Temporal Patterns of EEG Connectivity Unveil Parkinson's Disease Progression: Insights from Machine Learning Analysis

Caroline L. Alves

caroline.lourenco.alves@gmail.com

Aschaffenburg University of Applied Sciences

Loriz Francisco Sallum

Universidade de São Paulo

Francisco Aparecido Rodrigues

Universidade de São Paulo

Thaise G. L. de O. Toutain

Federal University of Bahia

Patricia Maria de Carvalho Aguiar

Hospital Israelita Albert Einstein

Michael Moeckel

Aschaffenburg University of Applied Sciences

Additional Declarations: No competing interests reported.

Temporal patterns of EEG connectivity unveil Parkinson's disease progression: Insights from Machine Learning analysis

Caroline L. Alves[1,*]**, Loriz Francisco Sallum**[2]**, Francisco Aparecido Rodrigues**[2]**, Thaise G. L. de O. Toutain**[3]**, Patrícia Maria de Carvalho Aguiar**[4,5]**, and Michael Moeckel**[1]

[1]Laboratory for Hybrid Modeling, Aschaffenburg University of Applied Sciences, Aschaffenburg, Germany
[2]Institute of Mathematical and Computer Sciences, University of São Paulo, São Paulo, Brazil
[3]Health Sciences Institute, Federal University of Bahia, Bahia, Brazil
[4]Hospital Israelita Albert Einstein, São Paulo, Brazil
[5]Department of Neurology and Neurosurgery, Federal University of São Paulo, São Paulo, Brazil
[*]Caroline.LourencoAlves@th-ab.de

ABSTRACT

Objectives: Parkinson's disease is a multifactorial neurodegenerative disorder whose progression remains complex despite extensive research. Our study presents an innovative approach to understanding PD progression through detailed analysis of electroencephalography signals. By segmenting patients based on disease duration, we uncover unique neural connectivity patterns corresponding to different duration of PD development. **Methods:** Employing advanced machine learning techniques, our methodology achieves exceptional accuracy rates in binary classification tasks compared to prior literature. Integration of Shapley Additive Explanations values enhances model interpretability, revealing critical brain regions and connectivity patterns implicated in PD pathophysiology. **Results:** Coherence emerges as a crucial metric for capturing synchronized signal behaviors, aiding in discriminating PD patients from controls. Further, our analysis suggests a continuum of neural connectivity patterns across disease duration, with early-stage PD resembling healthy brain function and advanced duration exhibiting distinct features indicative of disease progression. **Conclusions:** These findings deepen our understanding of PD pathogenesis, laying the groundwork for personalized diagnostic and therapeutic approaches tailored to different disease duration. Our study contributes significant insights into the complex interplay between neural dynamics, disease progression, and age-related changes in PD, offering potential for future research and clinical applications.

Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder marked by dopaminergic neuron loss in the substantia nigra pars compacta, leading to motor symptoms like bradykinesia, tremors, rigidity, and postural instability[1–3]. Non-motor symptoms, such as cognitive impairment, depression, and autonomic dysfunction, also significantly affect patients' quality of life[4,5]. While PD's exact etiology remains elusive, evidence suggests a complex interaction of genetic susceptibility, environmental factors, and aging[6,7].

Various studies employ machine learning (ML) and deep learning algorithms to differentiate between individuals with PD and healthy control (HC), as evidenced in Table 1. The predominant dataset utilized across these studies is the San Diego dataset, which we used in our analysis. PD presents significant challenges across its various duration and stage of the condition, highlighting a critical gap in current diagnostic methodologies[8–10]. While considerable efforts have been directed toward discerning patterns and markers indicative of PD, the complexities inherent in the disease's progression still need to be addressed[11].

Our methodology extends beyond conventional approaches by delving into the nuances of PD progression through a refined segmentation of the patient cohort. Unlike previous studies that commonly compare PD patients with medication, without medication, and controls, we adopt a novel stratification strategy based on disease duration. Specifically, we categorize individuals into two distinct groups: those with a disease duration of 1-3 years and those with a 6-12 years duration. This tailored segmentation offers a more nuanced perspective on PD progression, aiming to capture potential variations in EEG signals corresponding to different duration and the stage of the disease trajectory. Building upon techniques demonstrated to be effective in preprocessing EEG time series data[12], we employ various pairwise statistical metrics to construct connectivity matrices. Notably, we explore the efficacy of these metrics in binary classification tasks to discern PD patients from controls, alongside their utility in multiclass classification to differentiate between subgroups within the PD cohort based on disease

Authors	Correlation metrics	ML method	AUC	Accuracy	Recall	Precision
[19]	Phase locking value	Deep learning (Multiscale CNN)	88.7	88.7	86.7	-
[20]	-	Least Square SVM	-	97.65	96.67	98.76
[21]	-	Deep learning (2D-CNN)	-	99.46	99.46	99.48
[22]	-	KNN	-	99.89	99.87	-
[23]	-	Deep learning (CNN)	98.96	97.90	97.87	98.0

Table 1. Overview of PD binary classification research that uses the identical publicly accessible PD dataset as this work.

duration. Furthermore, akin to our prior works[12–16], we enhance the interpretability of our machine learning models using Shapley Additive Explanations (SHAP) values[17]. SHAP values provide valuable insights into the most influential features within our models, elucidating critical brain areas and connectivity patterns. This methodology, proven effective in prior research endeavors[13, 14, 18], facilitates a deeper understanding of the underlying neural dynamics in PD, complementing our multiclass classification framework. By integrating these advanced techniques, our study contributes to refining diagnostic approaches. It lays the groundwork for personalized therapeutic interventions tailored to the distinct needs of PD patients at different duration of disease progression.

Results

0.1 Binary classification

Considering the AUC metric, the performance of each connectivity metric is illustrated in Figure 1. The Sync measure has the best performance for the test set, equal to 0.980 for the mean AUC, 0.981 for precision, 0.979 for recall, and 0.980 for accuracy. The Sync. measure between signals i and j, denoted as $\text{Sync}[i, j]$, is computed using the average coherence across all frequency points, as defined by Equation 1:

$$\text{Sync}[i, j] = \frac{1}{N} \sum_{k=1}^{N} \frac{|P_{xy}(f_k)|^2}{P_{xx}(f_k) \cdot P_{yy}(f_k)} \tag{1}$$

where:

- $Sync[i, j]$ is the synchronization between signals i and j,

- N is the number of frequency points,

- $P_{xy}(f_k)$ is the cross-spectral density between signals i and j at frequency f_k,

- $P_{xx}(f_k)$ is the auto-spectral density of signal i at frequency f_k,

- $P_{yy}(f_k)$ is the auto-spectral density of signal j at frequency f_k,

- f_k represents the sampling frequency, assumed to be 50 Hz.

Then, we tested all ML algorithms that resulted in Figure 2; the best classifiers were CNN and MLP. CNN performance for the test set was equal to 0.999 for the mean AUC, 0.990 for precision, 0.990 for recall, and 0.990 for accuracy. MLP performance for the test set was equal to 0.990 for the AUC, 0.989 for the precision, 0.991 for the recall, and 0.990 for the accuracy. Considering the performance and the computational cost, we selected the MLP classifier for use in the next section.

To determine the ideal number of features crucial for achieving optimal performance, we utilized a Recursive Feature Elimination (RFE) analysis, illustrated in Figure 3. This method, widely employed in predictive modeling in the medical data field[24–27], systematically removes less impactful features step by step to evaluate their effect on the model's performance. This iterative approach allows for the identification of the most relevant features. As depicted in Figure 3, superior accuracy is achieved with 136 features. Hence, it becomes evident that employing the entire comprehensive feature set is unnecessary for optimal efficacy.
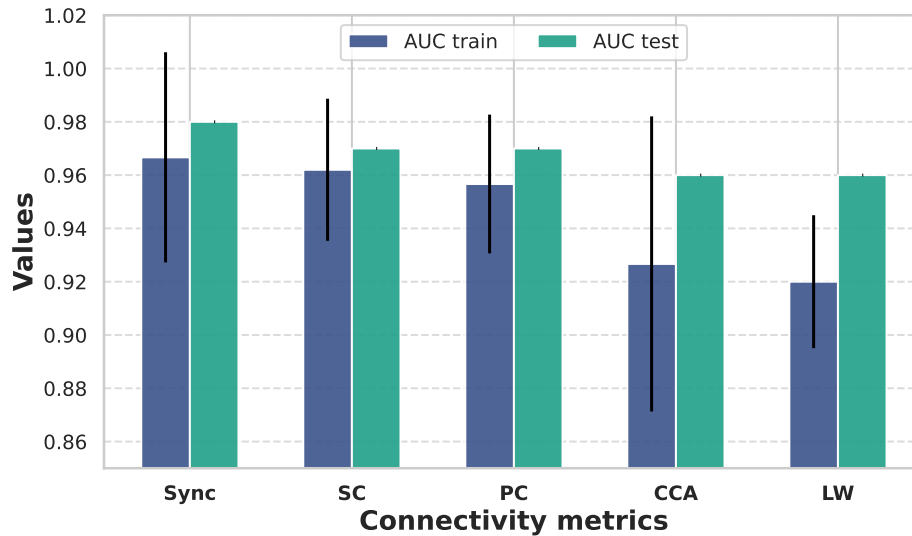
Figure 1. Selection of optimal correlation metrics with SVM. In blue is the AUC of the train set; in green is the AUC of the test set. The best performance was the Sync.

