Journal of Physics:
Complexity

PAPER • **OPEN ACCESS**

Dynamical patterns of EEG connectivity unveil Parkinson's disease progression: insights from machine learning analysis

To cite this article: Caroline L Alves *et al* 2025 *J. Phys. Complex.* **6** 035006

Journal of Physics: Complexity

PAPER

Dynamical 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, SãoPaulo, Brazil
[*] Author to whom any correspondence should be addressed.

E-mail: Caroline.Alves@th-ab.de

Keywords: brain network, Parkinson disease, age stage, machine learning

OPEN ACCESS

RECEIVED
25 March 2025

REVISED
26 June 2025

ACCEPTED FOR PUBLICATION
29 July 2025

PUBLISHED
7 August 2025

Abstract

Parkinson's disease (PD) is a multifactorial neurodegenerative disorder with complex progression. This study aims to analyze electroencephalography (EEG) connectivity patterns to better understand PD progression and stage of the disease using machine learning. Resting-state, eyes-closed EEG recordings were acquired from 31 individuals: 16 healthy controls (HCs) and 15 PD patients. The PD group was stratified by disease duration into early-stage (1–3 years, $n = 9$) and advanced-stage (6–12 years, $n = 6$). EEG was recorded using a 32-channel Biosemi Active-Two system (512 Hz), with signals segmented into non-overlapping 10 s windows. Functional connectivity matrices were constructed using multiple metrics, including coherence, Pearson, Spearman, canonical correlation, and Ledoit–Wolf shrinkage. Machine learning models were applied for both binary (PD vs HC) and multiclass (HC vs early vs advanced PD) classification. Interpretability was achieved using Shapley Additive Explanations (PD) methodology , and the most discriminative neural connections were statistically validated using the Wilcoxon test with Bonferroni correction. Our approach achieved high accuracy in classifying PD stages, with coherence emerging as the optimal metric for capturing synchronized neural activity. SHAP values revealed critical brain regions and connectivity patterns associated with disease progression. Statistical validation confirmed the significance of these connections across disease stages. Early-stage PD exhibited neural connectivity patterns similar to HCs, while advanced stages showed distinct connectivity changes. The findings highlight the utility of EEG connectivity and machine learning in staging PD, offering insights into PD pathogenesis and progression. SHAP-enhanced model interpretability ensures reliable identification of key neural connections, supporting personalized diagnostics and therapeutic strategies.

1. 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 [15, 76, 82]. Non-motor symptoms, such as cognitive impairment, depression, and autonomic dysfunction, also significantly affect patients' quality of life [87, 101]. While PD's exact etiology remains elusive, evidence suggests a complex interaction of genetic susceptibility, environmental factors, and aging [46, 69].

On the other hand, electroencephalography (EEG)-based connectivity measures enable the characterization of functional brain network disruptions in neuropsychiatric and neurodegenerative conditions [2, 31, 107]. These approaches model the brain as a complex system by computing statistical or

Table 1. Overview of PD binary classification research using the same publicly accessible PD dataset as this study. 'Not reported' indicates that the corresponding information was not explicitly provided in the referenced study. Most of these studies rely on direct time-series input rather than on derived connectivity matrices, which explains the absence of correlation metric information.

Authors	Correlation metric	ML method	AUC	Accuracy	Recall	Precision
[79]	Phase locking value	Deep learning (Multiscale CNN)	88.7	88.7	86.7	Not reported
[48]	Not reported	Least squares SVM	Not reported	97.65	96.67	98.76
[61]	Not reported	Deep learning (2D-CNN)	Not reported	99.46	99.46	99.48
[3]	Not reported	KNN	Not reported	99.89	99.87	Not reported
[84]	Not reported	Deep learning (CNN)	98.96	97.90	97.87	98.0

spectral associations between EEG signals from different regions, producing functional connectivity matrices that can reveal altered communication pathways [108]. For example, EEG-based connectivity characterizes altered network segregation in Alzheimer's disease [13] and atypical integration patterns in attention-deficit/hyperactivity disorder [14], highlighting its potential as a non-invasive diagnostic support tool.

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 [49, 77, 95]. Although considerable efforts have been made to discern patterns and markers indicative of PD, the complexities inherent in the disease's progression still need to be addressed [10].

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 by capturing potential variations in EEG signals that correspond to different durations and stages of the disease.

Importantly, functional brain connectivity patterns also change across the lifespan, even in healthy individuals [19, 27, 30]. For example, aging is associated reduced synchronization in resting-state EEG networks [26, 92]. Therefore, when comparing patient and control groups using EEG-based connectivity, it is crucial to control for age effects. In this study, we ensured that age ranges of the PD and HC cohorts were largely overlapping, with statistical confirmation of age comparability across groups (see figure 2(b)).

In addition to activity pattern analysis, we use pairwise statistical metrics to construct connectivity matrices, based on techniques proven effective for preprocessing EEG time series data [8]. In particular, we explore the efficacy of these metrics in binary classification tasks to distinguish PD patients from controls, along with their utility in multiclass classification to differentiate between subgroups within the PD cohort based on disease duration. Furthermore, similar to our previous work [4, 6–9], we enhance the interpretability of our ML models using Shapley Additive Explanations (SHAP) values [64]. The SHAP framework identifies which features–in this context, specific neural connections–contribute most to the model's decision-making process [64]. By elucidating the most influential connectivity patterns, this approach provides a deeper understanding of the underlying neural dynamics of PD. The integration of this explanatory framework with our high-performing classifier not only refines current diagnostic approaches, as proven in previous research [1, 6, 7], but also establishes a basis for developing personalized therapeutic interventions that account for PD patients with different durations and disease progression.

2. Methods

Figure 1 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.

2.1. Data and EEG preprocessing

EEG signals were recorded with a 32-channel Biosemi Active-Two system at a sampling rate of 512 Hz, using the international 10–20 electrode placement re-referenced to the common average, and a high–pass filter at 0.5 Hz was applied during acquisition to suppress slow drifts to enhance signal clarity [85]. Furthermore, summary statistics of participants' age, MMSE scores, and gender distribution are provided in figure 2.

Here, offline preprocessing was conducted using MNE-Python [38]. We applied a band-pass finite impulse response filter (0.5–50 Hz, Hamming window) to retain neurophysiologically relevant delta to low-gamma activity while reducing noise [72, 91]. Independent component analysis was employed to

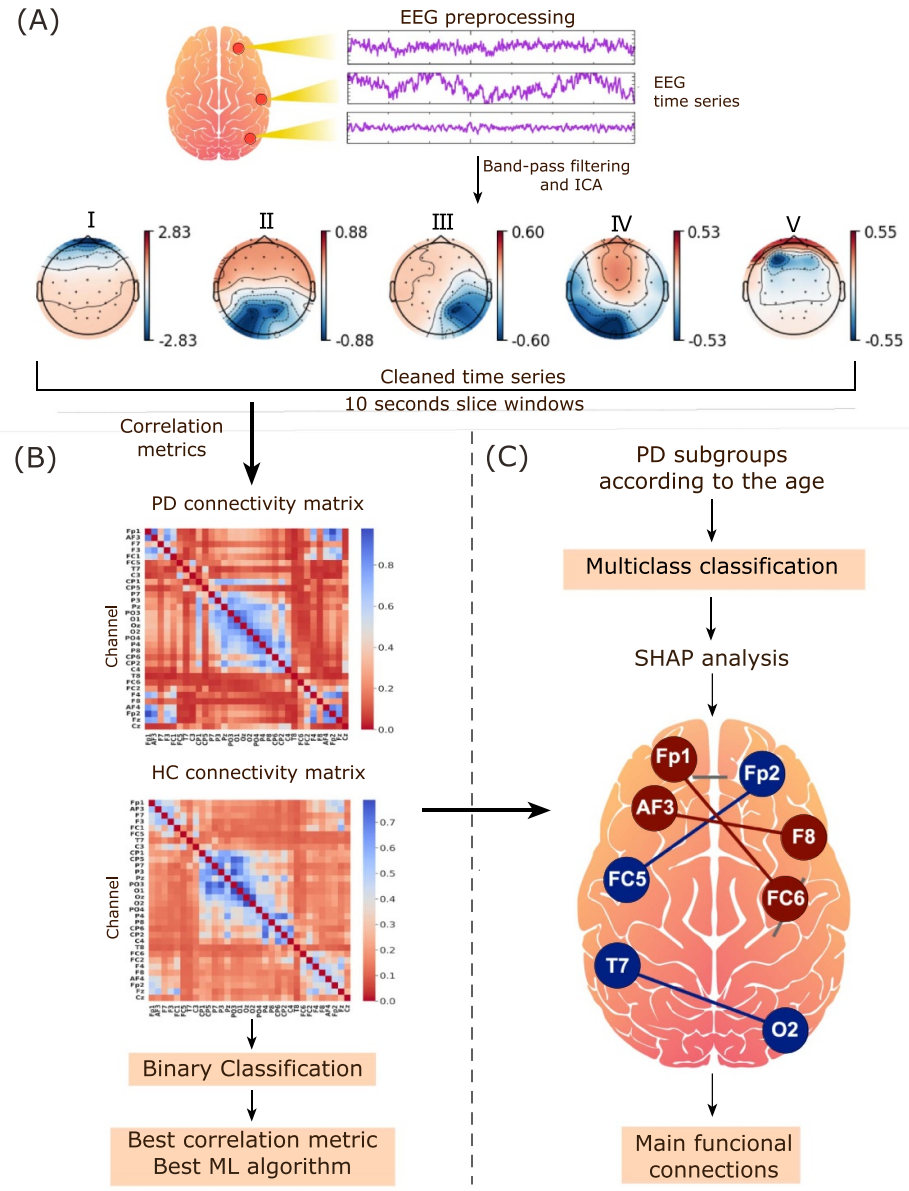


Figure 1. Summary of the methodology. In (A), which corresponds to the EEG preprocessing described in section 2.1, the EEG time series passed for a band-pass filter and to an independent component analysis. The Figure illustrates, as an example, a decomposition of the EEG signal of a PD patient into 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 s of slicing windows as a data augmentation technique. In (B), which corresponds to the binary classification described in section 2.2, 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 section 3.1. With the best correlation metrics 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 the stage of the disease, depicted in part (C) and described in section 2.2. Further, in (C), the SHAP values methodology was used to find the best connections. The results of the part (C) are in the section 3.2.

identify and remove ocular and other artifacts [103], with each component visually inspected and excluded for each participant (see section 3.2(A)).

