ECE496 Proposal

EEG Based Alzheimer Risk Assessment With Machine Learning

Project Number: 2024016

Supervisor: Berj Bardakjian

Administrator: Stewart Aitchison

Members: Xi Ping Chen, Xiangyu Chen, Bavya Mittal, Dhrumil Patel

Date of Submission: September 17, 2024

1. Background and Motivation

Alzheimer's Disease (AD) is one of the leading causes of dementia, affecting over 55 million people globally, and this number is expected to triple by 2060, making it a critical health concern worldwide [1]. AD leads to progressive cognitive decline, impairing memory, thinking, and behavior, with a significant impact on individuals' quality of life. Currently, no cure exists, and early detection remains crucial for effective management and intervention. However, traditional diagnostic methods, such as MRI, PET, and cerebrospinal fluid (CSF) analysis, are either invasive, expensive, or insufficient for early detection [2], [3].

In contrast, EEG is a cost-effective and non-invasive method, able to reflect changes in brain function even before clinical symptoms appear [4]. Non-invasive methods such as electroencephalography (EEG) offer an accessible and cost-effective alternative for diagnosing AD, as EEG can detect changes in brain rhythms that may indicate neurodegeneration. EEG-based diagnostic techniques have demonstrated potential for identifying patterns linked to AD [5].

Recent research has emphasized the importance of analyzing Cross-Frequency Coupling (CFC) which has been shown to reflect critical cognitive processes impaired by Alzheimer's, such as memory and learning [6]. Studies suggest that disruptions in these oscillatory couplings can serve as early biomarkers for AD, offering a new avenue for diagnosis [7][8].

The gap in utilizing CFC features in EEG data for Alzheimer's detection presents a significant opportunity. Integrating machine learning (ML) with EEG-based CFC analysis has shown promising results. A study using ML models to classify EEG signals has demonstrated that combining CFC features with ML algorithms such as support vector machines (SVMs) and convolutional neural networks (CNNs) can improve diagnostic accuracy, making this a promising area of research for non-invasive early detection tools [9 , 10].

Our project is a continuation of the thesis project of Sofya Smirnova done in April 2024. The machine learning model developed for Alzheimer's Disease (AD) detection from EEG data achieved mixed performance. While training results showed promise with validation accuracies approaching 80%, the final model using stratified k-fold cross-validation yielded an accuracy of 69%, with a sensitivity of 66.67% and specificity of 72.59%. The model did not outperform baseline models that used other features such as time-frequency EEG data [11].

Sofya's thesis demonstrated the potential of using EEG for early detection of Alzheimer's Disease, providing insights into neurophysiological changes linked to the disease. However, the current model did not surpass the diagnostic accuracy of existing methods in literature. We hope to extend her research by potentially exploring other machine learning models and using CFC patterns [11].

2. Problem Statement

We hope to create a non-invasive, cost-effective diagnostic tool for early Alzheimer's detection, enabling timely interventions to slow disease progression.

3. Project Goal

This project aims to develop a machine learning model that surpasses the results achieved in Sofya's thesis by utilizing CFC patterns and exploring various machine learning models.

4. Scope of Work

This project aims to create an AI model that distinguishes between healthy individuals and those with Alzheimer's using EEG signals. We will use data from Nemar [12], and if it is insufficient, possibly additional data from other sources. Key challenges lie in preprocessing the raw EEG data, extracting CFC images, choosing and building an adequate model architecture, and tuning hyperparameters to meet and exceed the accuracy of [11] and current state-of-the-art methods.

4.1 Original Contributions

As part of feature engineering, we will convert raw EEG data into phase amplitude cross-frequency coupling (CFC) data using custom scripts. This will allow us to better extract features. Preprocessing steps will include noise removal and artifact correction. For neural network development, we plan to design a model that integrates various types of neural networks, experimenting with both raw EEG data and processed CFC data. The network will be optimized for accuracy in classifying Alzheimer's patients. If necessary, additional datasets will be used to enhance performance.

4.2 Use of Existing Resources

The project will leverage the Nemar [12] dataset and may use other EEG datasets. We will use libraries such as TensorFlow and PyTorch for model development and other python libraries for processing the dataset. A limitation to our work is the availability of EEG datasets.

4.3 Outside the Scope

Brainwave related research, real time EEG measurement and hardware development are beyond the scope of this project. The emphasis will be on training and validating the model using pre-existing data.

