DEEP LEARNING FOR EEG-BASED ALZHEIMER'S DISEASE DETECTION USING A CNN-BILSTM-ATTENTION MODEL

A PROJECT REPORT

Submitted by

MEGHAVARSHINI M [RA2211026050054]
MANU SEBASTIAN [RA2211026050005]
MRISHIKA D [RA2211026050061]

Under the guidance of

Dr. P.K.A Chitra, Ph.D
(Assistant Professor, Department of Computer Science and Technology)

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND TECHNOLOGY WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

of

FACULTY OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE & TECHNOLOGY
(Deemed to be University u/s 3 of UGC Act, 1956)
TIRUCHIRAPPALLI

SRM Nagar, Trichy-Chennai Highway, Tiruchirappalli - 62115

MAY 2025

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY TIRUCHIRAPPALLI

BONAFIDE CERTIFICATE

Certified that this project report titled "**DEEP LEARNING FOR EEG-BASED ALZHEIMER'S DISEASE DETECTION USING A CNN-BILSTM-ATTENTION MODEL**" is the bonafide work of "**MEGHAVARSHINI M [RA2211026050054], MANU SEBASTIAN [RA2211026050005],MRISHIKA D[RA2211026050061]**", who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Dr. P.K.A Chitra, Ph.D
GUIDE
Assistant Professor
Department of Computer Science and Technology

SIGNATURE

Dr. Balaji Ganesh
HEAD OF THE DEPARTMENT
Department of Computer Science and Technology with specialization in Artificial Intelligence & Machine Learning

Signature of the Internal Examiner

Signature of the External Examiner

ABSTRACT

Alzheimer's Disease (AD) diagnosis using Electroencephalogram (EEG) signals offers a non-invasive and accessible approach for early detection. In this work, we propose a novel deep learning framework integrating Convolutional Neural Network (CNN)s, Bidirectional Long Short-Term Memory (BiLSTM) networks, and an **attention!** (**attention!**) mechanism for classifying EEG recordings to identify Alzheimer's disease. Our approach utilizes a feature-rich input representation derived from 18 EEG channels, where each channel is characterized by 11 statistical and nonlinear features (including Singular Value Decomposition (SVD) Entropy, Detrended Fluctuation Analysis (DFA), Zero-Crossing Rate (ZCR), Higuchi Fractal Dimension (HFD), Log Band Power (LBP), Standard Deviation (STD), Variance (VAR), Kurtosis (KUR), Absolute Energy (AE), Root Mean Square (RMS), and L2 Norm (NO)). The CNN layers learn spatial patterns across channels, the BiLSTM captures sequential dependencies among the processed channel features, and the **attention!** mechanism allows the model to focus on the most discriminative parts of the sequence. Experimental evaluation demonstrates the model's potential in distinguishing AD-related neural activity from healthy controls, highlighting the effectiveness of combining spatial and temporal learning with attention for enhanced classification performance.

Keywords: Alzheimer's Disease, EEG, Deep Learning, CNN, BiLSTM, Attention Mechanism, Feature Extraction

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my guide, Dr.P.K.A Chithra for her valuable guidance, consistent encouragement, personal caring, timely help and providing me with an excellent atmosphere for doing research. All through the work, in spite of her busy schedule, she has extended cheerful and cordial support to me for completing this research work.