To ensure methodological rigor and avoid information leakage during classification, the raw EEG recordings were first split split into training (75%) and test (25%). Further, data were segmented into non-overlapping 10 s epochs to enhance our dataset's size and facilitate comprehensive analysis of EEG signals. This technique involves partitioning EEG signals into smaller time windows [56], typically lasting 10 s [43, 66, 70], which effectively increases the number of instances available for analysis. We acknowledge that the small number of patients may impose limitations on the generalizability of our findings.

All processed segments were then used to compute connectivity matrices with Pearson correlation (PC) [16], Spearman correlation (SC) [63], Sparse canonical correlation analysis (CCA) [42], and Ledoit–Wolf

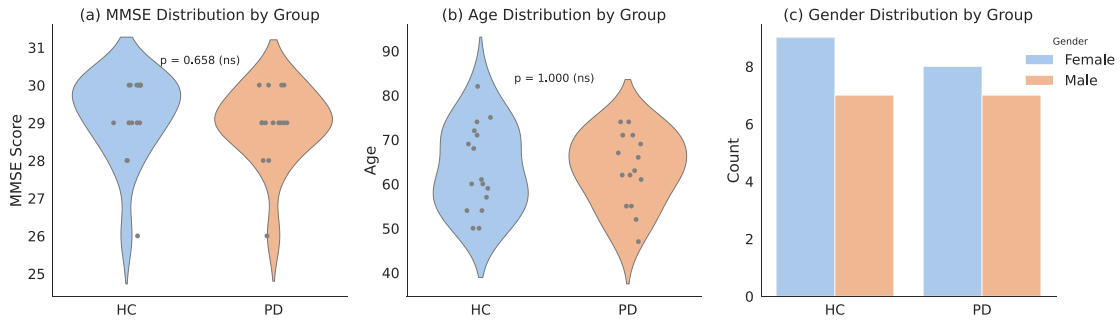


Figure 2. Demographic characteristics of the dataset. (a) Violin and strip plot of MMSE scores showing cognitive performance distribution across groups. (b) Violin and strip plot of participant age by group. (c) Gender distribution per group. Statistical comparisons for panels (a) and (b) were performed using the two-sided Mann–Whitney U test with Bonferroni correction for multiple comparisons. No significant differences were found between groups in MMSE ($p = 0.658$, ns) or age ($p = 1.000$, ns). Note: PD = Parkinson's disease, HC = healthy controls; ns = not significant.

shrinkage (LW) [57]. We emphasize coherence (Sync) [23, 33] for its unique ability to measure synchronization in pairwise EEG signals, enhances the dataset analysis, providing a nuanced exploration of EEG signals that were previously unexplored in our prior work [4, 6–9].

The Sync. measure between signals i and j, denoted as Sync$[i,j]$, is computed using the average coherence across all frequency points, as defined by equation (1):

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

where:

- Sync$[i,j]$ is the synchronization between signals i and j,
- N is the number of frequency points,
- $P_{ij}(f_k)$ is the cross-spectral density between signals i and j at frequency f_k,
- $P_{ii}(f_k)$ is the auto-spectral density of signal i at frequency f_k,
- $P_{jj}(f_k)$ is the auto-spectral density of signal j at frequency f_k,
- f_k represents the set of sampled points with a constant sampling frequency, assumed to be 50 Hz.

2.2. Binary and multiclass classification

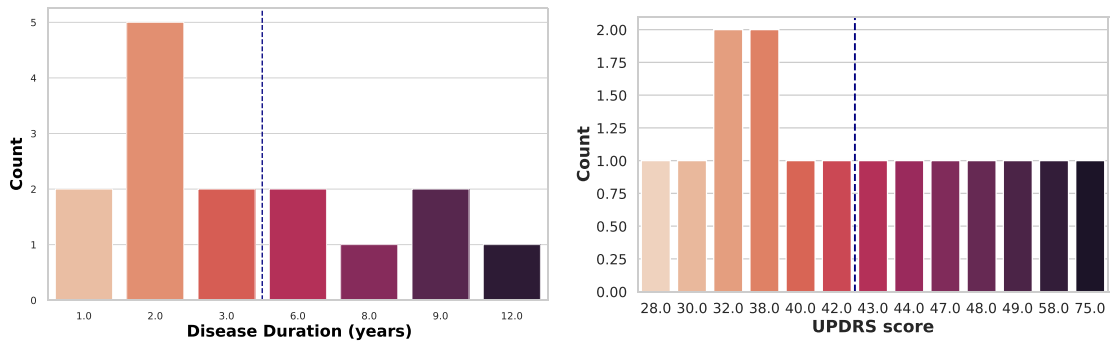
To prepare the data for model training and evaluation, we applied standardization and cross-validation procedures aimed at improving model generalizability and avoiding data leakage. Feature standardization was performed using the function StandScaler from the Scikit-learn library python package [75], which transforms features to have zero mean and unit variance [99]. This step ensures consistent scaling across all features and enhances the robustness of the classifiers to varying feature magnitudes and outliers [36, 80].

For model evaluation, we employed stratified 10-fold cross-validation with shuffling [25, 55, 65] in the training set. This approach preserves class proportions (PD and HC) across folds and provides a robust estimate of model performance by ensuring that each observation is used for both training and validation [18, 53]. The use of shuffling within each fold further mitigates the effects of temporal dependencies introduced by sliding window segmentation [22, 102].

By combining feature standardization, stratification, and data shuffling, we minimize the risk of information leakage and improve the fairness of evaluation [12, 21, 25, 36, 50, 55, 62, 65, 80, 89, 94]. This methodology enables the model to be tested on unseen data, thereby strengthening the validity and reproducibility of our findings [47, 88].

To classify PD patients and HCs (HC, with 16 individuals), we used a multiclass ML framework. The PD group was segmented based on disease duration into two subgroups: 1–3 years (PD 1–3, 9 individuals) and 6–12 years (PD 6–12, 6 individuals), as shown in figure 3(a). This division was chosen primarily to achieve a more balanced class distribution, given the limited sample size, and to allow meaningful training and evaluation of the classification models.

Furthermore, in order to assess not only the temporal changes attributable to 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 C. We divided the dataset into two



(a) Disease duration (years) (b) UPDRS

Figure 3. Distribution of the disease duration and UPDRS in PD group. In (a) 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). In (b) the blue line is split according to the labels used for the multiclass classifier: those with a disease stage of 20–40 (UPDRS1) and those with a disease stage of 43–75 (UPDRS2).

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 3(b). UPDRS1 typically represents milder stages of PD, while UPDRS2 indicates the more advanced stages, reflecting the progression of the disease severity between participants.

The current work builds upon our previous research, which have established a foundation for employing ML techniques in the classification of PD and HC groups [4–9]. Initially, we utilize the support vector machine (SVM) to select connectivity metrics, leveraging its lower computational cost and effectiveness in binary classification [8]. Afterwards, various ML algorithms, including logistic regression, random forest, multilayer perceptron (MLP), long short-term memory neural networks, and convolutional neural networks (CNN), are tested with the selected connectivity metric, as detailed in the section 3.1. Noteworthy, leveraging 10 s slicing windows data augmentations, a total dataset of 400 connectivity matrices is used, comprised of 200 datapoints representing HC, 100 PD in an early stage, and 100 PD in an progressed stage.

Hyperparameter optimization techniques, such as grid search and random search, are further employed to fine-tune the ML algorithms for optimal performance, with evaluation metrics including accuracy, precision, recall, receiver operating characteristic (ROC) curves, and area under the ROC curve (AUC) [6–9, 11, 20, 24, 29, 32, 44, 51, 54, 58, 59, 67, 73, 74, 81, 86, 100, 106, 109]. Additionally, the SHAP value technique is employed for medical interpretation, providing insights into the predictive importance of specific traits in the multiclass classification setting, as discussed in section 3.2.

