Diagnosis of Alzheimer's Disease via Resting State EEG - Integration of Spectrum, Complexity, and Synchronization Signal Features

Folasewa Maryam Abdulsalam (B02294068), Pratit Kandel (10459709)

Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by progressive cognitive decline, severely impacting the lives of the patients and their families. Recent research on AD has identified three key signal characteristics in the brain regions of AD patients: diffuse slowing of brain activity, reduced signal complexity, and decreased synchronization between the neural areas. However, traditional diagnostic methods often detect AD only at advanced stages, when significant neuronal degeneration has already occurred. This highlights the urgent need for improved diagnostic approaches to leverage these signal characteristics to enable earlier and more reliable detection.

In this study, the paper explored the signal preprocessing steps of the EEG recordings, followed by the integration of multi-domain EEG features: spectrum, complexity, and synchronization metrics. After feature extraction, the supervised learning classification methods: decision tree, random forest, and SVM, were used as the classifiers to develop a robust classification model for distinguishing AD patients from healthy individuals, with the highest accuracy of 95.86% for Random Forest. This approach demonstrates the need to combine signal processing with machine learning in order to enhance the diagnostic accuracy of AD progression.

Link: https://www.frontiersin.org/journals/aging-neuroscience/articles/10.3389/fnagi.2023.1288295/full

Methods

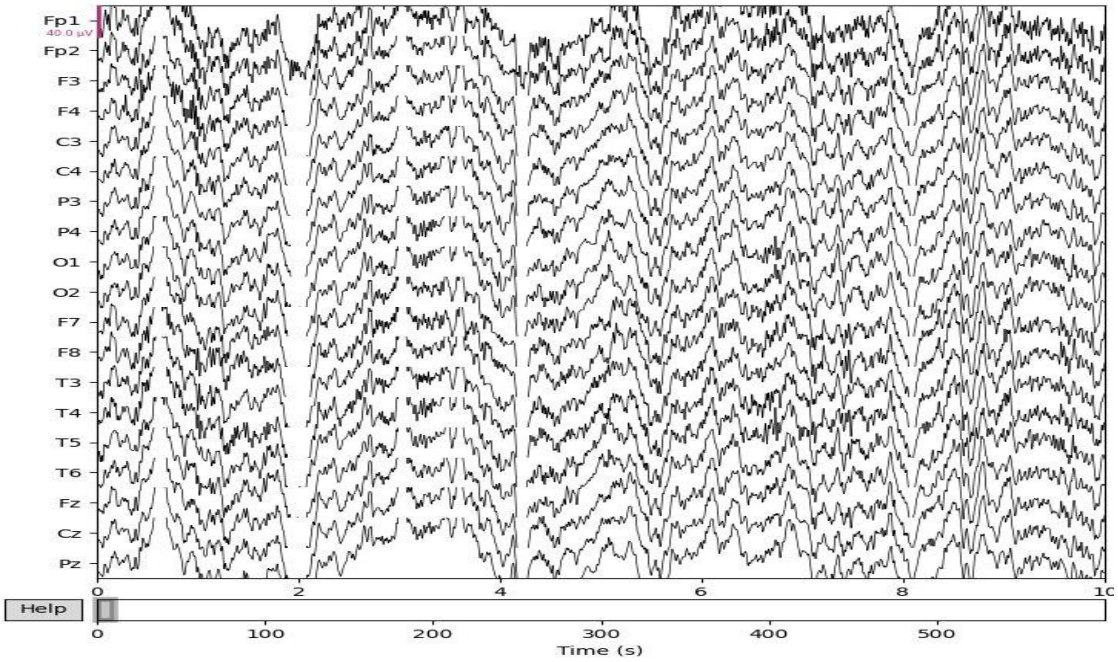
Database Description

We utilized a publicly available EEG resting-state dataset from OpenNeuro. This dataset included EEG recordings from 88 subjects, out of which 36 were diagnosed with Alzheimer's disease (AD group), 23 with Fronto-temporal Dementia (FTD group), and 29 as healthy controls (CN group). For this study, the FTD group was excluded and EEG recordings from only 65 subjects were utilized. EEG data were recorded using 19 scalp electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2), with an average recording duration of 13.5 minutes for the AD group and 13.8 minutes for the CN group. The data were sampled at a frequency of 500 Hz.

