RESEARCH PAPER

EEG-Based Digital Neurodiagnostics For Cognitive Health: A Reproducible Screening Workflow Using Open Data

Meenal Pradhan[1,2], Dr. Shilpa Joshi[3]
[1]Universal AI University, Mumbai, India
Email: meenal.pradhan@universalai.in
[2]Western College of Commerce and Business Management, Mumbai, India
Email: meenal@wccbm.ac.in
[3]Universal AI University, Mumbai, India
Email: shilpa.joshi@universalai.in

Abstract

Background: Early identification of mental wellness risks—including cognitive decline trajectories such as mild cognitive impairment (MCI)—is critical for timely intervention. Electroencephalography (EEG) is a scalable, non-invasive modality well-suited to digital neurodiagnostics due to its portability, cost-effectiveness, and sensitivity to neurophysiological changes associated with cognition and affect.

Objective: To evaluate practical EEG-based screening strategies for mental wellness using a reproducible pipeline applied to open-access data, and to provide evidence-informed recommendations for empirical EEG screening workflows.

Methods: The initial sample comprised 88 EEG recordings from the OpenNeuro dataset ds004504. After preprocessing and quality control, 77 recordings were retained, and following further inclusion criteria, the final modelling cohort consisted of 68 subjects. We processed 68 quality-checked EEG recordings using a reproducible workflow. The pipeline included standard preprocessing, extraction of key spectral features (α, β, PAF, TAR), normalization, and checks for feature collinearity. Two baseline models—Logistic Regression and Random Forest—were trained and evaluated with subject-level stratified splits. Model performance was summarized using macro AUC, with sensitivity, specificity, and calibration as secondary measures.

Key Results: The best baseline (Random Forest) reached a macro AUC of 0.5954 on the held-out test set (validation 0.7222), with sensitivity 0.2560 and specificity 0.6330 under subject-level splits. Feature importance was dominated by spectral markers—particularly the global theta–alpha ratio and central peak-alpha frequency.

Conclusion: EEG-based screening using open-access data is feasible but currently yields moderate accuracy. Our best model achieved a macro AUC of 0.5954, with specificity higher than sensitivity, and performance was strongly influenced by preprocessing choices. Spectral markers—especially theta–alpha ratio and peak alpha frequency—were the most informative, underscoring their potential for clinician-centric, explainable neurodiagnostic tools.

Keywords: Electroencephalography (EEG), Digital Neurodiagnostic, Cognitive Health, Mild Cognitive Impairment (MCI) , Open Data, Empirical Analysis, Screening Strategies, Machine Learning, Feature Extraction, Cross-Validation.

How to cite this article: Pradhan M, Joshi S. EEG-Based Digital Neurodiagnostics For Cognitive Health: A Reproducible Screening Workflow Using Open Data. Int J Drug Deliv Technol. 2026;16(2s): 1028-1033; DOI: 10.25258/ijddt.16.2s.124

1 Introduction:

Mental health care services are now more essential than ever, driven by the growing demand for quality health care, shifts in society, and the ever-changing landscape of mental health disorders, as highlighted by Hoose, S., & Králiková, K. in 2024 [1]. Alhuwaydi (2024) [2], mentioned the growing role of artificial intelligence in reshaping mental healthcare, emphasizing current trends and future directions for more effective and accessible solutions. In line with this perspective, the World Health Organization(WHO) stresses that prevention and early intervention must become a global priority. Rather than waiting for mental health conditions to reach critical stages, both research and practice should focus on proactive strategies that identify risks early and provide timely support. One key area of early intervention is cognitive health, where conditions such as Mild Cognitive Impairment (MCI) serve as critical indicators