Meghavarshini M

Manu Sebastian

Mrishika D

TABLE OF CONTENTS

ABSTRACT **iii**

ACKNOWLEDGEMENTS **v**

LIST OF TABLES **viii**

LIST OF FIGURES **ix**

ABBREVIATIONS **x**

LIST OF SYMBOLS **xi**

1 INTRODUCTION **1**

2 LITERATURE SURVEY **2**

3 METHODOLOGY **4**

3.1 Datasets 4

3.1.1 Dataset Overview 4

3.2 Feature Extraction 6

3.3 Proposed Deep Learning Model: CNN-BiLSTM-Attention 7

3.4 Training and Evaluation 8

3.5 Feature Importance Analysis (Optional but Recommended) 8

4 RESULTS AND DISCUSSIONS **9**

4.1 Experimental Setup 9

4.2 Performance Evaluation 9

4.3 Discussion of Results 10

4.4 Feature and Attention Insights 10

4.5 Limitations 13

5 CODING AND TESTING **14**

5.1 Development Environment and Libraries 14

5.2 Implementation Details 14

5.3 Testing Procedures 15

6 CONCLUSION **16**

7 FUTURE ENHANCEMENT **17**

A EEG PREPROCESSING AND FEATURE COMPUTATION DETAILS **18**

A.1 Preprocessing Pipeline 18

A.2 Feature Computation 18

A.3 Data Formatting for Model 19

B CNN-BiLSTM-ATTENTION MODEL HYPERPARAMETERS **20**

LIST OF TABLES

LIST OF FIGURES

3.1 **Architecture Diagrams** 5

4.1 **Training and Validation Performance Curves** 10

4.2 **Feature Importance Analysis Results** 11

4.3 **Exploratory Feature Analysis Results** 11

4.4 **[Caption describing power ratio comparison figure]** 12

4.5 **[Caption describing mean LPB figure]** 12

ABBREVIATIONS

AD	Alzheimer's Disease
AE	Absolute Energy
Attention	Attention Mechanism
BiLSTM	Bidirectional Long Short-Term Memory
CN	Cognitively Normal
CNN	Convolutional Neural Network
DFA	Detrended Fluctuation Analysis
EEG	Electroencephalogram
EMD	Empirical Mode Decomposition
HFD	Higuchi Fractal Dimension
KUR	Kurtosis
LBP	Log Band Power
MMSE	Mini-Mental State Examination
MRI	Magnetic Resonance Imaging
NO	L2 Norm
RMS	Root Mean Square
SVD	Singular Value Decomposition
ZCR	Zero-Crossing Rate
STD	Standard Deviation
VAR	Variance

LIST OF SYMBOLS

N	Number of samples in the signal
VAR	Variance of the signal
x_i	The i-th sample of the signal
μ_x	Mean of the signal x

CHAPTER 1

INTRODUCTION

AD is a progressive neurodegenerative disorder and the primary cause of dementia, impacting approximately 50 million individuals worldwide, with projections indicating a tripling of prevalence by 2050 due to aging populations. The disease significantly impairs memory, cognition, and daily functioning, imposing a considerable burden on healthcare systems and families. Early diagnosis of AD remains challenging due to overlapping symptoms with normal aging and other dementias. Conventional diagnostic tools, such as Magnetic Resonance Imaging (MRI) and the Mini-Mental State Examination (MMSE), are effective but often expensive, time-intensive, or subject to interpretive variability. EEG emerges as a promising alternative due to its affordability, accessibility, and ability to detect subtle brain activity changes associated with AD.

The analysis of EEG signals for AD detection often involves extracting relevant features that capture alterations in brain dynamics. Machine learning and deep learning techniques have shown considerable success in classifying AD based on these features. While CNNs are effective at learning spatial hierarchies and **LSTM!** (**LSTM!**)s excel at capturing temporal dependencies, a hybrid approach that leverages the strengths of both, combined with an **attention!** mechanism to focus on salient information, holds significant potential for improving classification accuracy and interpretability in complex EEG data. This study proposes and evaluates a novel deep learning framework that integrates CNN, BiLSTM, and **attention!** layers for EEG-based AD classification using a rich set of extracted features.

CHAPTER 2

LITERATURE SURVEY

Numerous studies have explored the use of EEG for the detection and diagnosis of AD and related cognitive impairments, leveraging both traditional machine learning and deep learning techniques. Kim et al. (2024) [2] proposed an EEG-based classification system during computer-based cognitive testing, achieving high accuracy in distinguishing AD spectrum disorders, emphasizing the role of task-specific EEG patterns. Xia et al. (2023) [1] introduced a deep pyramid CNN model for AD diagnosis using EEG signals, reporting competitive performance by capturing hierarchical features, though their approach relied heavily on complex architectures without metadata integration. Similarly, an EEG-based clinical decision support system using Empirical Mode Decomposition (EMD) and deep learning, as reported in Frontiers in Human Neuroscience (2023) [3], demonstrated the efficacy of signal decomposition techniques, but lacked scalability due to computational demands.

Systematic reviews, such as those in Bioengineering (2023) [4] and The Open Bioinformatics Journal (2020), have synthesized machine and deep learning trends in EEG-based AD detection, highlighting accuracies ranging from 80% to 90% with features like spectral power and entropy. These reviews underscore the potential of kurtosis-based de-noising, as explored in a 2015 study (Biomedical Signal Processing and Control), which improved EEG signal quality for AD analysis, though it focused on preprocessing rather than classification.

Integrated approaches combining EEG with neuropsychological data, as seen in Applied Sciences (2022) [6], and deep learning methods for MCI detection, as in Sensors (2020) [5], have further advanced the field by incorporating multimodal data, achieving accuracies up to 89%. However, these studies often rely on epoch-level analysis or larger datasets, limiting their applicability in resource-constrained settings. A review of EEG-based machine learning for MCI and AD (Bioengineering, 2019) and a recent study (Biomedical Signal Processing and Control, 2024) emphasize the growing trend of deep learning, with models like LSTM and

CNNs outperforming traditional classifiers. Hybrid models combining CNNs and **LSTM!**s have shown promise in capturing both spatial and temporal features in EEG [**?**]. Furthermore, the incorporation of **attention!** mechanisms has demonstrated improved performance by allowing models to weigh the importance of different input segments [**?**]. Despite these advances, challenges remain, including handling feature complexity, integrating multi-channel information effectively, and improving model interpretability. Our study aims to address these gaps by proposing a CNN-BiLSTM-Attention Mechanism (Attention) model that processes a comprehensive set of engineered features from multiple EEG channels, offering a robust and potentially more interpretable approach for AD detection.

