Alcoholic Addiction Classification EEG

Yiran Yao

December 2023

1 Abstract

This study focuses on the application of machine learning models on EEG data for classifying alcohol addiction. This study utilizes the EEG dataset from the UCI Machine Learning Repository, comprising data from participants in control and alcoholic groups exposed to various visual stimuli. We explore the efficacy of different machine learning models in distinguishing between these groups based on their EEG signals. The preprocessing pipeline involves data cleaning, transformation from the time domain to the signal domain, and the generation of RGB images to represent EEG signals. The models included in this study are: a basic neural network, a Convolutional Neural Network (CNN), a ResNet model, and an autoencoder. Our results highlight the potential of using EEG data and advanced machine learning techniques to accurately classify individuals with alcohol addiction, offering a robust tool for clinical diagnostics and personalized treatment strategies.

2 Dataset

The data set used in this study is from the UCI Machine Learning Repository a well-known repository for machine learning data sets[1]. Particular in this data set, the participants were given 3 scenarios, S1 obj, S2 match, and S2 no-match. Specifically, the participants in S1 are shown 1 picture and in S2 are shown 2 pictures. The S2 matching group are shown the same picture for 2 times, while the S2 no-match are shown 2 different picture. And EEG datas are collected from these participants, and put into the dataset. The participants were divided into alcoholic group and control (non alcoholic) group.

2.1 History of EEG

EEG has being used for studying addiction as early as late 19th century, despite during the time, only simple pattern in EEG waves are being inspected. The use of EEG on studying addition started around 1950s, although facing challenges later in 1970s, scientists continued to study EEG patterns in addicts, often focusing on potential treatments for addiction. Over the past 3 decades, more and more studies were being implemented on using Substance Use Disorders (SUDs).[2]

2.2 Why EEG

Electroencephalography (EEG), which shows brain's electron activity, are a pivotal tool in studying the Substance Use Disorders (SUDs). Due to its direct measurement of brain's neuro activities, EEG has an natural advantage over other tools in studying SUDs including alcoholic addiction. In addition, EEG would also provide quantitative measurements that would be ideal for precise classifications.[3]

2.3 Explanatory Data Analysis

After close inspection and data analysis on the data set, the possibility of implementing machine learning models were proven. The following figures are the comparison between alcoholic groups and non alcoholic groups.

In Figure 1, we can see several graphs about the S1 dataset. Looking at the 3D graphs of the EEG datas from the data, figure 1a and figure 1b, we can see that despite generally the patterns of EEG data for alcoholic group and the control group are the same, the alcoholic group has higher peaks and lower troughs. This can also be seen in figure 1c and figure 1d. In addition, there are also significant different between the two groups in terms of their correlation within the points.

As shown in Figure 2, the S2 Match group possess some different characteristics than S1 group. From Figure 2a and Figure 2b, it's obvious that different from S1 group, for S2 group, the control group and the Alcoholic Group processes completely different pattern. In addition, the difference between the correlation of points in S2 match is even bigger than S1, as shown in Figure 2e.

The difference in pattern for S2 no-match is even more significant compare to S2 match, as shown in 3. Alcoholic Group only as a few peaks, as shown in Figure 3b, which is significant different than the control group, Figure 3a. In addition, the Correlation graph, Figure 3e, shows that the difference in correlation between the two groups are also very significant.

From the above 3 analysis of the dataset, it's easy to tell that there exists significant differences between alcoholic groups and control groups in terms of EEG datas. So implementing machine learning algorithm would be a possible way to classify further EEG signals into the 2 groups.

3 AI models

As mentioned in the previous section, the trend in the dataset showed a possibility of implementing machine learning algorithm on the dataset. There are in total 4 models that are trained on the dataset.

3.1 Data preparations

For the 4 models, parts before AI models generally remains the same, as shown in Figure 4.

