Open Access

Research Article

PRNANO

Precis. Nanomed. 2025, 8(2):1464-1472

ANDOVER HOUSE

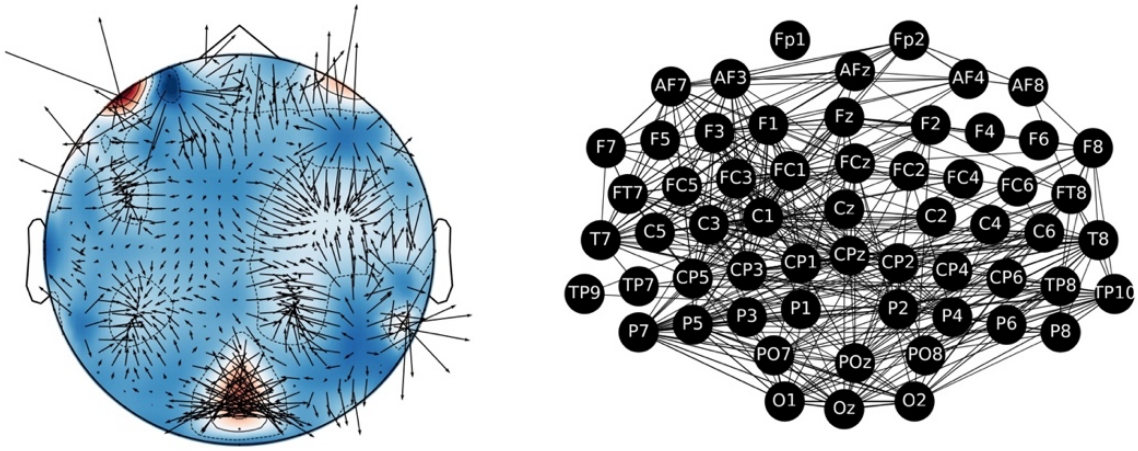
Diagnosis of Alzheimer's Disease, Parkinson's Disease, Frontotemporal Dementia, and Paranoid Schizophrenia via Complex Network Analysis of EEG Data

Miguel Angel Vargas Cruz*

Grupo Alianza Empresarial, Montecito 38 Piso 25 Oficina 15, Colonia Nápoles CP 03810, Ciudad de México.

Submitted: February 13, 2024 Accepted: March 27, 2024 Published: April 1, 2024

Graphical Abstract



Abstract

A novel diagnostic method that employs complex network analysis using electroencephalogram (EEG) data is presented, which achieves exceptional classification accuracy across a range of mental disorders. Our method demonstrates 96% accuracy for paranoid schizophrenia, 96% for frontotemporal dementia, 97% for Alzheimer's disease, and 96% for Parkinson's disease, highlighting the robustness and versatility of this approach. The method's efficacy lies in its ability to handle various data sets, including diverse channel configurations such as the 10-20 extended system for Parkinson's disease, thereby ensuring broad applicability. Despite utilizing different data formats and sizes, the approach consistently achieves high precision. The method's simplicity, computational efficiency, and scalability offer a significant advancement in neurodiagnostic applications.

Keywords: Alzheimer's disease, Parkinson's disease, frontotemporal dementia, schizophrenia, diagnosis, EEG, complex networks, discrete Fourier transform, KMeans, machine learning, classification accuracy, computational efficiency, random forest, GridSearchCV

Purpose, Rationale, and Limitations

This study explores the potential of a versatile EEG-based method, enhanced with KMeans clustering, for diagnosing various mental health conditions. Previous work demonstrated a 93% accuracy in diagnosing schizophrenia[1] using a similar framework. Incorporating KMeans clustering has now further elevated this accuracy to 96%. The method has been successfully

* Email: miguelangel@grupoalianzaempresarial.com

License: CC BY-NC 4.0

applied to Alzheimer's disease (AD), paranoid schizophrenia (PS), frontotemporal dementia (FTD), and Parkinson's disease (PD), providing unprecedented diagnostic precision across multiple diseases. The primary limitation of the current study is the relatively small size of the datasets; however, these sets have been utilized in numerous other studies and have been proven effective in distinguishing between patients and controls. Despite the small sample sizes, we aim to maximize the method's generalization by applying a rigorous, subject-by-subject validation approach. While the results are promising, larger and more diverse datasets are needed to confirm these findings.

