**Jaw Clenching Detection**

Alroomi, Noura

Behdarvadian, Prarin

Dode, Pragati  
Rehman, Khaleel

### **Overview**

Find out the particular data pattern which represents jaw clenching by analyzing data collected from EEG tests.

### **Propose**

In this project, we are planning to tackle the problem in three phases: Data labeling (jaw clenching and rest classes), data wrangling, and data modeling. In the first phase, Data labeling, we plan on using single processing algorithms to analyze changes in the behavior of the signal to identify the parts that represent jaw clenching. In the second phase, the data wrangling, we plan on filtering out noise and extracting signal features from the data, such as wavelength, frequency, etc. Additionally, we will apply Principal Component Analysis (PCA) coupled with Singular value decomposition (SVD) to reduce data dimensionality. Finally, in the last phase, we will be experimenting with classifiers to build the model to classify jaw clenching in the test EEG dataset.

### **Technical Approach**

The main challenge of this project is to divide the EEG data into two categories of clenched jaw or rest. We have been able to break down the problem into smaller steps to make it more approachable. The three main steps will be data labeling (jaw clenching and rest classes), data wrangling, and data modeling.

For the data labeling, after we collected the data, the software produces EDF files (European Data Format) that is a standard format used to store multi-channel signals in scientific medical fields. To work with data, we’ll convert the EDF into CSV files (comma separated values) to represent the data in a tabular format by taking data from different channels into separate columns and joining their values into rows based on the elapsed time. Afterwards, filters will be applied to remove noise and detect the abrupt changes in the EEG signals’ segmented amplitudes and labeling those as the positive data behavior (clenched jaw).

For Data wrangling, we’ll be looking into the data distributions along with filtering out any noise and taking the signal features like the amplitude and frequency of the data collected. Principal Component Analysis (PCA) and Single Value decomposition (SVD) will be used to map the data into new dimensions (Principal Components) that captures most data variances and reduces the noisy features by looking at dependencies and correlations of features.

For data modeling we are using Bayes Discriminant and Decision tree to perform classification on our data classes: Clenched Jaw and Rest Phase.

### **Parameters and variables**

The parameters for this project are items required by the team to gather data effectively and efficiently. The EEG machine is the most important parameter as that is how the data is collected, this in turn will also require a laptop that contains the PowerPoint slides for the prompt and the DSI streamer software. A controlled environment consisting of a quiet room for the subject is required to collect data where there are almost no distractions present while the data is being collected. Finally, a volunteer test subject is required and will wear the EEG machine and react to the prompt. For the variables, these will be the observations recorded by the EEG machine.

### **Assumptions**

We are assuming data we are collecting from the EEG machine is accurate and captures the data in the form of signals. Also, we are assuming that there is no external noise as we are asking the subject to keep still and follow instructions on the prompt. And finally, we are assuming the high amplitude in this wave represents jaw clenching only as the subject is following all instructions correctly.

### **Noise Sources**

Though we are recording data in a controlled environment, there can be some noise. The accuracy of EEG sensors depends on how closely they are attached to the scalp. We might get lower data feeding quality in case of a sensor not being placed properly. Also, if the subject is doing other movements during the test such as rapid blinking or distracted by surrounding noise, which can introduce noise to the data.

### **Area of application**

The area of application for this project is analyzing the EEG records and being able to classify them when the data is showing the patient is clenching their jaw. What makes this project interesting is that because the brain is a nonlinear system, the recorded data from the EEG machine will not be as simple as only having the two outcomes of jaw clenching vs non-clenching. The data could also be in rare cases recording the heartbeat of the patient as an example, which creates a challenge for how to read, analyze, and

What makes this project useful is that while there are applications of this topic made in the medical field, there is not much being done to add in any computational machine learning or data science. Potentially this gives the opportunity to create something new that has not been tried before and presents a challenge given the limited research and resources available.

### **Goals / Success**

The main goal of this project is to be able to classify the EEG data into two categories of clenched jaw or rest. One of the goals is to add timestamps to the data, by being able to identify the time intervals in EEG where jaw clenching is happening it will allow the labeling of the data of when the subject is clenching their jaw or at rest. Another goal will be to apply Machine Learning algorithms to model the labeled data to correctly classify jaw clenching in EEG data in our test dataset.

### **EEG**

An electroencephalogram is a record that measures electrical activity in the brain using small, metal discs (electrodes) attached to the scalp. Brain cells communicate via electrical impulses and are active all the time, even during sleep. This activity shows up as wavy lines on an EEG recording.

EEG records many types of brain waves, with Alpha, Beta, Theta, and Delta are the basic brain waves.

1. Alpha waves - Relaxation and Attention

They are present when a person is awake with eyes closed. Disappears when the person opens eyes and pays attention.

1. Beta wave - Awake

It doesn't matter if a person's eyes are open or closed. Drugs like Sedatives can influence these waves.

1. Theta waves - Sleep

These are the slow waves which are normal for all ages during sleep.

1. Delta waves- Also related to sleep

These waves are normal in adults who are in deep sleep and in young children.

There are two ways EEG can be abnormal

1. Brain activity suddenly interrupted and changed.
2. Abnormal result - Abnormal frequency, height, or shape of wave.

EEG performed for

The measurement given by EEG are used to confirm or rule out various conditions including

1. Seizure disorder
2. Head injury
3. Encephalitis (inflammation of brain)
4. Brain tumor
5. Encephalopathy (disease that causes brain dysfunction)
6. Sleep disorder
7. Stroke
8. Dementia

This test is also used to monitor brain activity of patients in coma or if a patient is undergoing surgery.

Factors interfere EEG result

Several types of movements can potentially cause “artifacts” on EEG that mimic brain waves.

1. Pulse and Heartbeat
2. Breathing
3. Sweating
4. Mouth movement
5. Muscle movements

Other factors that can influence EEG reading

1. Low blood sugar
2. Bright or flashing light
3. Medication like sedatives
4. Caffeine consumption
5. Oily hair or hair spray

### **Data collection**

The data collected for this project is a set of EEG tests recorded by Dr. Erika Parsons. The procedure to gather data from the subject entails that the subject will wear the EEG machine and they will be put in a dimly lit room where there would be no distractions and given a PowerPoint presentation that runs for about 1 minute and 30 seconds where the subject will be given a prompt to clench their jaw for a certain amount of time, and this test is run three times per subject. All three tests have a set timer and a prompt that will tell the subject when to clench their jaw and when to stop. Currently the data being gathered is from students from the 590 class. The data will be generated through a software called DSI streamer which will be connected to the EGG machine and taken in readings overtime.

The reason the current focus of the project is jaw clenching is because compared to other facial gestures, Jaw clenching has a significantly greater amplitude when using the EEG machine.

