MID-TERM PROJECT REPORT

EEG- Based Speller System

Name- Pari Singla
Roll no.- 240730

PROJECT SUMMARY

The aim of this project is to design and develop an EEG-based speller system that can interpret human brain signals to identify intended characters. By using a brain–computer interface (BCI), the system seeks to establish a direct communication pathway between the human brain and an external computing device.

The proposed workflow includes the acquisition of EEG signals, followed by preprocessing steps to eliminate noise and unwanted artifacts. Relevant features are then extracted from the processed signals and analyzed using machine learning and signal processing techniques to determine the user's intended character. The identified character is displayed on a graphical interface, allowing the formation of words and sentences.

Overall, the project focuses on implementing a complete EEG-based speller system and assessing its performance in accurately detecting user-intended characters from EEG responses. Emphasis is placed on developing a reliable processing pipeline and evaluating the effectiveness of various signal processing and classification methods for accurate character selection.

PROGRESS TILL MID-TERM

Assignment 0: Python Fundamentals

In Assignment 0, we revisited essential Python programming concepts, including array operations using NumPy, data manipulation and cleaning with Pandas, and data visualization using Matplotlib. These activities helped build practical skills in working with data structures, performing numerical computations, preprocessing datasets, and effectively presenting data through graphical representations.

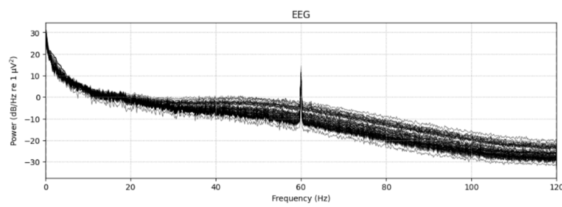
Assignment 1: Machine Learning Concepts

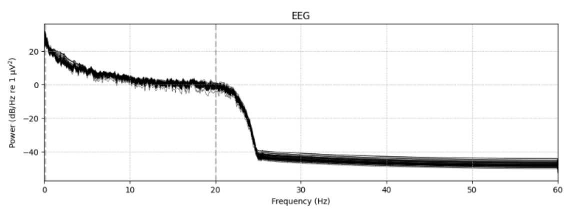
Assignment 1 focused on strengthening our theoretical understanding of core machine learning principles. The assignment explored the functioning and reasoning behind various learning algorithms, covering topics such as ensemble techniques (bagging, boosting, and random forests), loss functions, and regularization methods applied in models like SVM, LASSO, and Ridge regression. It also emphasized concepts such as the bias–variance trade-off, overfitting and underfitting, and the importance of splitting criteria and tree depth in decision tree learning. Overall, the assignment helped clarify the assumptions, constraints, and performance factors that influence machine learning models.

Assignment 2: EEG Signal Processing

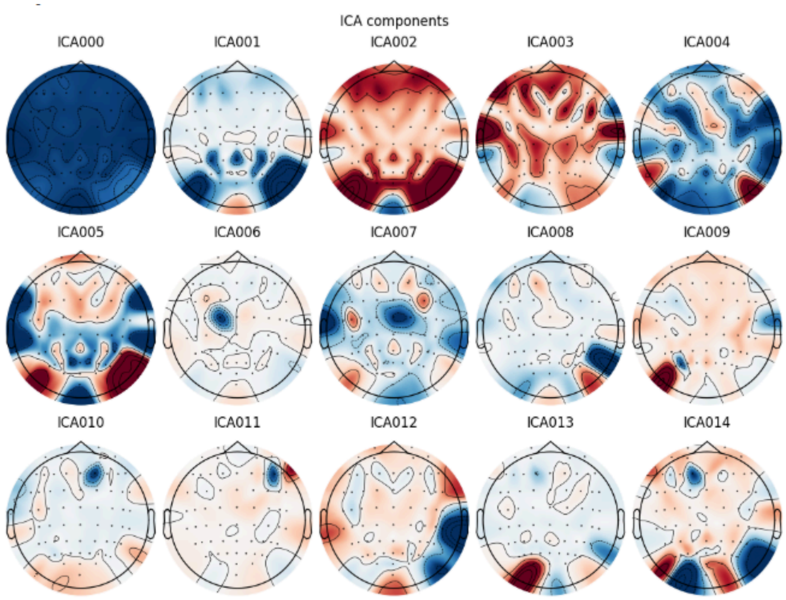
In Assignment 2, we carried out signal processing on EEG data using a real-world P300 speller dataset from the BCI Competition III, with the goal of understanding the complete preprocessing and analysis workflow. The major steps involved in the assignment are summarized below:

- Raw EEG data from the P300 speller dataset were loaded and reshaped, and event markers were generated using stimulus and flashing information to ensure accurate event alignment.