0.2 Multiclass classification

Regarding the temporal changes due to PD, the general performance for the MLP in a multiclass classifier for the test set was equal to 0.983 for the mean AUC, 0.989 for precision, 0.977 for recall, and 0.983 for accuracy. Figure 4 displays the ROC curve (Figure 4-(a)), the learning curve (Figure 4-(b)), and the confusion matrix (4-(c)), respectively. From Figures 4-(a) and (b), the confusion matrix and ROC curve, respectively, showed that the group most difficult to distinguish was the PD 6-12 (in the confusion matrix, the PD 6-12 group has some mistaken classification with the group HC).

The visual representation of the learning curve illustrates the impact of varying the number of training instances on the model's predictive accuracy[28]. Figure 4-(c) shows that all the data was required to converge the model.

The methodology employed for SHAP value analysis yielded results depicted in Figure 5, illustrating the primary connections found across the groups PD 1-3 and PD 6-12. Due to its high computational cost, we utilized 300 features for the SHAP value methodology. This number was chosen based on the results of recursive feature elimination, which identified 136 features as relevant. Therefore, selecting 300 features ensures adequate coverage for the analysis. The best connection found for the PD 1-3 group according to SHAP value methodology in this order of importance were lower synchronization values to FP2 and FC5 and lower synchronization for O2 and T7. For the PD 6-12 group, the best connections found in this order of importance were higher synchronization of F8 and AF3 and higher synchronization of FC6 and Fp1.

Further, SHAP values matrices for each class are concatenated into a single array, with rows representing samples and columns representing features. A cluster map, depicted in Figure 6, is generated through hierarchical Clustering along both rows (samples) and columns (features). Hierarchical Clustering, a method of cluster analysis, builds a cluster hierarchy by merging the closest pairs of clusters iteratively until all points belong to one cluster, typically using metrics like Euclidean or correlation distance. The resulting hierarchy can be visualized using a dendrogram. In Figure 6, clusters in the cluster map group together rows and columns with similar SHAP value patterns, with heatmap colors representing SHAP value magnitudes (cooler color for negative, warmer color for positive, and white for values near zero).

Our analysis revealed interesting clustering behavior wherein the SHAP values for PD duratuin 1-3 exhibited proximity to those of the HC class, forming a distinct cluster. Furthermore, this cluster was observed to be connected with the SHAP values corresponding to PD duratuin 6-12, indicating a potential continuum or progression in the underlying features contributing to the model predictions across different disease duration.

Further, as a complement to the temporal changes results due to PD, the differentiates changes evaluated UPDRS score, the general performance for the MLP in a multiclass classifier for the test set was equal to 0.969 for the mean AUC, 0.963 for precision, 0.955 for recall, and 0.967 for accuracy. Figure 7 displays the ROC curve (Figure 7-(a)), the learning curve (Figure 7-(b)), and the confusion matrix (7-(c)), respectively. From Figures 7-(a) and (b), the confusion matrix and ROC curve, respectively, showed that the group most difficult to distinguish was the UPDRS1 (in the confusion matrix, the UPDRS1 group has some mistaken classification with the group HC). This difficulty likely stems from the lower UPDRS scores associated with the UPDRS1 group, implying an earlier stage of PD progression.

To assess the impact of correlation metrics on the results of the SHAP values methodology, we employed the MLP in
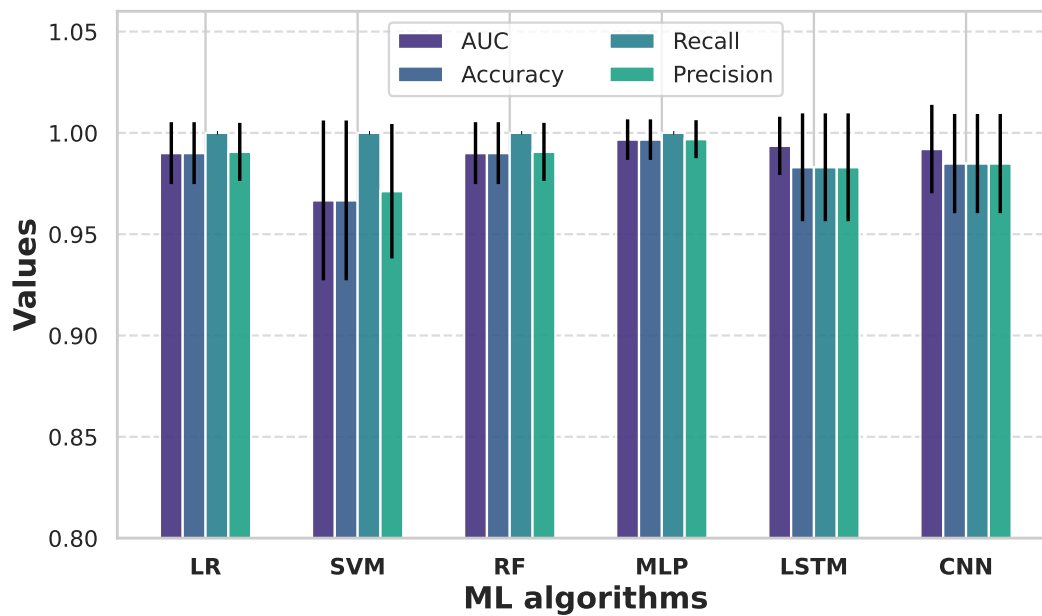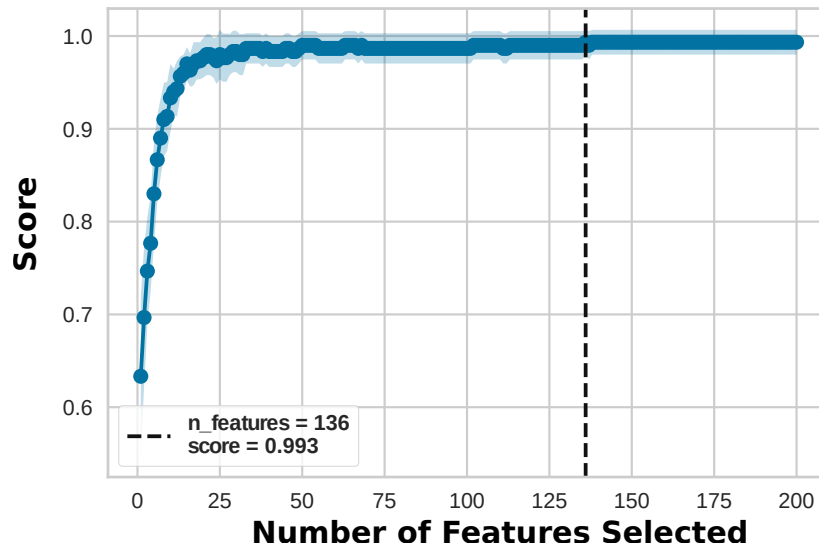
(a) Train set results
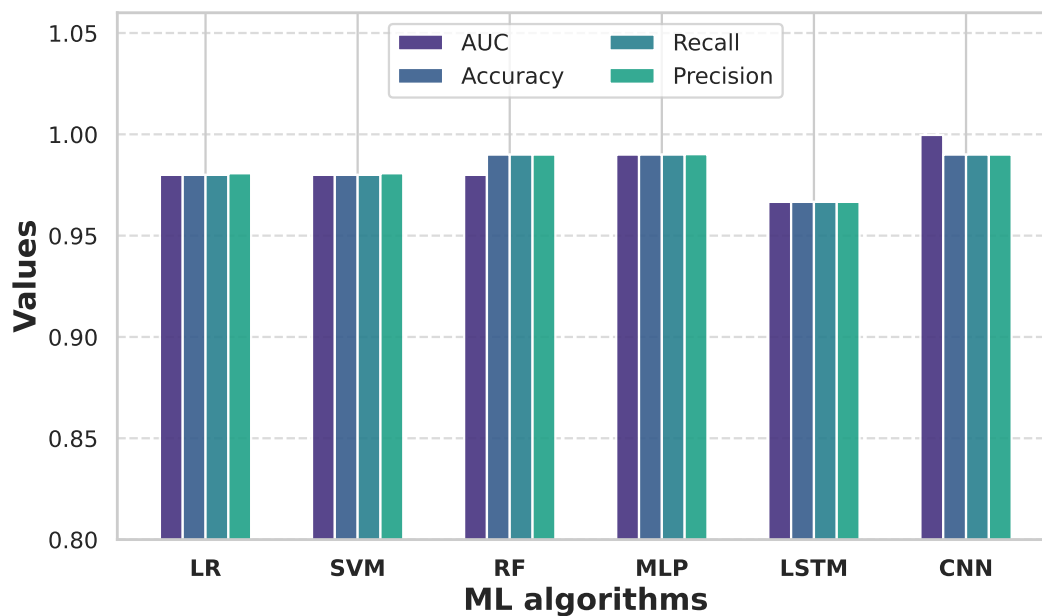
(b) Test set results

Figure 2. Results for method selection: best performance by MLP. In (a) the results refer to the train set with the error bar due to the 10-fold stratified cross-validation and (b) the results of the test set.