Introduction

This method works by modeling electroencephalogram (EEG) signals as complex networks, exploring the causal relationships between brain regions, and detecting subtle abnormalities specific to each condition. Notably, the approach is not limited to classification but also provides deeper insights into the functional connectivity of the brain, offering opportunities for further research into the pathophysiology of these diseases. Its low computational cost, flexibility, scalability, and broad applicability make it a promising tool for the clinical and research setting. Furthermore, the method's success is not confined to disease diagnosis alone. It has been applied to cognitive performance analysis in preliminary tests, distinguishing individuals with high and low mathematical proficiency. Although this application is promising, we opted to focus on the diagnostic potential in this paper.

Classification accuracies for EEG-based disease detection vary widely in the literature, depending on the pathology and the chosen technique. For instance, reported accuracies for AD have reached approximately 83.28% using convolution-transformer architectures[3] on the dataset employed in our study. Parkinson's disease classification has been validated at 85.7% with certain linear predictive coding approaches[4]. In the case of PS, we conducted a detailed comparison[1] to demonstrate the precision of the proposed method, and finally, there are no specific methods for frontotemporal dementia. While we acknowledge that some studies report high accuracies, many such results are difficult to verify due to the absence of open-source code, publicly available datasets, or reproducible methodological details. Therefore, we avoid drawing direct comparisons with unverifiable approaches. Instead, we emphasize the robust performance of our pipeline across multiple disease groups, validated through subject-by-subject analysis using publicly available data and code. This transparency and reproducibility facilitate meaningful comparisons and underscore the potential translational impact of our findings.

Most of the methods were already described in the first publication on PS [1], but here, we present a significant (albeit simple) improvement. The previous article introduced a key element that provides a deeper understanding and also aids in causal search. The previous method required more channels and had lower precision, but it also pointed to areas worth investigating. While our current method only needs F7 and Fz to achieve 96% accuracy, it's still helpful to examine O1, O2, P4, T5, Pz, and T6, as was done in the original method. In fact, T5 was crucial in developing the original method because it exhibited very specific behavior in some cases that were hard to diagnose using other methods, both those we replicated and those we developed. We can reach similar results through other methods we developed, but they come with added complexity.

To give more context on how the method is evolving, the complex networks approach has been very helpful in clustering and modularity. Spotting certain differences can be easier or harder depending on the threshold used to visualize correlations between channels. This is why T5 was so valuable in the schizophrenia analysis – it acted as a pivot, helping identify relevant behaviors and enabling discrimination. With a 50% threshold, it served as an important discriminator, but if we rely solely on visual analysis, it falls short of our goal. Therefore, it's necessary to turn to methods that may be less useful in causal search, although not completely black-box like others where large datasets are input to extract patterns to obtain accuracy. This approach helps identify specific correlations, enabling us to diagnose and assess how a treatment works in great detail. This allows us to design specific experiments and treatments that can track correlation changes.

License: CC BY-NC 4.0

In terms of modularity, we're not just looking for the densest connections; we can also investigate those that aren't as dense. This can provide many new leads in the search for causality and be applied to almost anything interesting. For example, we can ask why there is more or less activity in misdiagnosed individuals, which is why T5 was crucial to the method.

Additionally, this study serves as a stress test for the methodology, now using EEG recordings with a 10-20 extended system and 59 channels for PD, compared to the 19 channels used for AD, FTD, and schizophrenia. The substantial number of channels has resulted in a much higher computational load, requiring 1681 analyses for the 59 channels, as opposed to the 161 needed for the 19 channels. This also opens up the possibility for more combinations but also demands significantly more testing.