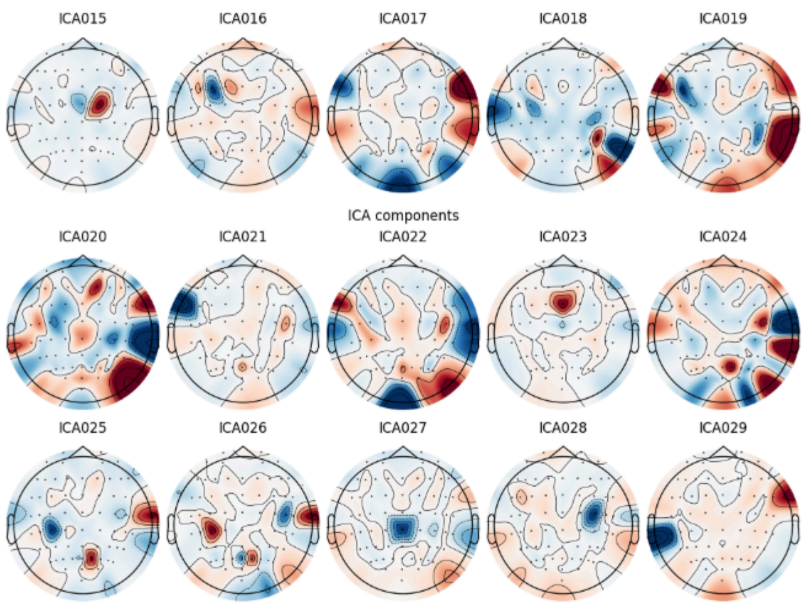




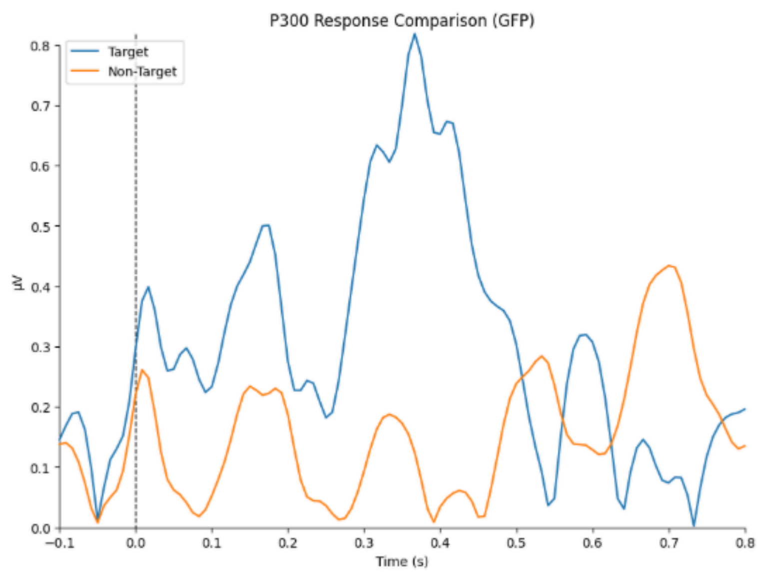
- An MNE-based preprocessing pipeline was developed, which included band-pass filtering, resampling, channel mapping to the standard 10–20 montage, and artifact removal using Independent Component Analysis (ICA). The spatial patterns of the extracted independent components are illustrated in Figures below



.



- Target and non-target epochs were extracted, baseline correction was applied, and averaged evoked responses were visualized at key electrode sites to examine the P300 component.



Overall, this assignment provided extensive hands-on experience with EEG data processing, including preprocessing, artifact removal, and event-related potential analysis using real BCI datasets. Exploring different methods and parameter settings helped reinforce theoretical concepts through practical implementation and demonstrated how these choices influence the final results.

