Predicting Alzheimer's Disease with Machine Learning and EEG Data

Magdalena Aretha Sunarko
Computer Science Department
School of Computing and Creative Arts
Bina Nusantara University
Jakarta, Indonesia 11480
magdalena.sunarko@binus.ac.id

Alrismany Abigail Sudjarwo
Computer Science Department
School of Computing and Creative Arts
Bina Nusantara University
Jakarta, Indonesia 11480
alrismany.sudjarwo@binus.ac.id

Ahmad Zaydan
Computer Science Department
School of Computing and Creative Arts
Bina Nusantara University
Jakarta, Indonesia 11480
ahmad.zaydan@binus.ac.id

Zhandos Yessenbayev
Computer Science Department
School of Computing and Creative Arts
Bina Nusantara University
Jakarta, Indonesia 11480
zhandos.yessenbayev@binus.ac.id

***Abstract*—This study follows a multi-stage workflow that includes data preparation, signal preprocessing, feature extraction, model optimization, and interface deployment. EEG recordings from the OpenNeuro ds004504 dataset were grouped by clinical condition and recording state prior to analysis. Each signal was detrended, band-pass filtered between 0.5 and 45 Hz, normalized, and trimmed to ensure consistent channel length. A structured feature vector was derived from every patient sample by computing relative spectral band powers across delta, theta, alpha, beta, and gamma ranges, along with entropy-based metrics, Hjorth mobility and complexity, and channel-to-channel connectivity measures. These features served as input to a supervised classification pipeline that tested multiple algorithms, including Logistic Regression, Support Vector Machines, Neural Networks, Random Forests, Gradient Boosting, and XGBoost. Hyperparameters for all models were tuned using Optuna Bayesian optimization with ROC-AUC as the objective. The best-performing model was then calibrated and evaluated using accuracy, F1-score, AUC, and confusion matrix outputs. Finally, the complete system was integrated into a Tkinter-based graphical interface that allows users to load patient EEG data, visualize raw and filtered signals, generate Alzheimer probability scores, and compare patient feature distributions through violin-plot visualizations.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

As humans grow older, parts of their biological function also degrade over time, whether it is appendages, inner organs, skeletal structure, or the nerve system. One of the most common types of nerve degradation is dementia. Dementia is a term used to describe the decline of a person's cognitive ability. The most common type of dementia is Alzheimer's disease (AD). AD is a progressive degenerative disorder, which means that it slowly damages the brain over time [1]. The percentage of people with AD increases with age, accounting for 3% of the patients aged 65–74 years and 17% of the patients aged 75–84 years. Despite advancements in neuroscience, early and accurate detection of Alzheimer's disease remains challenging. Traditional diagnostic methods, such as neuroimaging and cognitive testing, are costly, time-consuming, and often fail to detect the disease in its early stages.

Early diagnosis of Alzheimer's disease can be helpful for prevention at a young age through therapy after the diagnosis. Using resting-state EEG data, this study creates a machine learning system to aid in the early identification of Alzheimer's disease. The system filters, normalizes, and extracts features from raw multi-channel EEG recordings. To find patterns that differentiate Alzheimer's patients from healthy controls, a number of machine learning models were trained and assessed. The findings demonstrate the potential of computational neuroscience techniques for enhancing Alzheimer's diagnosis by demonstrating how EEG-based biomarkers, in conjunction with machine learning, might allow useful and affordable early screening.

II. Related Works

Researchers from Cambridge University made a machine learning tool that predicts whether people with early symptoms of cognitive degradation will stay stable or progress to AD [2]. This was done using non-invasive routine data, which includes cognitive tests and a structural MRI scan. The model was trained on more than 400 individuals and tested on additional cohorts, including real-world memory clinic data from the UK and Singapore. By using cognitive tests and MRi scans, the model could predict progression within less than three years and manage to result in 82% progressors and 81% non-progressors. This research did not implement EEG, which does not correlate with the topic of this research.

A webinar done by Dr. Fiona Harrison discussed that through altered EEG activity and memory task performance, it is possible to observe glutamate uptake or clearance, and change neural signaling, which are caused by dietary deficiency and toxin exposure [3]. This shows that EEG is able to detect changes in neural signaling relevant to memory impairments associated with early Alzheimer's-related changes.

