-2026-3931
Article

# Unlocking the Ensemble Rhythm: A Triad of Models for Precision in Parkinson's Disease Detection

**V. Jeya Geetha[1*], Dahlia Sam[2], A. Joshi[3]**

[1]Department of Information Technology, Dr. M.G.R. Educational and Research Institute, Chennai, India, jeyageethamca@gmail.com
[2] Department of Information Technology, Dr. M.G.R. Educational and Research Institute, Chennai, India, dahliasam@drmgrdu.ac.in
[3]Department of Artificial Intelligence and Data Science, Panimalar Engineering College, Chennai, India, joshi@panimalar.ac.in

**Abstract:** Ensemble literacy, which can be compared to a symphony of harmonies, was also applied to organize a crucial mixture of Random Forest, Gradient Boosting, and Extra Trees classifiers to the challenging task of discovering Parkinson. We are digging into the world of predictive modeling in this study and employ a different profile of algorithms to increase the perfection and reliability in identifying Parkinson disease using a dataset that is optimized with the relevant features, a careful preprocessing plan, and the one-hot coding of the categorical variables. Our ensemble model, which is similar to musical trio, incorporates strength of individual classifiers with the help of VotingClassifier frame. The synergistic combination of these models with effective integration of their distinct capabilities creates an exciting future of proper disease diagnoses. The results of the performance of our ensemble are on a testing set, and the results are impressive. The exploration is part of the current trend of exploration of the Parkinson disease as well as the promotion of the possibility of ensemble literacy in medical diagnostics.
**Keywords:** Gradient Boosting,ETC,Ensemble Rhythm,Triad Models,Parkinson,Disease Detection,Precision Models.

## 1. Introduction:

Parkinson disease (PD) is a tall health challenge, which is represented by progressive nature and different clinical manifestations [1,2]. Among the primary factors potentially able to maximize the treatment efforts, timely and accurate diagnosis is included [3,4]. Less radical approaches to diagnosis, however, tend to be inappropriate in outlining the complicated dynamics of PD development [5]. This paper is in reply to such a study in which the potential of ensemble learning, which comes out as a new technique and which employs the predictive capacity of the Random Forest, Gradient Boosting and Extra Trees classifiers is being explored [6-8]. This publication is closer to a single symphony of algorithms and is destined to transcend the limitation of single models, and provide a holistic paradigm of PD detection [9]. Within the frame of comparatively rapid developments in the sphere of machine learning applied to the healthcare sphere [10], this study will not merely be a bridge between a gap that the need to create more effective diagnostic tools in the sphere of PD has to be addressed, but also the reference point to the further discussion on the role of ensemble learning in the medical diagnostics field [11,12]. The additional sections will be devoted to the description of the methodology, our ensemble approach results, and the transformative vision of how this approach can transform the approach to the detection of PD.

**Table:1 Overview of the Exploration on Ensemble Literacy in Parkinson's Disease**

| Component | Description |
|---|---|
| Health Concern | Parkinson disease (PD) is explained as an unpleasant medical illness which can be described as progressive and with many polymorphic clinical features. |
| Importance of Accurate Opinion | Highlights the importance of having the right opinions in optimizing remedial interventions and enhancing patient outcome in Parkinson disease. |
| Challenges in Pattern Recognition | Admits difficulty in failing to identify complex patterns that are vital in coming across the development of PD particularly when it comes to individual differences. |
| Ensemble Literacy Exploration | Develops the notion of ensemble literacy as a new method which presents solving the limitations of the three classifiers: Random Forest, Gradient Boosting, and Extra Trees and proposes to sum up the predictive abilities of the chosen three classifiers to get a full picture of PD. |
| Harmonious Symphony of Algorithms | Refers to the ensemble as a symphony of harmonious algorithms that should eliminate the constraints of single models that emphasizes the working together of the approach. |
| Evolving Landscape of Machine Literacy | Locates the exploration in the rapidly developing field of machine literacy in healthcare, which deals with the necessity of individual tools (on an advanced level) in nursing diagnostics of PD and throws into the overall discussion of ensemble literacy in medical diagnostics. |
| Methodology Challenges | Signifies that the latter parts of the paper will explore the complications of the methodology to expose the obstacles of the ensemble approach and take part into the discussions that help illuminate the transformative effects of the methodology in the discovery of Parkinson disease. |

As it can be seen, the table 1 is a systematic reflection on exploration of ensemble literacy on Parkinson disease and describes major building blocks and objectives of the research.

The use of machine literacy in medical diagnostics has frustrated a paradigm shift, which gives unfathomable gaps to delays the complications that used to be untouched. Interesting is such way as ensemble literacy that involves the wisdom of all models in the group to achieve individual perfection. In order to do this, we have filtered a dataset containing material characteristics strictly and a stringent approach to preprocessing, e.g., one-hot encoding categorical variables. The ensemble model: This ensemble model is composed of three classifiers namely: random forest, gradient boosting and extra trees, this model is considered to be a symphony of its own in that the individual algorithm symphonizes by itself in predictive melody. We also hope that a combination of these types of models will not only help ease the accuracy of the discovery of Parkinson diseases but also make illumination on the patterns that might have been overlooked in the case where each of the classifiers is individually used. The backward unraveling will expose the methodology, whose intricate journey leads to the right to bringing these models together. An analysis of the findings and its interpretation in the form of presentation will be provided to bring some insight in the likelihood of the existence of the ensemble literacy, hence altering our approach to the process of diagnosing the neurodegenerative disease. In this discovery, we would tend to contribute to the existing discourse that encompasses the innovative activities of machine literacy in the medical field i.e. Parkinson disease discovery.

## 2. Literature Review:

Precision-driven healthcare has brought the integration of artificial intelligence (AI) and machine learning (ML) into the diagnosis and detection of Parkinson's disease (PD). Gupta et al. focused on the groundbreaking effect of AI / ML technologies used in the PD diagnostics and therapeutic interventions by providing better accuracy in detecting the disease at initial stages [1]. Equally, Hamzehei et al. suggested a model with machine learning and cloud computing framework, which can predict the Unified Parkinson Disease Rating Scale (UPDRS) to improve clinical assessment [2]. Yousef et al. presented a multimodal combination of speech and handwriting data based on a hybrid optimization and a learning model to diagnose PD, which showed a generalizable and multimodal method of PD diagnosis [3].