Further, we conducted comprehensive statistical analyses on these connections to further validate the ML findings, particularly the most significant cortical connections identified through the SHAP values methodology. Specifically, we selected the connection strength values between pertinent cortical regions and evaluated them across three distinct groups: HC, PD 1–3, and PD 6–12.

We employed the Wilcoxon rank-sum test for each connection to compare the distributions between groups. The Bonferroni correction was applied to account for multiple comparisons and control the family-wise error rate. The following symbols denote the levels of statistical significance:

- ns: $5.00e-02 < p \leqslant 1.00e+00$
- *: $1.00e-02 < p \leqslant 5.00e-02$
- **: $1.00e-03 < p \leqslant 1.00e-02$
- ***: $1.00e-04 < p \leqslant 1.00e-03$
- ****: $p \leqslant 1.00e-04$.

This rigorous statistical approach ensures that the observed differences in connection strengths are not attributed to random chance, thereby reinforcing the validity of the ML-derived insights.

3. Results

3.1. Binary classification

Considering the AUC metric, the performance of each connectivity metric is illustrated in figure 4. 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.

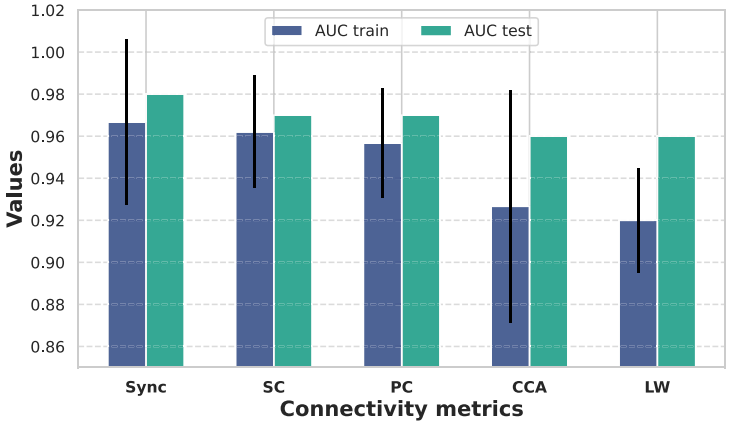
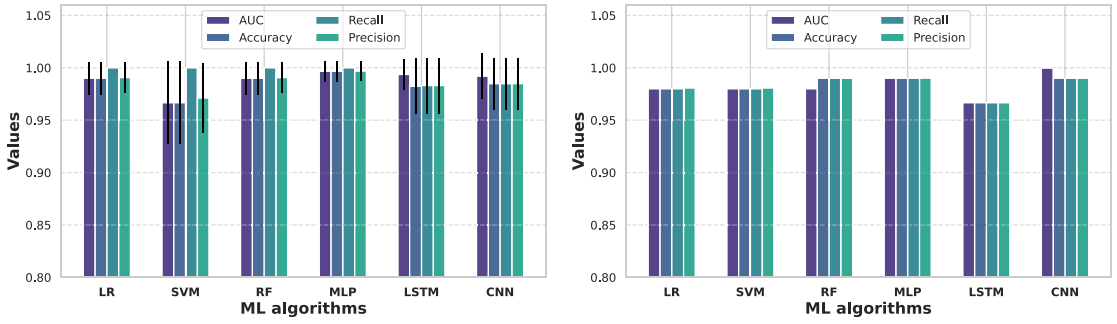


Figure 4. Selection of optimal correlation metrics with SVM. In blue is the AUC of the train set; in green is the AUC of the test set. The best performance was the Sync.



(a) Train set results (b) Test set results

Figure 5. 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.

Then, we tested all ML algorithms that resulted in figure 5; 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 optimize model performance, we performed feature selection using recursive feature elimination (RFE), an iterative method designed to identify the subset of features that maximizes classification accuracy. RFE is well-established in medical data analysis [34, 83, 104, 105] and works by progressively removing the least important features and reevaluating the model. Here, the RFE analysis was implemented with the Yellowbrick library [17] and revealed that peak performance was achieved with a subset of 136 features (figure 6). Consequently, this optimized feature set was used for all subsequent modeling, as the analysis confirmed that the complete feature set was unnecessary for achieving maximal efficacy.

3.2. 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 7 displays the learning curve (figure 7(a)) and the confusion matrix (7(c)), respectively. From figures 7(a), the confusion matrix 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 [78]. Figure 7(c) shows that all the data was required to converge the model.

For the interpretability of the model and the decision-making process, we performed a SHAP analysis. Although RFE identified 136 features as optimal for predictive performance, we intentionally used a broader set of the top 300 features for the SHAP analysis. This expanded set, chosen despite its high computational

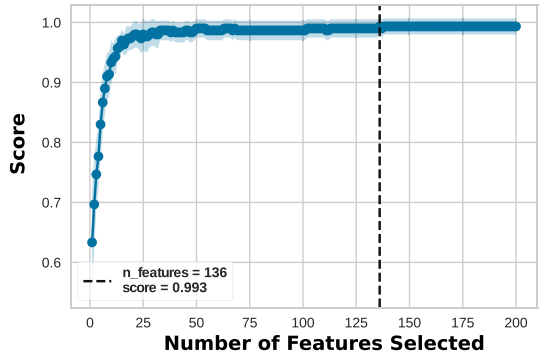
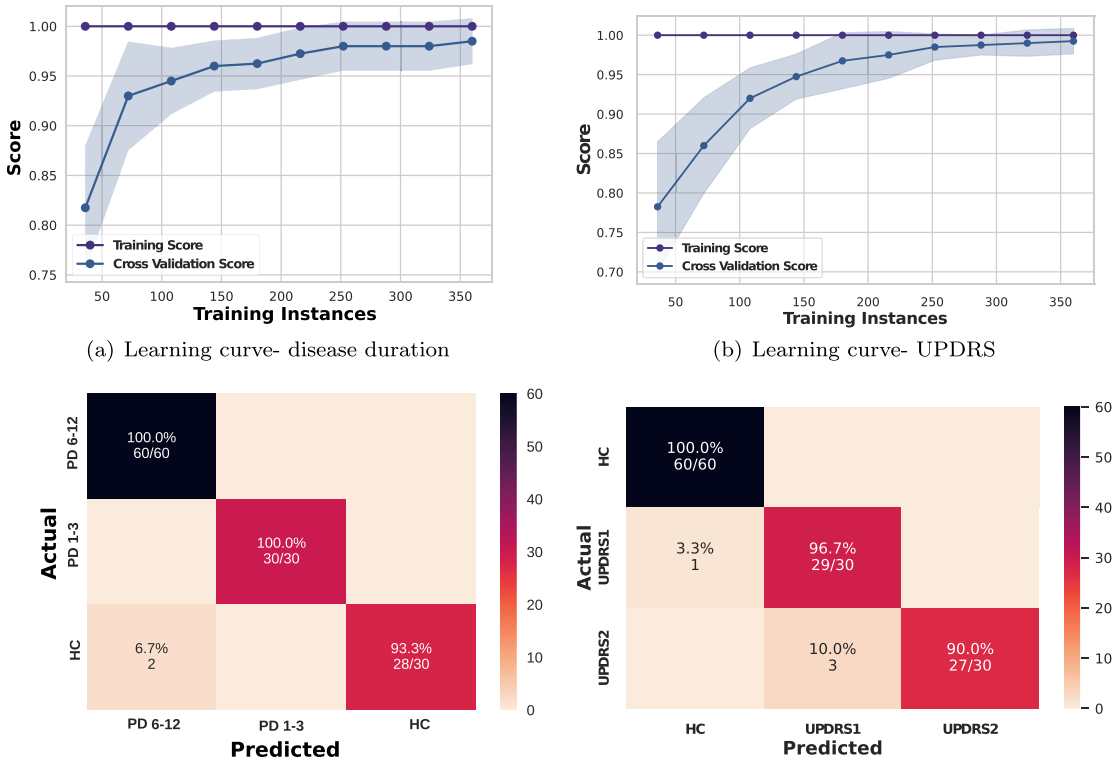


Figure 6. Recursive feature elimination (RFE). The optimal performance is attained with 136 features, so utilizing the entire comprehensive feature set is unnecessary.



(a) Learning curve- disease duration

(b) Learning curve- UPDRS

(c) Confusion matrix- disease duration

(d) Confusion matrix- UPDRS

Figure 7. The test sample ML results from connectivity matrices using disease duration and UPDRS. (a), (b) Learning curves for disease duration and UPDRS, showing training (purple) and test (blue) accuracies. (c), (d) Confusion matrices for disease duration and UPDRS, where diagonal elements represent true positives (TP), highlighting the classification performance for different classes.

cost, allows for a more comprehensive and robust investigation of the features ranked by their global contribution. The analysis identified distinct connectivity patterns driving the classification of early-stage versus late-stage PD (figure 8). In these plots, the feature value represents the connectivity strength, where red indicates high synchronization and blue indicates low synchronization. For the PD 1–3 group, model predictions were most influenced by decreased synchronization between the FP2-FC5 and O2–T7 electrode pairs. Conversely, for the PD 6–12 group, the most decisive features were increased synchronization between F8–AF3 and FC6–Fp1.