Arikan et al. proposed an Alzheimer's detection framework based on EEG that transforms the multi-channel EEG signals into time-frequency representations and extracts the features with a lightweight CNN architecture with dual-attention mechanisms, followed by classification using an SVM model optimized with Optuna. The result shows that the delta and beta frequency bands provide the most discriminative information, achieving high classification accuracy on a small dataset of 48 subjects.

Vidushi et al. proposed evaluating traditional machine learning algorithms such as SVM, Random Forest, k-NN, and Logistic Regression to detect Alzheimer's using engineered biomedical features. The results indicate that classical classifiers can achieve reasonable performance when combined with appropriate feature extraction.

Ghazal et al. apply transfer learning by utilizing deep convolutional neural networks that were pretrained to improve the performance of Alzheimer's disease classification on limited datasets. This approach reduces the training time compared to manually training deep models from scratch and manages to improve the accuracy.

Pirrone et al. proposed a framework based on EEG that combines frequency-domain signal processing with supervised machine learning to classify Alzheimer's disease, mild cognitive impairment, and healthy controls. This research highlights the importance of low-frequency EEG components and reports strong classification performance.

Shukla et al. summarize the existing Alzheimer's disease detection with machine learning approaches based on EEG, reinforcing the idea that EEG is a low-cost and non-invasive diagnostic tool while noting consistent findings such as EEG slowing in Alzheimer's patients. This research reviews the most commonly used preprocessing techniques, feature extraction methods, and classifiers.

Xia et al. suggested an Alzheimer's disease, mild cognitive impairment (MCI), and healthy controls classification with a deep learning framework using the resting-state EEG signals. This method extracts the frequency-domain features through FFT and applies an overlapping sliding-window strategy to simulate limited EEG data before classification with a deep pyramid convolutional neural network (DPCNN).

Rezaei et al. focus on capturing spatiotemporal and brain activity characteristics that are frequency-related. This research demonstrates that deep representations can improve classification between Alzheimer's patients and healthy individuals by using neural network architectures designed to model complex EEG patterns. The results prove that EEG is a meaningful modality for Alzheimer's research.

The main limitations of these studies are that many EEG-based research studies have adopted deep learning techniques, and although these approaches managed to achieve high accuracy, they often require significant computational resources, large amounts of training data, and complex preprocessing pipelines. Most studies in EEG-based AD detection use small datasets, which raises the concern for overfitting and limiting confidence in model generalization across different populations, recording devices, and clinical settings.

A. Gaps and Improvement

Many recent EEG-based studies employ deep learning techniques, including convolutional neural networks, attention mechanisms, transfer learning, and image-based EEG representations such as spectrograms or scalograms. Although these methods achieve high accuracy, they often require significant computational resources, large amounts of training data, and complex preprocessing pipelines. This complexity reduces interpretability and makes deployment in lightweight or real-time clinical environments challenging.

To address the identified gaps, this study adopts a pipeline-first and benchmarking-oriented approach to Alzheimer's disease prediction using machine learning.

First, emphasis on accessibility and scalability. Rather than relying on expensive neuroimaging modalities, this work focuses on EEG-inspired feature representations and machine learning methods that are computationally lightweight, aligning with the goal of developing scalable and low-cost screening tools.

Second, systematic comparison of multiple machine learning algorithms. Instead of proposing a single optimized model, this study benchmarks several standard classifiers, including logistic regression, support vector machines, random forests, and ensemble-based methods, under a unified preprocessing and evaluation protocol. This allows clearer insight into the relative strengths and weaknesses of different algorithms for Alzheimer's-related classification tasks.

Third, validation of the machine learning pipeline through prototyping. A prototype framework is developed using simulated feature data to validate the end-to-end machine learning workflow, including data splitting, feature scaling, cross-validation, metric computation, and visualization. This ensures that the pipeline is robust and reproducible before integration with real EEG datasets.

III. Methodology

A. Dataset collection

This research is conducted to evaluate machine learning algorithms for Alzheimer's disease (AD) classification using feature representations based on EEG.