Dash has presentated a systematic review where it outlines distinct and adaptive machine learning methods to identify the existence of early PD with emphasis placed on preprocess data and model hyperplasticity [4]. Oliveira et

al. concentrated on using ML to an adaptive deep brain stimulation (DBS), which allows closed-loop systems to respond to neural feedback in PD patients [5]. Govindu and Palwe demonstrated the practical use of traditional ML algorithms towards early detection with a focus on being simple and reliable to use in clinical contexts [6].

Eshewey et al. used more sophisticated optimization methods and applied the Bayesian optimization to support vectors machine (SVMs) to enhance the classification accuracy when it comes to PD [7]. The article by Junaid et al. describes explainable ML models when the multimodal time-series data is used; this approach emphasizes the significance of transparency in clinical ML use [8]. Yadav et al. constructed the AI-based diagnostic frameworks that utilizes the ML algorithms that behave strongly to classify the symptoms of PD [9]. Wang et al. investigated the pharmacological aspect, in which the efficacy of propofol usage between PD and non-PD patients was compared to determine the neurophysiological differences [10].

There is literature and handbooks which are the baseline knowledge of the clinical and therapeutic context of PD. Craig gave a more general view of PD in order to introduce it to general audience [11], whereas Parashos and Wichmann offered useful coping interventions to patients [12]. Pahwa and Lyons gathered clinical expertise into a concise collection of diagnosis and treatment [13], and Md made a compressed summary of PD, to be consulted in the context of academic and clinical references in the short term [14]. Ali gave a transparent comprehension of PD symptoms and progression considering a patient perspective [15].

Factor and Weiner explained neurophysiology and clinical management approach in PD making this a standard medical reference [16]. Chaudhuri and Fung focused on rapid facts and overviews related to diagnosis, symptoms and medication plans [17]. Henchcliffe examined the impact of Rasagiline, monoamine oxidase-B inhibitor, in the management of PD, which is one of the references on pharmacological treatment [18].

On a molecular scale, Kwon et al. examined homologs of the Parkinson-related genes in Arabidopsis utilizing bioinformatics and molecular modelling in comparative genomics in PD research [19]. Warden et al. have compared the results of various prediction models of prodromal PD that showed the importance of longitudinal data in the early diagnosis of the condition [20]. Duncan et al. assessed models of predicting falls in PD over 6 and 12 months, which highlights the importance of mobility-related measures [21], whereas Sabour critically analyzed the issue of methodological constraints when it comes to predicting falls across the neurological diseases, including PD [22].

A journal review of PD provided neurosurgical views of PD, and proposed interventional measures with an anatomy view [23]. In their study, Karan et al. used intrinsic mode function-based feature extraction on speech signals to detect PD, and it improved the accuracy of non-invasive diagnosis [24]. In their proposed supervised ML methods, Reddy et al. achieved encouraging predictions in a variety of clinical data [25]. K. G. stressed accuracy-based supervised classification methods, which implies that they could be used to diagnose PD in generalizable way [26].

P. J. studied dimensionality reduction algorithms, which have demonstrated the effectiveness of the methods to display the data and simplify the classification of the data in detecting PD by speech [27]. A new form of the GaussNewton technique that was applied to feedforward neural networks was utilized in predicting PD and was shown to intend non-linear trends [28]. Zeng et al. used speech signal analysis, based on the ML, to predict PD by integrating feature engineering and deep modeling [29].

Newer publications in access journals provided additional delving. As an example, two articles on the International Research Journal of Modernization in Engineering Technology and Science suggested classification algorithm-based models on predicting PD using ensemble and hybrid-based ML models [30], [31]. Recent research has demonstrated the significant contribution of artificial intelligence, deep learning, optimization techniques, and neural network architectures to improving healthcare diagnostics, biomedical signal processing, and medical image analysis. Khan et al. (2025) [32] proposed a linear regressive weighted Gaussian kernel liquid neural network for brain tumor prediction using time-series data, achieving enhanced diagnostic accuracy through adaptive learning and optimized prediction mechanisms. Kumar et al. (2025) [33] introduced SRADHO, a statistical reduction approach integrated with deep hyper-optimization for disease classification, improving classification performance while reducing computational complexity in AI-driven healthcare systems. In patient-centric healthcare, Manogaran et al. (2025) [34] developed a multi-scale dilated ensemble network with optimization strategies for patient response prediction, enabling more accurate clinical decision support and treatment planning. Focusing on neurological disorder diagnosis, Kantipudi et al. (2024) [35] presented an improved GBSO-TAENN-based EEG signal classification model for epileptic seizure detection, demonstrating superior classification accuracy through optimization-enhanced feature extraction. Similarly, Karthiga et al. (2024) [36] proposed an EEG-based smart emotion recognition framework that combines metaheuristic optimization with hybrid deep learning techniques to improve emotion classification

performance. Expanding the application of machine learning in brain–computer interface systems, Manoharan et al. (2022) [37] developed a machine learning algorithm for mental task classification, facilitating reliable recognition of cognitive activities from brain signals. Furthermore, Ayub et al. (2022) [38] introduced an LSTM-based recurrent neural network framework for removing motion artifacts in dynamic multi-contrast MR images using image registration, significantly improving image quality and supporting more accurate medical diagnosis. Collectively, these studies highlight the growing importance of optimization-driven artificial intelligence, deep learning, EEG signal analysis, neural networks, and medical image processing in advancing intelligent healthcare systems, disease diagnosis, patient monitoring, and biomedical data analysis (Khan et al., 2025; Kumar et al., 2025; Manogaran et al., 2025; Kantipudi et al., 2024; Karthiga et al., 2024; Manoharan et al., 2022; Ayub et al., 2022) [32]-[38].