The EEG datasets consist of the brain electrical activity reading of seven electrodes channels signals over time:

- Left Ear channel reading: LE channel. The Left Ear channel reads signals from the cortex which corresponds to higher processes of brain activity such as memory, thinking, learning, reasoning, and functions related sensory perception responses.

- Readings from the Frontal lobe region of the brain: F4 and F3 channel. The Frontal lobe reading corresponds to control of voluntary movement such as muscle movement, which includes jaw clenching.

- Readings from the Parietal lobe of the brain: P4 and P3 channels The Parietal lobe reading corresponds to sensory perception, which includes sounds and visual perceptions.

- Readings from the central region of the brain: C4 and C3 channels, the central region is Central Sulcus, which is the area separating the Parietal lobe from the Frontal lobe, thus its reading corresponds to both control of voluntary movement and sensory perception.

It's also worth mentioning the channels with suffix 3 correspond to readings from the left side of the brain, which controls speech and abstract thinking, while channels with suffix 4 readings are from the right side of the brain, which controls sensory perception activity, spatial thinking, and movement.

### **Preprocessing**

The data sheet that is retrieved from the EEG machine has 7 columns corresponding to each channel where the data was in a raw format. From there, we applied a High Pass filter on each channel’s signal to remove lower frequencies, thus denoising the data. After that, we applied Gabor Transform (developed by Luca Raad) on the filtered channels signals. Gabor Transform applies the Fast Fourier transform (FFT) on a sliding window. FFT transforms the signal from time domain into frequency domain by decomposing the signal into its frequency components. Then we calculated the area under the curve of the Gabor Transform windows and took the mean of the areas for each channel. Finally, we compared the calculated area Gabor Transform windows against the mean if the area is above the mean, then label it a jaw clench, this was done on each channel then we took the majority voting of the channel as the final label.

From there, a script is created to take the data from the labeled EEG records and is converted to these four features per channel: Mean, Peak Value, Standard Deviation, and Signal to Noise Ratio (SNR).

For the four features listed above, there will be one for each channel. For example, there will be a Mean feature for the LE channel and a Mean feature for the F4 channel. This will expand the dataset to have 28 columns to accommodate the features. The features will be separated by channel because each channel is reading in from a different area of the brain and outputting what is assumed to be unique values.

The features were extracted using the signalTimeFeatureExtractor function in MATLAB. The full dataset is given to the function where it will produce an array of 2D arrays per feature requested per channel. Like the labeling, the dataset would be viewed in the frame size of 300 in order to condense the data to show an observation per second and the values outputted would be calculated based on those 300 framed windows. The mean and the standard deviation will be calculated based on the frame size, while Peak Value will generate the absolute maximum value of the frame.

### **Observations**

In this section, we will analyze our dataset trends and clusters through plots, histograms, and scatter plots.

**Independence**

For the independence of the data, the assumption being made is that if the covariance is within the range of 0.10 and -0.10, we can assume independence. In the case of the data, there are some features like LE-Mean and LE-Standard Deviation that are in the range which are independent of each other. We also have features that are very dependent on each other, for example F3-Peak and F3-SNR. It can be assumed that features from the same channel will be dependent on each other. Another assumption would be that features from different channels will be independent from each other, however there are features from different channels like C4-Peak and F3-Standard Deviation that show some dependency.

### **Histograms**

In this section we will be plotting histogram plots for both the clench and non-clench classes for each feature in our data set. Such plots will allow us to view the data distributions shapes and ranges and aid us in identifying data features normality.

### **Channels Mean**

Statistics Table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Channel** | **Min** | **Max** | **Mean** | **Median** | **Variance** |
| **Jaw Clench** | **LE** | -0.581 | 0.745 | 0.039 | 0.037 | 0.033 |
| **F4** | -0.995 | 0.851 | 0.008 | -0.011 | 0.06 |
| **C4** | -1.051 | 1.262 | 0.15 | 0.119 | 0.124 |
| **P4** | -0.356 | 0.438 | -0.003 | -0.007 | 0.013 |
| **P3** | -0.551 | 0.886 | -0.001 | 0.003 | 0.014 |
| **C3** | -2.429 | 2.44 | 0.024 | 0.024 | 0.142 |
| **F3** | -0.886 | 0.808 | -0.073 | -0.059 | 0.068 |
| **Non-Clench** | **LE** | -0.607 | 0.38 | 0.027 | 0.031 | 0.004 |
| **F4** | -0.38 | 0.377 | 0.016 | 0.01 | 0.016 |
| **C4** | -0.358 | 1.225 | 0.087 | 0.048 | 0.023 |
| **P4** | -0.197 | 0.236 | -0.005 | -0.004 | 0.003 |
| **P3** | -0.243 | 0.171 | -0.004 | -0.005 | 0.005 |
| **C3** | -0.561 | 0.539 | 0.012 | 0.008 | 0.012 |
| **F3** | -0.602 | 0.509 | -0.046 | -0.038 | 0.035 |

Histograms:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  | |

From the channels mean histograms and the statistics table, overall, we can see that the value ranges are larger for jaw clench class compared to the non-clench class. When comparing the channels we see that the data is spread is larger in Channels C3 and C4 , which correspond to the readings from the Central sulcus, followed by Channels F3 and F4, which correspond to the readings from the Frontal lobe, while the other channels: P3 P4 and LE, which corresponds to the readings from the Parietal lobe and left ear, had lower data spread.

Also, we can see from the classes’ values ranges that the non-clench class range lies entirely within the range of the jaw clench class where both the ranges are centered around zero. This is due to the signal property of oscillating around zero and the jaw clench class having higher amplitudes compared to the non-clench class.

The channels means’ histograms exhibit a normal distribution, which tells us that we can assume normality of the channels mean values.

### **Channels Standard Deviation**

Statistics Table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Channel** | **Min** | **Max** | **Mean** | **Median** | **Variance** |
| **Jaw Clench** | **LE** | -11.72 | -1.46 | -6.571 | -6.678 | 3.316 |
| **F4** | -9.582 | 1.752 | -6.127 | -6.252 | 4.042 |
| **C4** | -10.103 | 5.679 | -6.411 | -6.683 | 4.22 |
| **P4** | -10.697 | 2.427 | -6.683 | -6.764 | 3.451 |
| **P3** | -10.231 | 0.075 | -6.235 | -6.517 | 3.304 |
| **C3** | -10.354 | 0.391 | -6.223 | -6.434 | 3.722 |
| **F3** | -10.508 | 1.305 | -6.024 | -6.139 | 3.466 |
| **Non-Clench** | **LE** | -10.966 | 7.152 | -3.856 | -4.146 | 9.098 |
| **F4** | -10.308 | 6.885 | -3.72 | -4.032 | 8.112 |
| **C4** | -10.589, | 7.62 | -4.07 | -4.586 | 10.028 |
| **P4** | -10.407 | 5.255 | -4.666 | -4.91 | 7.279 |
| **P3** | -10.343 | 5.18 | -3.945 | -4.264 | 6.563 |
| **C3** | -10.343 | 10.883 | -4.372 | -4.842 | 9.012 |
| **F3** | -10.641 | 9.046 | -3.69 | -3.894 | 7.304 |

Histograms:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  | |

From the standard deviation’ histograms and the statistics table for the channels, we can see that among all the classes, the distributions and ranges of the channels are the same. However, we can see that the range of the jaw clench class lies entirely within the lower part of the non-clench class data range. The larger values of the jaw clench class are due to the signals of non-clench Class signals being more sensitive to noise than the jaw clench class signals, which in term enlarges the signals data spread around the mean.

