NeuroPAC: A Pre-Processing, Analysis, and Classification Pipeline Platform for EEG Signals

Selin Uygun

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Software Engineering

Department of Software Engineering

Mugla Sitki Kocman University

Abstract

(Abstract text goes here...)

Contents

List of Tables

List of Figures

Chapter I

Introduction

I.1 Problem Statement

The core difficulty in Electroencephalography (EEG) data analysis lies in the trade-off between power/flexibility and usability/accessibility. Here are some limitations of two existing and most popular state of the art tools that are used for EEG pre-processing and analysis:

- **Fieldtrip, Complexity of Code-Based Toolboxes:** A high-performance, flexible toolbox like FieldTrip is essential for advanced analysis but rely on MATLAB code and scripting. This necessitates a high level of programming expertise, creating a significant barrier to entry and slowing down the research process for those whose primary focus is neuroscience or clinical application rather than coding and data structuring.
- **EEGLAB, Existing GUI Complexity:** While EEGLAB offers a Graphical User Interface (GUI), this interface can present a dense, overwhelming array of parameters, menus, and advanced options to a novice. This complexity often fails to make the fundamental, necessary workflow—preprocessing, processing, and classification—feel less intimidating or more welcoming to beginners.

There is a clear and persistent need for a dedicated, simplified interface that explicitly targets the essential and initial steps of EEG analysis, thereby democratizing the process.

- **Need for Simplified Workflow (Engineering Goal):** The study addresses the need for a streamlined, intuitive Graphical User Interface (GUI) designed to guide users through the crucial, initial steps of preprocessing, processing, and machine learning classification in a less intimidating manner. By building on the powerful computational back-end of toolboxes like FieldTrip, this GUI will abstract away the code, presenting only the core configuration options needed for a reproducible, foundational analysis.

- **Need for Clinical Validation and Application (Research Goal):** The efficacy of this simplified pipeline must be demonstrated in a real-world, high-stakes context. There is a need to validate this new streamlined analysis tool by successfully applying it to challenging neurological datasets, specifically in generating robust and clinically meaningful machine learning and deep learning classification results for conditions such as Parkinson's Disease and mild Traumatic Brain Injury (mTBI).

In summary, the basic difficulty is the lack of a user-friendly tool that simplifies the essential EEG analysis pipeline (preprocessing, feature extraction analysis, and classification). Existing tools are either highly code-dependent (FieldTrip) or still too complex for beginners (EEGLAB). The perceived need is thus the development of a simplified, welcoming user interface that leverages existing powerful toolboxes to make foundational EEG analysis accessible, efficient, and reproducible, with its utility confirmed through successful application and validation in classifying clinical cohorts like Parkinson's Disease and mTBI patients.

I.2 Purpose of the Study

The purpose of this study is two-fold: to address the usability gap in existing neuroscientific software and to validate the resulting platform through compelling clinical application.

- **Engineering Goal (Product):** To develop a user-centric Graphical User Interface (GUI) using Python that provides a streamlined, accessible, and intuitive workflow for the fundamental steps of EEG data analysis (preprocessing, processing, and classification). This GUI will abstract the complexities of code-based toolboxes like FieldTrip, thereby lowering the technical barrier to entry and welcoming new users to the field of neuroscience and EEG analysis.

- **Research Goal (Practical Outcomes):** To validate the utility and efficacy of the developed GUI and its integrated machine learning classification pipeline by generating meaningful, publishable findings. This includes providing example EEG data outcomes, specifically robust analysis results and accurate classification metrics for clinical datasets concerning Parkinson's Disease and mild Traumatic Brain Injury (mTBI).

The ultimate goal is to deliver a practical product that enhances the efficiency of the EEG analysis pipeline and produces high-quality, reproducible research outcomes.

I.3 Research Questions

- **Usability and Accessibility:** Can a dedicated Python-based Graphical User Interface (GUI) be successfully developed to simplify the complex, code-based commands of established EEG toolboxes (e.g., FieldTrip), thereby providing an intuitive and streamlined workflow for the fundamental steps of preprocessing and processing?

- **Comparative Advantage:** How does the proposed streamlined GUI workflow compare to existing EEG analysis interfaces (e.g., EEGLAB) in terms of reducing complexity and improving user efficiency for individuals new to neuroscience and EEG data analysis?

- **Classification Efficacy (Parkinson's Disease):** What level of classification accuracy can the integrated machine learning classification pipeline achieve in reliably distinguishing EEG data from individuals with Parkinson's Disease compared to healthy controls?

- **Clinical Differentiation (mTBI):** Can the complete EEG analysis and classification platform successfully extract meaningful neurophysiological features and generate differentiating results to classify patients with mild Traumatic Brain Injury (mTBI) from control subjects?

- **Data Outcome Utility:** Does the developed platform effectively output clear, reproducible analysis results and visualization products (e.g., graphs, classification metrics) that aid neuroscientists and clinicians in interpreting the EEG data outcomes from clinical populations?

I.4 Definition of Terms

(Content...)

I.5 Assumptions and Limitations of the Study

- **Essential Software Environment:** The core functionality of the GUI relies on proprietary and third-party software being present. It is fundamentally assumed that the end-user has a valid license for MATLAB and that both MATLAB and the FieldTrip toolbox are correctly installed, configured, and accessible within the operating system environment. Without these prerequisites, the computational functions handled by the GUI cannot execute.

- **FieldTrip and MATLAB Reliability:** It is assumed that the underlying algorithms and functions within the FieldTrip toolbox are accurate, well-tested, and functioning correctly as documented by their developers. The study focuses on the interface and workflow, not on re-validating the internal processing logic of these established tools.

- **Data Quality:** It is assumed that the clinical EEG datasets used for validation (Parkinson's Disease and mTBI) were collected using standardized, ethical procedures and possess sufficient signal quality to yield meaningful, artifact-reduced data after the initial preprocessing steps.

- **Machine Learning Framework Stability:** It is assumed that the Python-based Machine Learning libraries (e.g., scikit-learn, etc.) that power the classification module will function stably and reliably throughout the study.

- **User Competency:** For any usability evaluation, it is assumed that users participating possess a basic familiarity with desktop computing and will apply a consistent effort when using the developed GUI compared to other analysis tools.

I.6 Overview

(Content...)

Chapter II

Literature Review

II.1 Introduction

(Content...)

II.2 Next Heading

(Content...)

II.3 Next Heading

(Content...)

II.4 Summary

(Content...)

Chapter III

Methodology

III.1 Introduction

(Content...)

III.2 Research Question(s)

(Content...)

III.3 Research Design and Procedures

(Content...)

III.4 Population and Sample

(Content...)

III.5 Instrumentation

(Content...)

III.6 Data Analysis Procedures

(Content...)

III.7 Limitations

(Content...)

III.8 Summary

(Content...)

Chapter IV

Results

IV.1 Introduction

(Content...)

IV.2 Data Analysis

(Content...)

IV.3 Summary

(Content...)

Chapter V

Summary, Conclusions, and Recommendations

V.1 Introduction

(Content...)

V.2 Summary of the Results

(Content...)

V.3 Conclusions

(Content...)

V.4 Recommendations

(Content...)

Appendix A

Assessment Activity

(Content...)

Appendix B

Assessment Rubric

(Content...)