**Dataset:**

In this experiment, there were 20 participants. The EEG signal was recorded twice from each participant. Therefore, we have 40 instances of EEG. Each instance of EEG is 8 minutes (480 seconds) long.

**Data preprocessing:**

We have raw single-channel EEG data of 20 people. The main point of data collection is manual labeling. At first, we learned how the signal looks like when the person is in a normal state. So, while taking the EEG reading from a person, we observed the EEG pattern. We took three types of data which are calculation, happy, and boring because these are very much related to a person’s attention level. While a person solving an interesting math problem, his EEG pattern changed abruptly. However, this pattern doesn’t persist throughout a whole session. This change in the EEG pattern means the attention level because the tasks need positive or negative attention levels.

While recording the EEG signals in each session, we kept track of the attention periods. We recorded the starting and ending time of the attention period. The attention period and the attention level depends on the nature of the task and the surrounding environment. After several tests, we observed that the attention period normally lasted for 1-2 seconds in our experiment. We want to classify a person’s states into 4 classes. They are normal, calculative, funny, and boring states. The attention state was manually segmented from the raw EEG signal. The sampling rate of the EEG signal was 1000Hz. The segment length is 2 seconds which consists of 2000 data points. Every session (120 seconds) is segmented into 20 slots (consists of 2000 data points). These slots represent an active attention zone.