The channels standard deviation’ histograms exhibit a normal distribution, which tells us that we can assume normality of the channels mean values.

### **Channels Peak Value**

Statistics Table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Channel** | **Min** | **Max** | **Mean** | **Median** | **Variance** |
| **Jaw Clench** | **LE** | 14.489 | 125.979 | 38.052 | 35.745 | 245.655 |
| **F4** | 15.649 | 104.827 | 51.154 | 47.994 | 250.965 |
| **C4** | 23.171 | 175.224 | 66.153 | 64.216 | 679.811 |
| **P4** | 6.106 | 60.676 | 22.371 | 19.127 | 182.84 |
| **P3** | 3.967 | 64.124 | 20.883 | 21.823 | 142.675 |
| **C3** | 13.868 | 231.962 | 71.201 | 71.035 | 1387.967 |
| **F3** | 17.106 | 88.325 | 43.474 | 40.883 | 172.105 |
| **Non-Clench** | **LE** | 3.203 | 32.107 | 5.678 | 5.019 | 9.652 |
| **F4** | 3.428 | 38.279 | 8.344 | 7.989 | 21.41 |
| **C4** | 2.806 | 58.086 | 8.287 | 6.207 | 57.516 |
| **P4** | 2.1 | 20.978 | 4.031 | 3.633 | 4.416 |
| **P3** | 2.72 | 21.649 | 4.366 | 3.792 | 4.562 |
| **C3** | 2.998 | 76.189 | 7.938 | 6.083 | 66.204 |
| **F3** | 3.329 | 41.227 | 9.216 | 10.102 | 20.809 |

Histograms:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  | |

For the channel’s peaks value histograms and the statistics table, there is a trend of having similarly shaped graphs with the data from the raise phase being closely packed together at the very left-hand side of the graph, where the values are low, with less frequent values moving right, where the values are high. The Jaw clench data has a more even spread of distribution with its range values being higher than the non-clench class range. The key difference between all the histograms is the ranges of the distributions with the highest range of distribution being from the C3 and C4 channels followed by the Left ear channel, then the F3 and F4, and lastly the P3 and P4 Channels.

Viewing the data, the Jaw Clench has a wider range than the non-clench data. The data for the Jaw clench is right skewed with a few outliers in the high values. A potential theory could be the subjects applying more pressure during the clench phase which adds more movement data to be collected by the sensors. For the non-clench data, it is more normal, but it has a slight skew to the right, some potential causes are the shift between non-clench to jaw clench, a potential noise caught by the sensors, or a movement from the test subject.

In terms of normality, the jaw clench data exhibits a more normal distribution compared to the non-clench data.

### **Channels Sound to Noise Ratio**

Statistics Table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Channel** | **Min** | **Max** | **Mean** | **Median** | **Variance** |
| **Jaw Clench** | **LE** | 44.11 | 451.3 | 137.174 | 128.833 | 3730.397 |
| **F4** | 59.062 | 482.351 | 189.248 | 178.838 | 4008.781 |
| **C4** | 66.6 | 686.44 | 242.361 | 236.556 | 9331.754 |
| **P4** | 18.783 | 339.284 | 85.746 | 73.14 | 2852.778 |
| **P3** | 11.318 | 213.126 | 75.775 | 77.342 | 1931.532 |
| **C3** | 48.621 | 806.621 | 260.053 | 143.651 | 17243.36 |
| **F3** | 56.496 | 321.304 | 155.885 | 143.651 | 2468.56 |
| **Non-Clench** | **LE** | 8.894 | 149.35 | 20.561 | 16.295 | 295.616 |
| **F4** | 9.136 | 217.855 | 31.06 | 27.11 | 518.591 |
| **C4** | 7.999 | 353.758 | 32.579 | 19.181 | 1716.576 |
| **P4** | 5.942 | 137.688 | 14.782 | 11.71 | 153.388 |
| **P3** | 7.499 | 105.197 | 14.659 | 11.957 | 116.78 |
| **C3** | 8.19 | 368.732 | 31.155 | 19.331 | 1838.868 |
| **F3** | 8.79 | 160.477 | 34.96 | 33.126 | 457.842 |

Histograms:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  | |

For the histograms of the Signal to Noise Ratio data and the statistics table, the shape is almost identical in terms of distribution with a key difference being the ranges of the data. The higher the signal, the better the quality of the data. In the case of the outlying data of the rest phase, jaw clench data masks it and keeps from being noticeable in the overall distribution of the data. With signal to noise ratio, it looks at the ratio between a signal and background noise. This will be able to help us in terms of checking the performance of the data being recorded vs any noise. In the case of this data, the Jaw clenches are the more powerful signal in the recording against any noise in the data.

In terms of normality, the jaw clench data exhibits a more normal distribution compared to the non-clench data.

A reason for overall data spread difference between those channels can be that F3 and F4 channels readings reflect signals of voluntary movements and jaw clench is a voluntary movement which explains why the reflects the change if the signal due to the jaw clench better than the P3, P4, and LE channels, which reflects the signals resulting from a sensory stimulus. As for the C3 and C4 channels, the channel readings reflect both the signals resulting from voluntary movements and sensory stimuli, this due to the position of the electrodes being in the middle between the Frontal and Parietal lobe, which explains why the readings of those channels are higher than the all the other channels.

### **Scatter Plots**

Scatter plots are primarily being used to observe the data spread & show relationships between all the features in a channel. The dots in a scatter plot not only report the values of individual data points, but also patterns when the data are taken as a whole. Identification of correlational relationships can be inferred using the scatter plots. In this section, we will be plotting the scatter plots for all the channels and the respective features with the two classes clench and non-clench of the dataset.

|  |
| --- |
|  |
|  |
|  |
|  |
|  |
|  |
|  |

Overall, non-clench class clusters are denser compared to the jaw clench data. This is because during the non-clench phase of the testing, the subject will be relaxing and, in a room, completely devoid of any type of distractions. The signals should be consistent and have less variability. Any type outlying data would be from either a distraction or potential noise, or it could be the phase changed between non clench and clench phases.