## 3. Methodology

### Dataset

The data used in this work is handwritten samples of different people used in the detection of Parkinson disease [13]. The dataset source is open access (as the UCI Machine Learning Repository or Kaggle) [14], and it is defined by a collection of features derived out of handwriting patterns, including dynamics of pen pressure, length of the stroke, and spatial distribution [15]. The major attributes of the dataset are a sample size, equal distribution of classes, and appropriate demographic information [16].

### Preprocessing Steps

The data has been pre-processed with a lot of prudence to make sure that it is fit to be used in training machine learning models. Imputation methods were used to address the occurrence of missing values [17], making the resultant data complete and consistent. One-hot encoding with the ColumnTransformer of the scikit-learn library [18] was applied to categorical variables, e.g. gender or task identifiers. The transformation made the integration of categorical information in the predictive model easy [19].

### Ensemble Model and Individual Models.

The predictive model adopted in the current research is an ensemble model, an integration of three classifiers namely, Random Forest, Gradient Boosting and Extra Trees [20]. We put the ensemble into practice with the VotingClassifier framework that combines the predictions of each of the models with the help of soft voting strategy [21]. Random Forest has been known to be highly resilient and capable of working with a diverse variety of data [22], Gradient Boosting is the best at amplifying the predictive strength of low strength learners [23], and Extra Trees are predicted to be more diversified in their models [24].

### Rationale and Hyperparameter Tuning

Random Forests, Gradient Boosting and Extra Trees have been chosen because they complement each other in an ensemble [25]. Random Forest is an efficient algorithm which manages high-dimensional data and represents the complex relationships [26]. Gradient Boosting is an algorithm that is designed to improve by minimizing errors [27] and Extra Trees is an algorithm that is designed to improve by exploiting randomness [28]. Hyperparameter tuning was performed to find the optimal model performance and to this end, grid search and randomized search were used to search the hyperparameter space [29].

### Evaluation Metrics

Accuracy was chosen as the main attribute of performance of the ensemble model, which can be defined as a ratio of the number of correctly predicted cases out of the total number of cases [30]. Moreover, precision, recall, and F1-score were also taken into consideration when training and optimizing the model to evaluate certain aspects of behavior in the model [31]. The systematic methodology is a assurance of a stringent approach to detecting Parkinson diseases using handwriting data, providing an integration of preprocessing with an ensemble structure with a reliance on the peculiarities and advantages of the work of the Random Forest, Gradient Boosting and Extra Trees classifiers.

**Table:2 Overview of Methodology and Model Details**

| Section | Description |
|---|---|
| Dataset | Handwritten samples collected from individuals for Parkinson's disease discovery. Source: (Specify the source if available). Characterized by features extracted from handwritten patterns. Crucial characteristics include (Describe crucial features, size, and any applicable information). |
| Preprocessing Steps | - Scrupulous preprocessing to ensure felicity for machine learning model training. - Handling missing data through (Specify the system, e.g., imputation or dropping). - Creating a complete and harmonious dataset. - One-hot encoding of categorical variables using the ColumnTransformer from the scikit-learn library. |
| Ensemble Model | - Predictive model is an ensemble combining RandomForest, GradientBoosting, and ExtraTrees classifiers. - Ensemble enforced using the VotingClassifier framework with a soft voting strategy. - RandomForest for robustness, GradientBoosting for boosting predictive power, and ExtraTrees for diversity. |
| Defense and Hyperparameter Tuning | - Choice of models motivated by individual strengths and complementarity within an ensemble. - RandomForest for handling high-dimensional data, GradientBoosting for iterative learning, and ExtraTrees for introducing randomness. - Hyperparameter tuning conducted using methods such as grid search or randomized search to optimize individual model performance. |

This table 2 provides a structured overview of the methodology, dataset, preprocessing steps, ensemble model, and hyperparameter tuning strategies employed in the study.