CHAPTER 3

METHODOLOGY

3.1 Datasets

We used publicly available EEG recordings from OpenNeuro for the development of our Alzheimer's disease classification model. These datasets were selected for their quality, richness of electrophysiological signals, and adherence to standardized experimental protocols. The recordings provide detailed insights into brain activity across multiple cognitive states, enabling effective discrimination between healthy controls and Alzheimer's-affected individuals.

3.1.1 Dataset Overview

Sources and Materials: The EEG data were sourced from OpenNeuro, a trusted repository for neuroimaging datasets. Specifically, we utilized recordings labeled for Alzheimer's diagnosis, which were collected using clinically approved EEG systems under controlled settings. These datasets contain multi-channel EEG signals sampled at consistent rates and annotated for clinical metadata including subject age, diagnosis, and cognitive status. **Structure:** Each dataset contains raw EEG recordings from multiple scalp electrodes arranged according to the international 10-20 system. The recordings cover both resting-state and task-based cognitive paradigms. Subjects were categorized into two groups Alzheimer's patients and healthy controls to facilitate binary classification. The dataset comprises 26,207 samples, with each sample representing a collection of features extracted from a single EEG epoch.

Format and Language: The original recordings were stored in EEGLAB-compatible formats (e.g., .eeg, .set). For this study, the data were processed and analyzed using Python tools, resulting in a structured dataset for machine learning input.

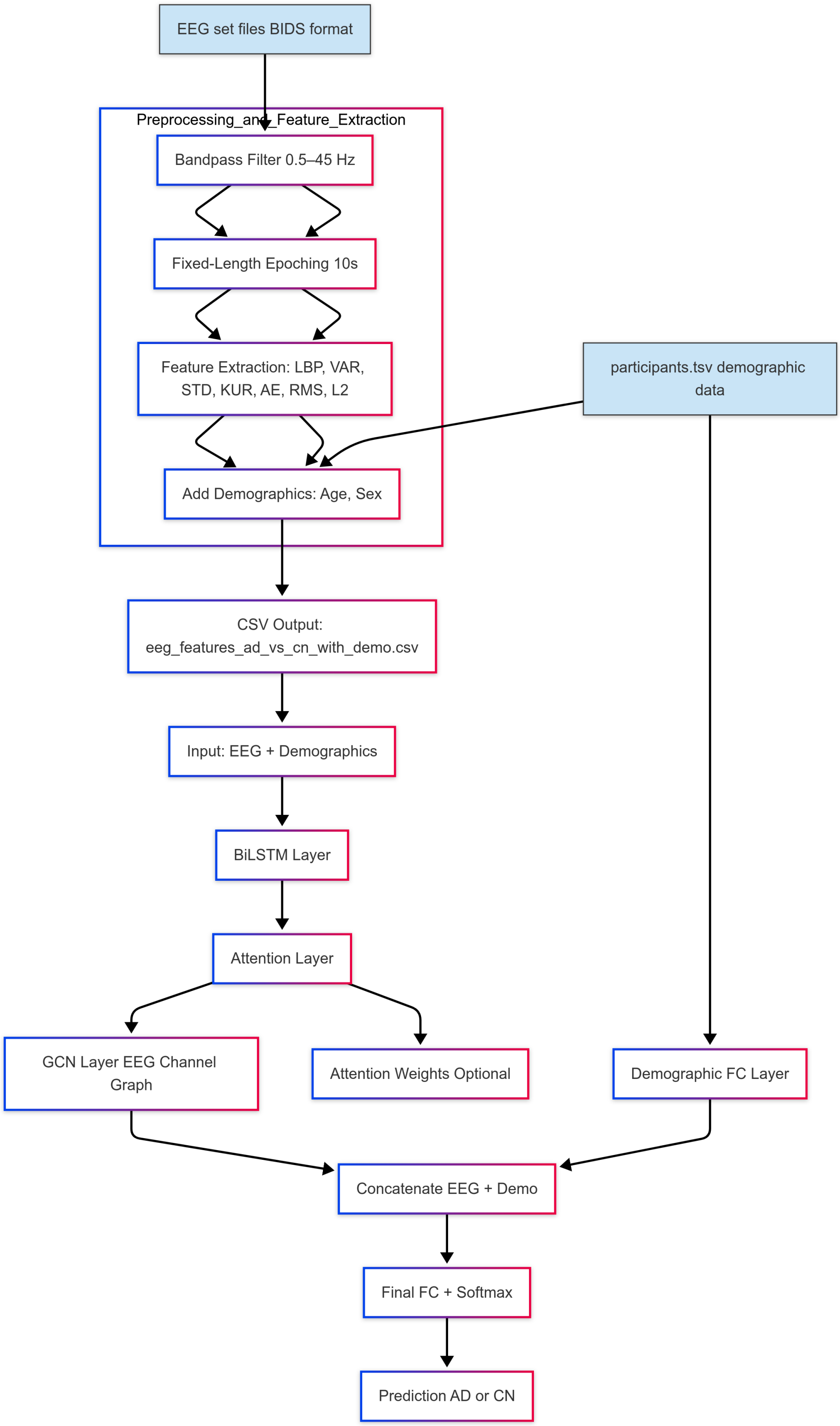

Figure 3.1: Architecture Diagrams

3.2 Feature Extraction

To quantify the electrophysiological characteristics of each EEG epoch, eleven statistical and nonlinear features were computed using a custom-built Python pipeline. The features were extracted from preprocessed EEG signals and aggregated across 18 standard EEG channels (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3/T7, T4/T8, T5/P7, T6/P8, Fz, Cz, Pz - adjust channel list if 18 are different from 19 standard) following the international 10-20 system, resulting in 18 channels * 11 features/channel = 198 features per sample.