When we look at the Mean and Peak Value scatter plots, we see that the data is forming separable clusters along the values of the y axis which correspond to the peak values while the classes data spread along the x-axis overlaps where the range of the non-clench class is encapsulated within the clench class.

For the Mean and Signal to Noise Ratio scatter plots, the plots exhibit the same behavior as the Mean and Peak Value scatter plots.

The Mean and Standard Deviation scatter plots do not exhibit any separable clusters behavior as the both the feature ranges highly overlap. This is like the histograms and exhibits the same reasoning.

As for the Peak Value and Standard Deviation scatter plots, we see that the data is forming separable clusters along the values of the y axis which correspond to the peak values while the classes data spread along the x-axis overlaps where the range of the Clench class is encapsulated within the non-clench class.

For the Standard Deviation and Signal to Noise Ratio scatter plots, the plots exhibit the same behavior as the Peak Value and Standard Deviation scatter plots but with reversed axes distribution.

As for the Peak Value and Signal to Noise Ratio scatter plots, we can see that the data is forming separable clusters along both axis’ distribution. And the behavior of the data shows a positive correlation between the two features. This is due to Signals of the clench class having both high amplitudes and high power that mask any noise vs non-clench class thus making the clench class values lie in the higher ranges of both axes.

The reasonings and conclusions interpreted by looking at the scatter plots match the reasonings that were described in the histogram section of this report.

Given the similarity between all the scatter plot graphs, this section will focus on any potential differences between the graphs and the reasonings of why they might be so. Our speculation as to why the results are so similar is because while the channels focus on different areas of the brain, they are still focused on one entity.

For the C3, the features, except for the Standard Deviation feature, are the most clustered together data compared to the features of the other channels followed by P3. This could be because they are both associated with the left side of the brain which corresponds to speech and abstract thinking.

When we look at the Standard Deviation feature, standard deviation for LE is the most condensed for the non-clench class and P3 for the clench class along the x-axis. We can observe that for C3 and C4 also following the same with little more scattered than LE and P3. For P4, F3 and F4 we can observe data is most scattered for this feature.

When we look at the F3 and F4 features scatter plots, we can see that both have similar clustering and data spread behavior. This can be because both read signals from the Frontal lobe which corresponds to the voluntary movement signals.

For P3 and P4 features scatter plots, we can see that the clench class data clusters spread behavior is slightly larger in P4 than P3 although both of their readings are from the Parietal lobe. This could be to which side of the brain they are placed.

# **Data Preprocessing and Feature analysis**

For data preprocessing we started by scaling the data using minimum and maximum approach:

After that we mean centered the data by using:

We also looked at the features means and standard deviations and their ratios (mean/standard deviation and standard deviation/mean) to identify useless and noisy features and we found that none of them were exhibiting such behavior. Finally, we balanced the data in order to have an equal amount of Jaw clench data and non-clench data.

# **Principal Component Analysis (PCA) & Data Analysis**

This section will be covering the Principal Component Analysis and the interpretation of the results derived from the PCs. The data variability of each channel will be viewed to determine which are more contributing to the data clustering.

### **Channels’ Mean Features**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

From the channels’ Mean features 3D plots above, we can see that the Non-clench data is closely clustered together which is understandable because the amplitudes for the rest phase will be at a constant state of change for the amplitude. The Jaw Clench means will mimic the behavior of the amplitudes where they will be higher and lower amplitudes and will contain more noise, so the means of the clench phase will be more surrounding the means of the Non-clench in almost a spherical cocoon shape to show that.

### **Channels’ Standard Deviation Features**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

From the channels’ Standard deviation features 3D plots above, we can see that the Non-clench data is scattered, and we don't see clear clusters of Clench and Non-clench data. Clusters are overlapping each other this is due to non-jaw clench data having more variation and jaw clench data variation lies within it.

### **Channels’ Peak Value Features**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

From the channels’ Peak Values features 3D plots above, we can see that the Non-clench data is clustered at the bottom of the graph while the jaw clench data is spreading out in a linear fashion which matches the pattern of the means graph where the amplitudes of the Clench class would be higher than the non-clench.

### **Channels’ SNR Features**

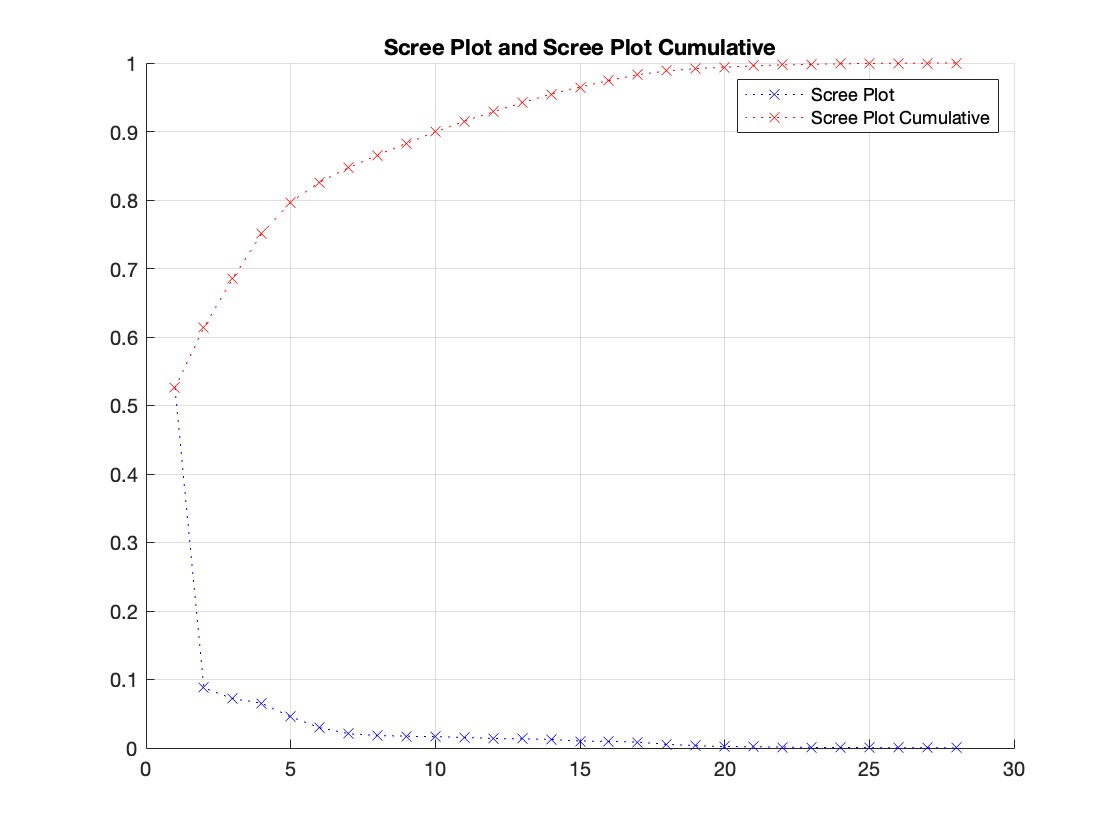
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

For the signal to noise ratio for the channel, the non-clench data is clustered tightly at the base of the graph, indicating the rest phase is more susceptible to noise while the clench data at the amplitudes get high so does the signal of the clenches which in turn masks the noise. This matches the pattern of the peak value where the clench data grows linearly with the strength of the clench.