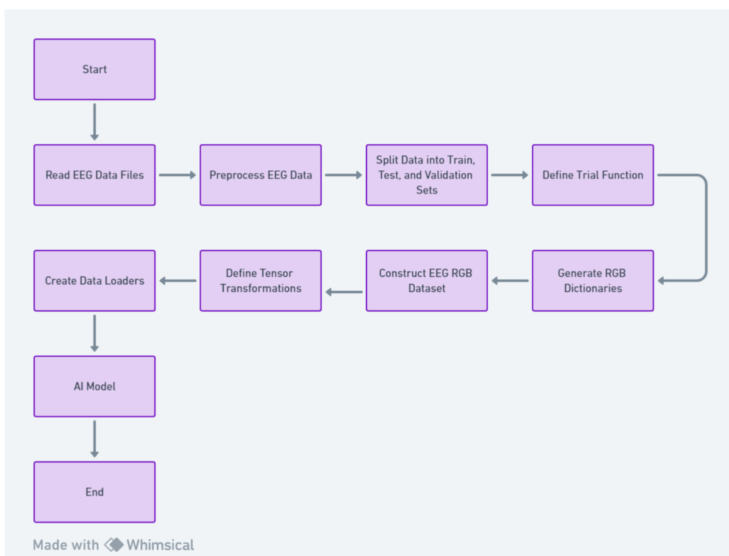


Figure 4: Overall Structure of the models

3.1.1 Packages used

Except for common used libraries, like PyTorch, NumPy, and Pandas, etc, this model also imported the package MNE, which could be used for further analyzing EEG datas.

3.1.2 Processing Data

The processing data stage of the code aims to prepare the EEG data for analysis and modeling. This stage generally includes a few parts: Data loading, data cleaning, and data splitting. After the file import, 2 separated Data frames are created - one for control group(non-alcoholic group), and one for alcoholic group. Then, irrelevant data are being removed, for example, 'Unnamed:0'. And then, the datas are being separated into training datas and testing datas, with a 4:1 ratio. This split would ensure the balance between the learning and validating the model.

3.1.3 Define Trail Function

The trail function is aimed to transform the originally time domain EEG data into signal domain data. And example of signal domain and time domain data is presented in Figure 7. This data transformation is essential for further feature extraction and analysis in the machine learning models. There are a few benefits of using signal domain than time domain data. As brain's signal, EEG data possess many complex patterns which would not be apparent if it is displayed in time domain format. In other words, this would allow for component separation to be applied to the data (In this case, it is Alpha, Beta and Theta). In addition, this transformation would help isolate the signal from noises, improving the clarity and quality of the data for analysis.

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \tag{1}$$

Figure 5: Fourier Transform Equation

$$PSD(f) = \lim_{T \to \infty} \frac{1}{T} |\hat{F}(f)|^2 \tag{2}$$

Figure 6: Power Spectral Density Equation

The trail function takes the EEG data as input and applies a Fourier transform to each segment(Figure 5). After the transformation, power spectral density(PSD, Figure 6) in various frequency bands(Alpha beta, Theta) is calculated, outputting a quantitative measure of signal strength.

3.1.4 RGB dataset

After separating the EEG datas into different waves(Alpha, beta, theata), the RGB dataset is being generated from the waves. Each waves corresponds one component in Red, Green and Blue. Red Channel: Represents the power in the Theta band. Green Channel: Represents the power in the Alpha band. Blue Channel: Represents the power in the Beta band. For each trail of EEG data, an RGB image is being generated, where each pixel's color represents the power values of the EEG signals at different electrodes and frequency bands. Noted that the power values are all normalized into a scale of 0-255 for each RGB channel. An example of RGB image generated would be Figure 8. There are generally a few advantages of generating RGB dataset. By converting the EEG datas into RGB datasets, this would enables more machine learning models that is related to image recognition and classification to be applied on the data set. In addition, the image would also help to visualize the signal domain data.

3.2 Models

After the data preparation, the model is ready to take input. There are in total 4 models that are trained on the dataset, and all the models are evaluated using the same metrics, accuracy and confusion matrix. The accuracy shows the model's percent of correct prediction on testing data, and confusion matrix shows the model's performance on two groups specifically. Figure 9 is an examle of confusion matrix. Ideally, we want (1,1) and (0,0) take up all the values, just like Figure

3.2.1 Base Model

Architecture The Base Model is a straightforward feedforward neural network designed for EEG data classification. Its architecture consists of the following layers:

- Input Layer: Accepts the EEG image data, which is flattened to suit the input requirements of the network.
- Hidden Layers: Comprises one or more layers (exact number not specified) that process the input data. Each layer typically involves a linear transformation followed by a non-linear activation function, like ReLU.
- Output Layer: Produces the final classification output, usually involving a sigmoid activation function for binary classification tasks.