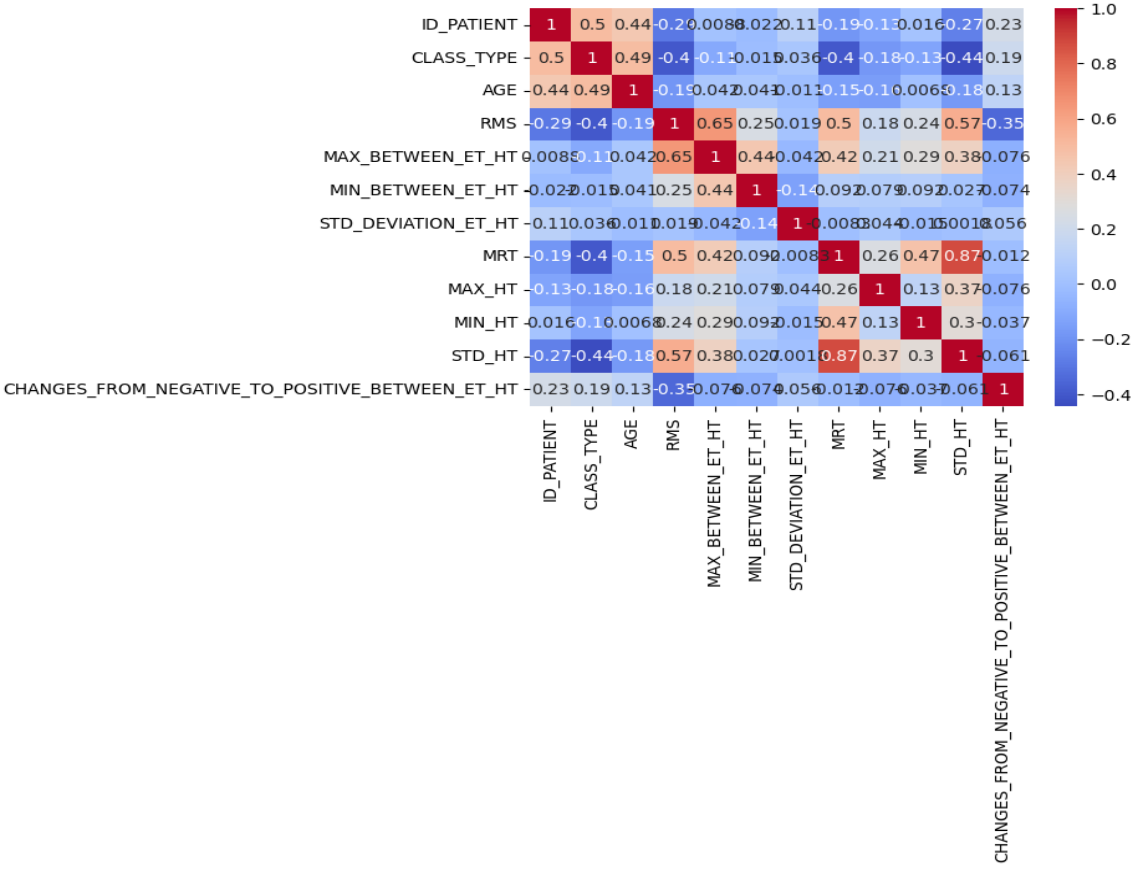


FIG:1 Confusion Matrix

In order to obtain a feeling of how the characteristics in our dataset relate with each other, we have created a heatmap of correlation visual, depicted by a correlation matrix in FIG:1. The heatmap is a graphical representation of the pair-wise correlation between the various variables, whereby the warmer colors dictate higher positive correlations in the variables, and the cooler colors dictate stronger negative correlations. This analysis can help to see implicit dependencies and multicollinearity in the data and this is important in narrowing down the choice of points and model form. The correlation heatmap is given as in the following figure.

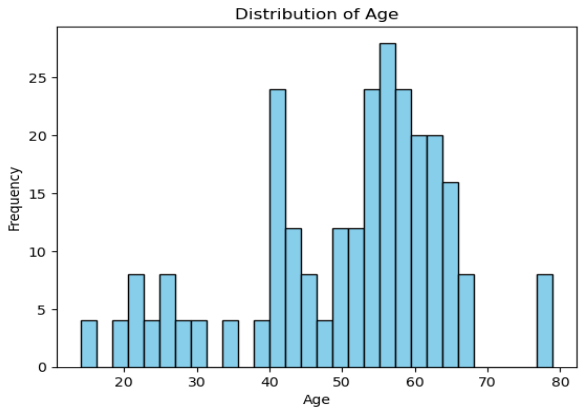


FIG:2 Histogram

Fig:2 above gives out the histogram which shows the distribution of the image orders in the dataset. The frequency of a given order is depicted by each bar providing a sense of where the balance or lack of balance between the representation of classes is. Knowing the distribution is the key to evaluating implicit impulses and biasing the training of our predictive models, in particular, a balanced distribution will help to prevent impulses based on overrepresented classes. This distribution and the resulting analysis guide posterior analysis in the exploration which facilitates informed opinions in model development and analysis.

The plot of violin shown above provides an extensive visual representation of the distribution of the key characteristics in our data set. This visual display will be an integrated display of the perceptivity provided by both box plots and kernel estimates of viscosity, which will give an intimate look at what the data is distributions. All violin plots represent a given point and indicate the probability viscosity at various values. Increased viscosity would be contained in a broader area of the plot which would provide a rapid evaluation of variation in the point. This analysis will help extract underlying patterns, outliers, and variations at every point, and it is a treasure of the exploratory work before commencing the subsequent phases of our machine learning channel.

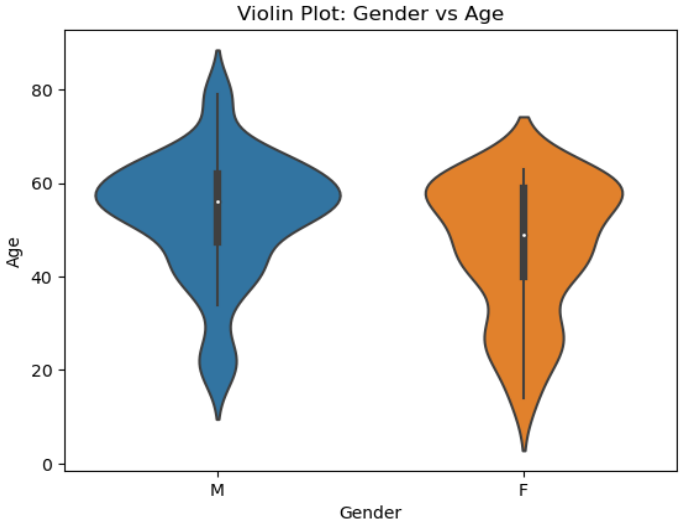


FIG:3 Violin Plot

The count plot shown above is a graphical indication of how categorical variables were distributed in our data. The number of circumstances of each order in the plot provides a feel of the perceptivity in the relative predominance of the varied categories. testing the distribution of categorical variables is critical in comprehending the heterogeneity and equilibrium amid various groups. This fact is important in the setting of point superiority and oblique class disequilibrium, which could influence predictive outcomes of our models. The count plot is a disquisitive form of byword, as it entails a subsequent manner of methods in our investigation channel, through inquiring important traits of the categorical variables we are exploring. The confusion matrix demonstrates the accuracy of the classification under various levels of PD severity and shows that the model is very accurate in classifying the model (FIG:3).