Figure 3. RFE. The optimal performance is attained with 136 features, so utilizing the entire comprehensive feature set is unnecessary.

conjunction with SHAP values to identify the five most significant connections for each correlation metric corresponding to the PD1-3 and PD6-12 classes, as illustrated in Figures 2 and 3, respectively.

## Discussion

0.3 Binary classification

The Sync was the best metric in a binary classification approach. As defined by[29], coherence emphasizes the synchronization or coupling between signals, which proved particularly relevant in our analysis of PD disease. This emphasis on the interrelationship and synchronization of data streams aligns well with the underlying dynamics of PD, where disruptions in neural signaling and coordination are vital features[30–33]. Therefore, we posit that coherence, with its focus on capturing the synchronized behavior of signals, offers unique insights and discriminatory power for distinguishing between PD from HC our machine learning framework.

Further, the best machine learning algorithms were the MLP classifier and CNN, which tested the AUC metric of 0.990 and 0.999, respectively. These performances are higher compared to the literature, according to Table 1, which studies using the same dataset of this study also in a binary approach comparing PD patients from HC.

0.4 Multiclass classification

Regarding the temporal changes due to PD, using the MLP, the obtained accuracies of 0.933 for PD 6-12 and 1.000 for both PD 1-3 (depicted in Figure 4-(c)) and HC suggest a promising discrimination capability of the machine learning model in distinguishing between different groups based on connectivity matrices. The perfect classification for PD 1-3 and HC groups could indicate distinct patterns of brain connectivity in these cohorts, reflecting the early duration of PD progression and healthy brain function, respectively. The slightly lower accuracy for PD 6-12 implies greater heterogeneity or subtler differences in connectivity patterns that become more complex or less distinct as the disease progresses among patients with longer disease duration despite medication intake. Nonetheless, the model can still distinguish between the two groups with high metric performance.

After employing the SHAP value methodology, we identified the most influential connections for the PD 1-3 group. Notably, lower synchronization values in regions such as FP2 and FC5 were found. The frontopolar cortex (FP2) is implicated in various higher-order cognitive processes, including decision-making, executive function, and social cognition[34,35]. On the other hand, the left frontal cortex (FC5) plays a crucial role in motor planning and execution, as well as language processing[36,37]. Thus,
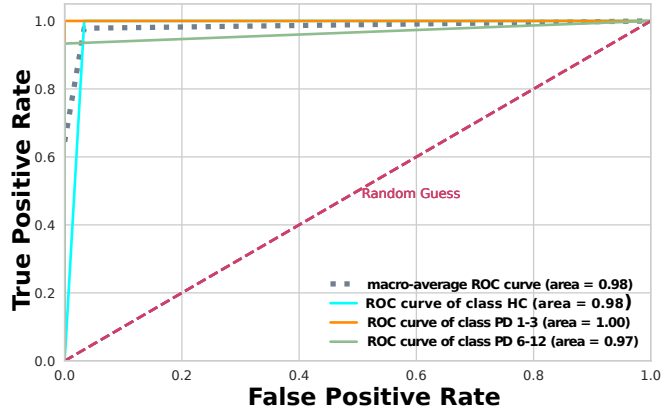
(a) ROC curve
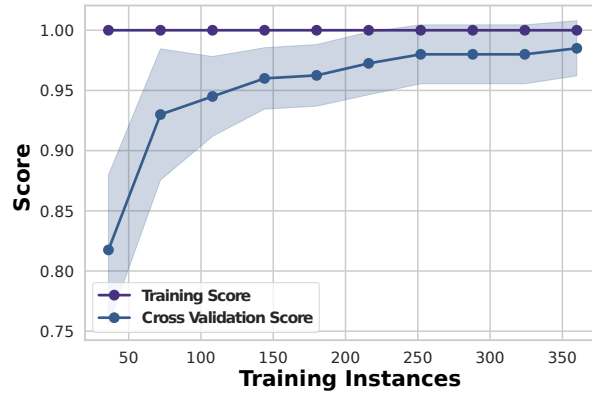
(b) Learning curve
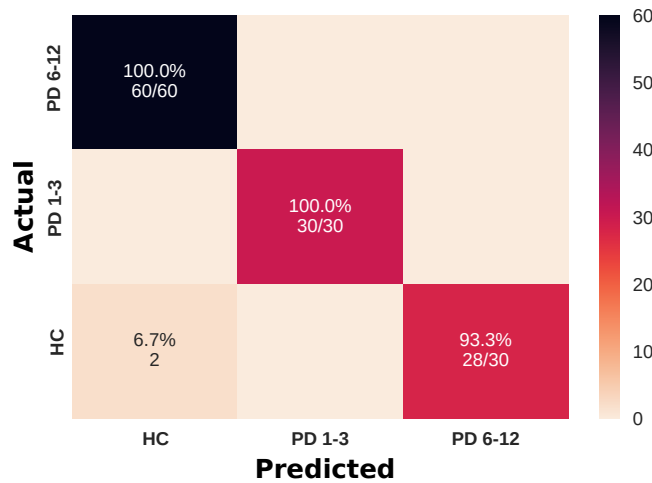
(c) Confusion matrix

Figure 4. The test sample ML results from connectivity matrices. (a) ROC curve for each ML algorithm. The dashed pink line represents the random choice classifier, the purple line the micro-average ROC curve, the gray line the macro-average ROC curve, the turquoise line the ROC curve referring to the HC class, the orange line the ROC curve referring to the PD 1-3 class, and the green line the ROC curve referring to the PD 6-12 class. (b) The learning curve for the training accuracy (purple) and test accuracy (blue). (c) Confusion Matrices depicting the performance of various ML algorithms on a test sample, the diagonal elements represent true positives (TP) values, showcasing each algorithm's accuracy in correctly identifying positive instances.
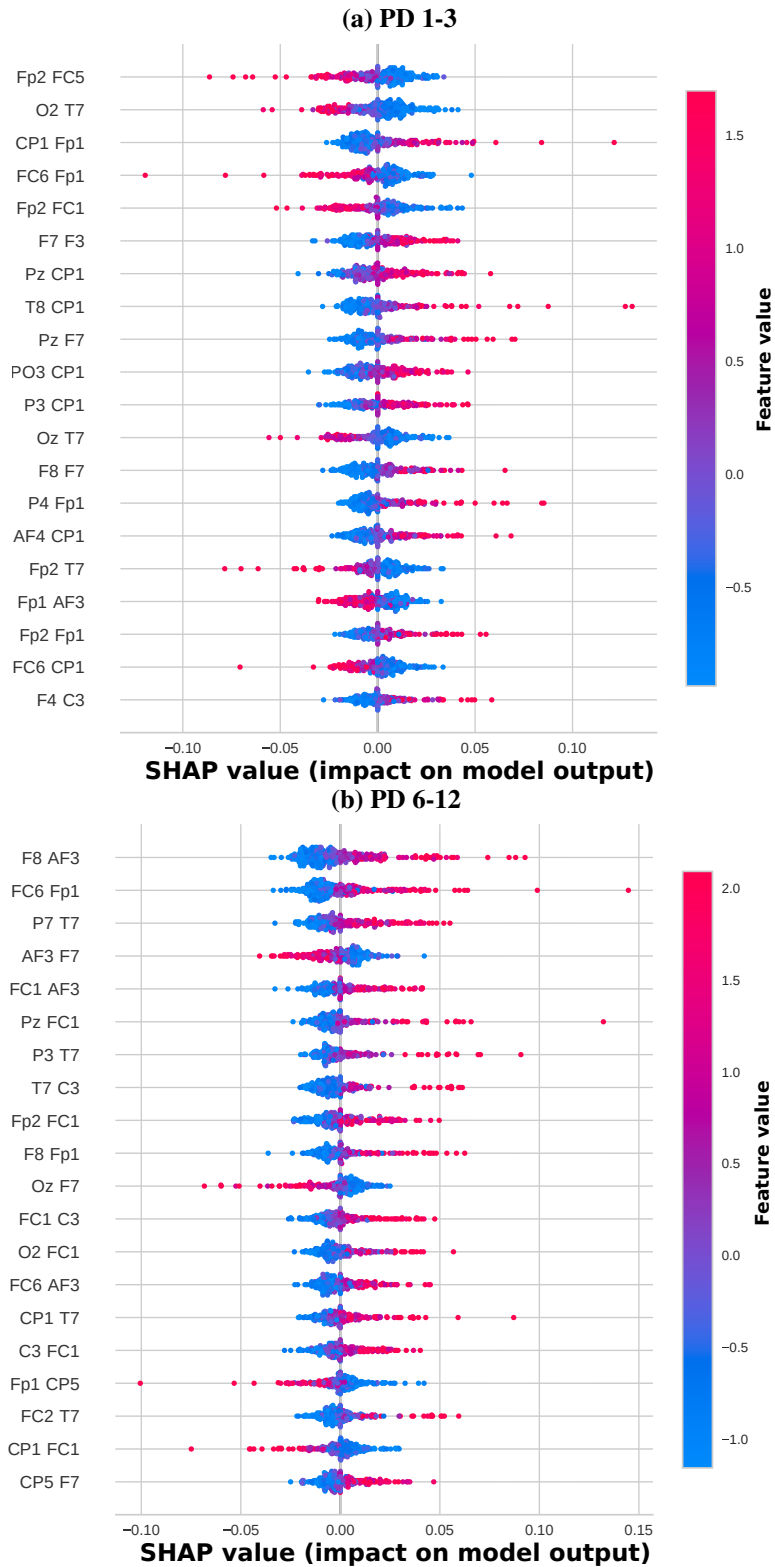
Figure 5. Feature importance ranking using the SHAP values methodology for the MLP classifier with brain EEG channels in descending order. (a) Feature importance ranking regarding PD 1-3 class. (b) Feature importance ranking regarding PD 6-12 class.
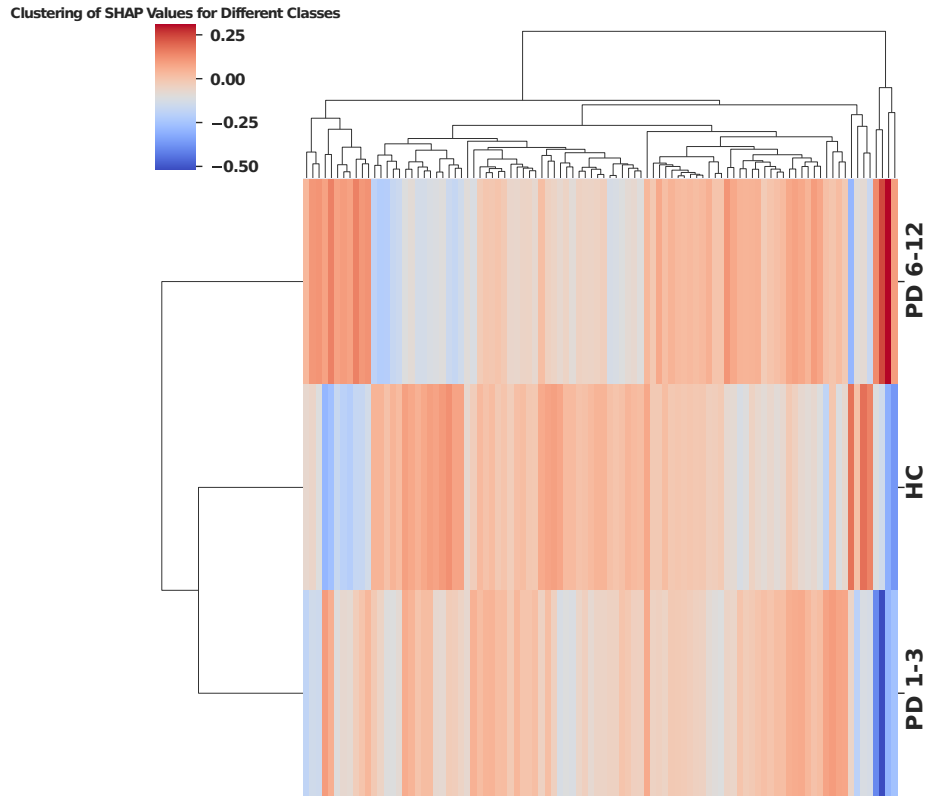
Figure 6. Cluster map generated through hierarchical clustering of SHAP values. The dendrogram of the superior part corresponds to the hierarchy clustering for each of the 300 SHAP value methodologies. The heatmap colors represent SHAP value magnitudes (cooler color for negative, warmer color for positive, and white for values near zero). The left side of the plot corresponds to the resulting clustering regarding the dendrogram SHAP values features wherein the groups PD 1-3, PD 6-12, and HC.