Performance The model achieved an accuracy of 57%. The confusion matrix for this model is as follows:

```
0 9 3
1 6 3
  0 1
```

This indicates a basic level of effectiveness in classifying EEG data, though the model's simplicity may limit its capability to capture the complex patterns inherent in EEG signals.

3.2.2 CNN Model

Architecture The Convolutional Neural Network (CNN) model for EEG data classification includes these components:

- Convolutional Layers: Extract features from EEG images using various filters.
- Pooling Layers: Follow the convolutional layers to reduce the spatial dimension of the extracted features, typically employing max pooling.
- Fully Connected Layers: Flatten the output from the convolutional layers and finalize the classification process.

Performance The CNN model's performance was significantly better, with an accuracy of 76%. Its confusion matrix is as follows:

```
0 10 2
1 3 6
  0 1
```

The improved accuracy compared to the base model highlights its enhanced capability in feature extraction from EEG images.

3.2.3 Resnet Model

Architecture The ResNet model, particularly the ResNet50 variant, is a deep convolutional neural network known for its residual learning framework. Key features include:

- Residual Blocks: Incorporates shortcut connections that skip one or more layers.
- Batch Normalization and ReLU Activation: Utilized in each residual block.
- Global Average Pooling: Positioned before the final fully connected layer.

Performance The ResNet model exhibited an accuracy of 47% in classifying EEG data. Its confusion matrix is:

```
0 8 4
1 7 2
  0 1
```

Despite its success in general image recognition tasks, the model's performance on EEG data is not optimal, indicating challenges in adapting pre-trained image models to the temporal nature of EEG signals.

3.2.4 Autocoder Model

Architecture The Autoencoder model for EEG data includes an encoder and a decoder:

- Encoder: Compresses the EEG data into a lower-dimensional latent space. It typically consists of convolutional layers followed by pooling layers.
- Decoder: Reconstructs the EEG data from the latent space. The architecture mirrors the encoder but with transposed convolutions or upsampling techniques.

Performance The Autocoder model exhibited an accuracy of 100% in classifying EEG data. Its confusion matrix is:

```
0 20 0
1 0 15
  0 1
```

Effectiveness of Encoder-Decoder Architecture for EEG Data: The choice of an encoder-decoder architecture for EEG data analysis in detecting alcohol addiction is grounded in several key advantages:

- Efficient Data Compression: EEG data is complex and high-dimensional. The encoder compresses this data into a more manageable, condensed representation. This compression is not merely about reducing size but about distilling the essence of the data, highlighting features most relevant to detecting alcohol addiction.
- Feature Extraction and Reconstruction: The encoder part of the model excels at extracting crucial features from EEG signals, while the decoder reconstructs these signals from the compressed representation. This process ensures that the model focuses on the most significant aspects of the data, aiding in more accurate classification.
- Handling Temporal Dynamics: EEG signals have intricate temporal dynamics. The encoder-decoder framework is adept at capturing these dynamics, essential for understanding brain activity patterns related to alcohol addiction. This is particularly important as these temporal patterns might be subtle and not immediately apparent.
- Customizability and Adaptability: This architecture allows for significant customization, enabling the model to be finely tuned to the specifics of EEG data. Unlike pre-trained models, which might not align well with EEG characteristics, the encoder-decoder can be adapted to better suit the data's nature, improving the model's performance and reliability.
- Robustness in Learning Representations: The autoencoder learns to generate a robust representation of EEG data, which is crucial for accurately

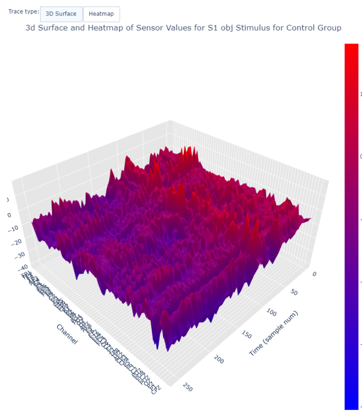
identifying patterns indicative of alcohol addiction. This robustness is key in ensuring the model's effectiveness across varied and potentially noisy real-world data

References

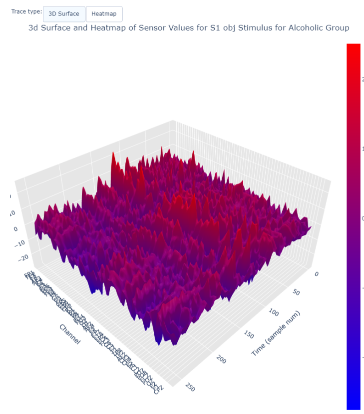
[1] UCI Machine Learning Repository. *EEG Database*. https://archive.ics.uci.edu/dataset/121/eeg+database. 2023.