SHAP value matrices for each class are concatenated into a single array, with rows representing samples and columns representing features. We averaged SHAP values across instances within each class for clarity. A cluster map, shown in figure 9, was generated using hierarchical clustering with cosine similarity as the distance metric, which captures the orientation of feature importance patterns rather than their magnitude differences [93]. The resulting dendrogram groups rows and columns with similar SHAP patterns, with

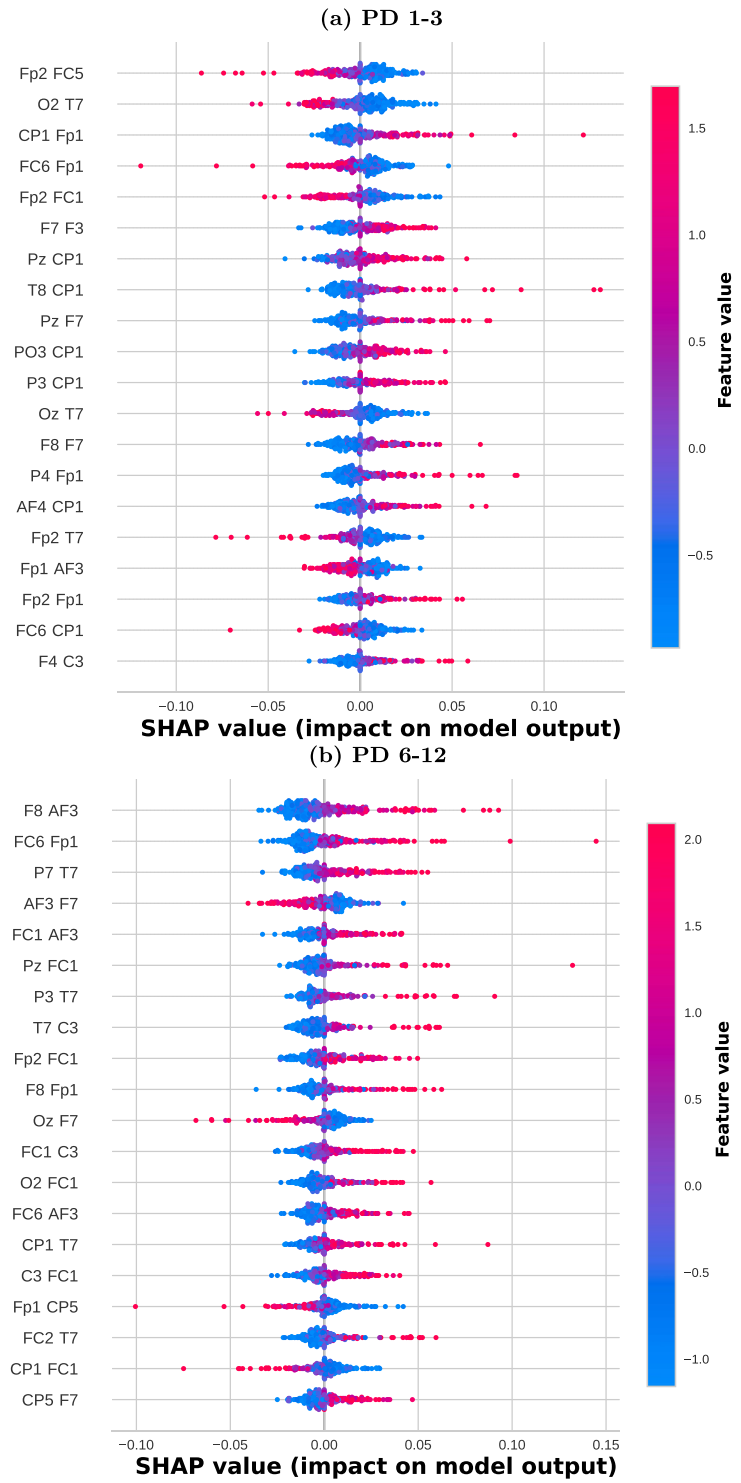


Figure 8. Feature importance ranking using the SHAP values methodology for the MLP classifier with brain EEG channels in descending order. The *y*-axis represents the most important EEG connections, ranked by mean absolute SHAP value. The *x*-axis represents the SHAP value, indicating the magnitude and direction of a feature's impact on the model's prediction for that class. The feature value represents the connectivity strength (red = high synchronization, blue = low synchronization).(a) Feature importance ranking regarding PD 1–3 class. (b) Feature importance ranking regarding PD 6–12 class.

heatmap colors indicating SHAP value magnitudes (blue color scale for negative, red color scale for positive, and white for values near zero).

The dendogram clearly shows that the PD 1–3 group clusters most closely with the HC group. This proximity indicates that the model identifies significant similarities in the neural connectivity patterns of early-stage PD and healthy individuals. More importantly, this combined group is directly linked to the late-stage PD group based on their Shap values. This indicates a potential continuum or progression of the

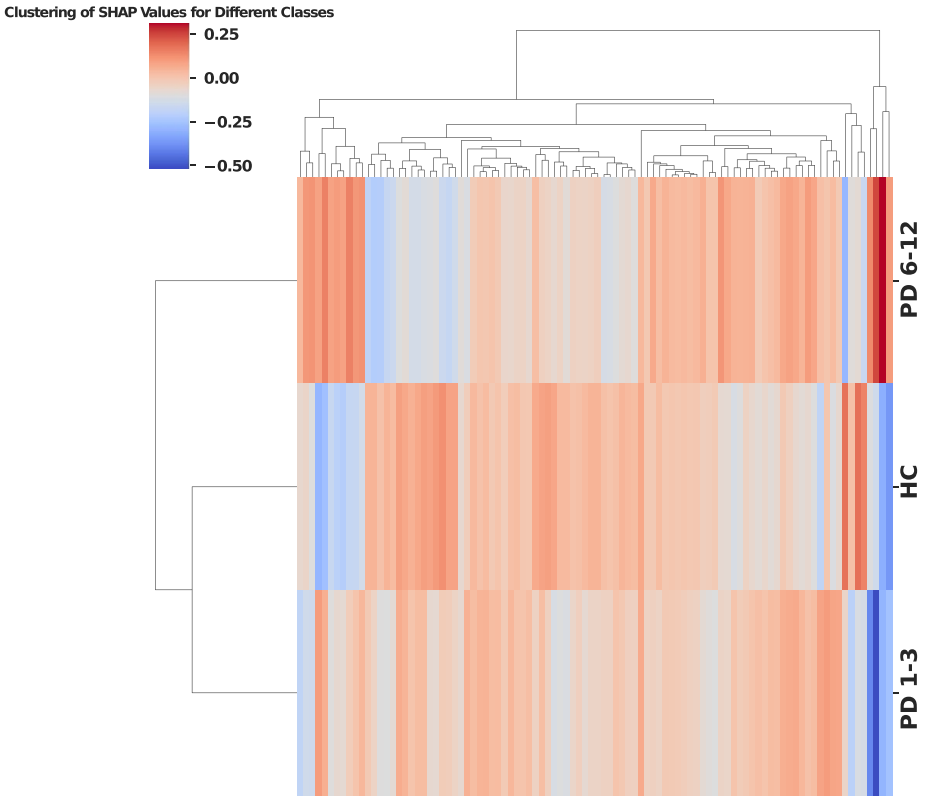


Figure 9. Cluster map generated through hierarchical clustering of SHAP values. The dendrogram of the superior part corresponds to the hierarchy clustering for each of the 300 features used in the SHAP value methodology. The heatmap colors represent SHAP value magnitudes (blue color scale for negative, red color scale 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.

features, reflecting a gradual transition between the groups. In essence, the model has identified a measurable path of change as the disease progresses.

Further, as a complement to the temporal changes results due to PD, the differentiates changes were evaluated using the UPDRS scores; 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 learning curve (figure 7(b)) and the confusion matrix (7(d)), respectively. From figure 7(d), the confusion matrix, 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.

We employed the MLP along with SHAP values to identify the five most significant connections for each correlation metric in the PD 1–3 and PD 6–12 classes (tables 2 and 3). This was done to assess the impact of the metrics, as their mathematical differences make them sensitive to distinct connectivity structures. For example, they capture different aspects of the signal, including linear (Pearson), rank-based (Spearman), frequency-domain (Synchronization), shrinkage-based (LW), and multivariate dependencies (CCA), which in turn leads to the identification of different sets of dominant connections.

Figure 10 illustrates the main EEG connectivity patterns identified for both PD 1–3 and PD 6–12 classes, based on the data from tables 2 and 3. In both classes, there is a concentration of significant connections in the frontal and prefrontal regions across multiple correlation metrics (Sync, SC, and CCA). Further, figure 10 shows that ongoing disease progression and the correlation of the frontal region (F7 and Fp2) with other areas is reduced.

In addition to the ML validation using SHAP values, we conducted rigorous statistical tests on the connectivity metrics, mainly focusing on the most significant connections identified through the SHAP values methodology. The results, depicted in figures 11 and 12, confirm that all connections found were statistically significant in at least one of the group comparisons (HC vs PD 1–3, HC vs PD 6–12, or PD 1–3 vs PD 6–12). These figures illustrate the strength of each connection across the groups, with significance levels indicated by the symbols described in the corresponding figure captions.