B. Data Preprocessing

To ensure the consistency and quality of the signal, all of the EEG recordings undergo the pipeline of standardized processing. Raw EEG signals are loaded channel-wise and shortened to a uniform length across channels to maintain alignment. A maximum of 10 EEG channels is kept per subject to ensure consistent dimensionality across samples. Then, each of the channels is preprocessed by removing the DC offset, followed by band-pass filtering between $0.5\,\mathrm{Hz}$

and 45 Hz using a fourth-order Butterworth filter, consistent with the standard practices of EEG analysis. Signals are then normalized according to their standard deviation to reduce inter-channelaplitude variability. This pipeline of preprocessing was applied to remove noise, suppresses frequency that are irrelevant, and ensure the numerical stability for the extraction and classification of downstream features.

C. Feature Extraction

The EEG signals are transformed into feature vectors that are compact and interpretable to capture spectral, statistical, and connectivity-related characteristics associated with AD. First, relative band power features are computed using Welch's method to estimate the power spectral density for each channel. Power is integrated within canonical EEG frequency bands, which are the δ (0.5–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30–45 Hz), and normalized by the total signal power. The features that are included from the relative band powers across channels are the mean and standard deviation.

Second, the features of signal complexity are extracted in the form of spectral entropy and Hjorth parameters, which quantify the EEG signal's irregularity and temporal dynamics. The mean and standard deviation of these measures across channels are computed to capture the inter-channel variability. Pearson correlation coefficients between EEG channels approximate the functional connectivity, and the global connectivity descriptor includes the mean absolute upper-triangular correlation value.

D. Machine Learning Models

This research evaluates multiple machine learning classification algorithms under one unified experimental framework. This was done to address the lack of standardized benchmarking in previous Alzheimer's detection studies based on EEG. The models include Logistic Regression, Support Vector Machines (SVM) with linear and radial basis function kernels, Multilayer Perceptron (MLP) neural networks, Random Forests, Gradient Boosting, and Extreme Gradient Boosting (XGBoost). The standardization of the feature is performed using the z-score normalization before the model is trained. Optuna-based hyperparameter optimization is employed to ensure the most optimal configurations, with the area under the ROC curve (AUC) used as the optimization objective. The model configuration that performs best is retrained and calibrated by utilizing the probability calibration to improve the reliability of predicted Alzheimer's probabilities.

E. Model Training and Evaluation

The dataset is divided into two subsets, which are training and testing, and to preserve class balance, a stratified 70:30 split was used. The performance of the model was evaluated using multiple complementary metrics such as accuracy, F1-score, ROC-AUC, and confusion matrices, in order to provide a comprehensive assessment of classification effectiveness. The probability calibration was applied to ensure meaningful confidence estimates reflected by the predicted class probabilities. To demonstrate the practical applicability of the proposed framework for subject-level inference, the trained model was applied to unseen subject data to estimate the individual Alzheimer's probability score.

F. Implementation and Libraries

All experiments are conducted using the Python programming language by utilizing scientific computing and machine learning libraries. Numpy was used for handling data and numerical operations, while signal processing functions such as filtering and power spectral estimation are implemented using SciPy. Scikit-learn provides the machine learning models and evaluation metrics, with XGBoost used for gradient boosting classification. The data visualization is carried out by Matplotlib and Seaborn. This implementation enables reproducibility and facilitates future extension.

IV. Result

A. Overall Model Performance

Using Optuna-based optimization on EEG-derived features, the performance of six machine learning models—Logistic Regression (LR), Support Vector Machine (SVM), Neural Network (NN), Random Forest (RF), Gradient Boosting (GB), and XGBoost (XGB)—was assessed. Accuracy, F1-score, and ROC-AUC were used to evaluate the model's performance; results for the eyes-open and eyes-closed EEG situations were presented separately.

The distribution of ROC-AUC, F1-score, and accuracy values for each model over Optuna trials is shown in Figures X–Y. The boxplots provide information about the resilience of the model under various hyperparameter setups by displaying both central tendency and variability.

B. Eyes-Closed EEG Results

Overall, ensemble-based models performed better for the eyes-closed situation. With continuously low interquartile ranges and the greatest ROC-AUC values, Random Forest and XGBoost demonstrated strong discriminative capacity and stable optimization behavior. Neural networks also demonstrated competitive performance, attaining high median ROC-AUC and F1-score values, but with somewhat greater trial-to-trial variability.

When modeling non-linear EEG patterns in the eyes-closed situation, Logistic Regression and SVM showed broader performance dispersion and lower median ROC-AUC values, indicating decreased robustness. Gradient Boosting demonstrated a reasonable level of performance but significant variability, especially in ROC-AUC, suggesting sensitivity to hyperparameter selection.