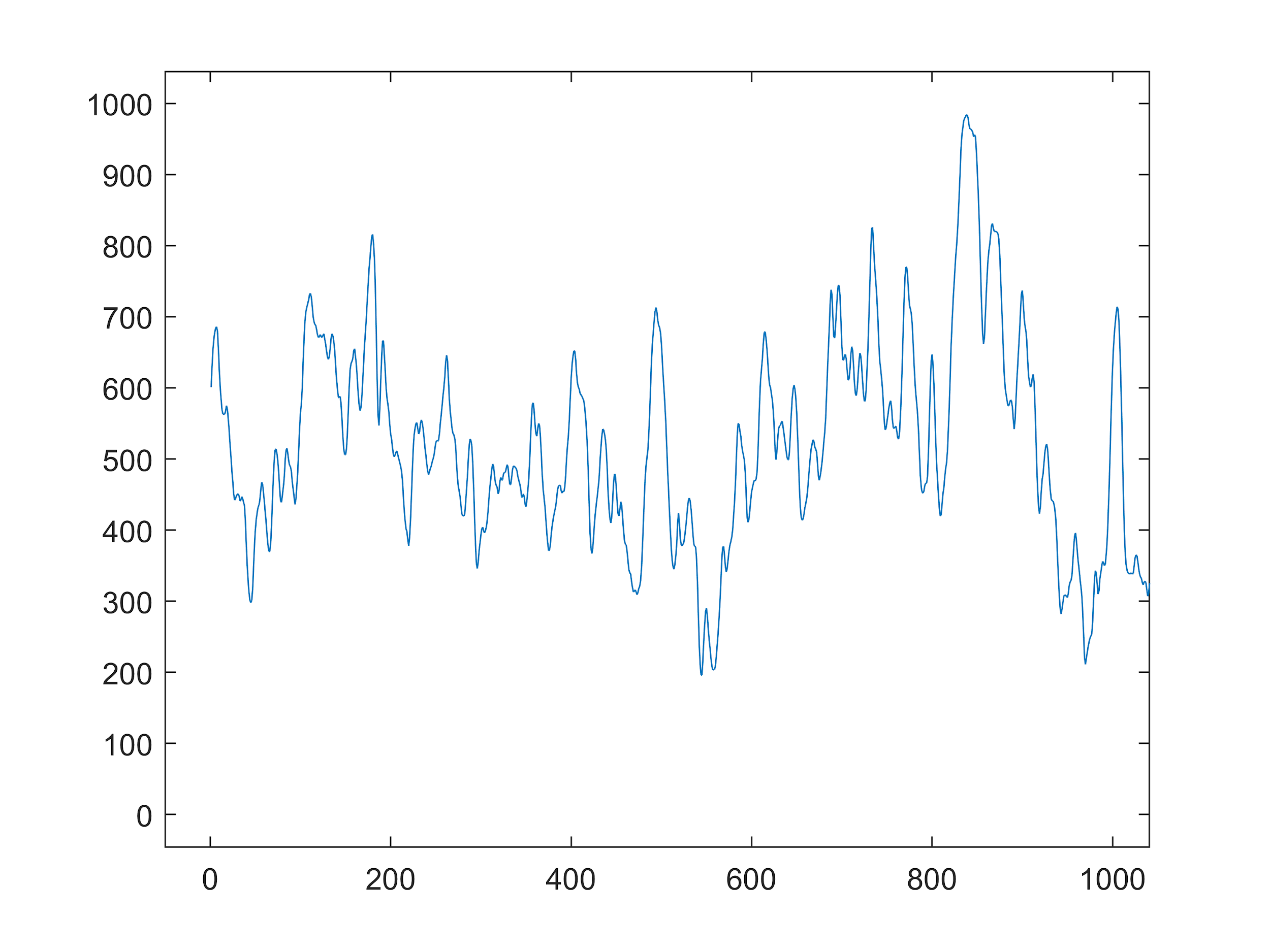


Figure: General state

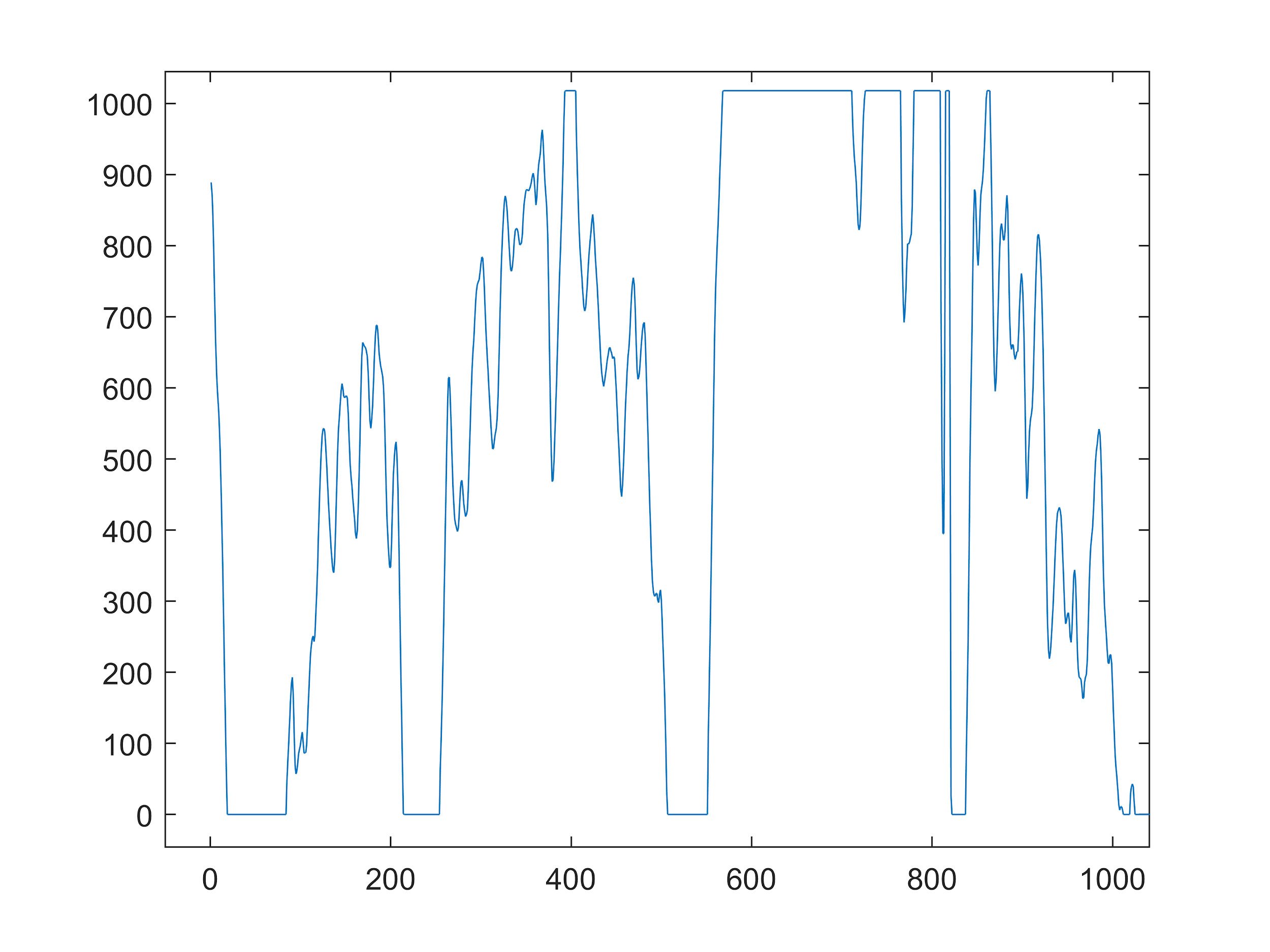


Figure: Calculation

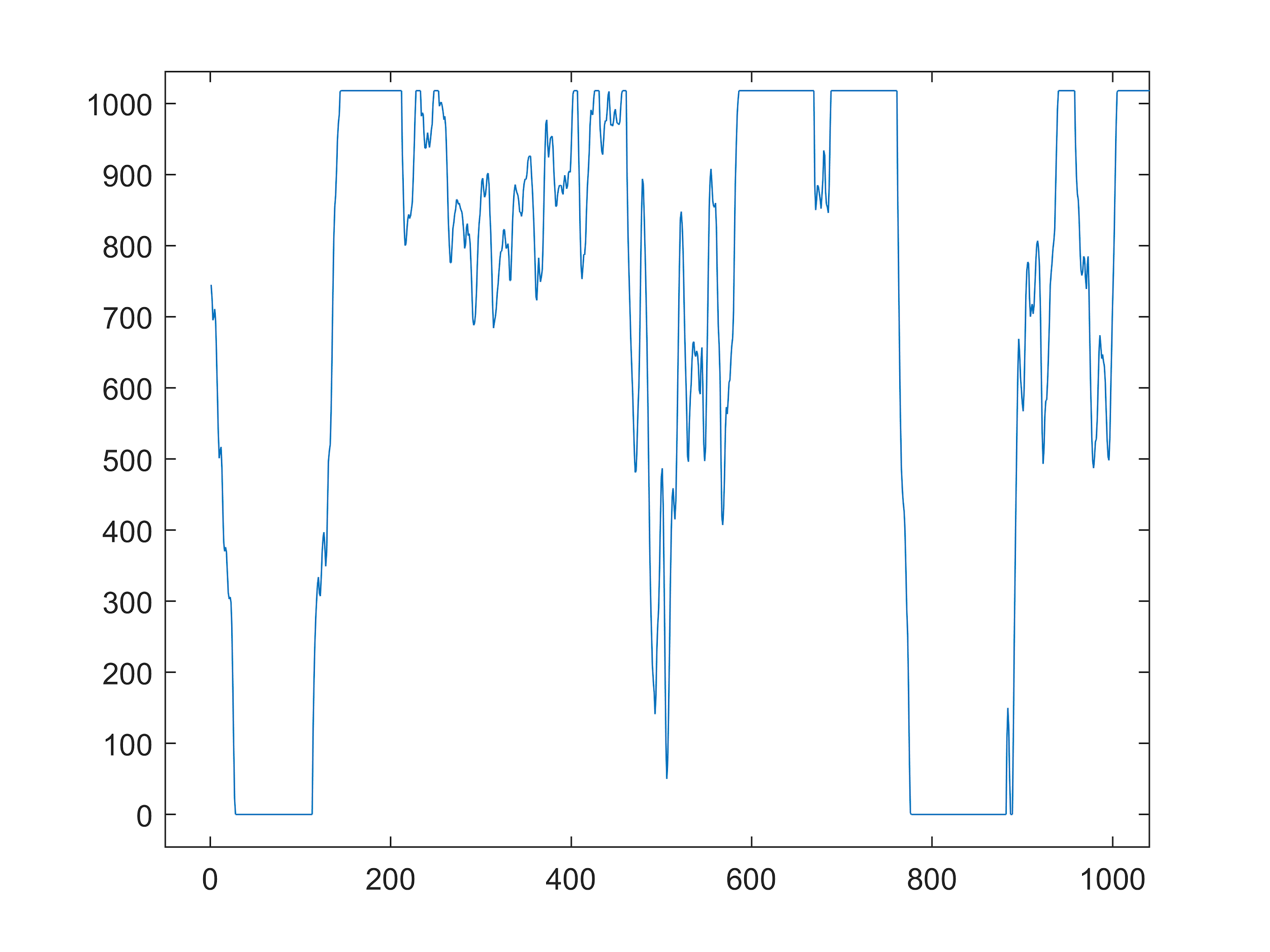


Figure: Happy

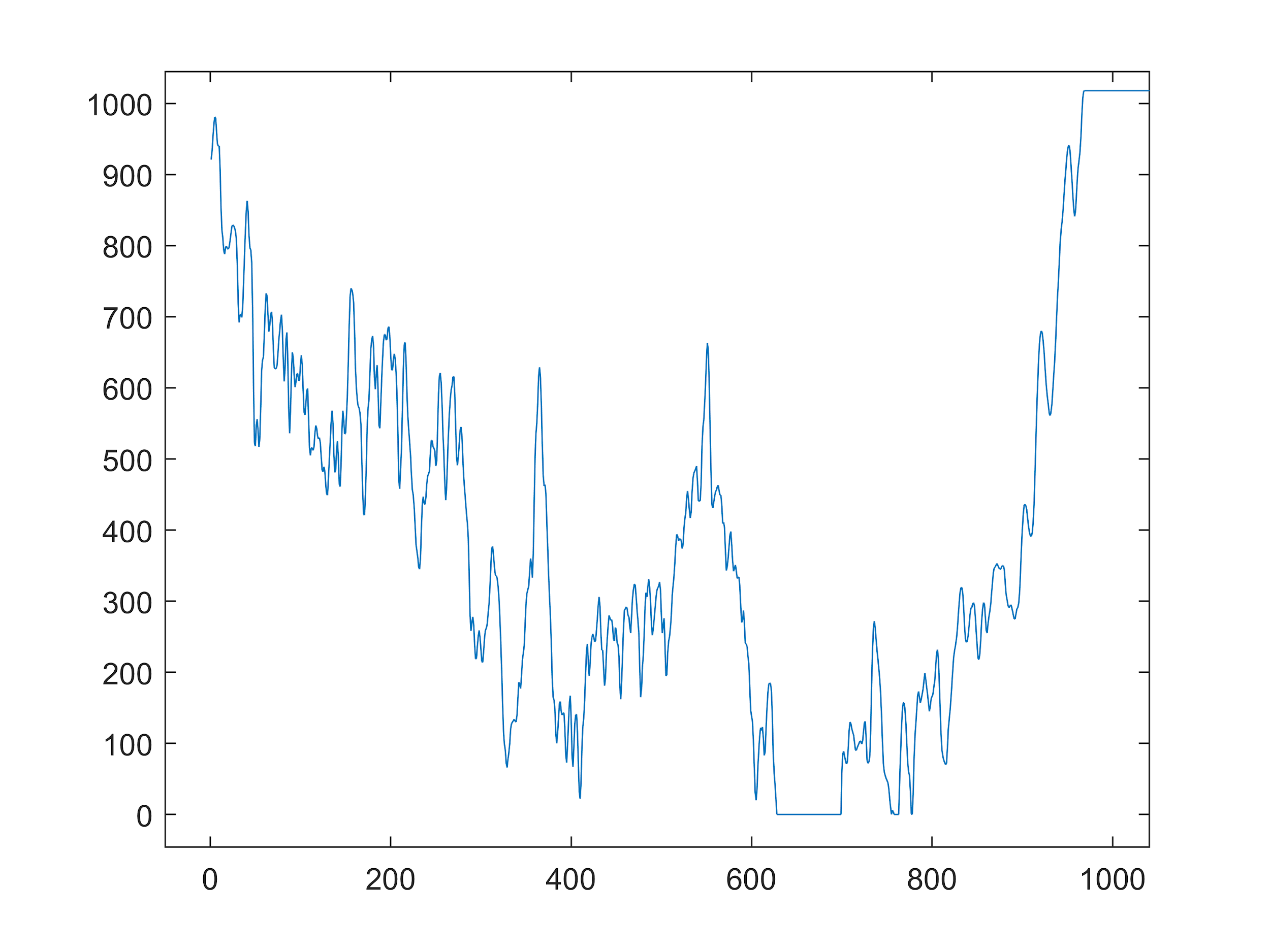


Figure: Boring

**Feature Extraction:**

There are different types of features are used to classify the EEG signals. The features can be classified into two types. One is time-domain features and the other is frequency-domain features. Time-domain features are different types of statistical features like mean, variance, power, peak to peak difference, etc. The frequency-domain features are related to the decomposition of the raw signals into sub-signals.