### **All Channels’ Features**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Channel** | **Mean, Peak Value, and SNR** | **Mean, Peak Value, and Standard Deviation** | **Mean, Standard Deviation, and SNR** | **Peak Value, Standard Deviation, and SNR** |
| **LE** |  |  |  |  |
| **F4** |  |  |  |  |
| **C4** |  |  |  |  |
| **P4** |  |  |  |  |
| **P3** |  |  |  |  |
| **C3** |  |  |  |  |
| **F3** |  |  |  |  |

When we look at the channels’ features 3D scatter plots, we can see that data classes show clear divided clusters in all plots around both SNR and Peak Values features axis. As we look at the Mean, Peak Value, and SNR features and Peak Value, Standard Deviation, and SNR features plots we can see that in both types of plots that the data exhibits a correlation behavior due the correlation that is portrayed in both Peak Value, and SNR features.



|  |  |
| --- | --- |
| PCs | Variability |
| First 5 | 80% |
| First 10 | 90% |
| First 15 | 96% |
| First 18 | 99% |
| First 20 | ~100% |

We can see from the scree plot that the first 5 PCs capture 80% of the dataset information (variability) which is a good starting point while PC 6 to 28 capture the remaining 20%.

|  |  |
| --- | --- |
| 52.5%  From PC1, we can see that in 52.5% of the data most contributing features are the channels' Peak Value and SNR, which are highly and positively correlated. This can be due to the fact that the high amplitudes of the jaw clench reading give the signal high peaks and higher power, thus a high SNR value. We can also see that both Peak Value and SNR are negatively correlated with the signals' standard deviations, this could be due to the high SNR signal meaning the data has less noise and thus less variation. | 8.7%  For PC2, about 8.7% of the overall data contribution comes from the prominent features of the standard deviation from each channel. The features are also all on the positive axis of the graph along with most of the other features. This can be due to the fact that this could be some noise during the Rest Phase of the data, or it could be during a transition phase from clench to rest or vice versa. LE\_Mean is insignificant in this loading vector while the other features are still visible in the graph. |
| 7.2%  From PC3, we can see that the Mean features for all channels are dominant and positively correlated except for LE\_Mean. Also, Mean and Standard deviation are positively correlated and comparatively, Standard deviation has a low contribution, this is because high amplitudes of the jaw clench reading give the signal high means and mask the noise thus having a low standard deviation. The negative correlation between LE\_Mean and the other channels’ Mean features could be due to how we observed the LE channel has the least affected channel by the jaw clench state. | 6.5%  In this PC4, the P4\_Mean and P3\_Mean are the most significant contributing features and are negatively correlated. While they are negatively correlated with the P3\_Mean which is the next significant feature. The low contribution of the F channel tells us that and the high contribution of the C and P channels tells us the data is in the rest phase which means that the signal is more susceptible to noise. The negative correlation between LE, C3, P3, and F3 with C4, P4, and F4 could be because the first group collects the data from the left side of the brain, while the second group collects the data from the right side. |
| 4.6%  In the PC5, the LE Mean has the highest significance and is positively correlated to F4 Mean and F3 Mean while they are negatively correlated to P4\_Mean P3\_Mean and the C3\_Mean. This could mean that this loading vector is taking in the data from when the subject is about to clench their jaw, the Mean for F3 and F4 is on the negative side of the graph indicating a decrease in Mean, while Peak Value and SNR are more apparent in F3 and F4 than the other channels which indicate that the other channels Peak Value and SNR may follow the trend being set as in the previous loading vectors, there is always a similar relationship between the features between channels. | |
| 2.9% 2.1%    1.8% 1.7%  1.7% 1.5%    1.4% 1.3%    1.3% 1%  Chart, bar chart  Description automatically generated  0.96% 0.86%  Chart  Description automatically generated  0.53% 0.35%    0.21% 0.18%  **Chart  Description automatically generated**  0.14% 0.09%    0.08%  For the remaining PCs, we see that in those loading vectors, all the correlations and the captured variations are due to noise in the data. | |
|
|
|
|
|
|
| 0.05%  In PC 25, the most contributing features are the Peak Value features, whereas Channel F3 is the most contributing feature. We can see the Peak Values are negatively correlated with the channels’ SNR features. This can be due to clench signal mask noise in the data at the jaw clench state during the transition phase from the jaw clench state. | 0.03%  In PC 26, the most significant features are LE\_Peak\_Value, and C4\_Peak\_Value (in order of significance). It can be seen that Peak Value from LE is negatively correlated to C4. This could be because of the noise recorded in the data. |
| 0.022%  In PC 27, the most contributing features are the Peak Value features. whereas Channel F4 is the most contributing feature we can see the Peak Values are negatively correlated with the channels’ SNR feature. This can be due to clench signal mask noise in the data at the jaw clench state. | 0.016%  This is the lowest contributing vector. C4, P4, P3, and C3 are the contributing channels in this vector. The channel's Peak Value features are the most contributing feature here. We can see that C4 and C3 and P4 and P3 are negatively correlated. We can conclude that as these observations form central and parietal regions this data represents sensory information. |

|  |  |
| --- | --- |
|  |  |
|  | **Chart, scatter chart  Description automatically generated** |
|  |  |
|  |  |
| **Chart, scatter chart  Description automatically generated** |  |

Above are the principal components for PCs 1 to 5, the first ten images have a clear divide between the two classes which are clustered amongst themselves. Of all the principal components and graphs, only when PC 1 is among the three PCs, does the separation behavior occur. This is due to the high level of contribution of variability from PC 1 which sits around 52.5%, PC 2 which holds 8.7% of the data variability doesn’t exhibit such behavior. On examining the loading vectors and the 3D feature plots, the separation of the data is based on the contributing variable per loading vectors.