Experimental Design

The method relies on the preprocessing and analyzing of EEG signals, creating adjacency matrices from the network of connections between selected brain regions. The feature extraction includes Discrete Fourier Transform (DFT) and spectral eigenvalues, incorporating KMeans clustering to enhance classification accuracy.

The Open Datasets we analyzed contain balanced groups of patients and controls and are publicly available:

PS dataset[2]. (19 channels, file format: .edf, selected channels: F7 and Fz, 299 seconds of EEG)

The study included 14 patients (seven males: 27.9 ± 3.3 years, seven females: 28.3 ± 4.1 years) diagnosed with PS, who were hospitalized at the Institute of Psychiatry and Neurology in Warsaw, Poland, and 14 healthy controls (seven males: 26.8 ± 2.9 years, seven females: 28.7 ± 3.4 years). The patients fulfilled the criteria for PS (F20.0) according to the International Classification of Diseases ICD–10. The Institute of Psychiatry and Neurology Ethics Committee approved the study protocol in Warsaw. All participants received a written description of the study protocol and gave their written consent to participate. Inclusion criteria included a minimum age of 18, an ICD-10 diagnosis of F20.0, and a medication washout period of at least seven days. Exclusion criteria were pregnancy, organic brain disorders, serious neurological conditions (such as epilepsy, AD, or PD), the presence of a general medical condition, and very early stages of schizophrenia, such as the first episode of the illness. The control group was age- and gender-matched with the 14 patients who completed the study.

AD and FTD dataset[3]

AD (19 channels, file format: .set, selected channels: O2 and Cz, 300 seconds of EEG) and FTD (19 channels, file format: .set, selected channels: F7 and F3, 300 seconds of EEG). A total of 36 participants were diagnosed with AD (AD group), 23 with FTD (FTD group), and 29 were healthy controls (CN group). Participants' cognitive and neuropsychological status was assessed using the Mini-Mental State Examination (MMSE), with scores ranging from 0 to 30, where lower MMSE scores indicate more severe cognitive decline. The median disease duration was 25 months, with an interquartile range (IQR) of 24–28.5 months. No dementia-related comorbidities were reported in the AD group. The average MMSE score for the AD group was 17.75 (SD = 4.5), for the FTD group, it was 22.17 (SD = 8.22), and for the CN group, it was 30. The mean age for the AD group was 66.4 years (SD = 7.9); for the FTD group, it was 63.6 years (SD = 8.2); and for the CN group, it was 67.9 years (SD = 5.4).

EEG recordings were obtained from the 2nd Department of Neurology at AHEPA General Hospital in Thessaloniki by a team of experienced neurologists. A Nihon Kohden EEG 2100 device was used for the recordings, employing 19 scalp electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) according to the 10-20 international system, as well as 2 reference electrodes (A1 and A2) on the mastoids for impedance checks. Each recording followed the clinical protocol, with participants seated and their eyes closed. Skin impedance was ensured to be below 5kΩ before each recording. The sampling rate was 500 Hz with a 10 μV/mm resolution. The recording montage used combined anterior-posterior bipolar and referential montages, with Cz as the common reference. The referential montage was included in the dataset. The amplifier's parameters for the recordings were sensitivity: 10 μV/mm, time constant: 0.3 s, and a high-frequency filter set to 70 Hz. Recording durations averaged 13.5 minutes for the

License: CC BY-NC 4.0

AD group (range: 5.1–21.3 minutes), 12 minutes for the FTD group (range: 7.9–16.9 minutes), and 13.8 minutes for the CN group (range: 12.5–16.5 minutes). In total, 485.5 minutes of AD data, 276.5 minutes of FTD data, and 402 minutes of CN data were collected.