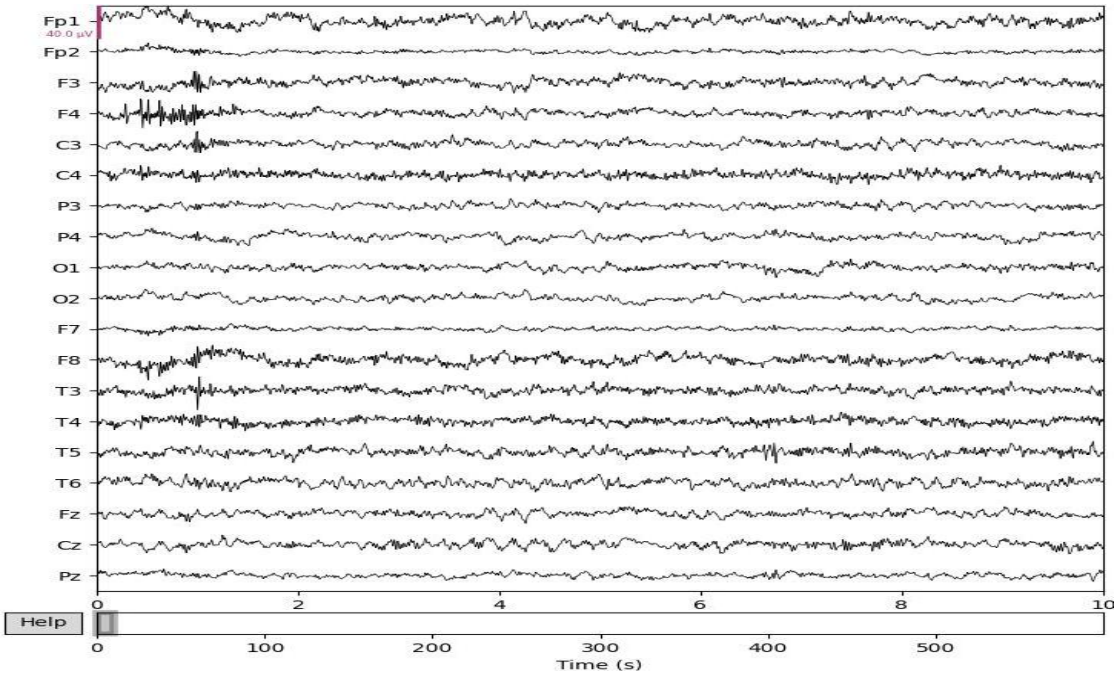
Data Preprocessing

EEG signal preprocessing involved re-referencing the EEG signals using an average referencing technique and applying a band-pass filter ranging from 0.5 Hz to 45 Hz. Noise levels were assessed by comparing the standard deviation of channel signals against a threshold of $6x10^{-5}$ Volts. Independent Component

Analysis (ICA) was applied when significant noise was identified. Artifact Subspace Reconstruction (ASR) was applied using parameters of 0.5 seconds duration and a threshold of 17 x 10^{-6}. The preprocessed data were stored for feature extraction. *Figure 1* shows the EEG recordings from 19 electrodes for one of the subjects *(a)* before and *(b)* after signal preprocessing.



(a)



(b)

Figure 1: EEG signal recordings (a) before and (b) after signal preprocessing

Feature Extraction

After the segmentation of EEG signals into 4-second epochs with 50% overlap, yielding 26,526 epochs (14,514 AD and 12,012 CN), three categories of features - spectrum, complexity, and synchronization were extracted from each epoch.

Spectrum Feature Extraction

The spectrum features were categorized into time-domain metrics and frequency-domain metrics. The time-domain metrics include mean, variance, and interquartile range (IQR), which were computed directly from the signal.