In PC 1, the most contributing features are the Peak Values and the Signal to Noise Ratio of each channel whereas in PC 2 it is the Standard Deviation of each channel. Because of the linear relationship of Peak Value and Signal to Noise Ratio feature, the clear separation of when the jaw is clenched and not clench is evident and clear, while mean less reliable because of the lack of separation between clenched and unclenched phases as seen in the 3D plots.

While for the remaining PCs 3D Plots we see that the class clusters are intertwined with no separation, this is because the PCs that are Graphed in those 3D Plots are PCs 2 to 5 where the most contusion features were the Channels Standard Deviation (for PC 2) and Mean (for PC 3 to 5). As we saw for the feature the scatter plots showed the classes clusters are intertwined with no separation between them.

|  |  |
| --- | --- |
|  | Chart, scatter chart  Description automatically generated |
|  | Chart, scatter chart  Description automatically generated |
|  | Chart, scatter chart  Description automatically generated |
| Chart, scatter chart  Description automatically generated | Chart, scatter chart  Description automatically generated |
|  | Chart, scatter chart  Description automatically generated |

3D data scatter plots for U, we observed that the clusters are almost the same but in different scales in U as compared to Ur. It is a smaller scale as compared to Ur. That is the reason Ur have more clarity in separating jaw clench and non-jaw clench class. This helps in viewing & separating the clusters.

## **Without the Mean**

During the classification analysis, it was discovered that with the removal of mean, the Pre-PCA data remind the same in terms of accuracy. In order to understand the data better, the PC Analysis was done a second time to the data without any Mean Features.

**Chart

Description automatically generated**

|  |  |
| --- | --- |
| PCs | Variability |
| First 1 | 69.42% |
| First 2 | 81% |
| First 5 | 89% |
| First 10 | 97% |
| First 13 | ~100% |

With the removal of Mean, We can see from the scree plot that the first 2 PCs captures 81% of the dataset information (variability) which is a good starting point while PC 3 to 21 capture the remaining 19%.

### **Loading Vectors**

|  |  |
| --- | --- |
| Chart, bar chart  Description automatically generated  69.42%  From PC1, we can see that in 69.42% of the data most contributing features are the channels' Peak Value and SNR, which are highly and positively correlated. This can be due to the fact that the high amplitudes of the jaw clench reading give the signal high peaks and higher power, thus a high SNR value. We can also see that both Peak Value and SNR are negatively correlated with the signals' standard deviations, this could be due to the high SNR signal meaning the data has less noise and thus less variation. | Chart, bar chart  Description automatically generated  11.50%  For PC2, about 11.50% of the overall data contribution comes from the prominent features of the standard deviation from each channel. The features are also all on the positive axis of the graph along with most of the other features. This can be due to the fact that this could be some noise during the Rest Phase of the data, or it could be during a transition phase from clench to rest or vice versa. |
| 3.19% | Chart, waterfall chart  Description automatically generated  2.68% |
| 2.19% | Chart, waterfall chart  Description automatically generated  2.08% |
| Chart, waterfall chart  Description automatically generated  1.84% | 1.75% |
| Chart  Description automatically generated  1.70% | Chart, bar chart  Description automatically generated  1.29% |
| Chart, bar chart  Description automatically generated  0.75% | 0.47% |
| 0.29% | Chart  Description automatically generated  0.26% |
| Chart, waterfall chart  Description automatically generated  0.18% | Chart, waterfall chart  Description automatically generated  0.12% |
| Chart, waterfall chart  Description automatically generated  0.10%  For the remaining PCs, we see that in those loading vectors, all the correlations and the captured variations are due to noise in the data. | |
| 0.07%  In PC 18, the most contributing features are the Peak Value features, whereas Channel F3 is the most contributing feature. We can see the Peak Values are negatively correlated with the channels’ SNR features. This can be due to clench signal mask noise in the data at the jaw clench state during the transition phase from the jaw clench state. | |
| 0.04%  In PC 19, the most significant features are LE\_Peak\_Value, and C4\_Peak\_Value (in order of significance). It can be seen that Peak Value from LE is negatively correlated to C4. This could be because of the noise recorded in the data. | 0.03%  In PC 20, the most contributing features are the Peak Value features. whereas Channel F4 is the most contributing feature we can see the Peak Values are negatively correlated with the channels’ SNR feature. This can be due to clench signal mask noise in the data at the jaw clench state. |
| 0.02%  This is the lowest contributing vector. C4, P4, P3, and C3 are the contributing channels in this vector. The channel's Peak Value features is the most contributing feature here. We can see that C4 and C3 and P4 and P3 are negatively correlated. We can conclude that as these observations form central and parietal regions this data represents sensory information. | |

|  |  |
| --- | --- |
|  |  |
|  |  |

Above are the principal components for PCs 1 to 4, the first three images have a clear divide between the two classes which are clustered amongst themselves. Of all the principal components and graphs, only when PC 1 is among the three PCs, does the separation behavior occur. This is due to the high level of contribution of variability from PC 1 which sits around 69.42%, PC 2 which holds 11.50% of the data variability doesn’t exhibit such behavior. On examining the loading vectors and the 3D feature plots, the separation of the data is based on the contributing variable per loading vectors.

In PC 1, the most contributing features are the Peak Values and the Signal to Noise Ratio of each channel whereas in PC 2 it is the Standard Deviation of each channel. Because of the linear relationship of Peak Value and Signal to Noise Ratio feature, the clear separation of when the jaw is clenched and not clench is evident and clear, while mean less reliable because of the lack of separation between clenched and unclenched phases as seen in the 3D plots.

While for the remaining PCs 3D Plots we see that the class clusters are intertwined with no separation, this is because the PCs that are Graphed in those 3D Plots are PCs 2 to 4 where the most contributing features were the Channels Peak Values and the Signal to Noise Ratio but with different correlation between channels. As we saw for the feature the scatter plots showed the classes clusters are intertwined with no separation between them.

|  |  |
| --- | --- |
|  |  |
|  |  |

3D data scatter plots for U, we observed that the clusters are almost the same but in different scales in U as compared to Ur. It is a smaller scale as compared to Ur. That is the reason Ur have more clarity in separating jaw clench and non-jaw clench class. This helps in viewing & separating the clusters.

# **Classification**

In this section we performed classification using Bayes Discriminant and Decision tree classification algorithms on our data classes: Jaw Clench and Non-Clench.

### **Data Separation**

To split the data into the training and testing sets, the data was randomized, balanced, and split into the respective sets with a 70% and 30% ratio for each class. The process of splitting the data was done to the original features (Pre-PCA) and also on the principal components data (Post-PCA, the Scores Matrix U and Regular Scores Matrix Ur). The same training and testing datasets have been used for all the models.