In our experiment, the raw EEG was decomposed into subbands using a 1-D wavelet decomposition function in Matlab. The signal was decomposed into 8 levels where each level represents different features. The detailed coefficient of levels 1 to 4 represents noise. So, these four bands were removed from the analysis. The rest of the detail coefficients of wavelet decompositions levels 5 to 8 represent gamma, beta, alpha, and theta sub-signals. The approximation coefficient of level 8 represents delta subbands. These subbands are very useful for EEG signal analysis.

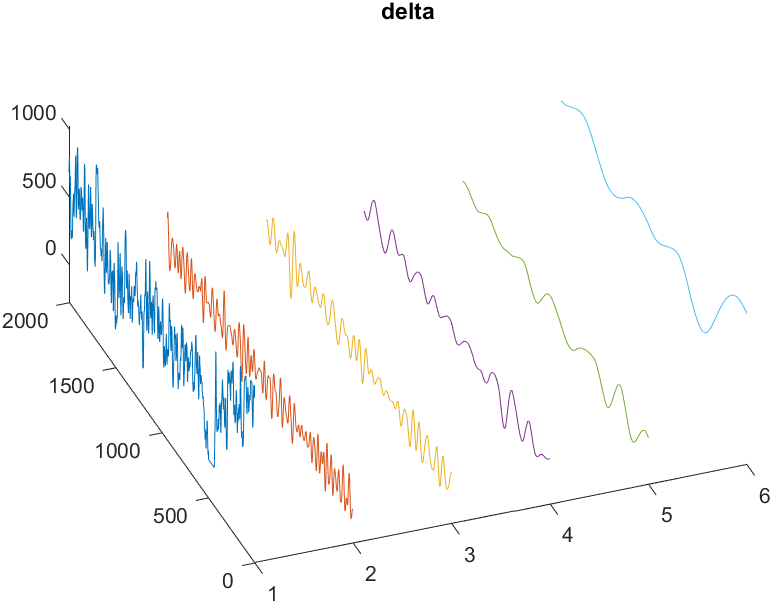


Figure: Raw EEG decomposition. 1-raw signal, 2-gamma, 3-beta, 4-alpha, 5-theta, 6-delta subband. The length of the signal is 2000 and the z-axis represents the amplitude of the signals.

Now we applied two functions on these sub-signals. One is Welch’s power spectral density estimate and another function is Spectrogram using a short-time Fourier transform.

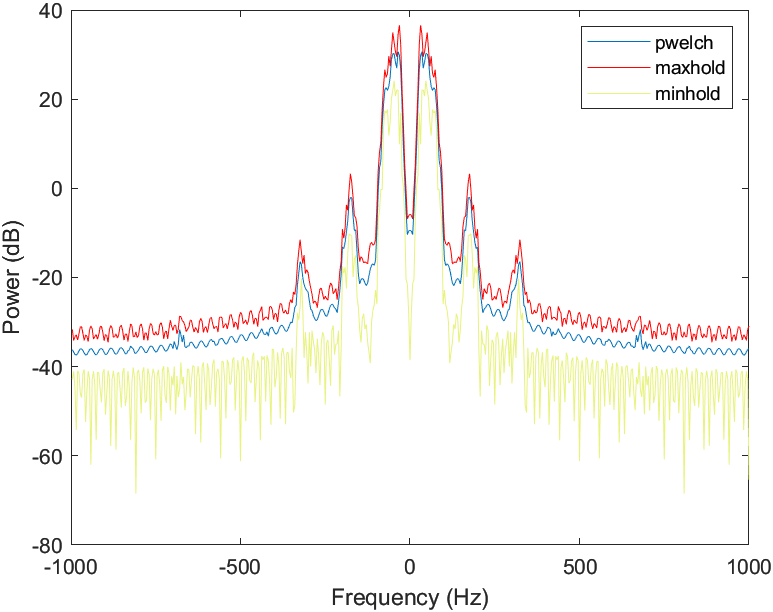


Figure: Welch’s power spectral density estimation

The length of each sequence was 2 seconds where there were 2000 data points as the sampling rate was 1000Hz. Now, a number of statistical features were calculated considering a fixed window size. The sliding window size is 100 milliseconds and the slide step size is 10 samples. Here, one sample is taken in one millisecond. Mean, Variance, Skewness, and Kurtosis were calculated from this signal using this sliding window configuration.