For a data segment x_j with length N, the mean, variance, and IQR metrics were calculated as shown in equations *(1), (2),* and *(3)* respectively.

$$\bar{x} = \frac{1}{N}\sum_{j=1}^{N} x_j \quad \text{.....(1)}$$

$$Var = \frac{1}{N-1}\sum_{j=1}^{N}(x_j - \bar{x})^2 \quad \text{.....(2)}$$

$$\text{IQR} = Q_3 - Q_1 \quad \text{.....(3)}$$

The frequency-domain spectrum metric is the power spectral density (PSD), calculated across the full frequency range (0.5–45 Hz) and further into the individual frequency bands (delta, theta, alpha, beta, and gamma). The relative band power (RBP) was calculated to normalize the PSD using equation *(4).*

$$RBP = \frac{Power\ in\ Specific\ Band}{Total\ Power} \quad \text{.....(4)}$$

These calculations were facilitated using libraries - **NumPy** and **SciPy**, and the results were saved in the CSV format.

Complexity Feature Extraction

Entropy measures the degree of complexity and predictability of a signal. For the complexity features, we extracted approximate entropy, permutation entropy, and sample entropy metrics as per the equations *(5), (6),* and *(7)* by using the following parameters according to the paper: Pattern length (m) = 2, R-factor (r) = 0.15, and Scale-factor (τ) = 5. We made use of the library - **Entropy Hub** for this task. After extracting the metrics for each epoch, the results were saved to the CSV file.

$$ApEn\ (m, r) = \Phi^m(r) - \Phi^{m+1}(r) \quad \text{.....(5)}$$

Where:

$$\Phi^m(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1} \ln C_i^m(r)$$

$$PermEn\ (m) = -\sum_{i=1}^{m!} P_i \log P_i \quad \text{.....(6)}$$

Where, P_i represents the probability associated with ordinal patterns i

$$SampEn\ (m,r) = -\ln\left(\frac{\Phi^m(r)}{\Phi^{m+1}(r)}\right) \text{.....(7)}$$

Synchronization Feature Extraction

For the synchronization features, the clustering coefficient, characteristic path length, global efficiency, and small-worldness metrics were extracted using equations *(8), (9), (10),* and *(11)* respectively.

$$C_i = \frac{2\ e_i}{k_i.(k_i-1)} \text{.....(8)}$$

$$L = \frac{1}{N.(N-1)} \sum_{i,j\in V, i\neq j} l_{ij} \text{.....(9)}$$

$$E_{global} = \frac{1}{N.(N-1)} \sum_{i,j\in V, i\neq j} \frac{1}{l_i} \text{.....(10)}$$

$$\sigma = \frac{\gamma}{\delta} \text{.....(11)}$$

To achieve this, for each epoch, we computed the connectivity matrix using the Pearson method, and then the result was applied a threshold value of 0.6 to take only the top 60% connections, and further the graph metrics were calculated. To achieve this, we made the use of **EntropyHub**, **Numpy,** and **NetworkX** libraries. After the metrics were extracted for each epoch, the results were again saved to the CSV file.

Classification Algorithm

Spectrum, complexity, and synchronization features were combined into a unified dataset with labels (1 for AD and 0 for CN). Supervised machine learning classifiers, including decision trees, random forests, Support vector machines (SVMs), and lightGBM were trained using GroupShuffleSplit cross-validation. This ensured the separation of subjects between training and testing datasets, thereby avoiding the overlap of subject-specific epochs across these sets. Metrics such as accuracy, sensitivity, and specificity were evaluated.

Results

1. Classification performance analysis

The classification done in our reference paper reported high classification accuracy for different models: Random Forest achieving 95.86% accuracy, Decision tree at 95.65%, and SVM at 88.54%. In contrast, the results from our study exhibited relatively lower classification performance, with the Random Forest achieving 70.09% accuracy, Decision tree at 68.68%, and SVM at 69.25%. Additionally, we incorporated LightGBM model in our study that achieved the classification accuracy of 66.82%.

Different evaluation metrics (accuracy, sensitivity and specificity) for these models were obtained as shown in *Table 1 (a)* according to the paper, and *(b)* as per our classification study.

Table 1 (a): Classification result according to the paper (with Leave-one-subject-out)

Models	**Accuracy**	**Sensitivity**	**Specificity**
Decision Tree	95.65%	95.91%	95.35%
Random Forest	95.86%	96.41%	97.40%
SVM	88.54%	94.72%	81.23%

Table 1 (b): Classification result according to our study (with GroupShuffleSplit)

Models	**Accuracy**	**Sensitivity**	**Specificity**
Decision Tree	68.68%	67.83%	69.77%
Random Forest	70.09%	75.33%	67.62%
SVM	69.25%	71.51%	69.07%
LightGBM	66.82%	76.66%	60.45%

Commonalities with the reference paper:

- Random Forest consistently performed the best in both studies, suggesting that it is well-suited for EEG-based AD classification.
- SVM had the lowest performance in both studies, reinforcing that linear separation may not fully capture the complexity of EEG features.