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, PD 1–3 and PD 6–12.

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 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, PD 1–3 and PD 6–12.

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					

4. Discussion

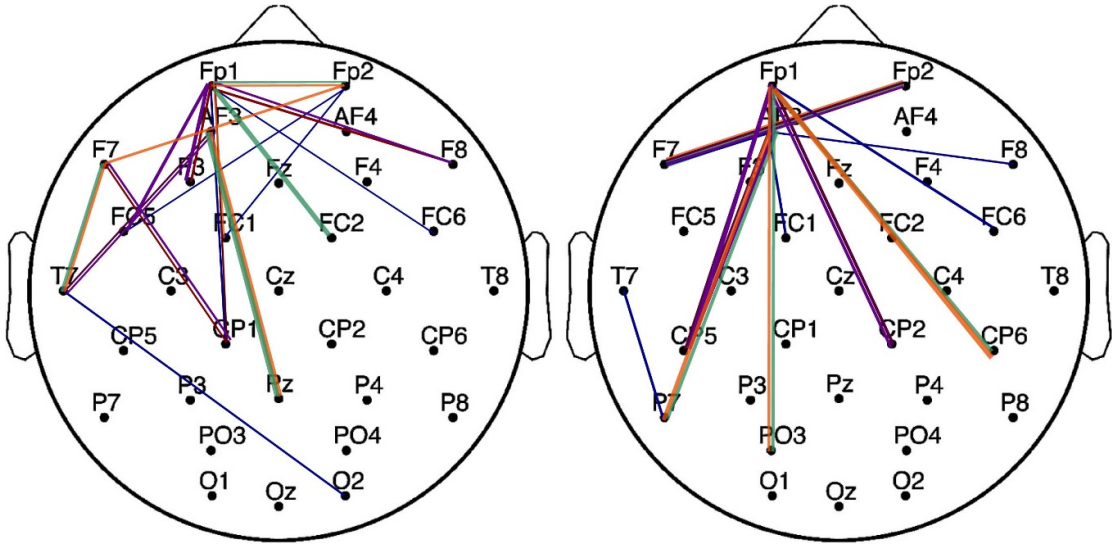
4.1. Binary classification

Among the evaluated connectivity metrics, coherence (Sync) yielded the best performance in the binary classification task. As defined by Bowyer [23], coherence captures both amplitude and phase coupling across a range of frequencies, providing a more comprehensive measure of signal synchronization. This makes it particularly suitable for analyzing brain connectivity in PD, a condition known to alter both spectral power and synchrony in multiple brain regions. The emphasis on interregional signal coupling aligns with the underlying pathophysiology of PD, where disruptions in neural signaling and coordination are hallmark features [28, 96–98]. Therefore, coherence emerges not only as a theoretically appropriate metric but also as an empirically superior one, offering unique insights and enhanced discriminatory power within our ML framework to distinguish PD patients from HCs.

Further, the best ML 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.

4.2. 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 7(c)) and HC suggest a promising discrimination capability of the ML 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



(a) Significant EEG connections in PD 1-3 class (b) Significant EEG connections in PD 6-12 class

Figure 10. Main EEG connections identified through SHAP value methodology and MLP classifier for each correlation metric in PD 1–3 (a) and PD 6–12 (b) classes. The color scheme denotes different metrics: Sync (blue), SC (red), PC (purple), CCA (orange), and LW (green).

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.

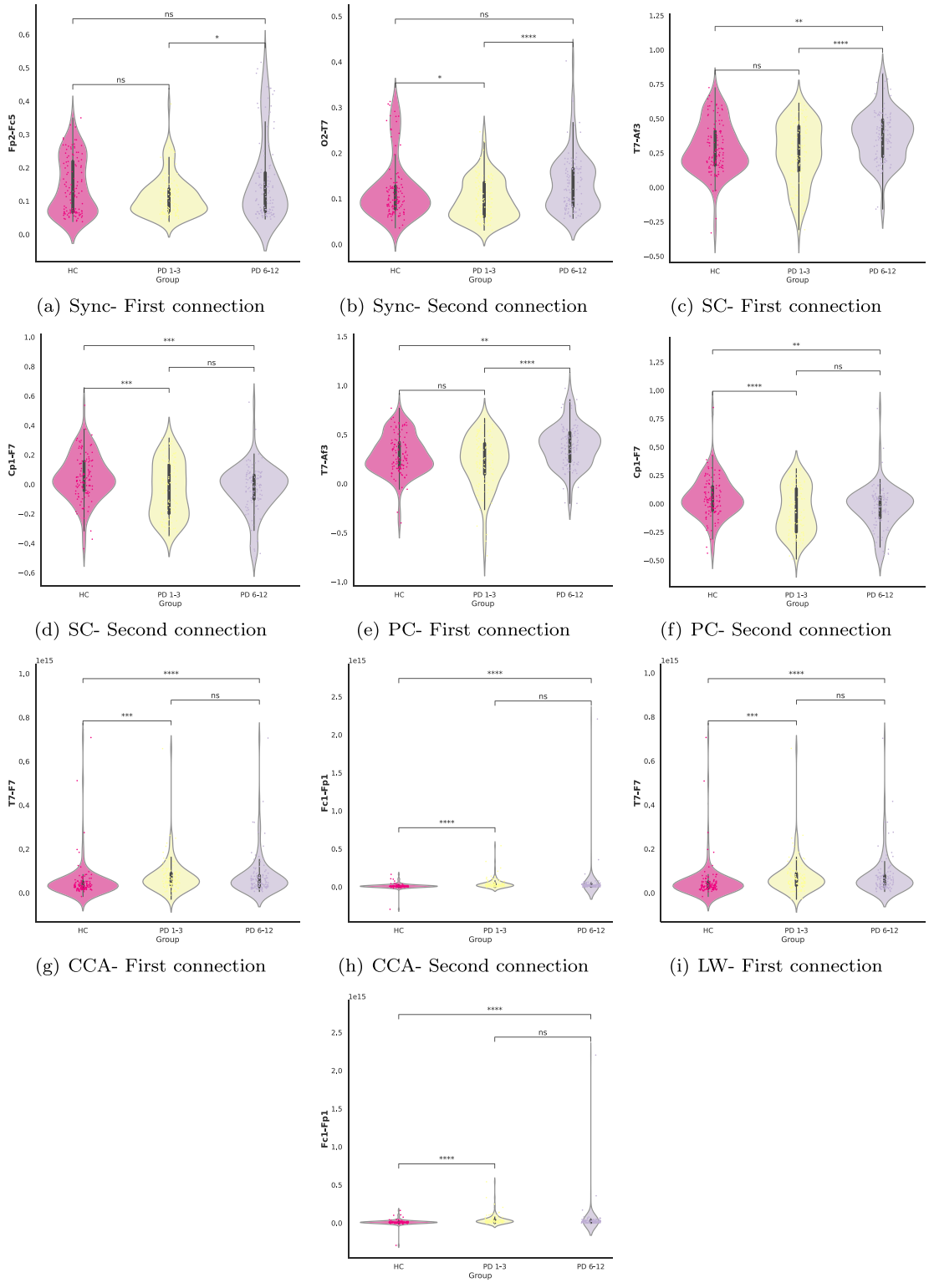
After employing the SHAP value methodology, we identified the most influential connections for the PD 1–3 group. Notably, lower synchronization values in regions such as FP2 and FC5 were found. The frontopolar cortex (FP2) is implicated in various higher-order cognitive processes, including decision-making, executive function, and social cognition [37, 52]. On the other hand, the left frontal cortex (FC5) plays a crucial role in motor planning and execution, as well as language processing [39, 60]. Thus, alterations in synchronization patterns within these regions may signify underlying deficits in both cognitive processing and motor coordination [45, 90]. 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 [41]. 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 stage-related changes in neural circuitry [71].

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 [40, 68].

These findings underscore the complex interplay between neural connectivity patterns, disease progression, and stage-related changes in PD.