The eleven features extracted per channel are:
SVD Entropy **DFA** **ZCR** **HFD** **LBP** (Log Band Power in the Alpha band)
STD **VAR** **KUR** **AE** (Absolute Energy) **RMS** **NO** (L2 Norm)

The extraction process involved loading preprocessed EEG data, segmenting into epochs (if not already done), and computing these 11 features for each channel within each epoch.

Preprocessing Steps

Prior to feature extraction, raw EEG signals underwent standard preprocessing steps, which typically include:

- Filtering (e.g., band-pass filtering to retain relevant frequencies and notch filtering to remove power line interference).
- Artifact removal (e.g., using ICA or other methods to mitigate ocular, muscle, or other noise).
- Referencing (e.g., re-referencing to a common average or linked mastoids).
- Epoching (segmenting continuous data into discrete time windows).

(Detailed preprocessing steps specific to your data should be provided in Appendix A).

Feature Computation Details

For each epoch and each channel, the 11 features were calculated. The formula for Variance (VAR) is given as:

$$\text{VAR} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2 \tag{3.1}$$

where N is the number of samples in the epoch/segment of the signal x for a given channel, x_i is the i-th sample, and μ_x is the mean of that signal segment. Similar computational definitions or references for the other 10 features can be included in Appendix A or B if space permits.

Data Structuring for Model Input

After feature extraction, the data for each sample (epoch) was structured into a tensor of shape ‘(num$_{c}hannels, num_{f}eatures_{p}er_{c}hannel)$‘, $whichis$‘$(18, 11)$‘$forinputintothedeeplearningmodel.Thiste$

3.3 Proposed Deep Learning Model: CNN-BiLSTM-Attention

The core of our framework is a hybrid deep learning model designed to capture the multi-faceted nature of EEG features. The architecture sequentially combines CNN, BiLSTM, and **attention!** layers.

The model architecture is as follows:

1. **Input Layer:** Accepts tensors of shape ‘(batch$_{s}ize, 18, 11)$‘.**Permute Layer:**$Rearrangesdimensionsf$
2. **CNN Block:**
   - **Conv1D Layer:** Applies 1D convolutions across the channel dimension (size 18) to learn local spatial patterns. Uses a kernel size (e.g., 3), padding, and outputs a defined number of filters (e.g., 64 or 128).
   - **Activation (ReLU):** Introduces non-linearity.
   - **Batch Normalization:** Stabilizes learning.
   - **MaxPooling1D:** Reduces the dimension corresponding to channels, making the model more robust to spatial variations.
3. **Permute Layer:** Rearranges dimensions back from ‘(batch$_{s}ize, filters, pooled_{c}hannels)$‘$to$‘$(batch_{s}ize, pool$ $lstm_{u}nits)$‘.

4. **Attention Mechanism:** Computes a weighted sum of the BiLSTM outputs, allowing the model to focus on the most relevant processed channel features. This generates a single context vector per sample. Output shape is ‘(batch$_{s}ize, 2 * lstm_{u}nits$)‘.
5. **Output Block:**
    - **Dropout Layer:** Applies dropout for regularization.
    - **Dense Layer:** A fully connected layer maps the context vector to the final output.
    - **Activation (Sigmoid):** Outputs a probability between 0 and 1 for binary classification.

This hybrid architecture allows the model to first learn spatially relevant features across channels (CNN), then model the sequence of these features (BiLSTM), and finally attend to the most informative parts (Attention) before making a classification.

3.4 Training and Evaluation

The model was trained using the Adam optimizer and a suitable loss function for binary classification (e.g., Binary Crossentropy or BCEWithLogitsLoss, depending on the framework's implementation). Training involved batches of data over a defined number of epochs. Hyperparameter tuning (e.g., using Optuna as indicated in your code) was performed to find optimal values for parameters such as the number of CNN filters, LSTM units and layers, learning rate, and dropout rate.

Evaluation was performed using standard metrics for binary classification, including Accuracy, Area Under the Receiver Operating Characteristic Curve (AUC), Precision, Recall, and F1-score. Techniques like k-fold stratified cross-validation can be employed to provide a more robust assessment of the model's generalization performance across different subsets of the data.

3.5 Feature Importance Analysis (Optional but Recommended)

Permutation importance or attention weight analysis can be used to understand which specific features or brain regions contribute most significantly to the model's classification decisions. Analyzing the attention weights from the attention layer can reveal which processed channel features the model focused on.

CHAPTER 4

RESULTS AND DISCUSSIONS

This chapter presents the experimental results obtained from training and evaluating the proposed CNN-BiLSTM-Attention model for EEG-based AD classification.

