Peer evaluation of Team 9's Project

Fabrice Gagnon (111 146 970)
Mathis Rezzouk (537 172 964)
Marc-Antoine Tremblay (536 801 905)

1. Team 9's Project Description

The project aims to detect Alzheimer's disease and frontotemporal dementia using EEG data, with a long-term goal of identifying disease stages. Various models, including EEG-NetV4, DTW, GMM, RNN (LSTM), and CNN, are being tested to compare their results and determine the most appropriate one for classification.

2. Strong aspects of the project

2.1. Well-Defined Scope with Clear and Valuable Objectives

The project has a clearly defined scope and a focused objective: detecting Alzheimer's and frontotemporal dementia, along with identifying disease stages. This goal addresses a critical need in the medical field and has the potential to provide significant value for early diagnosis and disease monitoring using EEG analysis.

2.2. Structured Vision and Anticipation of Future Steps

The project outlines a clear roadmap for future improvements, including hyperparameter tuning, advanced data preprocessing, and addressing class imbalance, showcasing a strong understanding of the challenges ahead.

2.3. Diversity of Models Used in EEG Classification

To achieve the objectives of EEG signal classification, their team explored and implemented a wide range of model families. This diversity of approaches allowed for a rigorous evaluation of model performance and generalization capabilities.

- **Deep Learning Models:** Modern neural network architectures, including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks, were employed to automatically learn representations from raw EEG signals. CNNs were particularly effective at capturing spatial patterns, while RNNs and LSTMs modeled the temporal dependencies of the signals.
- **Hybrid Architectures:** To leverage both spatial and temporal information, hybrid models combining CNNs and RNNs were implemented. These architectures showed promising results in improving classification performance by extracting multi-dimensional features.

3. Weak aspects of the project

3.1. Absence of state-of-the-art machine learning models

The absence of simpler state-of-the-art machine learning models limits the opportunity for baseline comparisons. Incorporating models such as Support Vector Machines, Random Forests, or Gradient Boosted Trees could provide valuable benchmarks, offering quicker insights and allowing the team to focus on refining deep learning models where they demonstrate clear advantages. Additionally, the team might benefit from reducing the number of deep learning models being implemented. Limiting the scope to a few well-chosen architectures would streamline the project and allow for more focused optimization and evaluation efforts.

3.2. **Inconsistent datasets**

The datasets are inconsistent, not only due to mixing preprocessed and unprocessed data but also because the subjects did not undergo the same tests or for the same durations. This heterogeneity introduces variability that can bias the models and complicate reliable comparisons. A potential improvement would be to focus on a single dataset, applying consistent preprocessing, and exploring data augmentation techniques to increase the dataset's size and diversity, thereby enhancing model robustness and comparability.

4. **Questions**

(1) What methods are you considering to address overfitting beyond hyperparameter adjustments?
(2) Why were the specific models (EEGNetV4, DTW, GMM, RNN, CNN) chosen for this project, and have alternative architectures like Transformers been considered?
(3) Have you considered using techniques like data augmentation or synthetic data generation to address the class imbalance?