4.4 Demonstration

For our demonstration, we will use pre-processed EEG signals as transforming data can take some time, and taking live EEG signals will not be possible. We will visualize preprocessed signals and then run our model. The focus will be on running the model and showcasing performance metrics such as accuracy and precision.

4.5 System Context Diagram

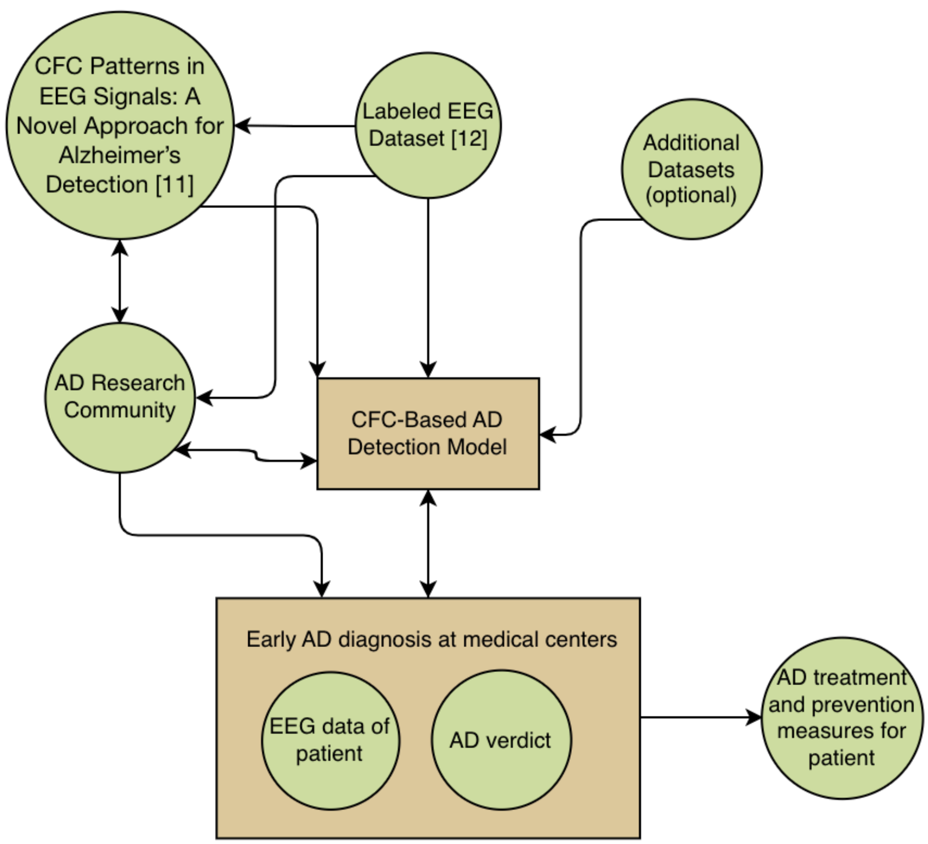

Figure 1: Project in context.

5 Requirements Specification (Preliminary)

ID	Project Requirement	Description
1	Achieve accuracy higher than 69.58% [11]	**Functional Requirement:** The final machine learning model should achieve this accuracy after testing
2	Achieve sensitivity greater than 66.67% [11]	**Functional Requirement:** Sensitivity is described as True Positives over True Positives + False Negatives. In medical related fields, False Negatives should be minimized as they can have harmful implications
3	Ensure correct EEG signals are displayed after processing data	**Functional Requirement:** After data processing, it is important to make sure that output is an EEG signal, as this will be the input to machine learning model and must accurately depict characteristic of EEG signals
4	The machine learning model must classify input as either True or False for Alzheimers	**Functional Requirement:** Model must predict given an input whether Alzheimer's is detected. This should be a true/false answer.
5	The raw EEG data must be pre-processed to remove noise	**Functional Requirement:** Any data from the dataset, should be pre-processed to remove noise and artifacts to ensure effective signal analysis
6	Machine Learning model must comply with legal regulations and false negatives should be minimized	**Constraint:** Since this is a tool for the medical industry, false negatives can result in patients with Alzheimer's falsely

		being identified, which can have legal implications
7	Budget of $200 [13]	**Constraint:** System should operate under this budget for computational resources
8	Data source should be Nemar [12]	**Constraint:** This is the only dataset that can be used for training this model as it has been verified and tested previously
9	System must be compatible with Python and Google Colab	**Constraint:** Most machine learning libraries are readily available in Google Colab and Python, so system must be compatible with both
10	Maximize detection accuracy while minimizing false positives and negatives	**Objective:** Increase the system's accuracy with few amount of false positives and negatives, as they can have implications on health
11	Minimize computational time (less than 5 min per one hour of EEG data)	**Objective:** Minimizing computational time allows for more convenience while using and can offer near real-time analysis
12	Improve scalability and generalization for large dataset	**Objective:** Generalization will allow the model to be used more frequently when diagnosing.
13	Enhance model results and interpretability	**Objective:** Having clear results and information can provide insights into the patterns that contribute to Alzheimer's detection