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(a)



(a) Sync- First connection (b) Sync- Second connection (c) SC- First connection

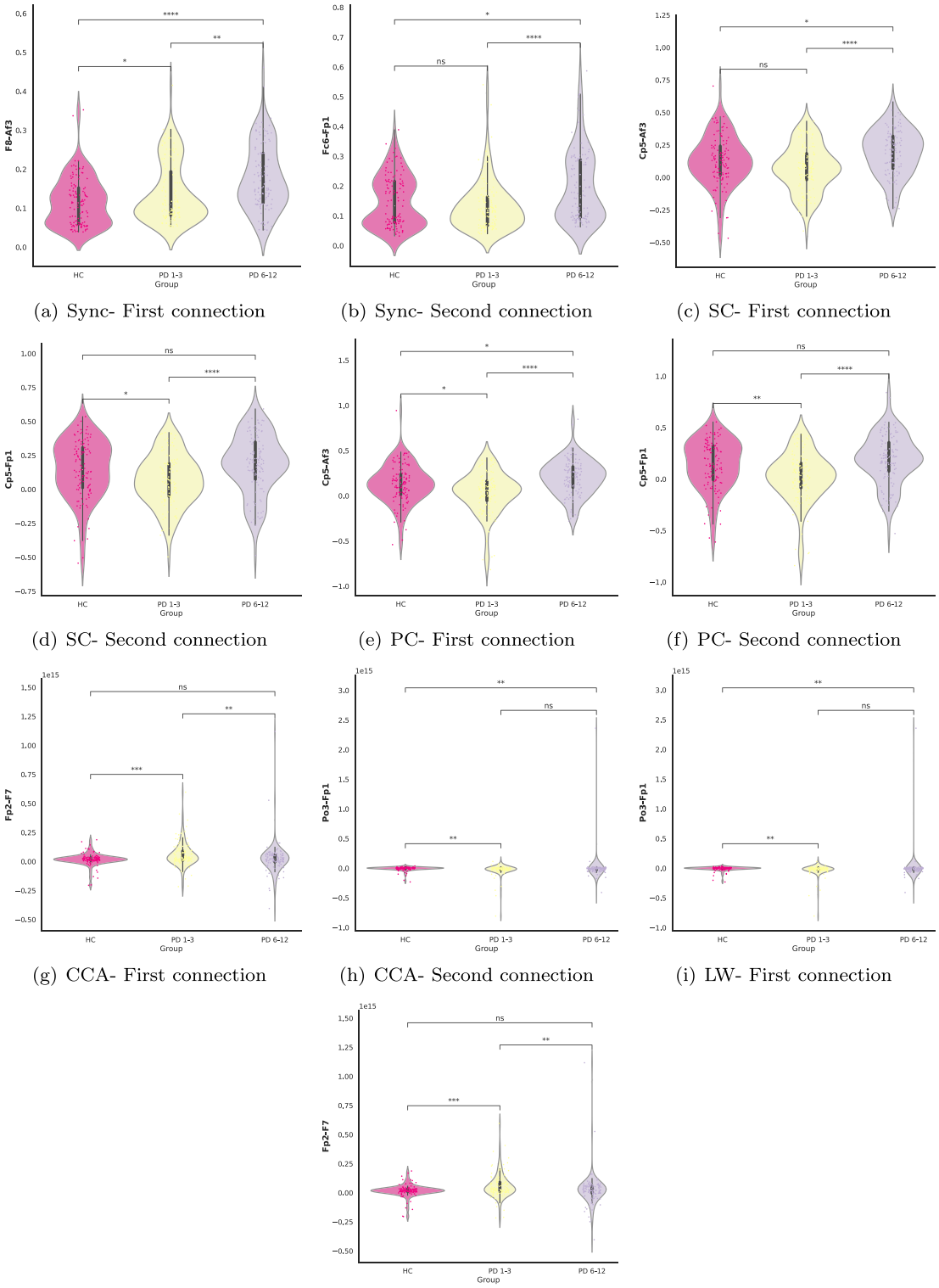
(d) SC- Second connection (e) PC- First connection (f) PC- Second connection

(g) CCA- First connection (h) CCA- Second connection (i) LW- First connection

(j) LW- Second connection

Figure 11. Statitical test of SHAP value methodology's main connection found for PD 1–3. Statistical analysis of the most significant connections identified through the SHAP values methodology for the PD 1–3 group. Each subplot shows the distribution of connection strengths between specific cortical regions for HC, PD 1–3, and PD 6–12. The following symbols indicate the statistical significance between the groups: ns ($5.00e-02 < p \leqslant 1.00e+00$), * ($1.00e-02 < p \leqslant 5.00e-02$), ** ($1.00e-03 < p \leqslant 1.00e-02$), *** ($1.00e-04 < p \leqslant 1.00e-03$), and **** ($p \leqslant 1.00e-04$).

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



(a) Sync- First connection (b) Sync- Second connection (c) SC- First connection

(d) SC- Second connection (e) PC- First connection (f) PC- Second connection

(g) CCA- First connection (h) CCA- Second connection (i) LW- First connection

(j) LW- Second connection

Figure 12. Statitical test of SHAP value methodology's main connection found for PD 6–12. Statistical analysis of the most significant connections identified through the SHAP values methodology for the PD 6–12 group. Each subplot shows the distribution of connection strengths between specific cortical regions for the groups HC, PD 1–3, and PD 6–12. The statistical significance between the groups is indicated by the following symbols: ns ($5.00e-02 < p \leqslant 1.00e+00$), * ($1.00e-02 < p \leqslant 5.00e-02$), ** ($1.00e-03 < p \leqslant 1.00e-02$), *** ($1.00e-04 < p \leqslant 1.00e-03$), and **** ($p \leqslant 1.00e-04$).

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).

Finally, validation of the ML findings was conducted through rigorous statistical tests on the connectivity metrics, mainly focusing on the most significant connections identified using the SHAP values methodology. These analyses revealed that all identified connections were statistically significant in at least one of the group comparisons (HC vs PD 1–3, HC vs PD 6–12, or PD 1–3 vs PD 6–12). This statistical significance across comparisons underscores the robustness of the connections identified by the ML model, further validating their importance in distinguishing between different stages of PD. This further validates the importance of these connections in distinguishing between different stages of PD.

5. 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 critical role of coherence in capturing the synchronized behavior of signals, which scontributes to distinguishing PD patients from controls. Moreover, employing ML 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 relevant connectivity patterns implicated in PD pathophysiology. An intriguing finding was the reduction of synchronization between EEG channels and the progression of the disease.

To ensure the robustness of our findings, we further validated the most significant connections using statistical tests. This validation confirmed that all identified connections were statistically significant in at least one group comparison.

Our analysis suggests changes in neural connectivity patterns across different disease stages. Notably, early-duration PD exhibits similarities with healthy brain function, while later stages manifest distinctive features indicative of disease progression.

However, it is essential to acknowledge the limitations of our methodology, particularly its reliance on small data. Furthermore, the San Diego dataset lacks detail regarding UPDRS assessments, such as whether they were conducted in ON or OFF medication states and the use of non-specialist raters. These factors may affect the reliability of disease staging. Future research should ensure clinically rigorous evaluations by trained professionals and include a broader range of disease durations to understand the impact of these interventions better.

Overall, our study contributes valuable insights into the complex interplay between neural dynamics, disease progression, and stage-related changes in PD, offering new avenues for further research and clinical application.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

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.

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.

Ethical approval

This study utilized the San Diego Parkinson's dataset, which is publicly available and fully de-identified, ensuring participant privacy. All analyses were conducted in accordance with ethical standards and the dataset's usage policies.

ORCID iDs

Caroline L Alves 0000-0003-4708-1330
Francisco Aparecido Rodrigues 0000-0002-0145-5571