Possible reasons for discrepancies:

- Feature extraction differences: Our analysis did not include one of the entropy metrics which could have contributed to higher accuracy results.
- Electrode referencing effects: Our use of average referencing versus the A1-A2 referencing done in the paper could have impacted the classification accuracy.
- Noise filtering differences: The paper did not specify a threshold value due to which we implemented based on our assumption, which might potentially remove the valuable EEG information.
- Model hyperparameters: Differences in tuning, dataset balancing, and feature selection could also contribute to the classification accuracy differences.

2. Statistical t-test results

We conducted independent t-tests on different EEG domains (time-domain, frequency-domain, complexity, and synchronization) and compared our findings to the referenced study as shown in the *Tables 2, 3, 4 and 5* respectively.

Despite sticking to the paper's methodology to a greater extent, some discrepancies were however observed in our results. These differences could potentially be due to variations in preprocessing techniques, noise filtering, electrode re-referencing, and missing feature extractions, which are discussed in the following sections.

Table 2: Comparison of Time-domain metrics between Paper and our Data

Features	**Trends in Paper**	**Trends in our Data**	**Significance**
Mean	Increase in AD	Decrease ❌	Not Significant
Variance	Increase in AD	Increase ✅	Not Significant
IQR	Increase in AD	Decrease ❌	Not Significant

Possible reasons for discrepancies:

- **Electrode re-referencing differences:** The A1-A2 referencing in paper versus average referencing in our study may have altered the time-domain signal amplitude.
- **Noise filtering threshold:** Since the paper did not define a strict noise threshold, it could potentially filter out the meaningful variance.
- **Variability in EEG Dataset:** Different preprocessing techniques might have led to differences in the amplitude scaling.

Table 3: Comparison of Frequency-domain metrics between Paper and our Data

Features	**Trends in Paper**	**Trends in our Data**	**Significance**
Delta	Increase in AD	Increase ✅	Significant ($p < 0.05$)
Theta	Increase in AD	Increase ✅	Highly Significant ($p < 0.001$)
Alpha	Decrease in AD	Decrease ✅	Highly Significant ($p < 0.001$)
Beta	Decrease in AD	Decrease ✅	Significant ($p < 0.05$)
Gamma	Decrease (insignificant)	Increase ❌	Not significant

Commonality with the reference paper:

- Our delta, theta, alpha, and beta trends align well with the study in the paper, reinforcing the hypothesis that AD causes EEG slowing.

Possible reason for discrepancies:

- **Gamma Power increment:** The paper reported a decrease in gamma, whereas we observed its increase. This might be due to:
  - Variations in EEG sensor placements between datasets.
  - Differences in filtering settings, affecting high-frequency noise.

Table 4: Comparison of Complexity metrics between Paper and our Data

Features	**Trends in Paper**	**Trends in our Data**	**Significance**
ApEn	Decrease in AD	Decrease ✓	Significant ($p < 0.05$)
PermEn	Decrease in AD	Increase ✗	Significant ($p < 0.05$)
SampEn	Decrease in AD	Decrease ✓	Significant ($p < 0.05$)

Possible reasons for discrepancies:

- **Lack of Multi-scale Entropy (MSE):** The study in the paper included MSE but this metric was not extracted in our analysis, potentially affecting overall complexity trends.
- **PermEn increment:** Contrary to the expectations, PermEn increased in AD rather than decrement. This could be influenced by:
  - Noise reduction techniques applied in preprocessing.
  - Threshold-based filtering may have introduced biases in entropy calculations.