4.1 Experimental Setup

The model was trained using the dataset described in Chapter 3. [Describe your train/validation/test split strategy or cross-validation setup here. Mention if SMOTE or other techniques were used to handle class imbalance]. Hyperparameter optimization was conducted to determine the best configuration for the model.

4.2 Performance Evaluation

The performance of the CNN-BiLSTM-Attention model was evaluated using [list the metrics you used, e.g., Accuracy, AUC, Precision, Recall, F1-score]. The model achieved the following results:

- Accuracy: [Your CNN-BiLSTM-Attention Model Accuracy]
- AUC: [Your CNN-BiLSTM-Attention Model AUC]
- Precision: [Your CNN-BiLSTM-Attention Model Precision]
- Recall: [Your CNN-BiLSTM-Attention Model Recall]
- F1-score: [Your CNN-BiLSTM-Attention Model F1-score]

[Optionally, include a table summarizing performance metrics, similar to the one in the previous IEEE draft.]

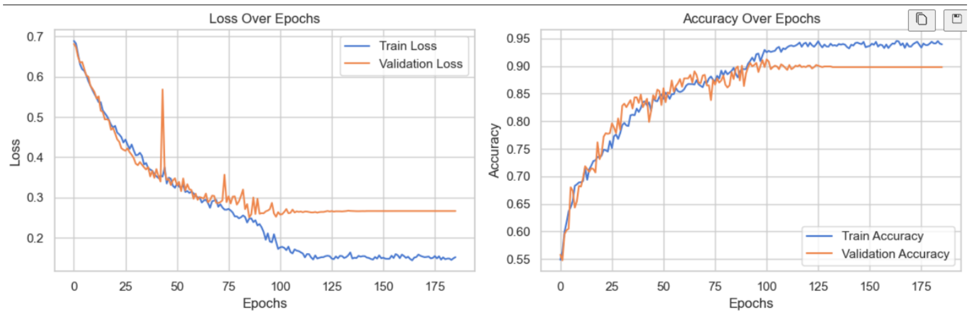

Figure 4.1: Training and Validation Performance Curves

[Include figures showing training/validation loss and accuracy curves over epochs.]

[Include figures showing the confusion matrix and ROC curve.]

4.3 Discussion of Results

The obtained results indicate that the CNN-BiLSTM-Attention model is capable of classifying AD from Cognitively Normal (CN) based on the extracted EEG features. [Discuss your results. How do they compare to the BiLSTM-only model results (if you ran both)? How do they compare to results from other studies in your literature review? Discuss the strengths and weaknesses observed.] The hybrid architecture likely benefits from the CNN's ability to capture spatial patterns across channels, the BiLSTM's strength in modeling the dependencies along the sequence of processed channels, and the **attention!** mechanism's ability to weigh the importance of different parts of the sequence.

4.4 Feature and Attention Insights

[Discuss findings from feature importance analysis if you performed one, e.g., permutation importance. If you analyzed the attention weights, discuss which channels or features the model typically focused on. Include relevant figures.]