References

[1] Al-Beltagi M 2021 Autism medical comorbidities *World J. Clin. Pediatr.* **10** 15
[2] Al-Ezzi A *et al* 2024 Disrupted brain functional connectivity as early signature in cognitively healthy individuals with pathological CSF amyloid/tau *Commun. Biol.* **7** 1037
[3] Aljalal M, Aldosari S A, Molinas M, AlSharabi K and Alturki F A 2022 Detection of parkinson's disease from EEG signals using discrete wavelet transform, different entropy measures and machine learning techniques *Sci. Rep.* **12** 22547
[4] Alves C, Wissel L, Capetian P and Thielemann C 2022a P 55 functional connectivity and convolutional neural networks for automatic classification of EEG data *Clin. Neurophysiol.* **137** e47
[5] Alves C L, Ciba M, Toutain T G d O, Porto J A M, de Sena E P, Thielemann C and Rodrigues F A 2024 On the advances in machine learning and complex network measures to an EEG dataset from DMT experiments *J. Phys. Complex.* **5** 015002
[6] Alves C L, Cury R G, Roster K, Pineda A M, Rodrigues F A, Thielemann C and Ciba M 2022 Application of machine learning and complex network measures to an EEG dataset from ayahuasca experiments *PLoS One* **17** e0277257
[7] Alves C L, Pineda A M, Roster K, Thielemann C and Rodrigues F A 2022c EEG functional connectivity and deep learning for automatic diagnosis of brain disorders: Alzheimer's disease and schizophrenia *J. Phys. Complex.* **3** 025001
[8] Alves C L, Toutain T G d O, de Carvalho Aguiar P, Pineda A M, Roster K, Thielemann C, Porto J A M and Rodrigues F A 2023 Diagnosis of autism spectrum disorder based on functional brain networks and machine learning *Sci. Rep.* **13** 8072
[9] Alves C L, Toutain T G d O, Porto J A M, Aguiar P M d C, Pineda A, Rodrigues F A, de Sena E P and Thielemann C 2023 Analysis of functional connectivity using machine learning and deep learning in different data modalities from individuals with schizophrenia *J. Neural Eng.* **20** 056025
[10] Amoroso N, La Rocca M, Monaco A, Bellotti R and Tangaro S 2018 Complex networks reveal early MRI markers of parkinson's disease *Med. Image Anal.* **48** 12–24
[11] Arcadu F, Benmansour F, Maunz A, Willis J, Haskova Z and Prunotto M 2020 Author correction: deep learning algorithm predicts diabetic retinopathy progression in individual patients *npj Dig. Med.* **3** 1–6
[12] Avuçlu E and Elen A 2020 Evaluation of train and test performance of machine learning algorithms and parkinson diagnosis with statistical measurements *Med. Biol. Eng. Comput.* **58** 2775–88
[13] Aydın S 2024 Alzhemimer's disease is characterized by lower segregation in resting-state eyes-closed EEG *J. Med. Biol. Eng.* **44** 894–902
[14] Aydın S, Cetin F H, Uytun M Ç, Babadagi Z, Gueven A S and Işık Y 2022 Comparison of domain specific connectivity metrics for estimation brain network indices in boys with ADHD-C *Biomed. Signal Process. Control* **76** 103626
[15] Balestrino R and Schapira A 2020 Parkinson disease *Eur. J. Neurol.* **27** 27–42
[16] Benesty J, Chen J, Huang Y and Cohen I 2009 Pearson correlation coefficient *Noise Reduction in Speech Processing* (Springer) pp 1–4
[17] Bengfort B and Bilbro R 2019 Yellowbrick: visualizing the scikit-learn model selection process *J. Open Source Softw.* **4** 1075
[18] Bengio Y and Grandvalet Y 2003 No unbiased estimator of the variance of k-fold cross-validation *Advances in Neural Information Processing Systems* p 16
[19] Betzel R F, Byrge L, He Y, Goñi J, Zuo X-N and Sporns O 2014 Changes in structural and functional connectivity among resting-state networks across the human lifespan *Neuroimage* **102** 345–57
[20] Bischl B *et al* 2023 Hyperparameter optimization: foundations, algorithms, best practices and open challenges *WIREs Data. Mining. Knowl. Discov.* **13** 1484
[21] Bisong E and Bisong E 2019 Introduction to scikit-learn *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners* pp 215–29
[22] Blum A, Kalai A and Langford J 1999 Beating the hold-out: bounds for k-fold and progressive cross-validation *Proc. 12th Annual Conf. on Computational Learning Theory* pp 203–8
[23] Bowyer S M 2016 Coherence a measure of the brain networks: past and present *Neuropsychiatr. Electrophysiol.* **2** 1–12
[24] Bracher-Smith M, Crawford K and Escott-Price V 2021 Machine learning for genetic prediction of psychiatric disorders: a systematic review *Mol. Psychiatry* **26** 70–79
[25] Brownlee J 2019 How to choose a feature selection method for machine learning *Machine Learning Mastery* vol 10
[26] Chow R, Rabi R, Paracha S, Hasher L, Anderson N D and Alain C 2022 Default mode network and neural phase synchronization in healthy aging: a resting state EEG study *Neuroscience* **485** 116–28
[27] Collin G and Van Den Heuvel M P 2013 The ontogeny of the human connectome: development and dynamic changes of brain connectivity across the life span *Neuroscientist* **19** 616–28
[28] Conte A, Khan N, Defazio G, Rothwell J C and Berardelli A 2013 Pathophysiology of somatosensory abnormalities in parkinson disease *Nat. Rev. Neurol.* **9** 687–97
[29] Dukart J, Weis S, Genon S and Eickhoff S B 2021 Towards increasing the clinical applicability of machine learning biomarkers in psychiatry *Nat. Human Behav.* **5** 431–2
[30] Edde M, Leroux G, Altena E and Chanraud S 2021 Functional brain connectivity changes across the human life span: from fetal development to old age *J. Neurosci. Res.* **99** 236–62
[31] Edlow B L, Claassen J, Schiff N D and Greer D M 2021 Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies *Nat. Rev. Neurol.* **17** 135–56

[32] Ferri C, Hernández-Orallo J and Salido M A 2003 Volume under the roc surface for multi-class problems *European Conf. on Machine Learning* (Springer) pp 108–20
[33] French C C and Beaumont J G 1984 A critical review of EEG coherence studies of hemisphere function *Int. J. Psychophys.* **1** 241–54
[34] Gadalla A A *et al* 2019 Identification of clinical and urine biomarkers for uncomplicated urinary tract infection using machine learning algorithms *Sci. Rep.* **9** 19694
[35] George J S, Strunk J, Mak-McCully R, Houser M, Poizner H and Aron A R 2013 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** 261–70
[36] Geron A 2022 *Hands-on Machine Learning With Scikit-Learn, Keras and Tensorflow* (O'Reilly Media, Inc.)
[37] Gilbert S J, Spengler S, Simons J S, Steele J D, Lawrie S M, Frith C D and Burgess P W 2006 Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis *J. Cogn. Neurosci.* **18** 932–48
[38] Gramfort A, Luessi M, Larson E, Engemann D A, Strohmeier D, Brodbeck C, Parkkonen L and Hämäläinen M S 2014 Mne software for processing meg and EEG data *Neuroimage* **86** 446–60
[39] Hagoort P 2014 Nodes and networks in the neural architecture for language: Broca's region and beyond *Curr. Opin. Neurobiol.* **28** 136–41
[40] Hammond C, Bergman H and Brown P 2007 Pathological synchronization in parkinson's disease: networks, models and treatments *Trends Neurosci.* **30** 357–64
[41] Hanganu A *et al* 2014 Mild cognitive impairment is linked with faster rate of cortical thinning in patients with parkinson's disease longitudinally *Brain* **137** 1120–9
[42] Hardoon D R and Shawe-Taylor J 2011 Sparse canonical correlation analysis *Mach. Learn.* **83** 331–53
[43] Hoffmann A, Jäger L, Werhahn K, Jaschke M, Noachtar S and Reiser M 2000 Electroencephalography during functional echo-planar imaging: detection of epileptic spikes using post-processing methods *Magn. Reson. Med.* **44** 791–8
[44] Ito Y, Unagami M, Yamabe F, Mitsui Y, Nakajima K, Nagao K and Kobayashi H 2021 A method for utilizing automated machine learning for histopathological classification of testis based on johnsen scores *Sci. Rep.* **11** 1–11
[45] Jones S R, Pritchett D L, Stufflebeam S M, Hämäläinen M and Moore C I 2007 Neural correlates of tactile detection: a combined magnetoencephalography and biophysically based computational modeling study *J. Neurosci.* **27** 10751–64
[46] Kaidery N A, Tarannum S and Thomas B 2013 Epigenetic landscape of parkinson's disease: emerging role in disease mechanisms and therapeutic modalities *Neurotherapeutics* **10** 698–708
[47] Kapoor S and Narayanan A 2023 Leakage and the reproducibility crisis in machine-learning-based science *Patterns* **4** 100804
[48] Khare S K, Bajaj V and Acharya U R 2021 Detection of parkinson's disease using automated tunable q wavelet transform technique with EEG signals *Biocybern. Biomed. Eng.* **41** 679–89
[49] Kilzheimer A, Hentrich T, Burkhardt S and Schulze-Hentrich J M 2019 The challenge and opportunity to diagnose parkinson's disease in midlife *Front. Neurol.* **10** 1328
[50] Kim D W, Lee S, Kwon S, Nam W, Cha I-H and Kim H J 2019 Deep learning-based survival prediction of oral cancer patients *Sci. Rep.* **9** 6994
[51] Kim J, Lee J, Park E and Han J 2020 A deep learning model for detecting mental illness from user content on social media *Sci. Rep.* **10** 1–6
[52] Koechlin E and Summerfield C 2007 An information theoretical approach to prefrontal executive function *Trends Cogn. Sci.* **11** 229–35
[53] Kohavi R *et al* 1995 A study of cross-validation and bootstrap for accuracy estimation and model selection *Ijcai* vol 14 pp 1137–45
[54] Krittanawong C, Virk H U H, Kumar A, Aydar M, Wang Z, Stewart M P and Halperin J L 2021 Machine learning and deep learning to predict mortality in patients with spontaneous coronary artery dissection *Sci. Rep.* **11** 1–10
[55] Kuhn M and Johnson K *et al* 2013 *Applied Predictive Modeling* vol 26 (Springer)
[56] Lashgari E, Liang D and Maoz U 2020 Data augmentation for deep-learning-based electroencephalography *J. Neurosci. Methods* **346** 108885
[57] Ledoit O and Wolf M 2012 Nonlinear shrinkage estimation of large-dimensional covariance matrices *Ann. Stat.* **40** 1024–60
[58] Li R C, Asch S M and Shah N H 2020 Developing a delivery science for artificial intelligence in healthcare *npj Digit. Med.* **3** 1–3
[59] Li Y, Nowak C M, Pham U, Nguyen K and Bleris L 2021 Cell morphology-based machine learning models for human cell state classification *npj Syst. Biol. Appl.* **7** 1–9
[60] Liakakis G, Nickel J and Seitz R 2011 Diversity of the inferior frontal gyrus—a meta-analysis of neuroimaging studies *Behav. Brain Res.* **225** 341–7
[61] Loh H W, Ooi C P, Palmer E, Barua P D, Dogan S, Tuncer T, Baygin M and Acharya U R 2021 Gaborpdnet: Gabor transformation and deep neural network for parkinson's disease detection using EEG signals *Electronics* **10** 1740
[62] Lopez-Martin M, Nevado A and Carro B 2020 Detection of early stages of alzheimer's disease based on meg activity with a randomized convolutional neural network *Artif. Intell. Med.* **107** 101924
[63] Lubinski D 2004 Introduction to the special section on cognitive abilities: 100 years after spearman's (1904)'general intelligence,'objectively determined and measured *J. Pers. Soc. Psychol.* **86** 96
[64] Lundberg S M and Lee S-I 2017 A unified approach to interpreting model predictions *Proc. 31st Int. Conf. on Neural Information Processing Systems* pp 4768–77
[65] Mahesh T *et al* 2023 The stratified k-folds cross-validation and class-balancing methods with high-performance ensemble classifiers for breast cancer classification *Healthcare Analytics* **4** 100247
[66] Matell M S, Meck W H and Nicolelis M A 2003 Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons *Behav. Neurosci.* **117** 760
[67] Mincholé A and Rodriguez B 2019 Artificial intelligence for the electrocardiogram *Nat. Med.* **25** 22–23
[68] Moazami-Goudarzi M, Sarnthein J, Michels L, Moukhtieva R and Jeanmonod D 2008 Enhanced frontal low and high frequency power and synchronization in the resting EEG of parkinsonian patients *Neuroimage* **41** 985–97
[69] Müller-Nedebock A C, Dekker M C, Farrer M J, Hattori N, Lim S-Y, Mellick G D, Rektorová I, Salama M, Schuh A F and Stoessl A J *et al* 2023 Different pieces of the same puzzle: a multifaceted perspective on the complex biological basis of parkinson's disease *npj Parkinson's Dis.* **9** 110
[70] Nakisa B, Rastgoo M N, Rakotonirainy A, Maire F and Chandran V 2020 Automatic emotion recognition using temporal multimodal deep learning *IEEE Access* **8** 225463–74
[71] Olde Dubbelink K T, Hillebrand A, Stoffers D, Deijen J B, Twisk J W, Stam C J and Berendse H W 2014 Disrupted brain network topology in parkinson's disease: a longitudinal magnetoencephalography study *Brain* **137** 197–207