Table 5: Comparison of Synchronization metrics between Paper and our Data

Features	**Trends in Paper**	**Trends in our Data**	**Significance**
Clustering Coefficient	Decrease in AD	Increase ✗	Not significant
Path Length	Increase in AD	Increase ✓	Not significant
Global Efficiency	Decrease in AD	Decrease ✓	Not significant
Small-worldness	Decrease in AD	Increase ✗	Not significant

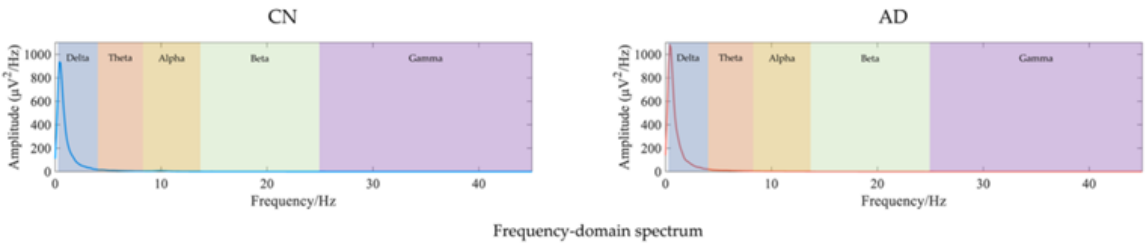
Possible reasons for discrepancies:

- **Electrode re-referencing effects:** Small-world properties and clustering coefficients depend on signal coherence, which is influenced by referencing techniques.
- **Graph reconstruction differences:** The method for computing network topology may have varied between the studies. Information was not explicitly provided in the paper.
  - Different thresholding techniques for adjacency matrices.
  - Different functional connectivity measures (e.g., phase lag index vs. correlation-based methods).

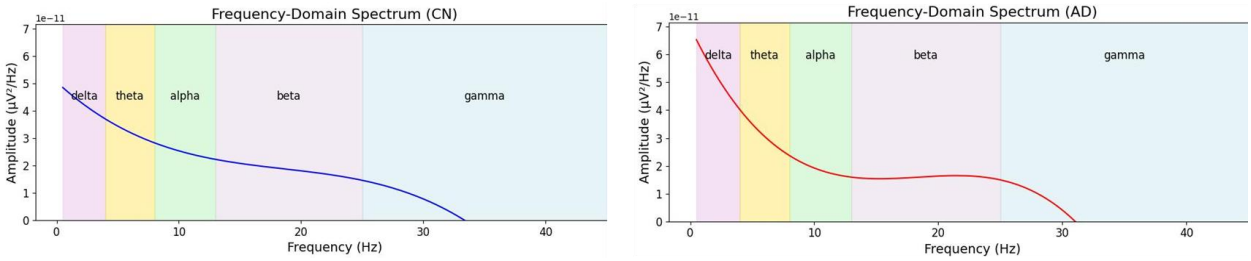
3. Visualizations

Frequency-domain spectrum analysis

Furthermore, we compared the frequency-domain spectrum plots of CN and AD subjects from the referenced paper to our study as shown in *Figure 2*.



(a)



(b)

Figure 2: Comparison of frequency-domain spectrum between (a) the paper, and (b) our study

Commonalities with the reference paper:

- The AD spectrum shows a shift toward lower frequencies (delta and theta) and a relative reduction in higher frequencies (alpha, beta, and gamma), consistent with the paper's findings.
- AD has higher delta power than CN in both plots, which is linked to cognitive decline in AD patients.

Discrepancy with the reference paper:

- The referenced paper's plot has a clear amplitude decay curve, while the plot in our study appears more uniform across frequencies.

Reason for discrepancies:

- **EEG Sensor Configuration:** The referenced study's dataset used a referencing scheme (A1-A2) different from what we did and this may have influenced spectral power distribution.
- **Use of UnivariateSpline:** This specific smoothing technique used in our study could have provided the plot which is different from the one in the paper.

Further, *Figure 3* shows the bar plot diagram for the comparison of PSD across different frequency bands between CN and AD groups. With compliance to the above findings, AD patients show significantly higher delta power compared to CN individuals. Increased delta activity is often associated with cognitive decline and neuronal dysfunction. Theta power is slightly higher in AD patients than CN, which aligns with the findings that suggest increased theta activity in neurodegenerative disorders. However, AD patients exhibit significantly lower alpha power compared to CN individuals which is linked to impaired cognitive processing and memory deficits. Likewise, the AD patients have comparatively lower beta and gamma power which may reflect disrupted neural synchrony and cognitive processing deficits. Hence, AD patients exhibit increased low-frequency (delta and theta) activity and reduced high-frequency (alpha, beta, and gamma) activity. This pattern is consistent with previous findings in EEG studies on Alzheimer's disease, indicating neural dysfunction and cognitive decline.