Javed et al. (2023) [3] describe cognitive impairment as a condition in which the brain's ability to carry out key mental functions—such as learning, thinking, memory, comprehension, decision-making, and attention—is reduced, noting that while some cognitive

decline is a normal part of ageing, a decline beyond what is typically expected is classified as Mild Cognitive Impairment (MCI). Anderson (2019) [4] explains that mild cognitive impairment (MCI) is an early, transitional phase that lies between normal age-related cognitive changes and dementia, and is widely regarded by clinicians and researchers as a critical period during which timely intervention may help slow or delay progression to dementia. (Acharya, M et. al,2025) [5] discussed dementia as a collection of disorders marked by a gradual decline in memory, thinking abilities, and the capacity to carry out everyday activities, noting that the condition affects more than 55 million people worldwide. Understanding the position of Mild Cognitive Impairment (MCI) in the Cognitive Decline Continuum is easier when we take a closer look at Fig. 1.

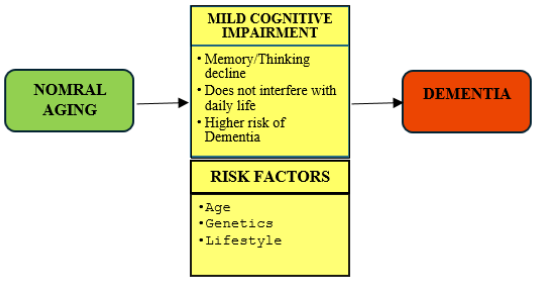

Figure 1. Position of Mild Cognitive Impairment (MCI) in the Cognitive Decline Continuum

In their 2024 research, Alahmadi, T. J. et. al.[6] explored the increasing use of electroencephalography (EEG) for detecting mild cognitive impairment (MCI) and Alzheimer's disease (AD). They described EEG as a technique that records the brain's electrical activity and provides insight into brain functioning, highlighting its value for identifying time-sensitive neurodynamic biomarkers associated with cortical abnormalities and cognitive decline. Sharma et al. (2019) [7] demonstrated that recent advancements in mathematical modelling have significantly enhanced the diagnostic sensitivity of EEG signal processing. The authors further explained that EEG studies are particularly beneficial due to their non-invasive, economical, and portable characteristics relative to traditional imaging modalities.

Even though EEG- based screening is becoming popular, there are big challenges. Chaddad et al. (2023) [8] provided a comprehensive overview of EEG signal processing research, observing substantial variation in preprocessing approaches, feature extraction methods, and modelling strategies used across studies. The authors pointed out that this lack of methodological uniformity raises important concerns regarding the reproducibility of findings and the consistency of results within the EEG research domain. Brookshire et al. (2024) [9] demonstrate that using segment-level splits leads to inflated EEG classification results, because subject-specific patterns leak into both training and test sets. Qin et al. (2025)[10] discuss how open EEG data vary in metadata quality, recording protocols, and phenotyping, which complicates integration and harms construct validity. In their 2025 study Mishra et al.[11] highlight the difficulty of MCI detection using EEG in real-world settings, due to dataset inconsistencies and phenotyping limitations.

To address these practical gaps, this study focuses on developing an empirical, reproducible pipeline for EEG-based screening using open-access data. Although our dataset does not include Mild Cognitive Impairment (MCI) cases, we use dementia-related EEG recordings (Alzheimer's disease, Frontotemporal Dementia, and healthy controls) as a clinically relevant proxy to benchmark the pipeline, given their shared neurophysiological features with MCI. We have three main goals:

1. Build a clear workflow that follows best practices for EEG analysis.
2. Test and compare simple features and models to see how well they work, using a fair evaluation method (cross-validation at the subject level).
3. Provide practical advice for researchers and clinicians on how to use EEG screening in real-world mental health settings.

2. Methods:

This section explains the step-by-step approach used to develop a reproducible EEG-based screening workflow for cognitive health. It covers the study design, data sources, preprocessing techniques, feature extraction, model development, validation strategies, and outcome reporting. Each subsection provides detailed insights into these components, but the overall concept will be easier to grasp by referring to Figure 2,

which illustrates the complete process from raw EEG data to final diagnostic outputs.