Connections	Sync.	SC	PC	CCA	LW
Fp2-FC5					
O2-T7					
Cp1-Fp1					
Fc6-Fp1					
Fp2-Fc1					
T7-Af3					
Cp1-F7					
F3-Fp1					
F8-Fp1					
Fc1-Fp1					
Fc5-Fp1					
T7- F7					
Fp2-Fp1					
Pz-Af3					
Fp2-F7					
Fc2-Fp1					

Table 2. Table depicting the most significant connections identified through the SHAP value methodology and MLP classifier for each correlation metric and the class PD 1-3. Connection importance is denoted by the intensity of red shading. In bold are the main connections that appear in both classes, PD1-3 and PD6-12.
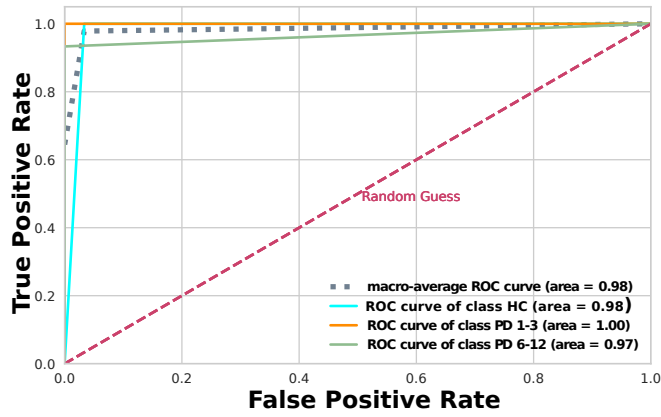
(a) ROC curve
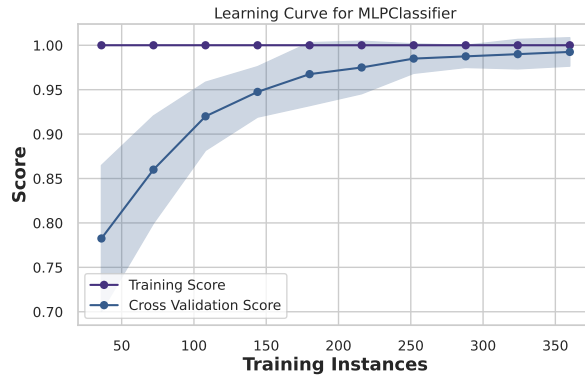
(b) Learning curve
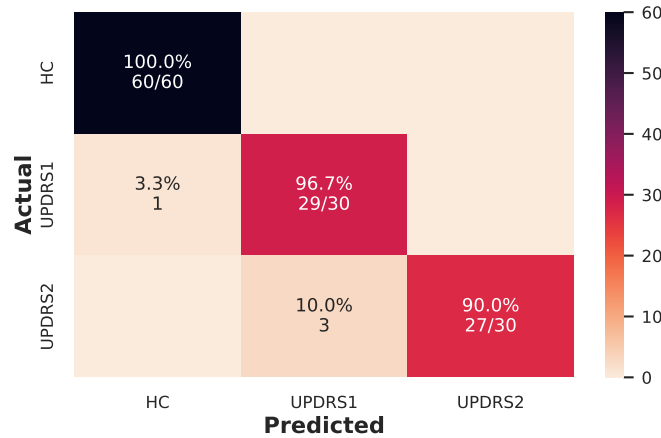
(c) Confusion matrix

Figure 7. The test sample ML results from connectivity matrices. (a) ROC curve for each ML algorithm. The dashed pink line represents the random choice classifier, the purple line the micro-average ROC curve, the gray line the macro-average ROC curve, the turquoise line the ROC curve referring to the HC class, the orange line the ROC curve referring to the UPDS1 class, and the green line the ROC curve referring to the UPDS2 class. (b) The learning curve for the training accuracy (purple) and test accuracy (blue). (c) Confusion Matrices depicting the performance of various ML algorithms on a test sample, the diagonal elements represent true positives (TP) values, showcasing each algorithm's accuracy in correctly identifying positive instances.

| Connections	Sync.	SC	PC	CCA	LW
F8-Af3					
Fc6-Fp1					
P7-T7					
Af3-F7					
Fc1-Af3					
Cp5-Af3					
Cp5-Fp1					
Fp2-F7					
Cp2-Fp1					
Fp1-Af3					
Po3-Fp1					
P7-Af3					
Cp6-Fp1					

Table 3. Table depicting the most significant connections identified through the SHAP value methodology and MLP classifier for each correlation metric and the class PD 6-12. Connection importance is denoted by the intensity of red shading. In bold are the main connections that appear in both classes, PD1-3 and PD6-12.

alterations in synchronization patterns within these regions may signify underlying deficits in both cognitive processing and motor coordination[38, 39]. Additionally, decreased synchronization in regions like O2 and T7 suggests possible impairment in sensory processing and attention, which are commonly affected in the early duration of PD[40]. Conversely, for the PD 6-12 group, heightened synchronization in regions like F8 and AF3 emerged as significant contributors, possibly reflecting potential compensatory mechanisms or adaptations occurring in response to disease progression and age-related changes in neural circuitry[41].

Furthermore, our investigation unveiled a significant association between the FC6 and FP1 electrodes in both PD 1-3 and PD 6-12 cohorts, showcasing diminished synchrony in the former and amplified synchrony in the latter. This observed increase in synchronization within frontal brain areas regarding PD is consistent with findings from other studies[42, 43].

These findings underscore the complex interplay between neural connectivity patterns, disease progression, and age-related changes in PD. These primary connections are summarized in Figure 8.