### **Bayes Discriminant**

After observing the data through the histograms, it was discovered that the features Mean, and Standard Deviation exhibited the shape of a normal distribution. Based on the shape, it was decided that a model to use the Bayes Discriminant model because it will allow us to view in a gaussian normal distribution.

|  |  |
| --- | --- |
| **Dataset without Mean** | **Dataset with Mean** |
| Accuracy: 97.25% | Accuracy: 98.62% |
| Accuracy: 95.87% | Accuracy: 99.54% |
| Accuracy: 98.62% | Accuracy: 96.78% |

In the Bayes Discriminant results, it is shown in the table that the Bayes Discriminant model had an accuracy of 98.62%, 96.78%, and 99.54% over the original features, Scores Matrix (U) and Regular Scores Matrix (Ur) respectively. From the result, it is shown that the model had the best performance on Regular Scores Matrix (Ur) dataset; this can be attributed to how PC 1 exhibited clear divided clusters between the two classes. The second-best result, with accuracy of 98.62%, was when the original features (Per-PCA) was used. The high model performance attributed to how the Channels Peak Value and Signal to Noise Ratio features distribution highly separate the two classes.

After we applied the PCA it was observed that the Mean feature had no contribution to PC1 and PC2 which represented more than 60% of the data variability so we decided to drop it and test the model performance.

After removing the Mean feature, it is shown in the table that the Bayes Discriminant model had an accuracy of 97.25%, 98.62%, and 95.87% over the original features, Scores Matrix (U) and Regular Scores Matrix (Ur) respectively. From the result, it is shown that the model had a better performance on the Scores Matrix (U) dataset than on Regular Scores Matrix (Ur). This can be due as Scores Matrix (U) observations are not scaled as Regular Scores Matrix (Ur) thus their class distribution will have less overlap.

It is also worth mentioning that in both original features and Regular Scores Matrix (Ur) data set they had a better performance on the data set with the mean. This can be due to the mean being dropped and the data variability being affected thus the class data distributions changed, which is something that Bayes Discriminant is very sensitive to.

### **Decision Tree Model**

Because the data classification only has two outcomes, it was decided to use a decision tree model in order the observe the level of importance for each feature and PC to see how the model would reach the conclusion on if an observation was a Jaw Clench or Non-Clench.

|  |  |
| --- | --- |
| **Dataset without Mean** | **Dataset with Mean** |
| Accuracy: 96.79% | Accuracy: 96.79% |
| Accuracy: 98.17% | Accuracy: 96.33% |
| Accuracy: 98.62% | Accuracy: 97.25% |

In the Decision Tree Model, using data with the Mean feature has an accuracy of 96.79%, 98.62% and 98.17% over Original Features (Pre-PCA), Scores Matrix (U) and Regular Scores Matrix (Ur) respectively, while for data without the Mean feature, the accuracies are 96.79%, 97.25% and 96.33% respectively.

When it comes to original features of both models (with the Mean features and without the Mean features), their performances were identical; this tells us that the model with Mean features did not consider the Mean as a deciding feature.

It is also observed that in both the Score Matrix (U) and the Regular Score Matrix (Ur) dataset the model performed better on the data set after removing the Mean Features and processing them. This indicates that the Mean Features allows and contributes to a high overlap between classes distributions, leading to the conclusion that by removing the Mean feature we made the data more separable.

Looking at models before and after removing the Mean Feature, its observed that the model performs better on Score Matrix (U) dataset than it does on Regular Score Matrix (Ur) dataset, this can be due scaling in Score Matrix (U) increase the overlap between the classes clusters, thus resulting in an increase in classification error.

|  |
| --- |
| Pre-PCA Data Decision Tree without the Mean    Pre-PCA Data Decision Tree with the Mean    The Tree model of the Pre-PCA data for both mean and without Mean gives insight on why the accuracy of both data sets remain equal even without the Mean feature of each channel being removed. As viewed by both trees, no Mean feature appears on any branch which indicates that Mean has no say on what the outcome of the classification is. Another thing to note is that the features that appear in the tree are the Peak Value and the Standard Deviation which will also be appearing in the Ur and U data. |
| Ur Data Decision Tree without the Mean    Ur Data Decision Tree with the Mean    With Ur data decision trees, the accuracy improves with the removal of mean. This is due to the precision and the structure of the trees and which PCs are being used at each crossroad to make a decision for the classification. By first observing the Ur Decision Tree with the Mean, the corresponding PCs all have a contributing feature that is not the Mean feature. Like the Pre-PCA data, these PCs have Peak Value and Standard Deviation as their highest contributions features. It stands to reason that the higher accuracy of the data with no Mean occurs because the Mean feature in the PCs is a potential noise or disruption that is affecting the data negatively.  The Ur Decision Tree without Mean is also heavily reliant on PC1 which has Peak Value and SNR as the highest contributing features and hold almost 70% of the data variability. |
| U Data Decision Tree without the Mean    U Data Decision Tree with the Mean    Unlike the Ur Decision Tree with Mean, the U Decision Tree with Mean does have PCs where the highest contributing feature is the Mean. However, the accuracy of the U data without the Mean is still higher than with Mean. Between the Ur and U, the idea that the Mean feature is noisy data becomes more apparent when a decision tree is being used. |

Overall, for the original features (Pre-PCA), Bayes Discriminant models had the better performance. Also, for the Post-PCA data (U and Ur datasets) without the Mean features, Decision Tree model had the better performance.

### **Future Steps**

A potential future project would be to check to see if there is a possibility to narrow the data to only use two channels. Because F3 and F4 has a focus on the frontal lobe which is associated with voluntary movement. The decision tree showed us that the accuracy of the Pre-PCA data to have increased with only using those two channels with only 6 features between them (Peak Value, Standard Deviation, and SNR). The reason of the accuracy may lie in the fact that because the F3 and F4 channels focus on the frontal lobe, the jaw clench frequency is stronger and that with that focus with the noisy from the other channels, the accuracy has now increased.

Additionally, by viewing the tree, it is observed that the path is reliable only on Peak Value without any contribution from the other features. Additionally, looking at the scree plots show that the data variability of PC 1 is now at 76%, which is an increase compared to using all of the channels.

|  |  |
| --- | --- |
| Loading Vector 1: 76.78% | Accuracy: 98.17% |
| Pre-PCA Decision Tree for F3 and F4 | |

A few other potential future steps would be to test and analyze the data without the SNR given how it is heavily correlated with Peak Value or adding new features such as Mean and Standard Deviation ratio.

Finally, a potential future step would be to create a new data set that focusses on another type of movement like neck movement or blinking. Neck movement would be a potential next step as it has a strong signal when using the EEG machine, while blinking is a more subtle movement and would be more difficult to capture.