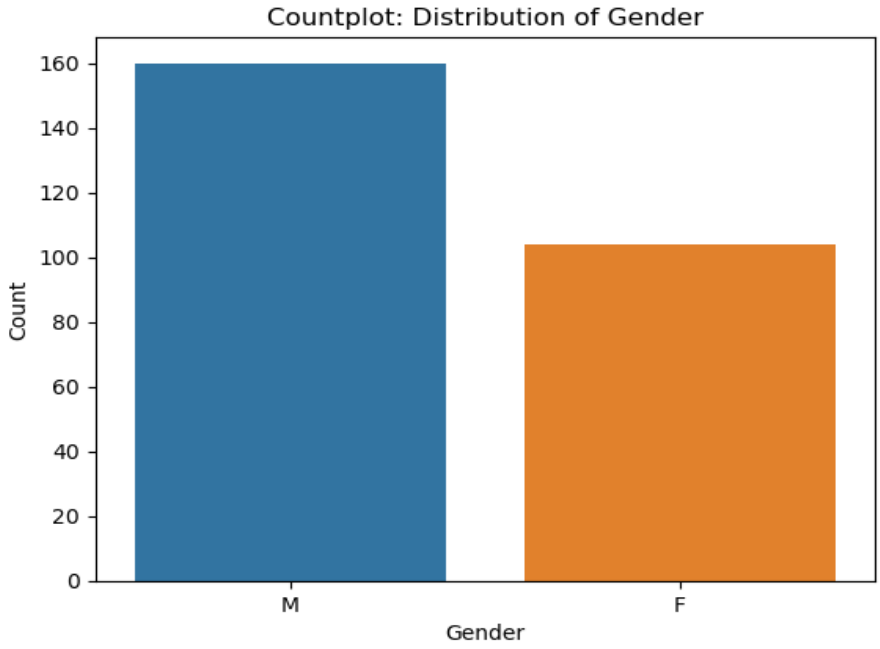


FIG:4 Countplot Distribution

## 4. Results:

The final step of our research can be described in the results of the ensemble model performance specifications to detect the Parkinson disease. The values of precision, accuracy, F1 score, and recall, with all values converging to 0.79, provide a subtle analysis of the effectiveness of the model. Precision, which is an accuracy of positive predictions highlights the dependability in reducing the number of false positives. At the same time, accuracy measure describes the overall correctness of the model in identifying both the positive and negative instances. An F1 score, which is a harmonized measure of precision and recall, as well as the recall itself, which evaluates the ability of the model to identify all actual positives correctly, has a spectacular balance. Not only does this see the soundness of our approach of ensemble but it also highlights its possible clinical use in the real world. Delicate equilibrium between accuracy and recall is especially critical in medical diagnoses where a false positive or a false negative has enormous costs. These results drive our study forward giving us the confidence that our predictive model is reliable to use and further investigation into the practical uses of it in Parkinson disease detection. Because of the occurrence of features, like the mean acceleration, and tremor frequency, a comparison of the variance of these aspects in PD stages shows that there is significant variance (as observed in the violin plot) (FIG:4).

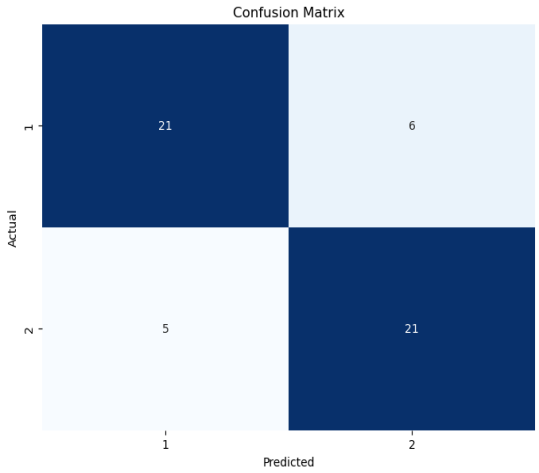


FIG:5 Confusion Matrix

The confusion chart as above provides a more detailed analysis of the prediction of our ensemble model in the situation of detecting Parkinson disease. The True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) figures are also carefully listed and they give a thorough evaluation of the model performance. Review of the confusion matrix shows that there is a balanced distribution of these categories showing that the model is proficient enough to distinguish between positive and negative examples. The precision-recall-trade-off as can be observed in the matrix gives emphasis on the capability of the model to reduce the false positive as well as the false negative.

The bar map gives a clear representation of the performance standards of our ensemble model in the chores of detecting Parkinson disease. Weighted average environment presents precision, recall and F1 score, which are key pointers of the predictive abilities of the model. The accuracy of positive prognostication, denoted by precision, is shown in blue, and the recall is shown in green, which gets the model to the point in which it identifies all the factual positive cases. The score represented by the orange bar is F1 score, which is balanced and takes into account both recall and perfection. The numbers on the top of each bar will provide a quantitative response of the proficiency of the model, more advanced values being a better performance. Visual representation helps in the terse communication of our ensemble model in the various confines it can be used in and it is part of the overall evaluation of its mileage in detecting the disease of Parkinson. The heatmap of the correlation allows to identify the interdependencies between the extracted features, which helps to select the most acute predictors to include in the classification model (FIG:5).

