Team 5: Final Proposal

[*Project Title*]

By: Aryan Pandey, Rachel Fox, Anoushka Bhat, Shaina Wang, Danielle Jaworski, Alexis Malveaux, Rushi Ravindran

I. Motivation:

Drowsiness is a naturally occurring human phenomenon, but with our ever-busy culture drowsiness can impede on daily tasks. One of the most dangerous ways that this manifests is when driving, as we see thousands of vehicle crashes annually due to drowsy driving. With our BCI we aim to address this by developing technology and methodology to read in brain data and provide real-time feedback to the user about their drowsiness levels. The goal of developing this project is to allow technology to assist in human productivity and reduce human tragedy in helping to monitor and keep people aware of their own drowsiness levels, allowing them to perform better in tasks otherwise inhibited by drowsiness.

II. Abstract:

Drowsiness is characterized by physiological signs such as increased eye blinking, decreased reflexes, and slow eyelid closure (Soares et. al, 2019), but electroencephalogram (EEG) data can be useful in objectively determining states of drowsiness as well. By analyzing individuals' drowsiness levels based on existing EEG data sets as well as our own experiments, we will develop a Brain Computer Interface (BCI) that implements a machine learning algorithm we will create that classifies drowsiness, then expand to varying degrees of drowsiness. We also will develop an innovation that provides real time feedback based on different scenarios to reduce drowsiness levels.

Key words:

Drowsiness, Machine learning, EEG, Theta Waves (4-8 Hz), Graded feedback

III. Theory:

Previous studies (Brown et. al, 2013) have shown that drowsiness does vary in different times of the day. The power spectrums of theta (4-7Hz) and alpha (8-13Hz) frequency waves have been observed to be an indicator of state of drowsiness, as the increase in the relative power of these frequencies indicated an increased likelihood of the subject experiencing drowsiness. Current methods to determine drowsiness include a psychological approach to evaluate drowsiness level, a video-based approach to monitor the subject, and a physiological approach that uses biomarkers to determine state of drowsiness (Ren et. al, 2021). EEG is an optimal approach because it is a non-invasive method to detect brain state and is applicable to machine learning networks that can classify different levels of drowsiness using only EEG signals (Jeong et. al, 2019). Classifiers such as feature selection algorithms and Support Vector Machine (SVM) have been implemented with high levels of accuracy in detecting drowsiness (Gangadharan et. al, 2022), utilizing a multitude of selected EEG features. Studies have analyzed drowsiness using features in the main areas of time-domain, frequency-domain, nonlinear features, entropies, spatiotemporal, and complex features (Stancin et. al, 2021), with the combination of feature categories being most effective in machine learning models. In our BCI, we plan to begin with a time-frequency analysis of the power bands of specific signals to create our model. To implement this classification, the data is collected and cleaned to remove extraneous noise and artifacts, and then epochs of selected features from the processed data are extracted and further analyzed to be used in the training and testing algorithm of the classifier (Gangadharan et. al, 2022).

IV. Methods/System Architecture:

Materials:

The materials we will be using for this project are Jupyter Notebooks, Python code, environments and libraries, and the Muse Headset. Jupyter Notebooks will be used in the form of Google Colab for our preliminary development of our drowsiness detection algorithm because it allows us to share code between team members as well as quickly test and edit our code. We will use the Python programming language for all of our code because it has the most amount of features that are useful in EEG analysis and machine learning. Specifically, we will be using the Python libraries numpy for creating and modifying arrays, matplotlib for visualizing output, sklearn for regression and classification analysis, and MNE for reading and interpreting the EEG data. The Muse Headset will be used in the demo portion of our project to test our algorithm on live EEG data that will be coming from the readings on the Muse Headset.

Dataset:

The dataset we will be using is the Sleep EDF Dataset Expanded (Kemp et. al, 2000) located on the open-source site Physionet. This dataset contains full-night PolySomnoGraphic (PSG) sleep recordings that contain EEG data as well as hypnogram files that are scored for sleep stages W, R, 1, 2, 3, 4, and M. For our project, we are going to be focusing on detecting the difference between stage W (wake), stage R (REM) and stage 1 sleep in order to determine when the subject is awake and when their EEG waveforms indicate drowsiness (ie: not in the wake stage). We felt that using this sleep stage dataset was beneficial because it provides annotations that we can use to train our model, a sufficient number of subjects to properly train the model, and sleeping and waking EEG epochs so the model can be more accurate.

Model Development:

To develop our model, we are going to begin with using a multiclass classification task where we use the MNE package to train the model using a certain amount of subjects from the dataset. We would be adapting the code from the MNE documentation (Chambon et. al, 2018) to use as a classifier for epochs. In addition to the basic classifier, we also plan to use feature extraction of EEG signals to make the model more accurate on EEG data. Brain waves at 4-8 Hz (Theta Waves) correlate with being in a drowsy or daydreaming state, while 13-38 Hz (Beta Waves) correlate with being in an alert or fully concentrated state. So, we plan to extract the relative power bands of each of those frequencies at each epoch in order to analyze whether the subject is in a drowsy state or not. We would train our feature extraction model with the provided sleep stage annotations in order to create a model that accurately predicts sleep stage based on the relative power of different EEG signals.

Model Testing:

Once we have finished creating the preliminary code for the model, we are going to test the model on unseen EEG data, first with data from other subjects in the EDF data set. We will then go back and make improvements to our model as necessary. If we want the accuracy score of our model to be higher, we may test other classification models such as the SVM detection algorithm to see if another form of classification will give better results.

Future Plans:

After we are satisfied with the accuracy of our model, we will learn how to integrate live EEG signals into our coding environment so we can test our model on live data. If there is time after we have finished these components, we plan to develop an application to go along with the live EEG collection that would notify the user if they are drowsy, meaning that their EEG signals indicate a higher amount of non-wake stage epochs. Upon further experimentation, we hope to also be able to classify varying levels of drowsiness. This could have great potential applications for driver fatigue detection, applications in hospitals for surgeons and doctors, and even be useful for chronic-fatigue syndrome patients.

I. **Timeline:**
   A. Week 7: Initial Proposal
   B. Week 8: Final Proposal
   C. Week 9: Final Proposal
      1. Monday/Tuesday: Final Proposal
      2. Wednesday/Thursday: Review and Planning along with committee formation
      3. Friday/Saturday/Sunday: Planning from each committee and beginning project
   D. Week 10: Check-in (Objective 1 accomplished)
   E. Spring Q Week 10: Project submission/demo