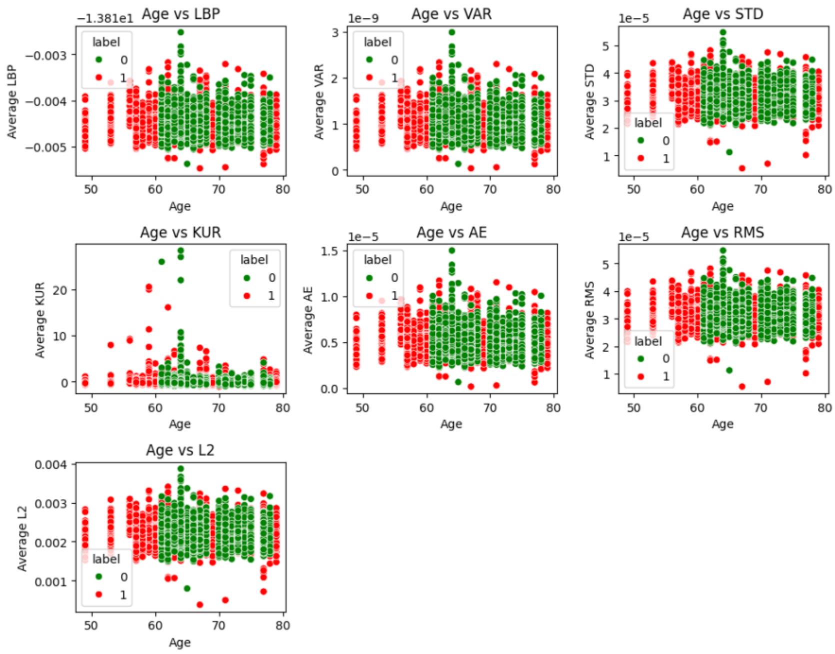

Figure 4.2: Feature Importance Analysis Results

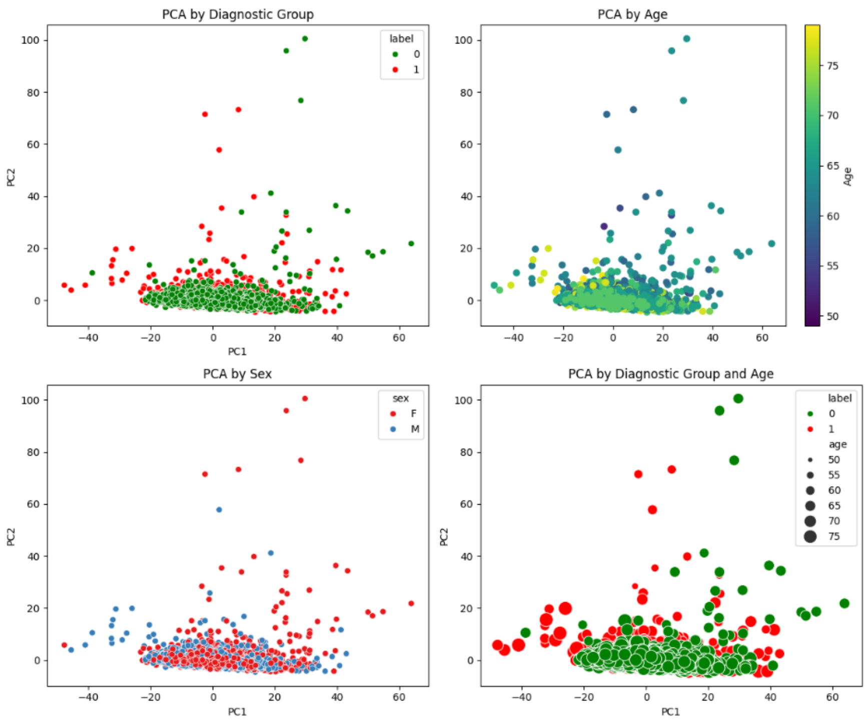

Figure 4.3: Exploratory Feature Analysis Results

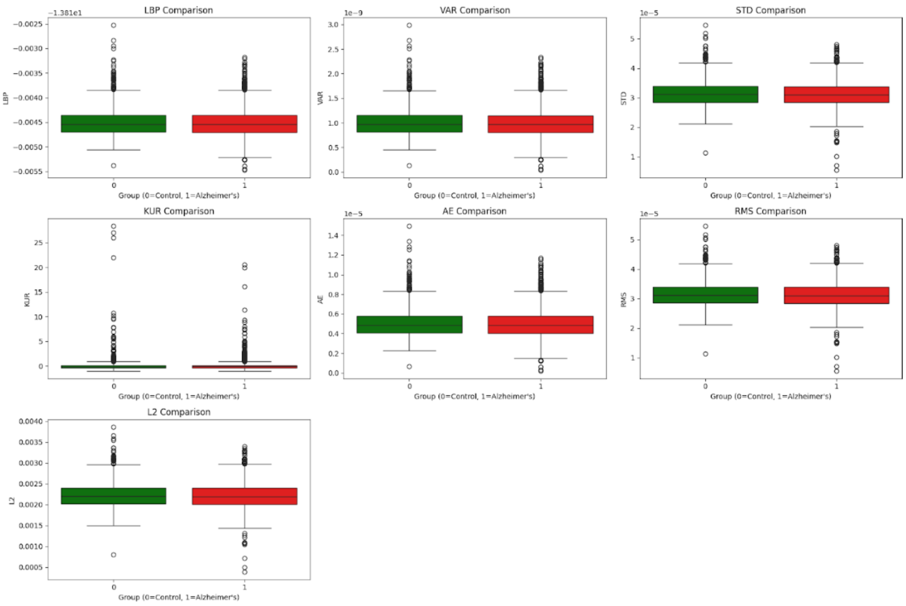

Figure 4.4: [Caption describing power ratio comparison figure]