The EEG recordings were initially exported in *.eeg* format and converted to the BIDS-compatible *.set* format for inclusion in the dataset. Automatic annotations from the Nihon Kohden EEG device marking artifacts (e.g., muscle activity, blinking, swallowing) were not included for language compatibility reasons (these annotations can be found in the preprocessed dataset in the "derivatives" folder). The raw EEG data is located in folders labeled "sub-0XX", and the preprocessed and denoised recordings are in corresponding folders within the "derivatives" subfolder. The preprocessing pipeline involved first applying a Butterworth band-pass filter (0.5–45 Hz) and re-referencing the signals to A1-A2. Then, the Artifact Subspace Reconstruction method, included in the EEGLab Matlab software, was used to remove artifact-contaminated data exceeding the 0.5-second window standard deviation of 17. Afterward, Independent Component Analysis (ICA) using the RunICA algorithm was performed, transforming the 19 EEG signals into 19 ICA components. ICA components identified as "eye artifacts" or "jaw artifacts" by the automatic ICLabel routine in EEGLab were rejected. Despite recordings being conducted under eyes-closed resting conditions, eye movement artifacts were still observed in some EEG recordings.

Parkinson's disease dataset[4] (59 channels, file format: .vhdr, selected channels: Fz and T8, 120 seconds of EEG).

While previous applications of the method benefited from longer EEG recordings, the PD dataset consists of shorter 120-second segments, which introduced additional challenges for generalization. The method is inherently more effective with extended EEG recordings, and the shorter duration requires careful attention.

EEG recordings were collected from 14 PD patients and 14 University of Iowa (UI) controls. Additionally, EEG data from OFF medication sessions were obtained for the 14 PD patients, recorded 12 hours after their last dopaminergic medication dose. Control participants were matched to PD patients by age and sex, and there were no differences in education or premorbid intelligence. Ethical approval was obtained from the Institutional Review Board of UI. All participants provided written informed consent and were compensated $40/hour. PD patients completed neuropsychological and questionnaire assessments in the ON state, and motor scores were assessed using the UPDRS.

Resting state EEG recordings for the Iowa participants were performed only in the eyes-open condition. The goal was to establish a reproducible and generalizable EEG protocol, with eyes-open being a standard clinical EEG assessment condition that is not influenced by specific abilities and produces diverse features for classification.

EEG data were recorded using 64-channel Brain Vision systems, with a sampling rate of 500 Hz and a frequency range of 0.1–100 Hz. For reference, Pz was used. Eye blink artifacts were removed through ICA. The Iowa dataset consisted of EEG recordings in the eyes-open condition, which were analyzed to determine the most effective variant for classification.

Hypothesis

The hypothesis underpinning this work is that integrating KMeans clustering to the original method[1] with complex network analysis will improve the diagnostic accuracy of EEG data for a broader range of mental health disorders. By applying this enhanced methodology across multiple conditions, we hypothesize that the approach will yield high precision and robustness, supporting its versatility for various clinical applications.

Methods

The approach adopted in this study follows a robust, multi-stage pipeline designed to handle electroencephalographic (EEG) data from various clinical conditions. Throughout all stages, emphasis is placed on reproducibility, scalability, and computational efficiency. A pivotal feature of this pipeline is that it remains largely uniform across all datasets, requiring only minor modifications (such as channel selection) to

License: CC BY-NC 4.0

accommodate different recording protocols and hardware setups.

Data access and preprocessing

Each dataset is accessed in its native file format—.edf, .set, or .vhdr—to preserve the original structure and metadata. Once the data files are loaded, we select the channels with better results, testing every channel combination to avoid bias. This selection typically focuses on frontal or parietal electrodes, as these areas often capture disease-specific signatures. Next, the data are partitioned into relatively long segments (e.g., 120 or 300 seconds), ensuring that potentially relevant transient and steady-state EEG dynamics are included. Each segment undergoes a standardized z-score scaling, removing the mean and scaling to unit variance across the samples in that segment. This normalization step mitigates amplitude-related variability among different subjects and channels, allowing subsequent connectivity estimates to be driven by underlying signal structure rather than raw amplitude differences.