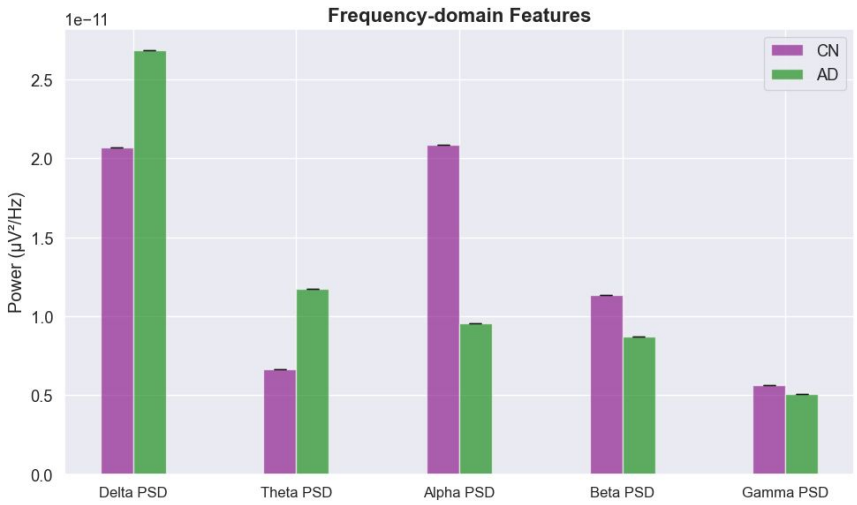


Figure 3: Comparison of frequency-domain features between CN and AD groups

Correlation matrix

From the *Figure 4,* the CN group exhibits more widespread high correlation values (red/yellow clusters) across different electrode pairs which suggests a more stable and synchronized neural network, consistent with findings that CN individuals have stronger global and local connectivity. However, the AD group shows more scattered and less pronounced connectivity patterns, with fewer high-correlation (red) areas. This aligns with disconnection syndrome in Alzheimer's disease, where functional connectivity weakens due to neuronal loss and synaptic dysfunction.

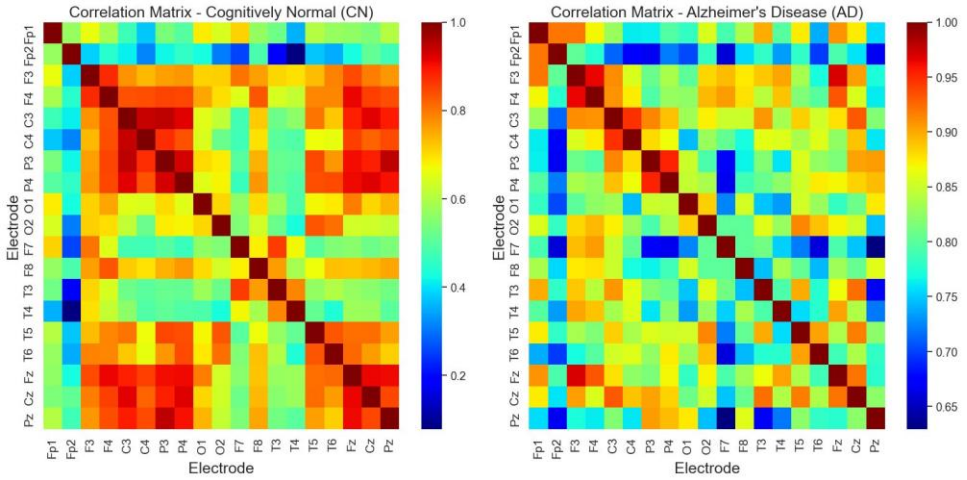


Figure 4: Correlation matrix for CN and AD groups

Classification models

In order to compare the accuracy, sensitivity and specificity of all four models for classifying CN and AD groups, the bar chart was drawn as shown in *Figure 5*. Although all the models have almost similar accuracy levels, Random Forest appears to be the best model with the highest accuracy percentage followed by SVM model. Similarly, LightGBM has the highest sensitivity compared to all other models

which is followed by Random Forest. Finally, Decision Tree and SVM models have the highest and almost similar specificity percentage whereas LightGBM has the least in comparison to all other models.