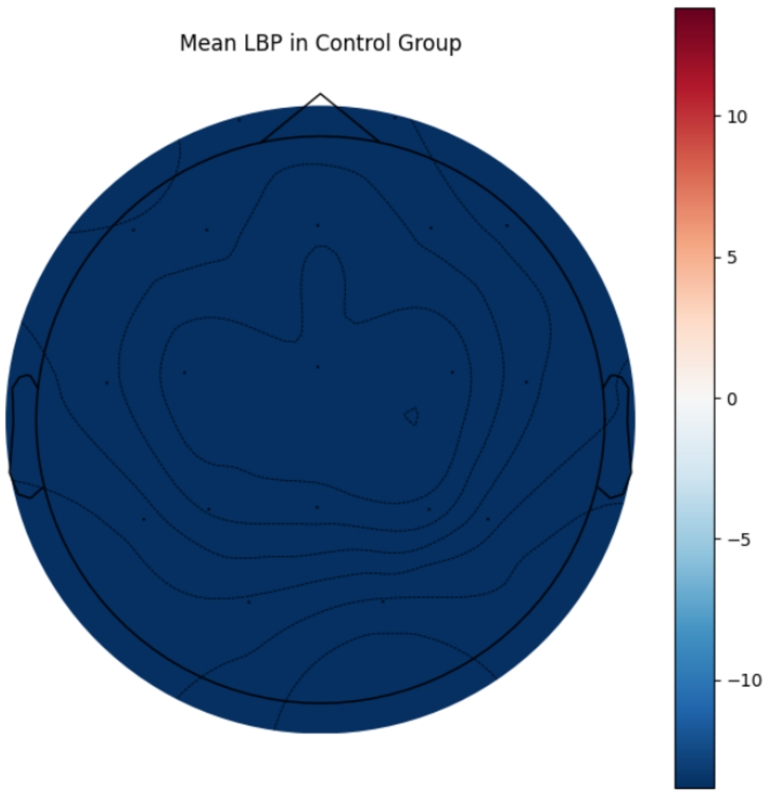

Figure 4.5: [Caption describing mean LPB figure]

4.5 Limitations

[Discuss limitations of your study, e.g., dataset size or characteristics (number of subjects, source), reliance on engineered features, lack of external validation, interpretability challenges, computational cost].

CHAPTER 5

CODING AND TESTING

In this chapter, the program coding related to your work using Python can be presented. You can describe the implementation details of your feature extraction pipeline, the CNN-BiLSTM-Attention model architecture, and the training and testing processes.

5.1 Development Environment and Libraries

- **Programming Language:** Python
- **Deep Learning Framework:** TensorFlow / Keras
- **Libraries:** NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, MNE-Python (for EEG processing), Optuna (for hyperparameter tuning - if used).
- **Development Environment:** Jupyter Notebooks, Google Colab, or other IDEs.

5.2 Implementation Details

1. **EEG Data Loading and Preprocessing:** Describe how you loaded the data (e.g., from .csv or other formats) and implemented the preprocessing steps (filtering, artifact removal, epoching) using libraries like MNE-Python or custom scripts. Explain how the 11 features were computed for each channel within each epoch.
2. **Data Preparation for Model Input:** Explain how the extracted features were structured into the ‘(batch$_{s}ize, 18, 11)‘tensorformat, includingtheinitialreshapeandtransposeoperationsreq$

3. **Training Script:** Describe the training loop. Mention the optimizer (Adam), the loss function (e.g., ‘BinaryCrossentropy(from$_{l}ogits = True)‘ifthelastdenselayerhasnoactivation, or‘BinaryCrossent$

5.3 Testing Procedures

- Data splitting strategy: Describe the method used (e.g., single train/validation/test split percentages, or k-fold stratified cross-validation) and how data leakage between splits was prevented.
- Metrics used for evaluation: List all metrics computed on the test set (Accuracy, AUC, Precision, Recall, F1-score, Confusion Matrix).
- Model checkpointing: If the best model based on validation performance was saved, mention this.
- Comparison with baseline models (if any).

CHAPTER 6

CONCLUSION

This study presented a deep learning framework combining CNN, BiLSTM, and **attention!** mechanisms for the classification of AD and CN subjects using a comprehensive set of 11 EEG-derived features from 18 channels. The proposed CNN-BiLSTM-Attention model achieved [Your CNN-BiLSTM-Attention Model Accuracy] accuracy and [Your CNN-BiLSTM-Attention Model AUC] AUC in distinguishing AD from CN.

The results demonstrate the potential of employing a hybrid deep learning architecture to capture both spatial and sequential patterns within engineered EEG features for automated AD detection. The inclusion of the **attention!** mechanism likely contributed by enabling the model to focus on the most discriminative feature representations.

[Briefly summarize key findings from feature/attention analysis if significant.]

While promising, the performance is subject to the characteristics of the dataset used. Further validation on independent and larger datasets is necessary to confirm the generalizability of the findings.

CHAPTER 7

FUTURE ENHANCEMENT

Future work can build upon this research to further enhance the performance and clinical applicability of EEG-based AD detection. Potential avenues include:

- Validating the CNN-BiLSTM-Attention model on larger, more diverse, and independent EEG datasets to assess its generalization capability.
- Exploring end-to-end deep learning approaches that process raw EEG signals directly, potentially reducing the reliance on hand-engineered features.
- Integrating multi-modal data sources, such as clinical scores (MMSE), demographic information (Age, Gender), or structural/functional MRI, alongside EEG features to improve diagnostic accuracy.
- Investigating advanced **attention!** mechanisms or transformer-based architectures better suited for capturing long-range dependencies in sequential data.
- Developing methods for model interpretability to understand which specific EEG features, channels, or time points are most influential in the model's decision-making process.
- Addressing class imbalance through more sophisticated data augmentation or loss weighting techniques tailored to EEG data.
- Conducting longitudinal studies to evaluate the model's ability to predict future cognitive decline or track disease progression.

APPENDIX A

EEG PREPROCESSING AND FEATURE COMPUTATION DETAILS

Provide detailed information about the steps taken to preprocess the raw EEG data and compute the 11 features per channel.