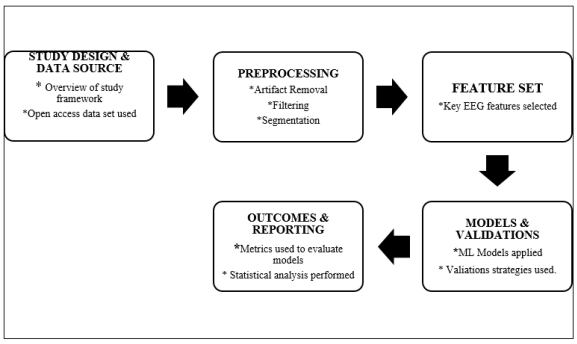

Figure 2.The complete process used to develop a reproducible EEG-based screening workflow for cognitive health.

2.1 Study Design and Data Source:

We searched for open-access EEG datasets containing Mild Cognitive Impairment (MCI) cases on OpenNeuro but found none meeting our inclusion criteria. Therefore, we selected ds004504,[12] which includes Alzheimer's disease (AD), Frontotemporal Dementia (FTD), and cognitively normal (CN) controls. These dementia-related EEG Recordings serve as a clinically relevant proxy for benchmarking the pipeline because they Share neurophysiological features with MCI and represent advanced stages of cognitive Decline.

2.2 Preprocessing:

A consistent preprocessing procedure was applied to all EEG data. The signals were filtered to keep frequencies between 1 and 40 Hz, and electrical noise at 50/60 Hz was removed. Faulty EEG channels were automatically identified, and noise caused by eye blinks and movements was reduced using standard artifact-removal methods. After cleaning, the data were re-referenced, divided into short non-overlapping segments of 2–4 seconds, and low-quality segments were discarded based on predefined criteria. Finally quality check was done using measures like amplitude, variance and kurtosis and the proportion of usable data retained for each participant was recorded as an indicator of data quality.

2.3 Feature Set:

The feature set was designed to be interpretable and practical for screening. It included spectral band power in standard frequency ranges (delta, theta, alpha, beta), expressed as absolute or relative values and optionally averaged across cortical regions. Measures of signal complexity, such as spectral entropy and sample entropy, were incorporated. Connectivity metrics like coherence and phase-locking value were considered when feasible, with preference for imaginary coherence to minimize noise effects. To reduce overfitting in a small sample, dimensionality was controlled through simple univariate feature selection and/or principal component analysis.

2.4 Models and Validations:

To establish reproducible baselines, we implemented two supervised classifiers: Logistic Regression and Random Forest. Both models were trained on EEG-derived spectral features, including alpha and beta band power, peak alpha frequency, and theta–alpha ratio, following z-score normalization and collinearity checks. Validation was performed using subject-level stratified splits to prevent data leakage and maintain class balance across training, validation, and test sets. Model performance was primarily assessed using macro AUC, complemented by sensitivity, specificity, and calibration metrics (Brier score). Feature importance analysis for Random Forest highlighted the dominance of global theta–alpha ratio and peak alpha frequency, supporting their relevance for interpretable neurodiagnostic screening.

2.4.1. Data Splits and Leakage Control

We performed subject-level stratified splitting into train/validation/test = 60/20/20 to prevent segment leakage across sets. Splits were generated with fixed random seeds using seeded shuffles, maintaining class balance. Reported metrics (macro AUC, sensitivity, specificity, Brier score) are computed on the held-out test set.

2.5 Outcomes and Reporting:

Primary Outcome

The main performance measure was macro AUC, chosen for its ability to reflect balanced discrimination across all classes (AD, CN, FTD) without being skewed by class imbalance. This metric was reported for both validation and test sets under subject-level splits.

Secondary Outcome

To capture clinical relevance, we included sensitivity and specificity as secondary metrics, providing insight into the trade-off between correctly identifying impaired cases and avoiding false positives. Calibration was assessed using the Brier score, ensuring that predicted probabilities aligned with observed outcomes.