[2] EEG Info. *Substance Abuse Addiction Research - Neurofeedback, EEG Biofeedback, Review papers and studies*. 2022. URL: https://www.eeginfo.com/research/addiction_main.jsp.

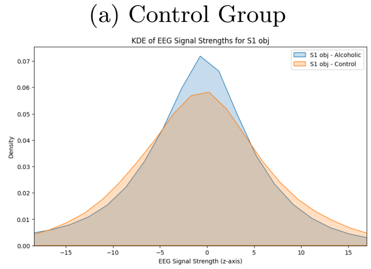
[3] Tarik S Bel-Bahar et al. "A scoping review of electroencephalographic (EEG) markers for tracking neurophysiological changes and predicting outcomes in substance use disorder treatment". In: *Frontiers in Human Neuroscience* 2022 (2022). URL: https://www.frontiersin.org/articles/10.3389/fnhum.2022.995534/full.



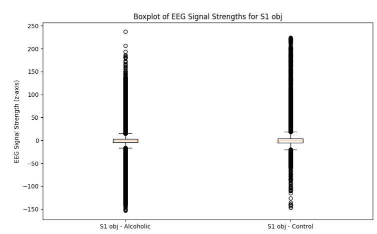
(a) Control Group



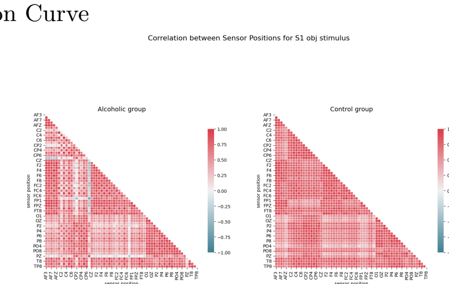
(b) Alcoholic Group



(c) Distribution Curve

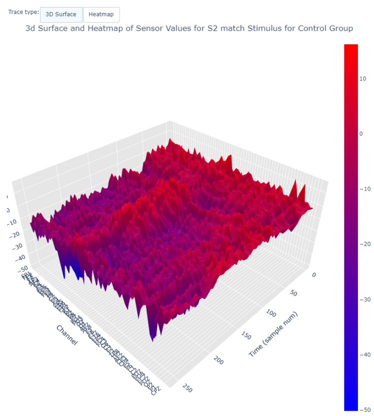


(d) Box plot

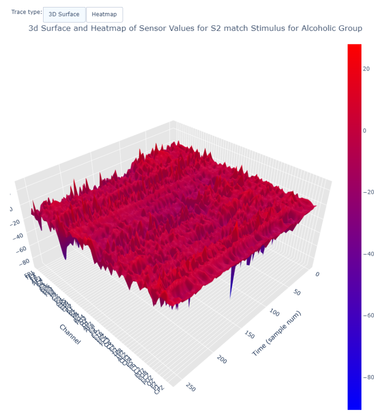


(e) Correlation between points

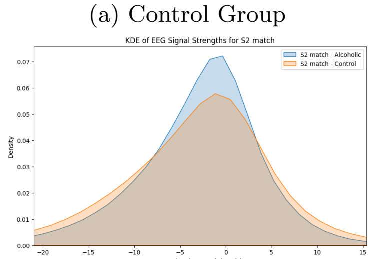
Figure 1: S1 obj: EEG data analysis



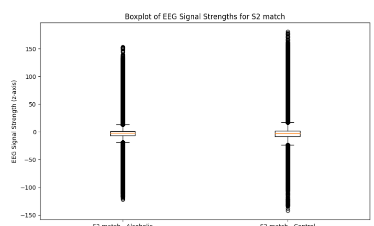
(a) Control Group



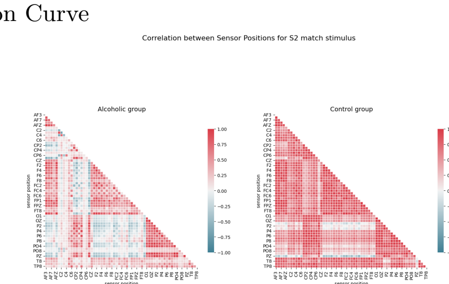
(b) Alcoholic Group



(c) Distribution Curve