A.1 Preprocessing Pipeline

- **Data Loading:** Describe how the EEGLAB .set/.eeg files were loaded (e.g., using MNE-Python).
- **Filtering:** Specify the type of filters (e.g., Butterworth) and cutoff frequencies used for band-pass and notch filtering.
- **Artifact Handling:** Detail the methods used for detecting and removing or correcting artifacts (e.g., ICA, regression, rejection based on amplitude thresholds).
- **Referencing:** Explain the re-referencing strategy applied.
- **Epoching:** Describe how continuous data was segmented into epochs, specifying the epoch duration and any overlap.
- **Channel Selection:** List the specific 18 channels retained for analysis.

A.2 Feature Computation

For each epoch and each of the 18 channels, the following 11 features were computed:

- **SVD Entropy:** [Brief description or formula/reference]
- **DFA:** [Brief description or formula/reference]
- **ZCR:** [Brief description or formula/reference]
- **HFD:** [Brief description or formula/reference]

- **LBP (Alpha Band):** [Brief description or formula/reference for power spectral density calculation and log transformation]
- **STD:** [Formula as in Methodology chapter, repeated or referenced]
- **VAR:** [Formula as in Methodology chapter, repeated or referenced]
- **KUR:** [Brief description or formula/reference]
- **AE:** [Brief description or formula/reference]
- **RMS:** [Brief description or formula/reference]
- **L2 Norm:** [Brief description or formula/reference]

Describe any libraries or custom code used for these computations.

A.3 Data Formatting for Model

Explain the exact steps to transform the computed features into the ‘(batch$_{s}ize, 18, 11)‘tensorinputformat, in$

APPENDIX B

CNN-BILSTM-ATTENTION MODEL HYPERPARAMETERS

Provide a detailed list of the final hyperparameters used for the trained CNN-BiLSTM-Attention model, particularly those determined through optimization.

- **Input Shape:** (18, 11)
- **CNN Layer 1:**
  - Filters: [Specify Number]
  - Kernel Size: [Specify Size, e.g., 3]
  - Stride: [Specify Stride, e.g., 1]
  - Padding: [Specify Padding, e.g., 'same' or 1]
  - Activation: ReLU
- **Batch Normalization (after Conv1D):** [Specify if parameters were learned or fixed]
- **MaxPooling1D:**
  - Pool Size: [Specify Size, e.g., 2]
  - Stride: [Specify Stride, e.g., 2]
- **BiLSTM Layer(s):**
  - Number of Layers: [Specify Number, e.g., 1 or 2]
  - Units per Layer: [Specify Number, e.g., 64 or 128]
  - Recurrent Dropout: [Specify rate if used]
- **Attention Mechanism:** [Describe the type of attention or refer to its implementation]
- **Dropout (before Output Layer):** Rate: [Specify Rate, e.g., 0.2]
- **Output Dense Layer:**
  - Units: 1
  - Activation: Sigmoid
- **Optimizer:** Adam
- **Learning Rate:** [Specify Final Value, e.g., 0.001 or value from schedule]

- **Loss Function:** [Specify, e.g., BinaryCrossentropy(from$_logits = True)]$**Batch Size:**$[SpecifySize]$

- **Epochs (Max):** [Specify Number]
- **Early Stopping:** Patience: [Specify Value if used]
- **L2 Regularization (Weight Decay):** [Specify Value if used]

REFERENCES

[1] W. Xia, R. Zhang, X. Zhang, and M. Usman, "A novel method for diagnosing Alzheimer's disease using deep pyramid CNN based on EEG signals," *Heliyon*, vol. 9, no. 4, p. e14858, 2023. [Online]. Available: https://doi.org/10.1016/j.heliyon.2023.e14858

[2] S.-K. Kim, H. Kim, S. H. Kim, J. B. Kim, and L. Kim, "Electroencephalography-based classification of Alzheimer's disease spectrum during computer-based cognitive testing," *Scientific Reports*, vol. 14, p. 5252, 2024. [Online]. Available: https://doi.org/10.1038/s41598-024-55656-8

[3] A. Alzahrani and A. Alzahrani, "EEG-based clinical decision support system for Alzheimer's disorders diagnosis using EMD and deep learning techniques," *Frontiers in Human Neuroscience*, vol. 17, p. 1190203, 2023. [Online]. Available: https://doi.org/10.3389/fnhum.2023.1190203

[4] M. Saeed and A. Alzahrani, "Machine and deep learning trends in EEG-based detection and diagnosis of Alzheimer's disease: A systematic review," *Bioengineering*, vol. 5, no. 3, p. 78, 2023. [Online]. Available: https://doi.org/10.3390/bioengineering5030078

[5] Y. Zhang and Y. Wang, "EEG-based classification of Alzheimer's disease using deep learning techniques," *Sensors*, vol. 20, no. 9, p. 2505, 2020. [Online]. Available: https://doi.org/10.3390/s20092505

[6] X. Li and Y. Chen, "A comprehensive review on EEG-based Alzheimer's disease detection using machine learning," *Applied Sciences*, vol. 12, no. 11, p. 5413, 2022. [Online]. Available: https://doi.org/10.3390/app12115413