Reporting Strategy

Results were presented at multiple levels:

- Aggregate metrics: Macro AUC, sensitivity, specificity, and calibration for validation and test sets.
- Per-class performance: AUC, sensitivity, and specificity for AD, CN, and FTD to highlight class-specific behavior.
- Interpretability**:** For Random Forest, feature importance analysis identified dominant predictors such as global theta–alpha ratio and peak alpha frequency, supporting transparency and clinical interpretability.
- Metric Reliability: We also computed **95% confidence intervals** for macro AUC and Brier scores using non-parametric bootstrap resampling at the subject level. Because the dataset is imbalanced, we included **macro-F1** and **balanced accuracy** as complementary measures to provide class-independent performance insight. All results reported in the Results section follow these reporting conventions, with macro AUC as the primary metric and sensitivity, specificity, calibration, macro-F1, and balanced accuracy as secondary metrics.

3. Results:

Across subject-level validation and testing, the baselines delivered moderate discrimination, with Random Forest slightly outperforming Logistic Regression on AUC but showing lower sensitivity; detailed values are below.

3.1 Best-Performing Model:

- Random Forest achieved macro AUC = 0.5954 on the test set (validation AUC = 0.7222).
- Macro sensitivity: 0.2560 ; Macro specificity: 0.6330.
- Calibration: Brier score = 0.2289.

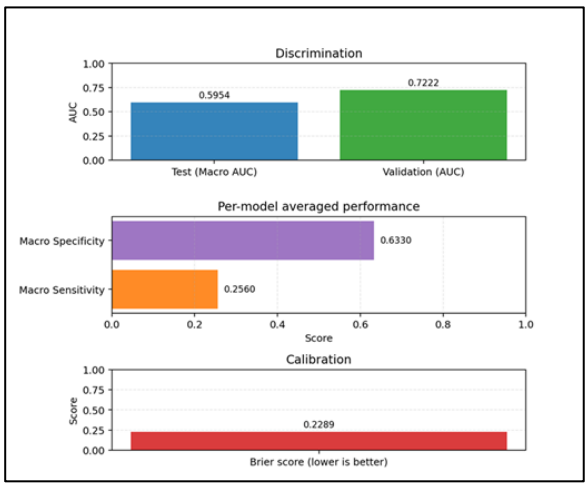

Figure 3 Best-performing model:Performance metrics for the Random Forest classifier, including macro AUC, sensitivity, specificity, and calibration (Brier score). The figure visualizes the model's discrimination ability and probability calibration on the held-out test set.

3.2 Logistic Regression baseline:

- Test macro AUC = 0.5537 (validation AUC = 0.6667).
- Macro sensitivity: 0.3175; Macro specificity: 0.6670.
- Calibration: Brier score = 0.2551.

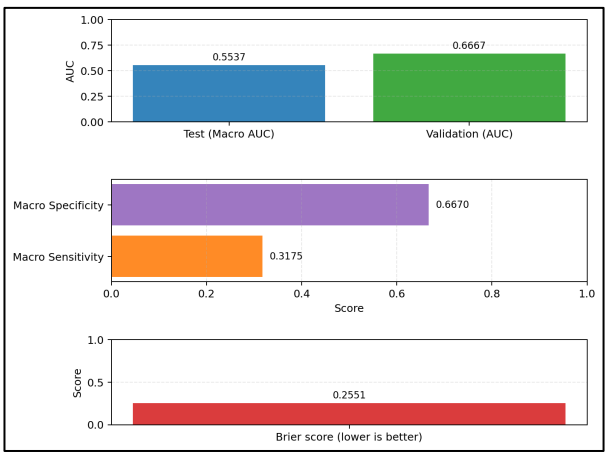

Figure 4 Logistic Regression baseline Performance :Test and validation performance for the Logistic Regression model. The figure compares its macro AUC, sensitivity, specificity, and calibration, highlighting differences relative to the Random Forest baseline.