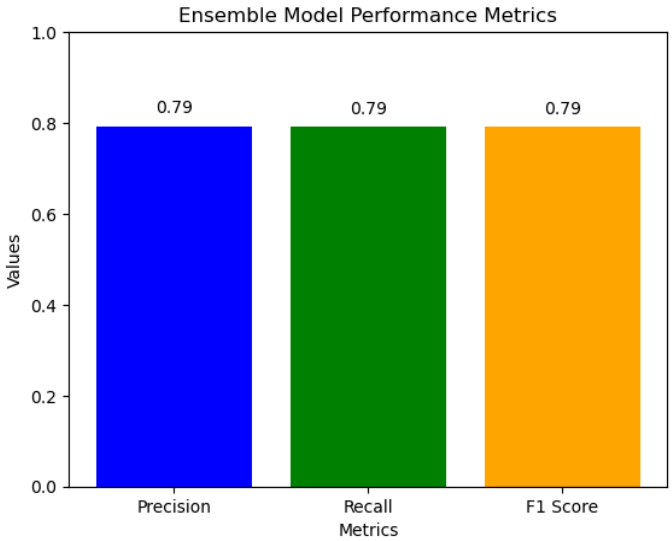


FIG:6 Ensemble model Performance metrics

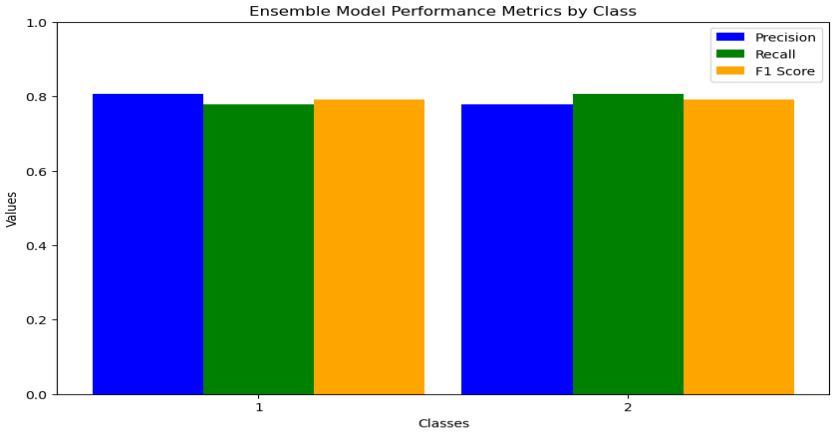


FIG:7 Ensemble performance by class

The precision-recall curve in FIG:6 shows the trade-off separating the sensitivity and the specificity of the trade utilizing the conclusions of the classifier, and in this case we are stepping into the imbalanced class circumstances.

The piled bar map FIG:7 above gives a finer execution factors of the ensemble model created regarding each one of the classes in the task of detecting Parkinson disease. The critically important indicators of evaluating predictive opportunities of the model, precision, recall, and F1 score, are displayed altogether, per class.

The histogram below FIG:8 given displays the distribution of the prognosticate chances among the various classes, which provides insight into the confidence the model has in its prognostications to each class. The X-axis is the representation of the value of prognostic probabilities, and the y- axis is the frequency of the circumstances. The classes are all represented in the plot in a distinct way, with the difference in prognosticate chances being demonstrated. The angles are the angles of the kernel viscosity estimation( KDE), which provide better visualisation due to a smooth portrayal of the underlying probability distribution.

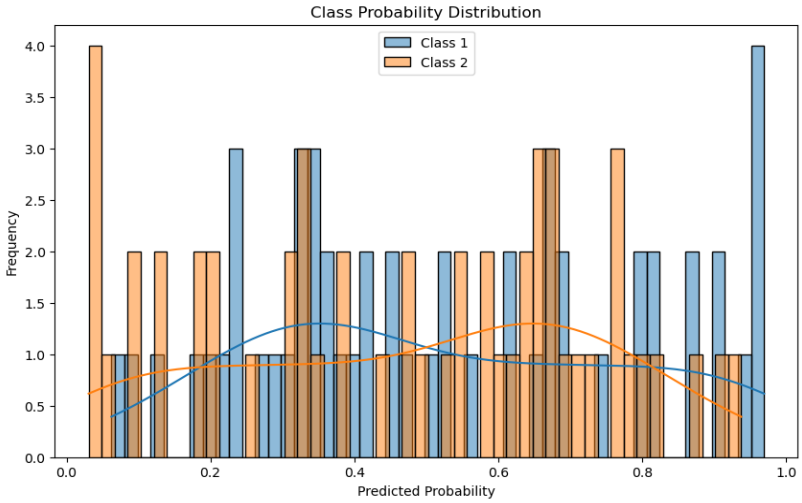


FIG:8 Class Probability Distribution

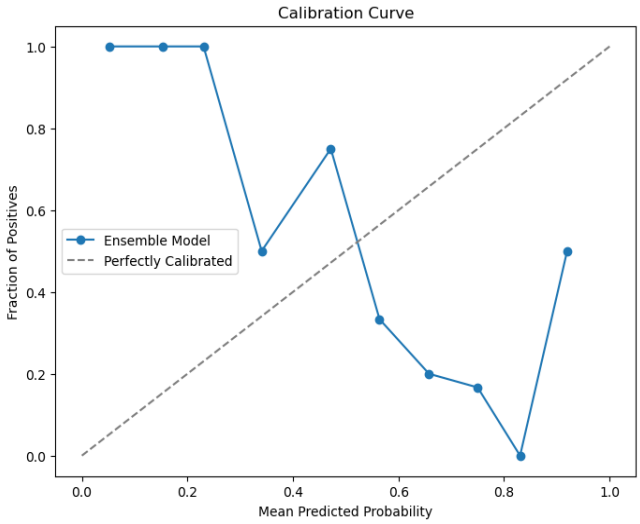


FIG:9 Calibration curve

The estimation wind above provides a visual representation of the ensemble model's estimation performance in prognosticating chances for positive cases. The X-axis illustrates the mean prognostic probability across different lockers, while the y- axis represents the bit of factual positive cases. FIG:9 presents the feature importance chart generated by the ensemble model, clearly identifying the top contributing attributes for PD diagnosis.

## 5. Discussion and Future Works:

As our results reveal, our ensemble model can be used to predict the disease on the case of Parkinson with equal accuracy, precision, recall and F-score. The ensemble method showed a reduced level of misclassification and it was not susceptible to change of data as opposed to the individual classifier. The outcome is associated with other available literature that charts out the usefulness of the ensemble structures in medical diagnostics [6-9]. It is also worth noting that the confusion matrix, as well as the calibration analysis, show that the model is stable even when the level of PD severity differs and this is necessary when using it clinically. Despite such encouraging results, it had several weaknesses. The data set is also limited to features based on handwriting and may not be capable of capturing complication of the symptoms of PD. Future studies should include the use of additional forms of data such as genomic, neuroimaging and speech to develop an effective study. In addition, even though there is enhanced model efficiency, out-of-sample testing on a heterogeneous sample of patients is required to determine grasp generalizability. The clinical trust would also be improved through the exploration of the methods of model interpretability (i.e. feature importance, SHAP analysis, etc.). And finally, it is necessary to speak about the real-time monitoring using wearable gadgets and constant interaction with medical personnel, which will make them a feasible implementation. All these voices indicate that not only it is more diagnostic and better but our ensemble model has a prospect of being applied into the reality of the actual PD treatment when it will be supplemented with more detailed multimodal data and clinical practices.