Complex network feature extraction

After preprocessing, each EEG segment is transformed into a network representation—a process rooted in graph theory. Specifically, we compute pairwise correlation coefficients among the selected EEG channels, using either Pearson or Spearman correlation, depending on the distribution and stationarity of the data. These coefficients populate an adjacency matrix, where each entry indicates the strength of functional coupling between two electrodes. The resulting matrix effectively captures the instantaneous connectivity pattern unique to each segment of brain activity.

To further elucidate the spectral properties of these connectivity patterns, we apply a Discrete Fourier Transform (DFT) to each adjacency matrix. Conceptually, the DFT translates the matrix from the time (or direct correlation) domain to a frequency-like domain, rendering certain network-level oscillations more explicit. We then flatten or summarize these matrices—for instance, by extracting mean values of key frequency bins—to reduce dimensionality.

In parallel, we compute the spectral eigenvalues of the transformed connectivity matrices. These eigenvalues, derived from a graph's Laplacian or from a direct matrix decomposition, help quantify how information might flow within the network and have shown strong discriminative power in prior complex network analyses. Consequently, each segment's final feature set integrates the correlation-based DFT profiles and the associated spectral descriptors.

KMeans-based feature enrichment

Building on the original complex network framework, we introduce an additional layer of unsupervised learning by applying KMeans clustering to the feature vectors obtained. KMeans is performed on the reduced connectivity features, partitioning them into two clusters—a choice informed by early exploratory analyses that suggested a binary partitioning often aligns loosely with "patient-like" vs. "control-like" patterns. The cluster labels become an extra categorical feature, effectively adding an unsupervised perspective to the overall classification process.

Dimensionality reduction

The combined feature vectors (encompassing DFT coefficients, spectral eigenvalues, and KMeans cluster assignments) are subsequently input into a Principal Component Analysis (PCA) step. PCA condenses the feature space by retaining the principal axes of maximum variance, thereby reducing noise and mitigating the risk of overfitting. Although the chosen number of principal components can vary depending on the dataset's size and complexity, a typical range involves keeping only those components that collectively explain the majority of the variance (often 90% or more).

Random forest classification and hyperparameter tuning

For classification, we employ a random forest—a well-established ensemble technique that trains multiple decision trees and aggregates their outputs. Random Forests are particularly robust to outliers and can capture nonlinear relationships in the data. To fine-tune model hyperparameters (such as the number of estimators or the depth of each tree), we perform a GridSearchCV. This grid search systematically evaluates different configurations through an inner cross-validation loop to maximize classification accuracy.

Subject-by-subject validation

A hallmark of this work is its subject-by-subject validation scheme, where each fold isolates all segments from a single individual in the test

License: CC BY-NC 4.0

set. This procedure ensures that the model never "sees" any data from the held-out subject during training, minimizing the likelihood of overfitting to idiosyncrasies in the dataset. Given the relatively small sample sizes—common in EEG-based clinical studies—this rigorous form of validation yields a more reliable measure of the pipeline's real-world diagnostic potential.

Although PD dataset have 59 channels in an extended montage, the pipeline remains fundamentally consistent. In such cases, only minor reconfigurations—such as selecting a different subset of electrodes for correlation analysis or adjusting the PCA threshold—are needed to accommodate the increase in data dimensionality. Empirically, these modifications scale well, with only modest increases in computational cost even when channel counts grow.

The entire process—from raw data loading to final classification—can be reproduced using the code provided in the Appendix. This openness fosters transparency and facilitates incremental improvements, where future researchers may refine individual components (e.g., by exploring different correlation metrics or advanced deep learning modules) while preserving the core methodology.