Our analysis revealed intriguing clustering behavior that sheds light on the underlying patterns of feature importance across different disease duration. Specifically, the SHAP values associated with PD 1-3 demonstrated a notable proximity to those of the HC class, forming a distinct cluster. This clustering suggests that the features contributing to the ML model's predictions for early-duration PD share similarities with those of healthy brain function. Moreover, we observed a connection between this cluster and the SHAP values corresponding to PD duration 6-12. This linkage implies a potential continuum or progression in the underlying neural connectivity patterns captured by the model, spanning from the early to later duration of PD. The hierarchical clustering highlights the distinctiveness of brain connectivity patterns in early-duration PD and healthy brains. It suggests a gradual shift towards features more characteristic of advanced PD as the disease progresses.

Regarding the results of PD stages, using the MLP in a multiclass classifier for the test set was equal to 0.969 for the mean AUC, 0.963 for precision, 0.955 for recall, and 0.967 for accuracy. Therefore, it was possible to distinguish PD patients from the temporal and stage degree of the disease. Further, Figure 7-(c) demonstrates that the model achieved convergence without necessitating the entirety of the dataset. This suggests that distinguishing between PD patients and HC subjects based on disease stage was relatively straightforward compared to predicting disease duration. Consequently, to accurately model disease duration, it was imperative to utilize the entire dataset for convergence

Our analysis regarding correlation metrics' influence on the SHAP values results, as illustrated in Figures 2 and 3, uncovered consistent patterns across various correlation metrics, notably in characteristic regions such as Fp1. However, distinct correlations emerged as significant within specific groups of metrics. Notably, linear correlation metrics like SC and PC clustered together, while CCA and LW formed another cohesive group. Additionally, Sync. stood out as an independent correlation metric. Importantly, these findings represent novel results that have not been evaluated in the existing literature to the best of our knowledge.

Further, the connections FC6-Fp1 and Fp2-F7 appear for both classes (highlighted in bold in Tables 2 and 3).
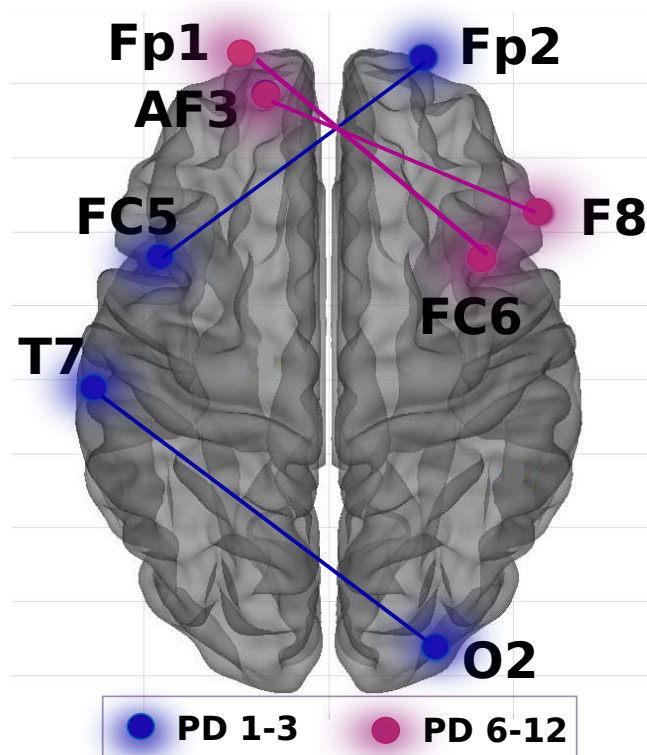
Figure 8. The most important connections found. Two-dimensional schematic (ventral-axis), where the most critical connection for PD 1-3 and PD 6-12 are highlighted in blue and pink, respectively. The brain plot was developed by the Braph tool[44], and each region was plotted according to[45,46].

Methods

Figure 9 summarizes the methodology used in the present work. The python code with the methodology used in this work is available at: `https://github.com/Carol180619/Paper-temporal-Parkinson.git`.

0.5 Data and EEG preprocessing

The study utilized the SanDiego dataset[47], comprising resting-state EEG recordings from both medicated and non-medicated PD patients alongside an HC group of individuals. Here, we considered only the non-medicated PD EEG time series. Demographic details such as age, gender, handedness, cognitive assessments, the mini mental state examination (MMSE), the North American adult reading test (NAART), and the Unified Parkinson's disease rating scale (UPDRS) were included[22, 48, 49]. PD severity varied based on Hoehn and Yahr scales, with an average disease duration of 4.5–3.5 years. Rigorous data preprocessing involved band-pass filtering (0.5–50 Hz) using an FIR filter and ICA with the infomax algorithm to enhance signal quality and mitigate interference. Artifacts, particularly from eye blinking, were identified and removed.

0.6 Binary and multiclass classification

After preprocessing EEG data using effective techniques demonstrated in our previous endeavors[12], we segmented it into 10-second windows and built connectivity matrices with Pearson Correlation (PC)[50], Spearman Correlation (SC)[51], Sparse Canonical Correlation Analysis (CCA)[52], and Ledoit-Wolf shrinkage (LW)[53]. We emphasize coherence (Sync)[29, 54] for its unique ability to measure synchronization in EEG signals, enhances the dataset analysis, providing a nuanced exploration of EEG signals that were previously unexplored in our prior work[12–16].

In our study, we adopted a slicing window approach to enhance our dataset's size and facilitate comprehensive analysis of EEG signals. This technique involves partitioning EEG signals into smaller time windows[55], typically lasting 10 seconds[56–58], which effectively increases the number of instances available for analysis. However, to address potential issues stemming from this augmentation, we integrated a stratified K-fold cross-validation methodology coupled with shuffling techniques to fortify our analyses' robustness and mitigate data dependencies arising from the slicing window method[59–61].

By employing a stratified K-fold approach alongside shuffling, we ensured that each fold of cross-validation retained the original class distribution (PD and HC) while simultaneously shuffling data within each fold[62]. This meticulous shuffling process minimizes the risk of data leakage, where inadvertent transfer of information from the training to testing sets could compromise model evaluations[63]. This precautionary measure upholds the integrity of the evaluation process by rigorously assessing the model's performance on unseen data, thereby preserving the generalizability of our findings and instilling confidence in our results[64].

This enhances dataset analysis, particularly in PD cases, shedding light on neural dynamics. Our approach differentiates patients and HC (with 16 individuals) using a multi-classifier, segmenting the PD EEG dataset based on the disease duration into two groups: 1-3 years (PD 1-3), with nine individuals, and 6-12 years (PD 6-12), with six individuals, mirroring the age distribution in Figure 10.

Furthermore, in order to assess not only the temporal changes attributable to Parkinson's Disease (PD) but also its stage, we stratified the PD EEG dataset based on the Unified Parkinson's Disease Rating Scale (UPDRS). The UPDRS is a widely used clinical rating scale that evaluates the severity of PD symptoms and functional impairment. Detailed information on UPDRS and the present dataset can be found in[49]. We divided the dataset into two groups based on the UPDRS scores: 20-40 (UPDRS1), comprising eight individuals, and 43-75 (UPDRS2), comprising six individuals. This segmentation mirrors the distribution of UPDRS scores illustrated in Figure 11. UPDRS1 typically represents milder stages of PD, while UPDRS2 indicates more advanced stages, reflecting the progression of the disease severity among the participants.

The current work builds upon our previous research endeavors, which have established a foundation for employing machine learning techniques in the classification of PD and HC groups[12–16, 65]. Initially, we utilize the support vector machine (SVM) to select connectivity metrics, leveraging its lower computational cost and effectiveness in binary classification[12]. Afterwards, various machine learning algorithms, including Logistic Regression (LR), Random Forest (RF), Multilayer Perceptron (MLP), Long Short-Term Memory neural networks (LSTM), and Convolutional Neural Networks (CNN), are tested with the selected connectivity metric, as detailed in the subsection 0.1. Noteworthy, leveraging 10 seconds slicing windows data augmentations, a set of 200 connectivity matrices from HC and 200 from individuals with PD, comprising 100 matrices from each subgroup of PD, were employed as inputs for the machine learning algorithm.

Standardization and data-splitting techniques were then employed to prepare the data for model training and evaluation. Features were standardized to ensure consistent scaling and resilience against outliers, while stratified cross-validation with shuffling is utilized to ensure unbiased model evaluation[66–76].