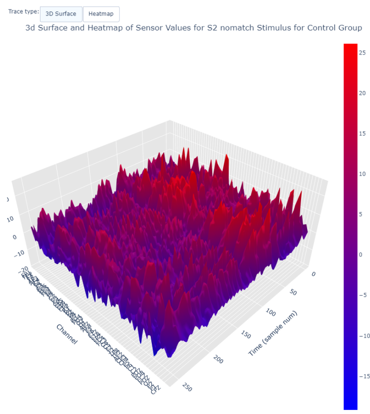


(d) Box plot

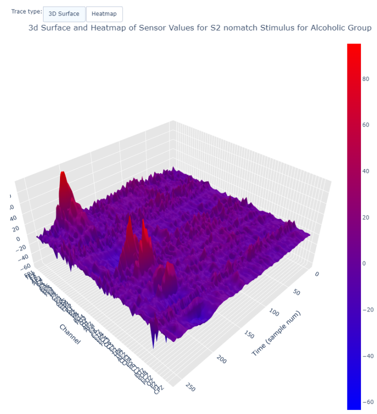


(e) Correlation between points

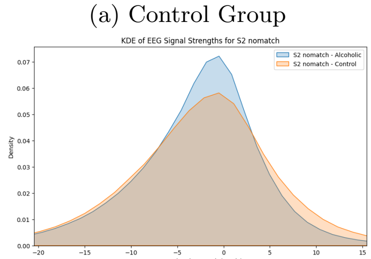
Figure 2: S2 match: EEG data analysis



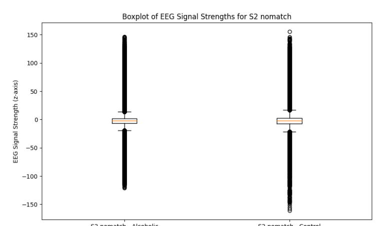
(a) Control Group



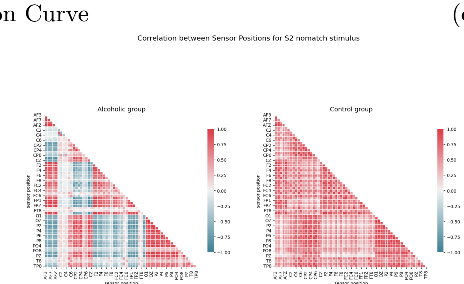
(b) Alcoholic Group



(c) Distribution Curve

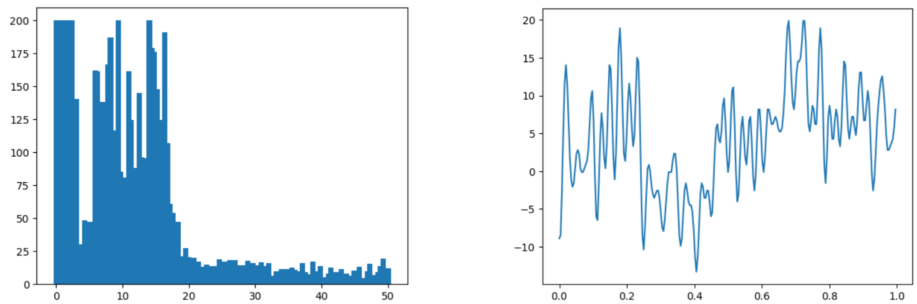


(d) Box plot



(e) Correlation between points

Figure 3: S2 no-match: EEG data analysis



(a) Signal Domain Data (b) Time Domain Data