Results

The topographic maps and graphs (see Appendix Graphs) reveal substantial variability in brain activity across subjects, complicating the visual identification of consistent patterns. This variability prompted the exploration of multiple approaches to enhance classification accuracy.

Schizophrenia: 96% accuracy using channels F7 and Fz.

	Precision	Recall	F1-Score	Support
Control	0.93	1	0.97	14
PS	1	0.93	0.96	14
Accuracy			0.96	28
Macro Avg	0.97	0.96	0.96	28
Weighted Avg	0.97	0.96	0.96	28

Confusion Matrix:

[[14 0]
[1 13]]

Alzheimer's Disease: 97% accuracy using channels O2 and Cz.

	Precision	Recall	F1-Score	Support
Control	1	0.93	0.96	29
AD	0.94	1	0.97	29
Accuracy			0.97	58
Macro Avg	0.97	0.97	0.97	58
Weighted Avg	0.97	0.97	0.97	58

Confusion Matrix**:**

[[27 2]
[0 29]]

License: CC BY-NC 4.0

Frontotemporal Dementia: 96% accuracy using channels F7 and F3.

	Precision	Recall	F1-Score	Support
Control	1	0.91	0.95	23
FTD	0.92	1	0.96	23
Accuracy			0.96	46
Macro Avg	0.96	0.96	0.96	46
Weighted Avg	0.96	0.96	0.96	46

Confusion Matrix**:**

[[21 2]
[0 23]]

Parkinson's Disease: 96% accuracy using channels Fz and T8.

	Precision	Recall	F1-Score	Support
Control	1	0.93	0.96	14
PD	0.93	1	0.97	14
Accuracy			0.96	28
Macro Avg	0.97	0.96	0.96	28
Weighted Avg	0.97	0.96	0.96	28

Confusion Matrix**:**

[[13 1]
[0 14]]

Discussion

The diagnostic precision achieved in this study surpasses the results presented in the original papers that utilized these datasets. The subject-by-subject validation scheme ensures a realistic evaluation and provides a more generalizable model. Despite the relatively small sample sizes found in the literature, the method is promising for widespread clinical adoption. The potential for further improvements lies in applying the method to larger and more diverse datasets, which would help confirm its robustness and scalability.

Conclusions

This work presents a significantly improved method for diagnosing mental health disorders using EEG data. The high classification accuracy achieved across multiple diseases demonstrates the approach's versatility and robustness. The method's simplicity and computational efficiency suit real-world clinical applications. Further research with larger datasets should be used to validate its generalization and improve its clinical applicability.

Funding: This research was not supported by grants.

Conflicts of Interest: The author declares no conflict of interest. For a signed statement, contact the Editorial Office.

Quote this article as Vargas Cruz, Miguel Angel. 2025. "Diagnosis of Alzheimer's Disease, Parkinson's Disease, Frontotemporal Dementia, and Paranoid Schizophrenia via Complex Network Analysis of EEG Data."*Precision Nanomedicine*, 8(2), 1464–1472, https://doi.org/10.33218/001c.133823

License: CC BY-NC 4.0

COPYRIGHT NOTICE ©The Author(s) 2024. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

[1] Vargas, M. (2025). Paranoid Schizophrenia Diagnosis via Complex Network Analysis on EEG Data. Precision Nanomedicine. DOI: 10.33218/001c.128586.

[2] Olejarczyk, E., Jernajczyk, W. (2017). Graph-based analysis of brain connectivity in schizophrenia. PLOS ONE. DOI: 10.1371/journal.pone.0188629.