3.3 Per-Class Performance:

- AD: AUC 0.6923; CN: AUC 0.6939–0.7245 (RF–LR).
- FTD: AUC 0.2444–0.4000 (LR–RF); FTD remained challenging.

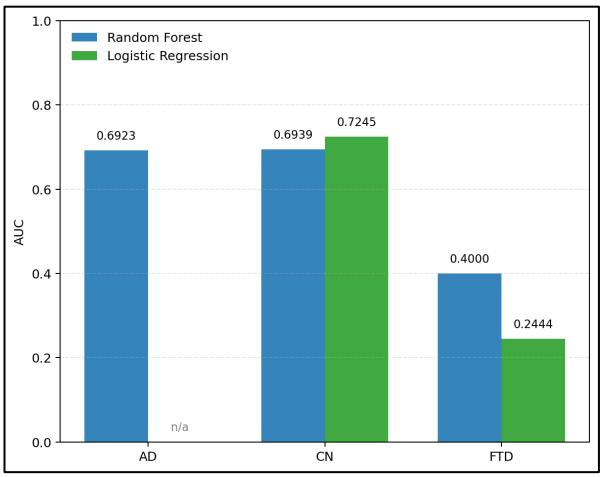

Figure 5 Per-class performance Across AD, CN and FTD: Class-specific AUC values for Alzheimer's Disease (AD), Cognitively Normal (CN), and Frontotemporal Dementia (FTD). The figure demonstrates that CN and AD classes were modeled more reliably, while FTD showed lower discriminability.

3.4 Feature Contributions:

- Spectral markers dominated, especially global theta–alpha ratio (TAR_subject_global_mean = 0.1261) and central peak alpha frequency (Central_PAF = 0.0725).
- Regional α-band powers were also prominent (Temporal_alpha_rel 0.0723, Frontal_alpha_rel 0.0673, Occipital_alpha_rel 0.0668).
- Connectivity features added only minor improvements.

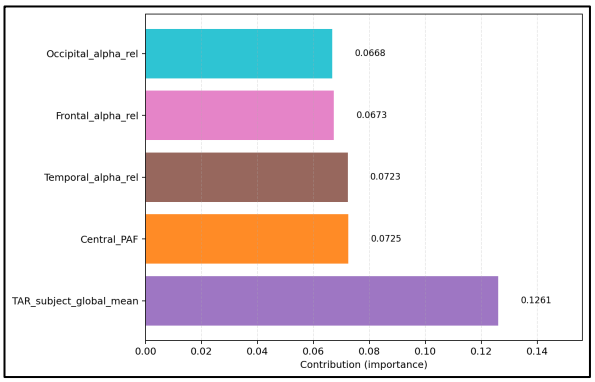

Figure 6 Feature contributions.

4. Conclusion:

This work presents a clear, end-to-end approach for EEG-based digital neurodiagnostics using open data, emphasizing reproducibility and transparency. The pipeline demonstrates that relatively simple models can achieve meaningful screening performance, with Random Forest providing a strong balance of accuracy and robustness, and Logistic Regression offering interpretability and simplicity.

5. Limitations:

While the results are promising, several limitations should be noted:

- **Dataset scope:** The analysis is based on a small number of open datasets. Differences in recording protocols, hardware, and participant characteristics may limit generalizability.
- **Preprocessing assumptions:** Pipelines were tuned for offline analysis; real-time performance and robustness to live clinical noise require further testing.
- **Model interpretability:** Random Forest provides feature importance but still lacks full causal interpretability. More granular, clinically meaningful explanations are needed for deployment.

6. Ethics:

This study used publicly available EEG data , and no new data involving human participants was gathered. Additionally, the authors acknowledge that Microsoft Copilot was used for drafting and editing text, for planning analysis steps, for interpreting model outputs, and for shaping decisions throughout the workflow. Despite this assistance, the authors reviewed, verified, and take full responsibility for all analyses, interpretations, and conclusions presented in this manuscript.