### **Team Assessment**

**Noura Alroomi**

* Created scripts to balance and split data.
* Worked on PCA update for the balanced data and data without Mean features.
* Regraphed 3D features Scatter plots
* Regraphed Scree plots, Loading Vectors for the balanced data.
* Data analysis for the features, loading vectors, and 3D graphs to interpret results for the balanced data and the balanced data without Mean.
* Added legend for loading vectors to identify each channel and its features.
* Classification results analysis.
* Documentation and editing.
* PowerPoint Documentation

**Prarin Behdarvandian**

* Created the script for the Decision Tree Model along with the confusion matrix for accuracy
* Regraphed the Post PCA (U and Ur) 3D scatter plots based on the balanced data and data without Mean features.
* Regraphed Scree plots, Loading Vectors for the for data without Mean features.
* Data analysis for the features, loading vectors, and 3D graphs to interpret results for the balanced data and the balanced data without Mean.
* Explored the Pre-PCA and scree plot data of only two channels, F3 and F4
* Classification Analysis
* Documentation and Editing
* PowerPoint Documentation

**Pragati Dode**

* PCA without Signal to Noise Ratio. Also used Decision Tree Model to check change in Confusion Chart and Model Accuracy after removing this feature. (we are not using this as we decided to go with removing the Mean feature.)
* PCA with Mean to Standard Deviation Ratio. Also used Decision Tree Model to check change in Confusion Chart and Model Accuracy after adding this feature. (We are not using this as we decided to go with removing the Mean feature)
* Data analysis for the features, loading vectors, and 3D graphs to interpret results for the balanced data and the balanced data without Mean.
* Scatter Plot Matrix without Mean feature for all channels
* Documentation and Editing
* PowerPoint Documentation

**Khaleel Rehman**

* I wrote a script to compute the Naive Bayes classifier for all datasets - balanced vs unbalanced, U, Ur, Pre PCA and after removing the mean feature. Also plotted the respective confusion charts for the analysis in the final report.
* I explored MATLAB’s Classification Learner tool which provides automation to the testing and training of many different classifiers and used the Decision trees (all its variations - fins, medium, coarse), the Naive Bayes Classifier from here and played around with data to see the accuracies (We omitted this tool for the final report and wrote the code for the classifiers instead).
* Performed short research on Gaussian vs KNN vs Naive Bayes to understand which suits our data the best. It turns out the decision tree suits most because it prunes the useless features besides the fact that data is categorical.
* Analysis and Interpretation of graphs & plots, classifier results in the report.
* Contributed to the Documentation and Editing part of the final report.
* Contributed to PowerPoint Documentation.

### **Appendix**

* EEG: is an instrument that measures and records the brain electrical activity.

### **References**

Blocka, K. (2018, September). *EEG (electroencephalogram): Purpose, procedure, and risks*. Healthline. Retrieved November 10, 2022, from https://www.healthline.com/health/eeg#risks

Browne, C. (2022, October). *Introduction to the Project*. Lecture, Bothell; University of Washington.

*Central sulcus - definition*. @neurochallenged. (n.d.). Retrieved November 19, 2022, from <https://neuroscientificallychallenged.com/glossary/central-sulcus>

*Gabor transform*. Gabor Transform - an overview | ScienceDirect Topics. (n.d.). Retrieved November 19, 2022, from <https://www.sciencedirect.com/topics/mathematics/gabor-transform>

Harris, K. D., & Thiele, A. (2011, August 10). *Cortical state and attention*. Nature reviews. Neuroscience. Retrieved November 19, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3324821/#:~:text=Cortical%20activity%2C%20even%20in%20primary,is%20determined%20by%20cortical%20state.

*Highpass Filter*. Highpass-filter signals - MATLAB. (n.d.). Retrieved November 19, 2022, from <https://www.mathworks.com/help/signal/ref/highpass.html>

*An interactive guide to the Fourier transform*. BetterExplained. (n.d.). Retrieved November 19, 2022, from https://betterexplained.com/articles/an-interactive-guide-to-the-fourier-transform/

Lewis, S. (2020, August 3). *Understanding your EEG results*. Healthgrades. Retrieved November 10, 2022, from https://www.healthgrades.com/right-care/electroencephalogram-eeg/understanding-your-eeg-results

Mayo Foundation for Medical Education and Research. (2021, June 3). *Frontal lobe seizures*. Mayo Clinic. Retrieved November 19, 2022, from <https://www.mayoclinic.org/diseases-conditions/frontal-lobe-seizures/symptoms-causes/syc-20353958#:~:text=The%20frontal%20lobe%20is%20important,is%20primarily%20responsible%20for%20vision>.

MediLexicon International. (n.d.). *Left brain vs. right brain: Characteristics, functions, and myths*. Medical News Today. Retrieved November 19, 2022, from <https://www.medicalnewstoday.com/articles/321037#:~:text=The%20left%20and%20right%20sides%20of%20the%20brain.&text=The%20brain's%20left%20half%20is,left%20side%20of%20the%20body>.

Mikkelsen, K. B., Kappel, S. L., Mandic, D. P., & Kidmose, P. (1AD, January 1). *EEG recorded from the ear: Characterizing the ear-EEG method*. Frontiers. Retrieved November 19, 2022, from <https://www.frontiersin.org/articles/10.3389/fnins.2015.00438/full>

*Parietal lobe*. Physiopedia. (n.d.). Retrieved November 19, 2022, from <https://www.physio-pedia.com/Parietal_Lobe#:~:text=Research%20suggests%20that%2C%20the%20more,receiving%20and%20processing%20their%20input>.

Parsons, E. (2022, October). Bothell; University of Washington.

Raad, L. (2022, October). *Introduction to Filtering the data*. Lecture, Bothell; University of Washington.

Sirven, J. et all (n.d.). *Electroencephalography (EEG)*. Epilepsy Foundation. Retrieved November 19, 2022, from https://www.epilepsy.com/diagnosis/eeg

*Streamline signal time feature extraction*. Streamline signal time feature extraction - MATLAB. (n.d.). Retrieved November 19, 2022, from https://www.mathworks.com/help/signal/ref/signaltimefeatureextractor.html

*A step-by-step explanation of principal component analysis (PCA)*. Built In. (n.d.). Retrieved October 19, 2022, from <https://builtin.com/data-science/step-step-explanation-principal-component-analysis>

Mayo Foundation for Medical Education and Research. (2022, May). *EEG (electroencephalogram)*. Mayo Clinic. Retrieved November 10, 2022, from <https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>

*What is signal to noise ratio and how to calculate it?* Cadence. (2022, October 13). Retrieved November 19, 2022, from https://resources.pcb.cadence.com/blog/2020-what-is-signal-to-noise-ratio-and-how-to-calculate-it