Hyperparameter optimization techniques, such as grid search and random search, are further employed to fine-tune the machine learning algorithms for optimal performance, with evaluation metrics including accuracy, precision, recall, Receiver Operating Characteristic (ROC) curves, and Area Under the ROC Curve (AUC)[12–15, 77–79, 79–94]. Additionally, the SHAP value
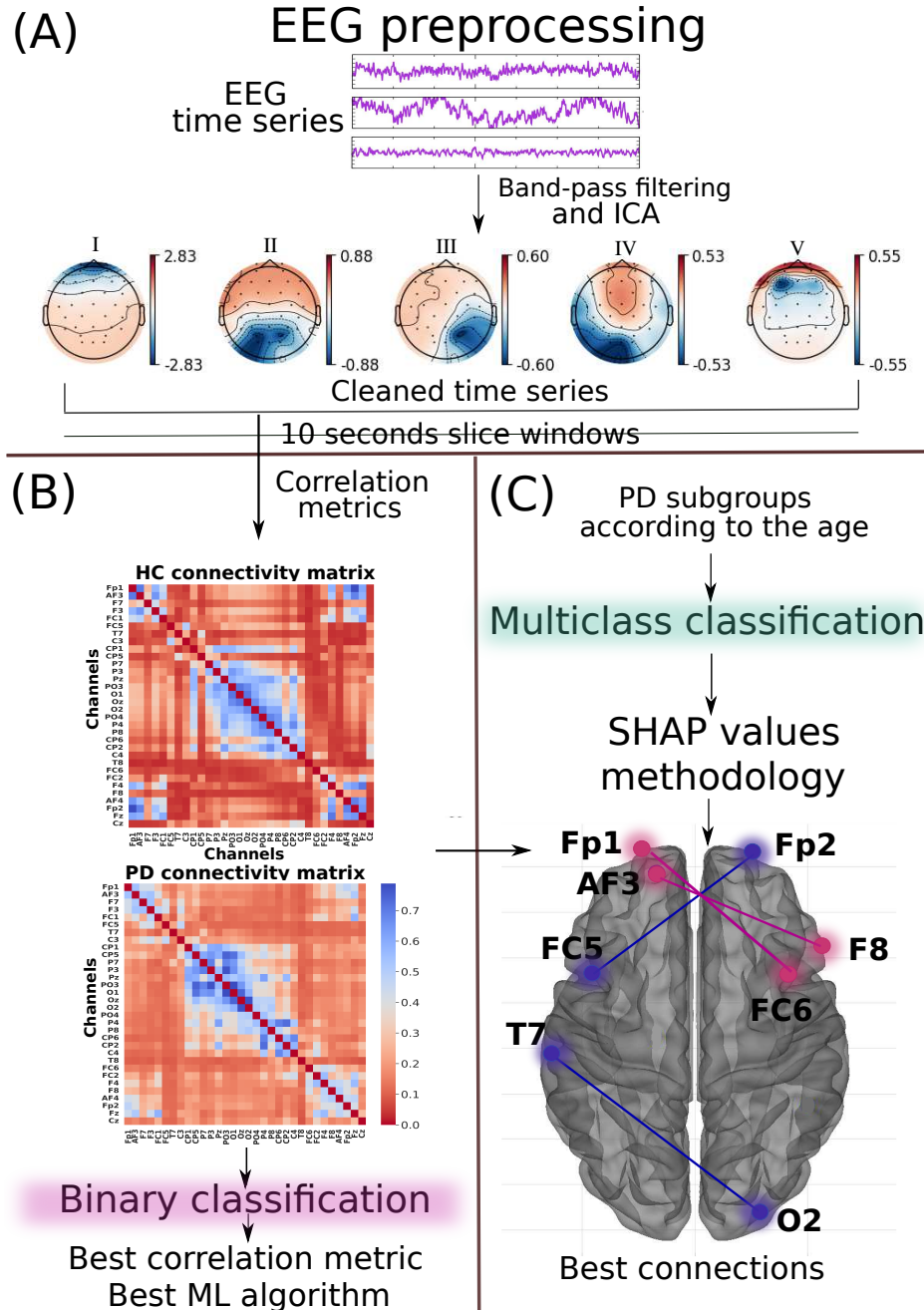
Figure 9. Summary of the methodology. In (A), which corresponds to the EEG preprocessing described in subsection 0.5, the EEG time series passed for a band-pass filter and to an independent component analysis (ICA). The Figure illustrates, as an example, a decomposition of the EEG signal of a PD patient into five components. In the Figure, as an example, the last component with higher frontal activity, indicating blink eye, was removed, and this process was made for all PD and control individuals. After preprocessing, we did 10 seconds of slicing windows as a data augmentation technique. In (B), which corresponds to the binary classification described in subsection 0.6, the cleaned and sliced EEG time series were used to compute various connectivity matrices with different correlation metrics. Using a support vector machine (SVM), the best correlation metric to distinguish PD from HC was found. Also, in (B), many ML algorithms were tested, and the best was selected. The results for part (B) can be found in subsection 0.1. With the best correlation metric's connectivity matrices and the best ML algorithm from part (B) used in a multiclass way to the connectivity matrices, the PD group was decomposed into two other groups according to their age, depicted in part (C) and described in subsection 0.6. Further, in (C), the SHAP values methodology was used to find the best connections. The results of the part (C) are in the subsection 0.2.
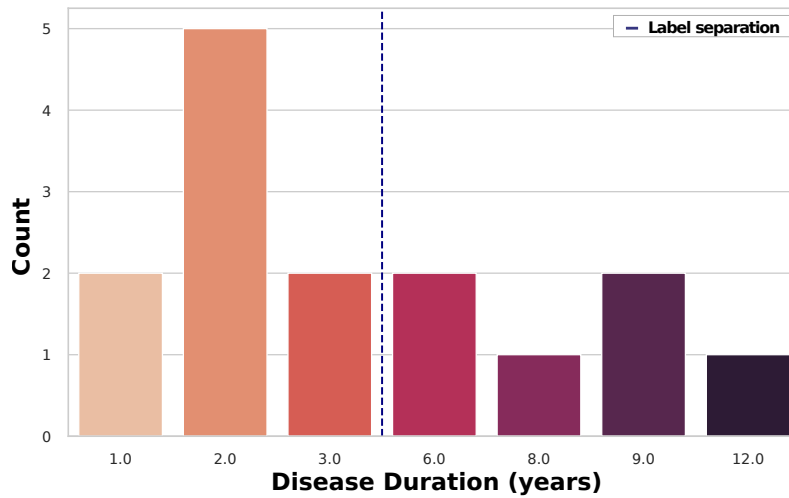
Figure 10. Distribution of the age in the PD group. The blue line is split according to the labels used for the multiclass classifier: those with a disease duration of 1-3 years (PD 1-3) and those with a disease duration of 6-12 years (PD 6-12).
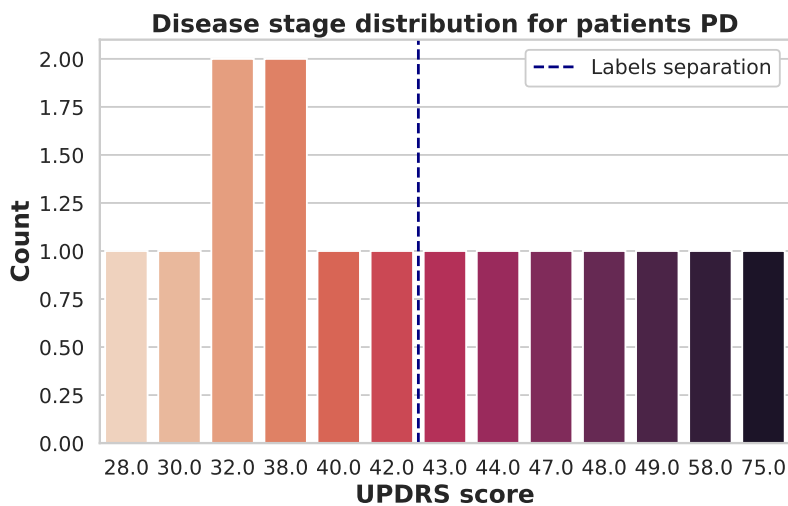
Figure 11. Distribution of the UPDRS in the PD group. The blue line is split according to the labels used for the multiclass classifier: those with a disease stage 20-40 (UPDRS1) and those with a disease stage 43-75 (UPDRS2).

technique is employed for medical interpretation, providing insights into the predictive importance of specific traits in the multiclass classification setting, as discussed in subsection 0.2.

1 Conclusions and future work

In conclusion, our study presents a novel approach to understanding PD progression through refined segmentation and advanced analysis of EEG signals. By stratifying patients based on disease duration, we reveal distinct neural connectivity patterns corresponding to different disease trajectory duration. Our findings underscore the importance of coherence as a critical metric in capturing the synchronized behavior of signals, which proved crucial in distinguishing PD patients from controls. Moreover, employing machine learning algorithms such as the MLP classifier and CNN, we achieved high accuracy rates compared to those reported in the literature, particularly in binary classification tasks. This indicates the robustness of our methodology in identifying subtle differences in connectivity patterns associated with PD progression and stage.

Furthermore, leveraging SHAP values enhances the interpretability of our models, revealing critical brain regions and connectivity patterns implicated in PD pathophysiology. An intriguing finding was the reversal of synchronization and physiological behavior between brain areas and the progression of the disease. While FP1 and FC5 showed low synchronization for PD1-3, FP2 and FC6 exhibited high synchronization for PD6-12. These findings point to a specific physiological behavior for these two PD duration. Our study revealed a notable correlation between FC6 and FP1 electrodes in both PD 1-3 and PD 6-12 groups, with decreased synchrony in the former and increased synchrony in the latter. This observed synchronization escalation over time may indicate a relationship with medication effectiveness, implying a progressive influence of treatment on neural connectivity.