[72] Pant A and Kumar A 2024 Hanning fir window filtering analysis for EEG signals *Biomed. Anal.* **1** 111–23
[73] Park Y and Kellis M 2015 Deep learning for regulatory genomics *Nat. Biotechnol.* **33** 825–6
[74] Patel D, Kher V, Desai B, Lei X, Cen S, Nanda N, Gholamrezanezhad A, Duddalwar V, Varghese B and Oberai A A 2021 Machine learning based predictors for covid-19 disease severity *Sci. Rep.* **11** 1–7
[75] Pedregosa F *et al* 2011 Scikit-learn: machine learning in python *J. Mach. Learn. Res.* **12** 2825–30
[76] Poewe W, Seppi K, Tanner C M, Halliday G M, Brundin P, Volkmann J, Schrag A-E and Lang A E 2017 Parkinson disease *Nat. Rev. Dis. Primers* **3** 1–21
[77] Postuma R B and Berg D 2019 Prodromal parkinson's disease: the decade past, the decade to come *Mov. Disorders* **34** 665–75
[78] Prottasha N J, Murad S A, Muzahid A J M, Rana M, Kowsher M, Adhikary A, Biswas S and Bairagi A K 2023 Impact learning: a learning method from feature's impact and competition *J. Comput. Sci.* **69** 102011
[79] Qiu L, Li J, Zhong L, Feng W, Zhou C and Pan J 2024 A novel EEG-based parkinson's disease detection model using multiscale convolutional prototype networks *IEEE Trans. Instrum. Meas.*
[80] Raschka S 2015 *Python Machine Learning* (Packt publishing ltd)
[81] Rashidi H H, Sen S, Palmieri T L, Blackmon T, Wajda J and Tran N K 2020 Early recognition of burn-and trauma-related acute kidney injury: a pilot comparison of machine learning techniques *Sci. Rep.* **10** 1–9
[82] Raza C and Anjum R *et al* 2019 Parkinson's disease: mechanisms, translational models and management strategies *Life sci.* **226** 77–90
[83] Richhariya B, Tanveer M, Rashid A H and Initiative A D N *et al* 2020 Diagnosis of alzheimer's disease using universum support vector machine based recursive feature elimination (USVM-RFE) *Biomed. Signal Process. Control* **59** 101903
[84] Rizvi S Q A, Wang G, Khan A, Hasan M K, Ghazal T M and Khan A U R 2023 Classifying parkinson's disease using resting state electroencephalogram signals and u^{EN}-pdnet *IEEE Access*
[85] Rockhill A P et al 2020 Uc san diego resting state EEG data from patients with parkinson's disease
[86] Sato M, Morimoto K, Kajihara S, Tateishi R, Shiina S, Koike K and Yatomi Y 2019 Machine-learning approach for the development of a novel predictive model for the diagnosis of hepatocellular carcinoma *Sci. Rep.* **9** 1–7
[87] Schapira A H, Chaudhuri K R and Jenner P 2017 Non-motor features of parkinson disease *Nat. Rev. Neurosci.* **18** 435–50
[88] Shorten C, Khoshgoftaar T M and Furht B 2021 Text data augmentation for deep learning *J. Big Data* **8** 1–34
[89] Singh V, Pencina M, Einstein A J, Liang J X, Berman D S and Slomka P 2021 Impact of train/test sample regimen on performance estimate stability of machine learning in cardiovascular imaging *Sci. Rep.* **11** 14490
[90] Smith N J and Kutas M 2015 Regression-based estimation of erp waveforms: i. the rerp framework *Psychophysiology* **52** 157–68
[91] Soltész F, Szűcs D, Leong V, White S and Goswami U 2013 Differential entrainment of neuroelectric delta oscillations in developmental dyslexia *PLoS One* **8** e76608
[92] Stam C J, Montez T, Jones B, Rombouts S, Van Der Made Y, Pijnenburg Y A and Scheltens P 2005 Disturbed fluctuations of resting state EEG synchronization in alzheimer's disease *Clin. Neurophysiol.* **116** 708–15
[93] Strehl A, Ghosh J and Mooney R 2000 Impact of similarity measures on web-page clustering *Workshop on Artificial Intelligence for web Search (AAAI 2000)* vol 58 p 64
[94] Su Q *et al* 2022 Faecal microbiome-based machine learning for multi-class disease diagnosis *Nat. Commun.* **13** 6818
[95] Subramanian I, Mathur S, Oosterbaan A, Flanagan R, Keener A M and Moro E 2022 Unmet needs of women living with parkinson's disease: gaps and controversies *Mov. Disorders* **37** 444–55
[96] Surmeier D J, Obeso J A and Halliday G M 2017 Selective neuronal vulnerability in parkinson disease *Nat. Rev. Neurosci.* **18** 101–13
[97] Swinnen S P 2002 Intermanual coordination: from behavioural principles to neural-network interactions *Nat. Rev. Neurosci.* **3** 348–59
[98] Takakusaki K, Takahashi M, Noguchi T and Chiba R 2023 Neurophysiological mechanisms of gait disturbance in advanced parkinson's disease patients *Neurol. Clin. Neurosci.;* **11** 201–17
[99] Thara D and PremaSudha B *et al* 2019 Auto-detection of epileptic seizure events using deep neural network with different feature scaling techniques *Pattern Recognit. Lett.* **128** 544–50
[100] Tolkach Y, Dohmgörgen T, Toma M and Kristiansen G 2020 High-accuracy prostate cancer pathology using deep learning *Nat. Mach. Intell.* **2** 411–8
[101] Tsuboi T, Satake Y, Hiraga K, Yokoi K, Hattori M, Suzuki M, Hara K, Ramirez-Zamora A, Okun M S and Katsuno M 2022 Effects of mao-b inhibitors on non-motor symptoms and quality of life in parkinson's disease: a systematic review *npj Parkinson's Dis.* **8** 75
[102] Varma S and Simon R 2006 Bias in error estimation when using cross-validation for model selection *BMC Bioinform.* **7** 1–8
[103] Vigário R N 1997 Extraction of ocular artefacts from EEG using independent component analysis *Electroencephalogr. Clin. Neurophysiol.* **103** 395–404
[104] Wu Z, Chen H, Ke S, Mo L, Qiu M, Zhu G, Zhu W and Liu L 2023 Identifying potential biomarkers of idiopathic pulmonary fibrosis through machine learning analysis *Sci. Rep.* **13** 16559
[105] Yang M, Wu Y, Yang X-b, Liu T, Zhang Y, Zhuo Y, Luo Y and Zhang N 2023 Establishing a prediction model of severe acute mountain sickness using machine learning of support vector machine recursive feature elimination *Sci. Rep.* **13** 4633
[106] Yu X, Pang W, Xu Q and Liang M 2020 Mammographic image classification with deep fusion learning *Sci. Rep.* **10** 1–11
[107] Zhang W, Sweeney J A, Bishop J R, Gong Q and Lui S 2023 Biological subtyping of psychiatric syndromes as a pathway for advances in drug discovery and personalized medicine *Nat. Mental Health* **1** 88–99
[108] Zhang Y *et al* 2021 Identification of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography *Nat. Biomed. Eng.* **5** 309–23
[109] Zhong Z, Yuan X, Liu S, Yang Y and Liu F 2021 Machine learning prediction models for prognosis of critically ill patients after open-heart surgery *Sci. Rep.* **11** 1–10