When compared to linear models, ensemble approaches maintained a good classification balance and fewer misclassifications. Accuracy and F1-score trends closely mirrored ROC-AUC behavior.

TABLE I
PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS ON EYES-CLOSED EEG DATA

Model	ROC-AUC	Accuracy	F1-score
RF	**0.999**	**0.964**	**0.980**
XGB	0.981	0.923	0.957
NN	0.935	0.955	0.974
LR	0.914	0.920	0.956
SVM	0.914	0.929	0.960
GB	0.844	0.940	0.965

C. *Eyes-opened EEG Result*

With ROC-AUC values close to unity and little variation across Optuna trials, Random Forest once more produced the best overall performance in the eyes-open condition. Strong discriminative performance was also shown by SVMs and neural networks, which produced consistent F1-scores and high ROC-AUC.

In comparison to Random Forest, XGBoost demonstrated competitive performance, albeit with somewhat less regularity. Among the models that were assessed, Logistic Regression had the lowest median performance, especially in ROC-AUC, which is indicative of its limits in capturing intricate EEG feature interactions.

EEG signals recorded during eyes-open resting states may provide more discriminative information for Alzheimer's disease categorization, as evidenced by the slightly higher and more stable performance metrics in the eyes-open condition compared to the eyes-closed condition.

TABLE II
PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS ON EYES-OPEN EEG DATA

Model	ROC-AUC	Accuracy	F1-score
RF	**0.997**	**0.964**	**0.980**
XGB	0.973	0.902	0.946
NN	0.968	0.929	0.958
SVM	0.968	0.921	0.957
LR	0.957	0.900	0.945
GB	0.740	0.929	0.962

D. *Limitations and Future Work*

Instead of using raw signal learning, this study uses handcrafted EEG characteristics and concentrates on a single EEG condition at a time. Furthermore, while Optuna offers effective hyperparameter optimization, wider search areas or different optimization techniques may yield different outcomes.

Subsequent research may investigate subject-independent cross-validation techniques, multimodal EEG states, and deep learning architectures that may learn directly from unprocessed EEG signals. Predicting Alzheimer's disease in its early stages may potentially be enhanced by including longitudinal EEG data.

V. CONCLUSION

This study examined the efficacy of many machine learning models for the classification of Alzheimer's disease using EEG-derived characteristics in both closed and open resting states. A fair and methodical comparison of the Logistic Regression, Support Vector Machine, Neural Network, Random Forest, Gradient Boosting, and XGBoost models was carried out using Optuna-based hyperparameter optimization.

According to experimental findings, ensemble-based methods—Random Forest in particular—consistently produced the best results in terms of ROC-AUC, accuracy, and F1-score under both EEG situations. While linear models had much poorer discriminative abilities, particularly in the eyes-closed condition, neural networks and XGBoost also demonstrated competitive performance. Furthermore, models trained on eyes-open EEG data typically performed slightly better and more consistently, indicating that this recording state offers richer discriminative information for the categorization of Alzheimer's disease.

Overall, the results emphasize how crucial model choice and EEG recording circumstances are to EEG-based Alzheimer's disease identification. The findings show that carefully designed EEG features in conjunction with ensemble learning techniques can successfully capture neurophysiological patterns associated with disease. Subsequent research may expand this methodology by adding multimodal EEG conditions, subject-independent validation techniques, and deep learning architectures that can learn straight from unprocessed EEG data.

REFERENCES

[1] A. Kumar, J. W. Tsao, J. Sidhu, and A. Goyal, "Alzheimer Disease," https://www.ncbi.nlm.nih.gov/books/NBK499922/, Feb. 2024, national Library of Medicine, Accessed: 12 Feb. 2024.

[2] University of Cambridge, "Artificial intelligence outperforms clinical tests at predicting progress of alzheimer's disease," https://www.cam.ac.uk/research/news/artificial-intelligence-outperforms-clinical-tests-at-predicting-progress-of-alzheimers-disease, 2024, accessed: 2025-03-08.

[3] InsideScientific, "Eeg monitoring approaches to predict learning and memory changes in early alzheimer's disease," https://insidescientific.com/webinar/eeg-monitoring-approaches-to-predict-learning-and-memory-changes-in-early-alzheimers-disease/, 2022, webinar, Accessed: 2025-03-08.