Assignment 3: EEG-Based Classification Workflow

Assignment 3 focused on designing and implementing a comprehensive EEG analysis workflow that combines multiple stages of signal processing and machine learning. The main goal was to understand how each stage of the pipeline contributes to accurately separating target and non-target neural responses using real EEG recordings.

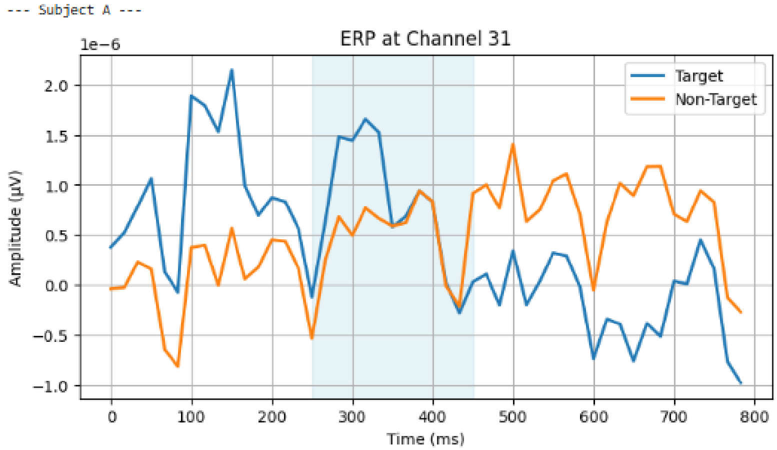
- EEG data were first loaded and preprocessed through steps such as band-pass filtering, resampling, and the removal of physiological and environmental artifacts.

- Signal segments were then extracted based on stimulus timing, followed by baseline normalization. Average evoked responses for both target and non-target conditions were computed to study event-related potential (ERP) characteristics.

- Meaningful features were derived from the EEG signals and used with traditional machine learning classifiers to assess their effectiveness in distinguishing between target and non-target brain activity.

Overall, this assignment combined EEG preprocessing, ERP evaluation, and classification into a single, coherent framework. By experimenting with different processing strategies, parameter settings, and classification methods, we gained valuable insight into the impact of each design choice on signal quality and model performance. This iterative approach strengthened our practical understanding of EEG-based analysis and provided a solid foundation for future work in feature enhancement, model optimization, and real-time EEG-based speller system development.

--- Subject A ---



```
=== LDA Evaluation ===
Accuracy: 0.5237937871777925
Confusion Matrix:
 [[1346 1175]
 [ 266  239]]
Classification Report:
               precision    recall  f1-score   support

         0.0       0.83      0.53      0.65      2521
         1.0       0.17      0.47      0.25       505

    accuracy                           0.52      3026
   macro avg       0.50      0.50      0.45      3026
weighted avg       0.72      0.52      0.58      3026

ROC AUC: 0.5241531531177711

=== Logistic Regression Evaluation ===
Accuracy: 0.5270984798413747
Confusion Matrix:
 [[1352 1169]
 [ 262  243]]
Classification Report:
               precision    recall  f1-score   support

         0.0       0.84      0.54      0.65      2521
         1.0       0.17      0.48      0.25       505

    accuracy                           0.53      3026
   macro avg       0.50      0.51      0.45      3026
weighted avg       0.73      0.53      0.59      3026

ROC AUC: 0.5266124946489096
```

```
Training Random Forest (n_estimators=100)...

=== Classical Model Comparison ===
SVM                  | Accuracy=0.5357 | F1=0.2343
Random Forest        | Accuracy=0.8334 | F1=0.0079
Gradient Boosting    | Accuracy=0.5681 | F1=0.2370
```

WORK DONE DURING THE PROJECT

- **Assignment 1: [Winter-projects-25-26/EEG-Based P300 Speller/assignments/assignment_1/240730_Pari_assignment1.txt at main · Pari-Singla/Winter-projects-25-26]**

- **Assignment 2:[Winter-projects-25-26/EEG-Based P300 Speller/assignments/assignment_2/Pari_240730_Assignment2.ipynb at main · Pari-Singla/Winter-projects-25-26]**

- **Assignment 3:[Winter-projects-25-26/EEG-Based P300 Speller/assignments/assignment_3/Pari_240730_Assignment3.ipynb at main · Pari-Singla/Winter-projects-25-26]**