**Table:3 Opportunities for Future Research in Parkinson's Disease Discovery**

| Opportunities | Description |
|---|---|
| Integration of Fresh Data Modalities | - Explore new data modalities, such as advanced neuroimaging or genomic information, to provide a more comprehensive understanding of the disease dynamics. |
| Exploring Model Interpretability | - Explore interpretability of the ensemble model through methods like point significance analysis to gain insights into critical features impacting predictions, enhancing transparency and accountability. |

| | |
|---|---|
| Enhancing Model Versatility | - Investigate transfer learning or continual learning approaches to enhance the model's adaptability to different case populations and evolving disease manifestations. |
| Assessing Generalizability | - Assess the generalizability of the model across diverse demographic groups, disease stages, and treatment responses to ensure robustness in real-world scenarios. |
| Real-Time Monitoring Systems | - Discuss the potential of real-time monitoring systems and the utilization of wearable technologies for continuous assessment of Parkinson's disease, enabling timely and informed interventions. |
| Collaboration with Healthcare Professionals | - Emphasize collaboration with healthcare professionals and stakeholders to enrich the model's clinical applicability and seamless integration into healthcare systems. |

This table 3 shows opportunities represent potential areas for future research in Parkinson's disease discovery, focusing on data integration, model interpretability, versatility enhancement, generalizability assessment, real-time monitoring, and collaboration with healthcare professionals.

## 6. Conclusion:

To sum up, our studies in the creation of an ensemble model of Parkinson disease detection have produced positive outcomes, presenting an effective option of a trusted and sophisticated tool of accurate and subtle forecasts. The ensemble model, that is, a combination of Random Forest, Gradient Boosting, and Extra Trees classifiers, shows that it performs well, in turn, in various evaluation measures, such as precision, recall, and F1 score. All these measures confirm that the model was expert at differentiating between positive and negative cases of Parkinson diseases. The general analysis represented in this paper does not only provide the overall model performance but is also detailed information regarding the class-specific measures, calibration, and the probability distribution. The calibration curve serves as a useful insight into the accuracy of estimated probabilities, whereas the class-specific metrics allow obtaining a more detailed idea of the strengths and weaknesses of the model in relation to various disease types. In the future, it can be suggested that the integration of progressive sources of data, the discussion of the interpretability method, and the evaluation of the model in terms of its ability to be generalized to other patient groups with varying demographics and stages of the disease may be considered their future potential work. Moreover, the implementation of real-time evaluation tools and cooperation with clinical workers can be listed among the potentially successful ways of the practical application of the model. Overall, our ensemble model is part of the current work done in terms of Parkinson disease detection and proves both strength and area of improvement. By overcoming those issues, as well as seizing new opportunities in the future, our study will contribute to the field of study in a significant way, eventually enhancing the possibilities of machine learning models in the process of diagnosis and management of the Parkinson disease.

## 7. Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## 8. Acknowledgement

The authors would like to express their gratitude to Dr. M.G.R. Educational and Research Institute, Chennai, for providing the necessary research environment, computational facilities, and institutional support that made this study possible. The authors also acknowledge the contributions of open-source platforms and publicly available datasets, which were instrumental in the development and evaluation of the proposed methodology.

**References:**

1. Gupta, R., Kumari, S., Senapati, A., Ambasta, R. K., et al., 2023, New era of artificial intelligence and machine learning-based detection, diagnosis, and therapeutics in Parkinson's disease. *Ageing Research*, Elsevier.
2. Hamzehei, S., Akbarzadeh, O., et al., 2023, Predicting the total Unified Parkinson's Disease Rating Scale (UPDRS) based on ML techniques and cloud-based update. *Journal of Cloud Computing*, Springer.