Figure 7: Signal domain vs time domain

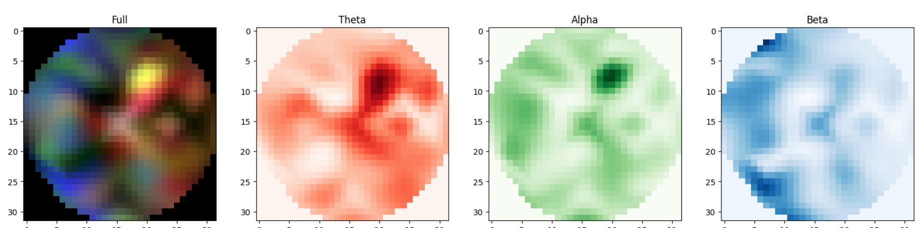


Figure 8: RGB image input

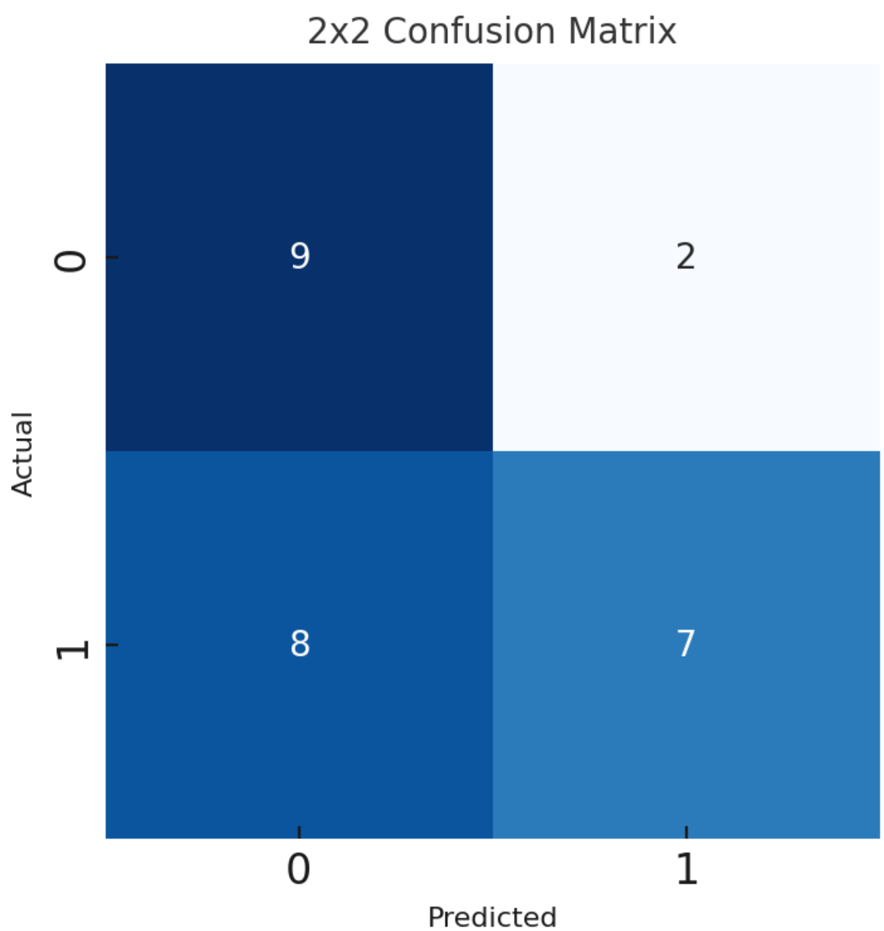


Figure 9: Confusion Matrix Example

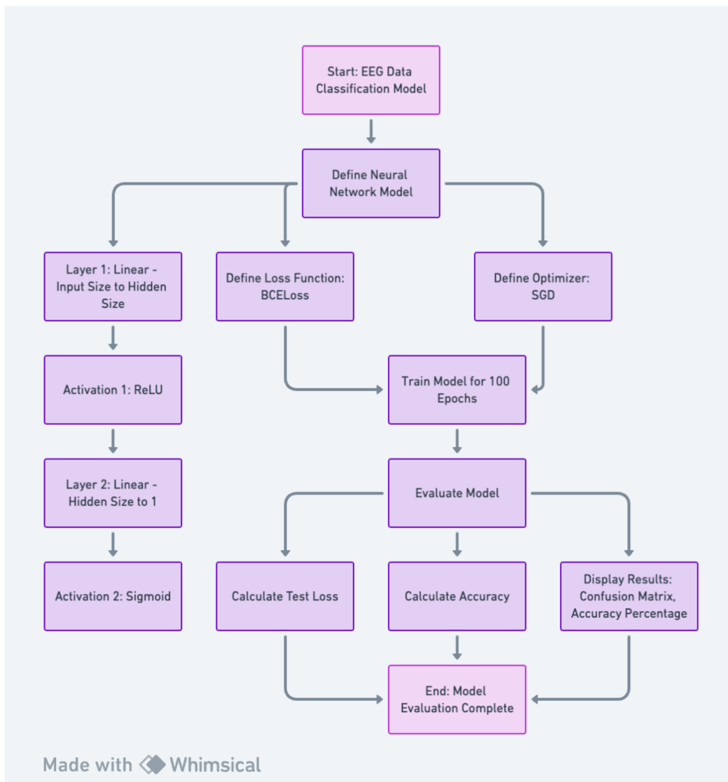


Figure 10: Base model

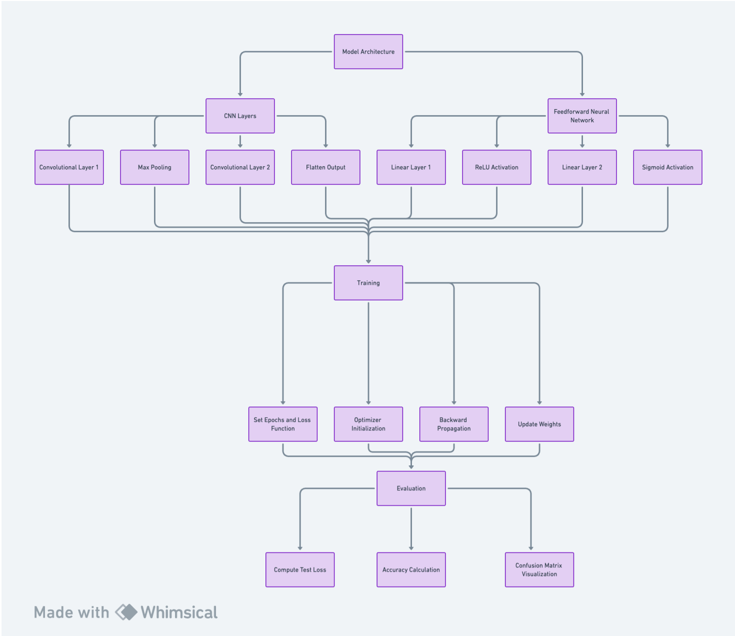


Figure 11: CNN model

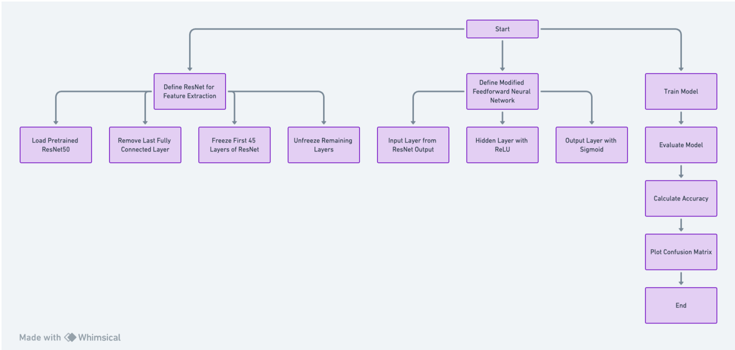


Figure 12: Resnet Model

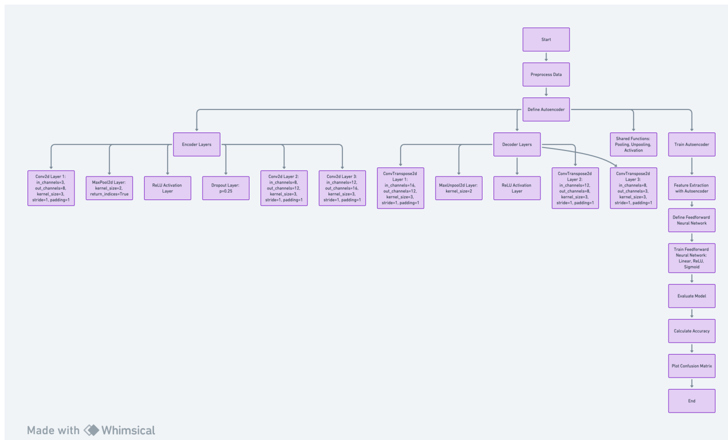


Figure 13: Autocoder Model

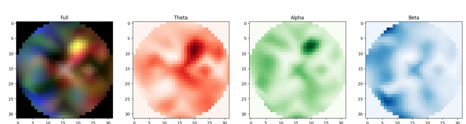


Figure 14: Autocoder