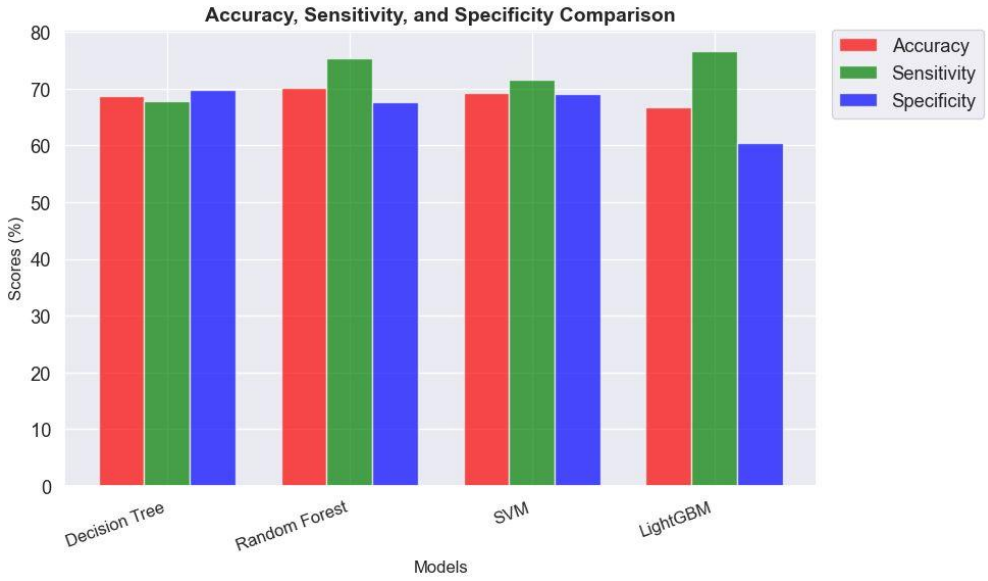


Figure 5: Comparison of performance metrics between different models

Discussion

Our results highlight both commonalities and discrepancies compared to the reference study. The classification performance in our study was lower, with the random forest achieving 70.09% accuracy compared to 95.86% in the reference study. These differences may be attributed to variations in feature selection, preprocessing methods, and dataset structure. Notably, the lack of one entropy metric and differences in EEG referencing techniques may have contributed to the reduced model performance.

The statistical analysis shows strong agreement in delta, theta, alpha, and beta bands, all aligning with expected AD-related EEG slowing. However, discrepancies in Gamma band trends and complexity metrics suggest differences in noise filtering and entropy computation techniques. Additionally, the path length increase in AD is consistent across studies, reinforcing the hypothesis that AD patients have less efficient brain networks.

Regarding visualizations, both studies show increased delta and theta power in AD and reduced alpha, beta, and gamma power, confirming the characteristic cortical slowing in AD. Also, from the correlation matrix, the AD group has more scattered connectivity patterns and fewer high-correlation areas which suggest that Alzheimer's disease is associated with disrupted neural synchronization and weakened functional connectivity, reflecting neuronal loss and synaptic dysfunction.

Overall, these findings emphasize the importance of standardizing EEG preprocessing pipelines and feature extraction methods to improve reproducibility across studies. Future work should explore alternative normalization techniques, noise filtering strategies, and dataset balancing methods to enhance classification accuracy and ensure better comparability with existing literature.

Conclusion

Our replication successfully captures some major trends observed in the reference paper, such as increased delta power and decreased entropy in AD, but exhibits weaker or inconsistent differences in other domains, including time-domain variability, frequency power distributions, and synchronization metrics. Also, the differences in performance metrics for different models can be observed in our classification study compared to that in the reference paper.

The observed discrepancies likely arise from differences in preprocessing, sample composition, and statistical methods. Future refinements should include ensuring consistent filtering, normalization, and statistical validation to strengthen comparability. Further dataset preprocessing refinements need to be taken into consideration to improve result alignment with the reference study.