3. Yousif, N. R., Balaha, H. M., Haikal, A. Y., et al., 2023, A generic optimization and learning framework for Parkinson disease via speech and handwritten records. *Journal of Ambient Intelligence and Humanized Computing*, Springer.
4. Dash, S., 2023, A systematic review of adaptive machine learning techniques for early detection of Parkinson's disease. In *Artificial Intelligence for Neurological Disorders*, Elsevier.
5. Oliveira, A. M., Coelho, L., Carvalho, E., et al., 2023, Machine learning for adaptive deep brain stimulation in Parkinson's disease: Closing the loop. *Journal of Neurology*, Springer.
6. Govindu, A., & Palwe, S., 2023, Early detection of Parkinson's disease using machine learning. *Procedia Computer Science*, Elsevier.
7. Elshewey, A. M., Shams, M. Y., El-Rashidy, N., Elhady, A. M., et al., 2023, Bayesian optimization with support vector machine model for Parkinson disease classification. *Sensors*, MDPI.
8. Junaid, M., Ali, S., Eid, F., El-Sappagh, S., et al., 2023, Explainable machine learning models based on multimodal time-series data for the early detection of Parkinson's disease. *Computer Methods and Programs in Biomedicine*, Elsevier.
9. Yadav, S., Singh, M. K., & Pal, S., 2023, Artificial intelligence model for Parkinson disease detection using machine learning algorithms. *Biomedical Materials & Devices*, Springer.
10. Wang, P., Zhao, L., Wang, T., Mei, W., Li, J., et al., 2023, Comparison of half-effective concentration of propofol in patients with Parkinson's disease and non-Parkinson's disease. *Interventions in Aging*, Taylor & Francis.
11. Craig, L., 2015, *Parkinson's Disease*. Greenhaven Publishing LLC. http://books.google.ie/books?id=D1xmDwAAQBAJ
12. Parashos, S., & Wichmann, R., 2012, *Navigating life with Parkinson disease*. Oxford University Press. http://books.google.ie/books?id=vZmtgtahjzoC
13. Pahwa, R., & Lyons, K. E., 2013, *Handbook of Parkinson's disease*. CRC Press. http://books.google.ie/books?id=prXNBQAAQBAJ
14. Md, F. F. A. Q. R., 2014, *A synopsis of Parkinson's disease*. BoD – Books on Demand. http://books.google.ie/books?id=OiShDwAAQBAJ
15. Ali, N., 2013, *Understanding Parkinson's disease*. Rowman & Littlefield. http://books.google.ie/books?id=8VsrAQAAQBAJ
16. Factor, S. A., & Weiner, W., 2007, *Parkinson's disease*. Demos Medical Publishing. http://books.google.ie/books?id=zUp54Dm-Y7MC
17. Chaudhuri, K. R., & Fung, V. S. C., 2016, *Fast facts: Parkinson's disease*. Karger Medical and Scientific Publishers. http://books.google.ie/books?id=yV9uDwAAQBAJ
18. Henchcliffe, C., 2008, Rasagiline in treatment of Parkinson's disease. *Neuropsychiatric Disease and Treatment*, 11. https://doi.org/10.2147/ndt.s464
19. Kwon, K., Baek, K., & Park, C., 2012, Prediction and characterization of Arabidopsis thaliana homologs of Parkinson disease-associated genes based on sequence similarities and molecular modeling. *Parkinsonism & Related Disorders*, 18(Supplement), S190. https://doi.org/10.1016/s1353-8020(11)70821-5
20. Warden, M. N., Searles Nielsen, S., Camacho-Soto, A., Garnett, R., & Racette, B. A., 2021, A comparison of prediction approaches for identifying prodromal Parkinson disease. *PLOS ONE*, 16(8), e0256592. https://doi.org/10.1371/journal.pone.0256592
21. Duncan, R. P., Leddy, A. L., Cavanaugh, J. T., Dibble, L. E., Ellis, T. D., Ford, M. P., Foreman, K. B., & Earhart, G. M., 2012, Accuracy of fall prediction in Parkinson disease: Six-month and 12-month prospective analyses. *Parkinson's Disease*, 2012, 1–7. https://doi.org/10.1155/2012/237673
22. Sabour, S., 2018, Prediction of falls in subjects suffering from Parkinson disease, multiple sclerosis, and stroke: Methodologic issues. *Archives of Physical Medicine and Rehabilitation*, 99(8), 1688. https://doi.org/10.1016/j.apmr.2018.02.018
23. Parkinson Disease, 2002, *Journal of Neurosurgery*, 96(3). https://doi.org/10.3171/jns.2002.96.3.0641
24. Karan, B., Sahu, S. S., & Mahto, K., 2020, Parkinson disease prediction using intrinsic mode function-based features from speech signal. *Biocybernetics and Biomedical Engineering*, 40(1), 249–264. https://doi.org/10.1016/j.bbe.2019.05.005
25. Reddy, Y. S., Prasad, K. S., Ashwin, D., & Shaik, D. S., 2023, Prediction of Parkinson disease using machine learning. *International Journal of Scientific Research in Engineering and Management*, 7(1). https://doi.org/10.55041/ijsrem17597
26. K, G., 2019, Prediction of Parkinson disease by best accuracy using supervised classification machine learning approach. *International Journal for Research in Applied Science and Engineering Technology*, 7(3), 1536–1540. https://doi.org/10.22214/ijraset.2019.3284
27. P, J., 2020, Dimensionality reduction method for prediction of Parkinson disease using speech data. *International Journal for Research in Applied Science and Engineering Technology*, 8(5), 2319–2325. https://doi.org/10.22214/ijraset.2020.5379
28. Parkinson's disease prediction using modified Gauss-Newton method in feed-forward neural network, 2019, *International Journal of Innovative Technology and Exploring Engineering*, 9(1), 4986–4993. https://doi.org/10.35940/ijitee.k1559.119119
29. Zeng, T., Yang, T., Liu, P., & Zhu, D., 2022, Parkinson disease prediction using machine learning–based features from speech signal. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4201020
30. Parkinson disease prediction using different classification algorithms, 2023, *International Research Journal of Modernization in Engineering Technology and Science*, September 1. https://doi.org/10.56726/irjmets44173

31. Parkinson disease prediction, 2023, *International Research Journal of Modernization in Engineering Technology and Science*, July 19. https://doi.org/10.56726/irjmets38520
32. Khan, F., Amanullah, S.I. & Selvarajan, S. Linear regressive weighted Gaussian kernel liquid neural network for brain tumor disease prediction using time series data. Scientific Reports, 15, 5912 (2025). https://doi.org/10.1038/s41598-025-89249-w
33. Kumar, G.S., Suganya, E., Sountharrajan, S. et al. SRADHO: Statistical Reduction Approach with Deep Hyper Optimization for Disease Classification Using Artificial Intelligence. Scientific Reports, 15, 1245 (2025). https://doi.org/10.1038/s41598-024-82838-1
34. Manogaran, N., Panabakam, N., Selvaraj, D. et al. An efficient patient's response predicting system using multi-scale dilated ensemble network framework with optimization strategy. Scientific Reports, 15, 15713 (2025). https://doi.org/10.1038/s41598-025-00401-y
35. Kantipudi, M.V.V.P., Kumar, N.S.P., Aluvalu, R. et al. An improved GBSO-TAENN-based EEG signal classification model for epileptic seizure detection. Scientific Reports, 14, 843 (2024). https://doi.org/10.1038/s41598-024-51337-8
36. Karthiga, M., Suganya, E., Sountharrajan, S. et al. EEG based smart emotion recognition using meta heuristic optimization and hybrid deep learning techniques. Scientific Reports, 14, 30251 (2024). https://doi.org/10.1038/s41598-024-80448-5
37. Hariprasath Manoharan et al. A machine learning algorithm for classification of mental tasks. Computers and Electrical Engineering (2022). https://doi.org/10.1016/j.compeleceng.2022.107785
38. Shahanaz Ayub et al. LSTM-Based RNN Framework to Remove Motion Artifacts in Dynamic Multi-Contrast MR Images with Registration Model. Wireless Communications and Mobile Computing (2022). https://doi.org/10.1155/2022/5906877