6 References

[1] “Dementia,” World Health Organization, Available: https://www.who.int/news-room/fact-sheets/detail/dementia (accessed Oct. 3, 2023).

[2] Alzheimer’s disease fact sheet — National Institute on Aging, https://www.nia.nih.gov/health/alzheimers-and-dementia/ alzheimers-disease-fact-sheet (accessed Apr. 15, 2024).

[3] F. A. Alonso, M. M. Lopez, and J. P. Rodriguez, "Aging and Neurodegeneration: Advances in Cross-Frequency Coupling Analysis for Early Detection of Alzheimer’s Disease," *Frontiers in Aging Neuroscience*, vol. 15, art. no. 1288295, pp. 1-14, Sep. 2023. [Online]. Available: https://www.frontiersin.org/journals/aging-neuroscience/articles/10.3389/fnagi.2023.1288295/full

[4] “What is alzheimer’s disease?,” Centers for Disease Control and Prevention, https://www.cdc.gov/aging/aginginfo/alzheimers.htm (accessed Apr. 15, 2024).

[5] P. Durongbhan et al., ”A Dementia Classification Framework Using Frequency and Time-Frequency Features Based on EEG Signals,” in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 5, pp. 826-835, May 2019, doi: 10.1109/TNSRE.2019.2909100. Available: https://ieeexplore.ieee.org/document/8681406 (accessed Oct. 2, 2023).

[6] V. Grigorovsky and B. L. Bardakjian, “Low-to-high cross-frequency coupling in the electrical rhythms as biomarker for hyperexcitable neuroglial networks of the brain,” IEEE Transactions on Biomedical Engineering, vol. 65, no. 7, pp. 1504–1515, 2018. doi:10.1109/tbme.2017.2757878. Available: https://tspace.library.utoronto.ca/bitstream/1807/79760/1/Low-to-High%20Cross-Frequency%20Coupling_Tspace.pdf (accessed Oct. 2, 2023).

[7] C. Lucasius, V. Grigorovsky, H. Nariai, A. S. Galanopoulou, J. Gursky, S. L. Moshe, and B. L. Bardakjian, ”Biomimetic Deep Learning Networks With Applications to Epileptic Spasms and Seizure Prediction,” IEEE Trans. Biomed. Eng., vol. 71, no. 3, pp. 1056-1067, Mar. 2024, doi: 10.1109/TBME.2023.3325762. Epub Feb 26, 2024.

[8] L. Cao et al., "Alzheimer’s disease diagnosis using resting-state EEG: exploring the clinical potential of new machine learning models," *Frontiers in Aging Neuroscience*, vol. 15, Article 1142085, Aug. 2023. doi: 10.3389/fnagi.2023.1142085.

[9] X. Zheng et al., "Diagnosis of Alzheimer’s disease via resting-state EEG: integration of spectrum, complexity, and synchronization signal features," *Frontiers in Aging Neuroscience*, vol. 15, Article 1288295, Nov. 2023. doi: 10.3389/fnagi.2023.1288295.

[10] J. Wang, Y. Fang, X. Wang, H. Yang, X. Yu, and H. Wang, "Enhanced Gamma Activity and Cross-Frequency Interaction of Resting-State Electroencephalographic Oscillations in Patients with

Alzheimer’s Disease," *Frontiers in Aging Neuroscience*, vol. 9, Article 243, July 2017. doi: 10.3389/fnagi.2017.00243.

[11] S. Smirnova, Bardakjian, “Cross-Frequency Coupling Patterns in EEG Signals: A Novel Approach for Alzheimer’s Detection,” B.A.Sc. thesis, Engineering Science, UT, Toronto, ON, Canada, 2024.

[12] “Data search: Detail A dataset,” Nemar, https://nemar.org/dataexplorer/detail?dataset_id=ds004504&processed=0 (accessed Jan. 20, 2024).

[13] “Plans & pricing to upgrade your cloud storage,” Google One, https://one.google.com/about/plans?g1_landing_page=0