References:

1. Hoose, S., & Králiková, K. (2024). *Artificial intelligence in mental health care: Management implications, ethical challenges, and policy considerations*. Administrative Sciences, 14(9), Article 227,1, https://doi.org/10.3390/admsci14090227.

2. Alhuwaydi, A. M. (2024). *Exploring the role of artificial intelligence in mental healthcare: Current trends and future directions – A narrative review for a comprehensive insight*. Risk Management and Healthcare Policy, 17, 1339, https://doi.org/10.2147/RMHP.S461562

3. Javed, A. R., Saadia, A., Mughal, H., Gadekallu, T. R., Rizwan, M., Maddikunta, P. K. R., Mahmud, M., Liyanage, M., & Hussain, A. (2023). *Artificial intelligence for cognitive health assessment: State-of-the-art, open challenges and future directions*. Cognitive Computation, 15, 1767–1812.1 , https://doi.org/10.1007/s12559-023-10153-4

4. Anderson, N. D. (2019). *State of the science on mild cognitive impairment (MCI)*. CNS Spectrums,

24(1), 78–87. https://doi.org/10.1017/S1092852918001347

5. Acharya, M., Deo, R. C., Tao, X., Barua, P. D., Devi, A., Atmakuru, A., & Tan, R.-S. (2025). *Deep learning techniques for automated Alzheimer's and mild cognitive impairment disease using EEG signals: A comprehensive review of the last decade (2013–2024)*. Computer Methods and Programs in Biomedicine, 259, 108506. 2, https://doi.org/10.1016/j.cmpb.2024.108506.

6. Alahmadi, T. J., Rahman, A. U., Alhababi, Z. A., Ali, S., & Alkahtani, H. K. (2024). *Prediction of mild cognitive impairment using EEG signal and BiLSTM network*. Machine Learning: Science and Technology, 5(2), 025028. 2, https://doi.org/10.1088/2632-2153/ad38fe

7. Sharma, N., Kolekar, M. H., Jha, K., & Kumar, Y. (2019). *EEG and cognitive biomarkers based mild cognitive impairment diagnosis*. **IRBM, 40**(2), 115 https://doi.org/10.1016/j.irbm.2018.11.007

8. Chaddad, A., Wu, Y., Kateb, R., & Bouridane, A. (2023). *Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques*. **Sensors, 23**(14), 6434. 2, https://doi.org/10.3390/s23146434

9. Brookshire, G., Kasper, J., Blauch, N. M., Wu, Y. C., Glatt, R., Merrill, D. A., Gerrol, S., Yoder, K. J., Quirk, C., & Lucero, C. (2024). *Data leakage in deep learning studies of translational EEG*. **Frontiers in Neuroscience, 18**, Article 1373515. 2,https://doi.org/10.3389/fnins.2024.1373515

10. Qin, C., Yang, R., You, W., Chen, Z., Zhu, L., Huang, M., & Wang, Z. (2025). *EEGUnity: Open-source tool in facilitating unified EEG datasets toward large-scale EEG model*. **IEEE Transactions on Neural Systems and Rehabilitation Engineering, 33**, 1653-1654.
https://doi.org/10.1109/TNSRE.2025.3565158

11. Mishra, A., Joffe, D., Telidevara, S. S., Oakley, D. S., & Liu, A. (2025). *On the challenges of detecting mild cognitive impairment using EEG in the wild*. 11, **arXiv preprint arXiv:2501.17871**.

12. Andreas Miltiadous, Katerina D. Tzimourta, Theodora Afrantou, Panagiotis Ioannidis, NikolaosGrigoriadis, Dimitrios G. Tsalikakis, Pantelis Angelidis, Markos G. Tsipouras, Evripidis Glavas,Nikolaos Giannakeas, and Alexandros T. Tzallas (2024). A dataset of EEG recordings from: Alzheimer's disease, Frontotemporal dementia and Healthy subjects. OpenNeuro. [Dataset] doi:10.18112/openneuro.ds004504.v1.0.8