Interestingly, our analysis unveils intriguing clustering behavior, suggesting a continuum of neural connectivity patterns across different disease duration. Notably, early-duration PD exhibits similarities with healthy brain function, while advanced duration manifest distinctive features indicative of disease progression. These findings deepen our understanding of PD pathogenesis and lay the groundwork for personalized diagnostic and therapeutic interventions tailored to the specific needs of patients at different duration of disease progression.

However, it is essential to acknowledge the limitations of our methodology, particularly its reliance on small data. Future work should aim to evaluate a dataset with a broader range of disease duration, allowing for a more comprehensive understanding of PD progression across its various duration. By incorporating larger datasets, we can further validate the robustness and generalizability of our findings, ultimately advancing the field towards more effective diagnostic and therapeutic strategies. Overall, our study contributes valuable insights into the complex interplay between neural dynamics, disease progression, and age-related changes in PD, offering new avenues for further research and clinical application.

Conflict of interest

The authors have no conflicts of interest to declare.

Author contributions statement

C.L.A.: conceptualization, formal analysis, investigation, methodology, visualization, validation, software and writing – original draft. L.F.S.: validation, writing writing – original draft. F.A.R.: validation, writing – review editing. T.L.G.L.O.T.: validation, writing – review editing. P.M.C.A: validation, writing – review editing. M.M: funding acquisition, project administration, resources, supervision, validation, writing – review editing.

Data availability

The authors have no conflicts of interest to declare. The study used the SanDiego dataset, openly available at https://openneuro.org/datasets/ds002778/versions/1.0.5.. The preprocessed matrices generated or analyzed during this study are included in this published article (and its Supplementary files).

References

1. Poewe, W. *et al.* Parkinson disease. *Nat. reviews Dis. primers* **3**, 1–21 (2017).

2. Balestrino, R. & Schapira, A. Parkinson disease. *Eur. journal neurology* **27**, 27–42 (2020).

3. Raza, C., Anjum, R. *et al.* Parkinson's disease: Mechanisms, translational models and management strategies. *Life sciences* **226**, 77–90 (2019).

4. Tsuboi, T. *et al.* Effects of mao-b inhibitors on non-motor symptoms and quality of life in parkinson's disease: A systematic review. *npj Park. Dis.* **8**, 75 (2022).

5. Schapira, A. H., Chaudhuri, K. R. & Jenner, P. Non-motor features of parkinson disease. *Nat. Rev. Neurosci.* **18**, 435–450 (2017).

6. Müller-Nedebock, A. C. *et al.* Different pieces of the same puzzle: a multifaceted perspective on the complex biological basis of parkinson's disease. *npj Park. Dis.* **9**, 110 (2023).

7. Kaidery, N. A., Tarannum, S. & Thomas, B. Epigenetic landscape of parkinson's disease: emerging role in disease mechanisms and therapeutic modalities. *Neurotherapeutics* **10**, 698–708 (2013).

8. Kilzheimer, A., Hentrich, T., Burkhardt, S. & Schulze-Hentrich, J. M. The challenge and opportunity to diagnose parkinson's disease in midlife. *Front. Neurol.* **10**, 1328 (2019).

9. Postuma, R. B. & Berg, D. Prodromal parkinson's disease: the decade past, the decade to come. *Mov. disorders* **34**, 665–675 (2019).

10. Subramanian, I. *et al.* Unmet needs of women living with parkinson's disease: gaps and controversies. *Mov. Disord.* **37**, 444–455 (2022).

11. Amoroso, N., La Rocca, M., Monaco, A., Bellotti, R. & Tangaro, S. Complex networks reveal early mri markers of parkinson's disease. *Med. image analysis* **48**, 12–24 (2018).

12. Alves, C. L. *et al.* Diagnosis of autism spectrum disorder based on functional brain networks and machine learning. *Sci. Reports* **13**, 8072 (2023).

13. Alves, C. L., Pineda, A. M., Roster, K., Thielemann, C. & Rodrigues, F. A. Eeg functional connectivity and deep learning for automatic diagnosis of brain disorders: Alzheimer's disease and schizophrenia. *J. Physics: complexity* **3**, 025001 (2022).

14. Alves, C. L. *et al.* Application of machine learning and complex network measures to an eeg dataset from ayahuasca experiments. *Plos one* **17**, e0277257 (2022).

15. Alves, C. L. *et al.* Analysis of functional connectivity using machine learning and deep learning in different data modalities from individuals with schizophrenia. *J. Neural Eng.* (2023).

16. Alves, C., Wissel, L., Capetian, P. & Thielemann, C. P 55 functional connectivity and convolutional neural networks for automatic classification of eeg data. *Clin. Neurophysiol.* **137**, e47 (2022).

17. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems*, 4768–4777 (2017).

18. Al-Beltagi, M. Autism medical comorbidities. *World journal clinical pediatrics* **10**, 15 (2021).

19. Qiu, L. *et al.* A novel eeg-based parkinson's disease detection model using multiscale convolutional prototype networks. *IEEE Transactions on Instrumentation Meas.* (2024).

20. Khare, S. K., Bajaj, V. & Acharya, U. R. Detection of parkinson's disease using automated tunable q wavelet transform technique with eeg signals. *Biocybern. Biomed. Eng.* **41**, 679–689 (2021).

21. Loh, H. W. *et al.* Gaborpdnet: Gabor transformation and deep neural network for parkinson's disease detection using eeg signals. *Electronics* **10**, 1740 (2021).

22. Aljalal, M., Aldosari, S. A., Molinas, M., AlSharabi, K. & Alturki, F. A. Detection of parkinson's disease from eeg signals using discrete wavelet transform, different entropy measures, and machine learning techniques. *Sci. Reports* **12**, 22547 (2022).

23. Rizvi, S. Q. A. *et al.* Classifying parkinson's disease using resting state electroencephalogram signals and u^{EN}-pdnet. *IEEE Access* (2023).

24. Richhariya, B., Tanveer, M., Rashid, A. H., Initiative, A. D. N. *et al.* Diagnosis of alzheimer's disease using universum support vector machine based recursive feature elimination (usvm-rfe). *Biomed. Signal Process. Control.* **59**, 101903 (2020).

25. Yang, M. *et al.* Establishing a prediction model of severe acute mountain sickness using machine learning of support vector machine recursive feature elimination. *Sci. Reports* **13**, 4633 (2023).

26. Wu, Z. *et al.* Identifying potential biomarkers of idiopathic pulmonary fibrosis through machine learning analysis. *Sci. Reports* **13**, 16559 (2023).

27. Gadalla, A. A. *et al.* Identification of clinical and urine biomarkers for uncomplicated urinary tract infection using machine learning algorithms. *Sci. reports* **9**, 19694 (2019).

28. Prottasha, N. J. *et al.* Impact learning: A learning method from feature's impact and competition. *J. Comput. Sci.* **69**, 102011 (2023).

29. Bowyer, S. M. Coherence a measure of the brain networks: past and present. *Neuropsychiatr. Electrophysiol.* **2**, 1–12 (2016).

30. Takakusaki, K., Takahashi, M., Noguchi, T. & Chiba, R. Neurophysiological mechanisms of gait disturbance in advanced parkinson's disease patients. *Neurol. Clin. Neurosci.* **11**, 201–217 (2023).

31. Swinnen, S. P. Intermanual coordination: from behavioural principles to neural-network interactions. *Nat. Rev. Neurosci.* **3**, 348–359 (2002).

32. Surmeier, D. J., Obeso, J. A. & Halliday, G. M. Selective neuronal vulnerability in parkinson disease. *Nat. Rev. Neurosci.* **18**, 101–113 (2017).

33. Conte, A., Khan, N., Defazio, G., Rothwell, J. C. & Berardelli, A. Pathophysiology of somatosensory abnormalities in parkinson disease. *Nat. Rev. Neurol.* **9**, 687–697 (2013).

34. Gilbert, S. J. *et al.* Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J. cognitive neuroscience* **18**, 932–948 (2006).

35. Koechlin, E. & Summerfield, C. An information theoretical approach to prefrontal executive function. *Trends cognitive sciences* **11**, 229–235 (2007).

36. Hagoort, P. Nodes and networks in the neural architecture for language: Broca's region and beyond. *Curr. opinion Neurobiol.* **28**, 136–141 (2014).

37. Liakakis, G., Nickel, J. & Seitz, R. Diversity of the inferior frontal gyrus—a meta-analysis of neuroimaging studies. *Behav. brain research* **225**, 341–347 (2011).

38. Jones, S. R., Pritchett, D. L., Stufflebeam, S. M., Hämäläinen, M. & Moore, C. I. Neural correlates of tactile detection: a combined magnetoencephalography and biophysically based computational modeling study. *J. Neurosci.* **27**, 10751–10764 (2007).

39. Smith, N. J. & Kutas, M. Regression-based estimation of erp waveforms: I. the rerp framework. *Psychophysiology* **52**, 157–168 (2015).

40. Hanganu, A. *et al.* Mild cognitive impairment is linked with faster rate of cortical thinning in patients with parkinson's disease longitudinally. *Brain* **137**, 1120–1129 (2014).

41. Olde Dubbelink, K. T. *et al.* Disrupted brain network topology in parkinson's disease: a longitudinal magnetoencephalography study. *Brain* **137**, 197–207 (2014).