[3] Miltiadous, A., Tzimourta, K. D., Afrantou, T., Ioannidis, P., Grigoriadis, N., Tsalikakis, D. G., Angelidis, P., Tsipouras, M. G., Glavas, E., Giannakeas, N., Tzallas, A. T. (2023). A Dataset of Scalp EEG Recordings of Alzheimer's Disease, Frontotemporal Dementia and Healthy Subjects from Routine EEG. Data, 8(6), 95. DOI: 10.3390/data8060095

[4] Anjum M., Dasgupta S., Mudumbai R., Singh A., Cavanagh J., Narayanan N. (2020). Linear predictive coding distinguishes spectral EEG features of Parkinson's disease. Parkinsonism. Relat. Disord. 79 79–85. DOI: 10.1016/j.parkreldis.2020.08.001.

[5] Acharya, M., Deo, R. C., Tao, X., Barua, P. D., Devi, A., Atmakuru, A., Tan, R. S. (2025). Deep learning techniques for automated Alzheimers and mild cognitive impairment disease using EEG signals: A comprehensive review of the last decade (2013-2024). Elsevier. DOI: 10.1016/j.cmpb.2024.108506.

[6] Cavanagh, J. F., Kumar, P., Mueller, A. A., Richardson, S. P., Mueen, A. (2018). Diminished EEG habituation to novel events effectively classifies Parkinson's patients. Elsevier. DOI: 10.1016/j.clinph.2017.11.023.

[7] Brown, D. R., Richardson, S. P., Cavanagh, J. F. (2020). An EEG marker of reward processing is diminished in Parkinsons disease. Elsevier. DOI: 10.1016/j.brainres.2019.146541.

[8] Jiao, B., Li, R., Zhou, H., Qing, K., Liu, H., Pan, H., Lei, Y., Fu, W., Wang, X., Xiao, X., Liu, X., Yang, Q., Liao, X., Zhou, Y., Fang, L., Dong, Y., Yang, Y., Jiang, H., Huang, S., Shen, L. (2023). Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimers disease using EEG technology. Nature. DOI: 10.1186/s13195-023-01181-1.

[9] Zyma, I., Seleznov, I., Popov, A., Chernykh, M., Shpenkov, O. (2018). EEG During mental arithmetic tasks. PhysioNet. DOI: 10.13026/C2JQ1P.

[10] Kim, J., Wilhelm, T. (2007). What is a complex graph?. Elsevier. DOI: 10.1016/j.physa.2008.01.015.

[11] Newman, M. (2003). The structure and function of complex networks. Society for Industrial and Applied Mathematics. DOI: 10.1137/S003614450342480.

[12] Bullmore, E., Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. Nature. DOI: 10.1038/nrn2575.

[13] Rubinov, M., Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. DOI: 10.1016/j.neuroimage.2009.10.003.

[14] Pearson, K. (1895). Note on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London. DOI: 10.1098/rspl.1895.0041.

[15] Cohen, M. (2014). Analyzing Neural Time Series Data: Theory and Practice. The MIT Press. DOI: 10.7551/mitpress/9609.001.0001.

[16] Spearman, C. (1904). The proof and measurement of association between two things. The American Journal of Psychology. DOI: 10.2307/1412159.

[17] Bruns, A. (2004). Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches?. Journal of Neuroscience Methods. DOI: 10.1016/j.jneumeth.2004.03.002.

[18] Lloyd, S. (1982). Least squares quantization in PCM. IEEE. DOI: 10.1109/TIT.1982.1056489.

[20] Tipping, M., and Bishop, C. (1999). Probabilistic principal component analysis. Journal of the Royal Statistical Society. DOI: 10.1111/1467-9868.00196.

[21] Ho, T. (1995). Random decision forests. IEEE. DOI: 10.1109/ICDAR.1995.598994.

[22] Smith, P., Ganesh, S., Liu, P. (2013). A comparison of random forest regression and multiple linear regression for prediction in neuroscience. Elsevier. DOI: 10.1016/j.jneumeth.2013.08.024.

[23] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. DOI: 10.5555/1953048.2078195.

[24] Vargas, M. (2025). References. https://www.miguelangelvargascruz.com/en/references

License: CC BY-NC 4.0