42. Hammond, C., Bergman, H. & Brown, P. Pathological synchronization in parkinson's disease: networks, models and treatments. *Trends neurosciences* **30**, 357–364 (2007).

43. Moazami-Goudarzi, M., Sarnthein, J., Michels, L., Moukhtieva, R. & Jeanmonod, D. Enhanced frontal low and high frequency power and synchronization in the resting eeg of parkinsonian patients. *Neuroimage* **41**, 985–997 (2008).

44. Mijalkov, M. *et al.* Braph: a graph theory software for the analysis of brain connectivity. *PloS one* **12**, e0178798 (2017).

45. Zhao, Z. *et al.* Classification of schizophrenia by combination of brain effective and functional connectivity. *Front. Neurosci.* **15**, 651439 (2021).

46. Yin, Y. *et al.* Fomo and the brain: Loneliness and problematic social networking site use mediate the association between the topology of the resting-state eeg brain network and fear of missing out. *Comput. Hum. Behav.* **141**, 107624 (2023).

47. Rockhill, A. P. *et al.* Uc san diego resting state eeg data from patients with parkinson's disease. *Available:* (2020).

48. Han, C.-X., Wang, J., Yi, G.-S. & Che, Y.-Q. Investigation of eeg abnormalities in the early stage of parkinson's disease. *Cogn. neurodynamics* **7**, 351–359 (2013).

49. George, J. S. *et al.* Dopaminergic therapy in parkinson's disease decreases cortical beta band coherence in the resting state and increases cortical beta band power during executive control. *NeuroImage: Clin.* **3**, 261–270 (2013).

50. Benesty, J., Chen, J., Huang, Y. & Cohen, I. Pearson correlation coefficient. In *Noise reduction in speech processing*, 1–4 (Springer, 2009).

51. Lubinski, D. Introduction to the special section on cognitive abilities: 100 years after spearman's (1904)"'general intelligence,'objectively determined and measured". *J. personality social psychology* **86**, 96 (2004).

52. Hardoon, D. R. & Shawe-Taylor, J. Sparse canonical correlation analysis. *Mach. Learn.* **83**, 331–353 (2011).

53. Ledoit, O. & Wolf, M. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals Stat.* **40**, 1024–1060 (2012).

54. French, C. C. & Beaumont, J. G. A critical review of eeg coherence studies of hemisphere function. *Int. J. Psychophysiol.* **1**, 241–254 (1984).

55. Lashgari, E., Liang, D. & Maoz, U. Data augmentation for deep-learning-based electroencephalography. *J. Neurosci. Methods* **346**, 108885 (2020).

56. Nakisa, B., Rastgoo, M. N., Rakotonirainy, A., Maire, F. & Chandran, V. Automatic emotion recognition using temporal multimodal deep learning. *IEEE Access* **8**, 225463–225474 (2020).

57. Hoffmann, A. *et al.* Electroencephalography during functional echo-planar imaging: detection of epileptic spikes using post-processing methods. *Magn. Reson. Medicine: An Off. J. Int. Soc. for Magn. Reson. Medicine* **44**, 791–798 (2000).

58. Matell, M. S., Meck, W. H. & Nicolelis, M. A. Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behav. neuroscience* **117**, 760 (2003).

59. Bengio, Y. & Grandvalet, Y. No unbiased estimator of the variance of k-fold cross-validation. *Adv. Neural Inf. Process. Syst.* **16** (2003).

60. Kohavi, R. *et al.* A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, vol. 14, 1137–1145 (Montreal, Canada, 1995).

61. Varma, S. & Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC bioinformatics* **7**, 1–8 (2006).

62. Blum, A., Kalai, A. & Langford, J. Beating the hold-out: Bounds for k-fold and progressive cross-validation. In *Proceedings of the twelfth annual conference on Computational learning theory*, 203–208 (1999).

63. Kapoor, S. & Narayanan, A. Leakage and the reproducibility crisis in machine-learning-based science. *Patterns* **4** (2023).

64. Shorten, C., Khoshgoftaar, T. M. & Furht, B. Text data augmentation for deep learning. *J. big Data* **8**, 1–34 (2021).

65. Alves, C. L. *et al.* On the advances in machine learning and complex network measures to an eegdataset from dmt experiments. *J. Physics: Complex.* (2024).

66. Bisong, E. & Bisong, E. Introduction to scikit-learn. *Build. Mach. Learn. Deep. Learn. Model. on Google Cloud Platform: A Compr. Guid. for Beginners* 215–229 (2019).

67. Raschka, S. *Python machine learning* (Packt publishing ltd, 2015).

68. Géron, A. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow* (" O'Reilly Media, Inc.", 2022).

69. Kim, D. W. *et al.* Deep learning-based survival prediction of oral cancer patients. *Sci. reports* **9**, 6994 (2019).

70. Su, Q. *et al.* Faecal microbiome-based machine learning for multi-class disease diagnosis. *Nat. Commun.* **13**, 6818 (2022).

71. Singh, V. *et al.* Impact of train/test sample regimen on performance estimate stability of machine learning in cardiovascular imaging. *Sci. reports* **11**, 14490 (2021).

72. Avuçlu, E. & Elen, A. Evaluation of train and test performance of machine learning algorithms and parkinson diagnosis with statistical measurements. *Med. & Biol. Eng. & Comput.* **58**, 2775–2788 (2020).

73. Kuhn, M., Johnson, K. *et al. Applied predictive modeling*, vol. 26 (Springer, 2013).

74. Brownlee, J. How to choose a feature selection method for machine learning. *Mach. Learn. Mastery* **10** (2019).

75. Mahesh, T., Geman, O., Margala, M., Guduri, M. *et al.* The stratified k-folds cross-validation and class-balancing methods with high-performance ensemble classifiers for breast cancer classification. *Healthc. Anal.* **4**, 100247 (2023).

76. Lopez-Martin, M., Nevado, A. & Carro, B. Detection of early stages of alzheimer's disease based on meg activity with a randomized convolutional neural network. *Artif. Intell. Medicine* **107**, 101924 (2020).

77. Zhong, Z., Yuan, X., Liu, S., Yang, Y. & Liu, F. Machine learning prediction models for prognosis of critically ill patients after open-heart surgery. *Sci. Reports* **11**, 1–10 (2021).

78. Sato, M. *et al.* Machine-learning approach for the development of a novel predictive model for the diagnosis of hepatocellular carcinoma. *Sci. reports* **9**, 1–7 (2019).

79. Krittanawong, C. *et al.* Machine learning and deep learning to predict mortality in patients with spontaneous coronary artery dissection. *Sci. reports* **11**, 1–10 (2021).

80. Arcadu, F. *et al.* Author correction: Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *NPJ digital medicine* **3**, 1–6 (2020).

81. Rashidi, H. H. *et al.* Early recognition of burn-and trauma-related acute kidney injury: a pilot comparison of machine learning techniques. *Sci. reports* **10**, 1–9 (2020).

82. Bischl, B. *et al.* Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **13**, e1484 (2023).

83. Mincholé, A. & Rodriguez, B. Artificial intelligence for the electrocardiogram. *Nat. medicine* **25**, 22–23 (2019).

84. Tolkach, Y., Dohmgörgen, T., Toma, M. & Kristiansen, G. High-accuracy prostate cancer pathology using deep learning. *Nat. Mach. Intell.* **2**, 411–418 (2020).

85. Dukart, J., Weis, S., Genon, S. & Eickhoff, S. B. Towards increasing the clinical applicability of machine learning biomarkers in psychiatry. *Nat. Hum. Behav.* **5**, 431–432 (2021).

86. Li, R. C., Asch, S. M. & Shah, N. H. Developing a delivery science for artificial intelligence in healthcare. *NPJ digital medicine* **3**, 1–3 (2020).

87. Park, Y. & Kellis, M. Deep learning for regulatory genomics. *Nat. biotechnology* **33**, 825–826 (2015).

88. Ito, Y. *et al.* A method for utilizing automated machine learning for histopathological classification of testis based on johnsen scores. *Sci. reports* **11**, 1–11 (2021).

89. Kim, J., Lee, J., Park, E. & Han, J. A deep learning model for detecting mental illness from user content on social media. *Sci. reports* **10**, 1–6 (2020).

90. Li, Y., Nowak, C. M., Pham, U., Nguyen, K. & Bleris, L. Cell morphology-based machine learning models for human cell state classification. *NPJ systems biology applications* **7**, 1–9 (2021).

91. Yu, X., Pang, W., Xu, Q. & Liang, M. Mammographic image classification with deep fusion learning. *Sci. Reports* **10**, 1–11 (2020).

92. Bracher-Smith, M., Crawford, K. & Escott-Price, V. Machine learning for genetic prediction of psychiatric disorders: a systematic review. *Mol. Psychiatry* **26**, 70–79 (2021).

93. Patel, D. *et al.* Machine learning based predictors for covid-19 disease severity. *Sci. Reports* **11**, 1–7 (2021).

94. Ferri, C., Hernández-Orallo, J. & Salido, M. A. Volume under the roc surface for multi-class problems. In *European conference on machine learning*, 108–120 (Springer, 2003).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- matricesfinalscientificreports.zip