Welch’s power spectral density estimation of all the sub-signals are taken as input features. The four statistical values are also taken as input features. The feature vector consists of 1 raw signal, 5 sub-signals, 24 statistical values (6 signals x 4 statistical formula), 6 Welch’s power spectral density estimation of each signal. In total there are 48 values in the feature vector. All these values are numerical in nature.

Spectrogram of a signal represents the Power distribution in both the time-frequency domain. The visual representation of the spectrogram is very important for signal analysis. Therefore, the spectrogram function was applied to 6 signals. Among them, the spectrogram of raw, beta, and alpha signal was selected for analysis. This selection was based on empirical reasons and simulation of the deep learning model.

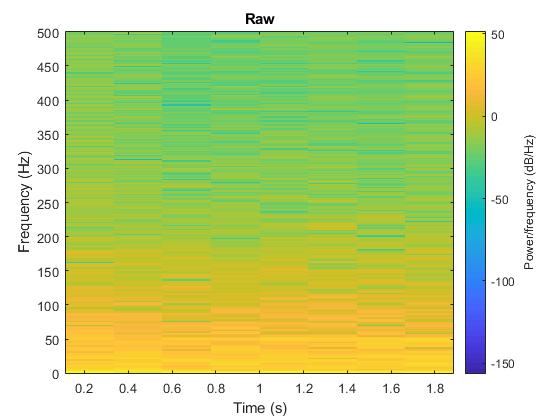


Figure: Spectrogram of Raw signal

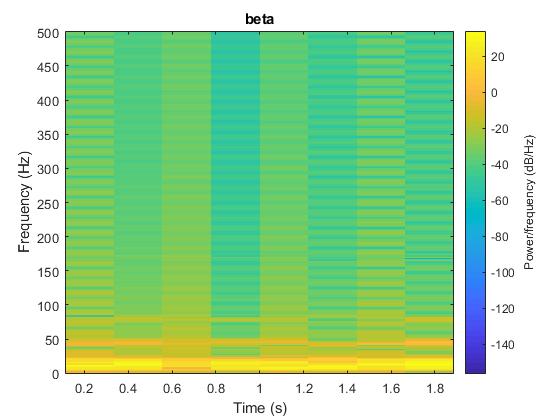


Figure: Spectrogram of Beta signal

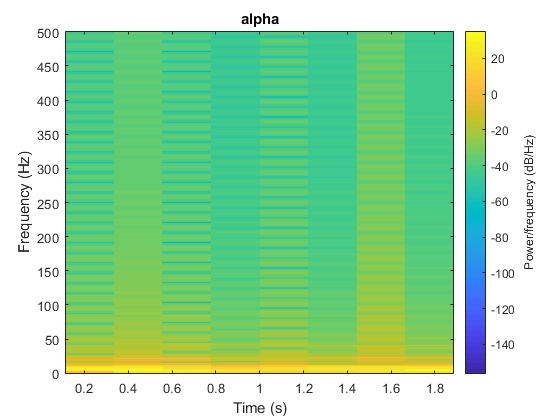


Figure: Spectrogram of the alpha signal

**Model:**

For having two types of input data, we build a hybrid deep learning model for better accuracy. The LSTM is used for sequence classification. The dataset consists of a sequences and images. Each sequence is 2000 points long and has 48 features. So, the input dimension of LSTM is (40, 2000, 48) where 40 represents the instances of EEG.

The visual dataset consists of Spectrogram of raw EEG, beta and alpha signal. All the Spectrogram images were converted into 64x64x3 dimensions where 64x64 represents the height and width and 3 represents the RGB layers.

Three CNN input layers take spectrograms of raw, beta, and alpha signal. And one LSTM input layer takes the sequential input. The output layer from each sub-models in concatenated into one layer followed by dense layers.

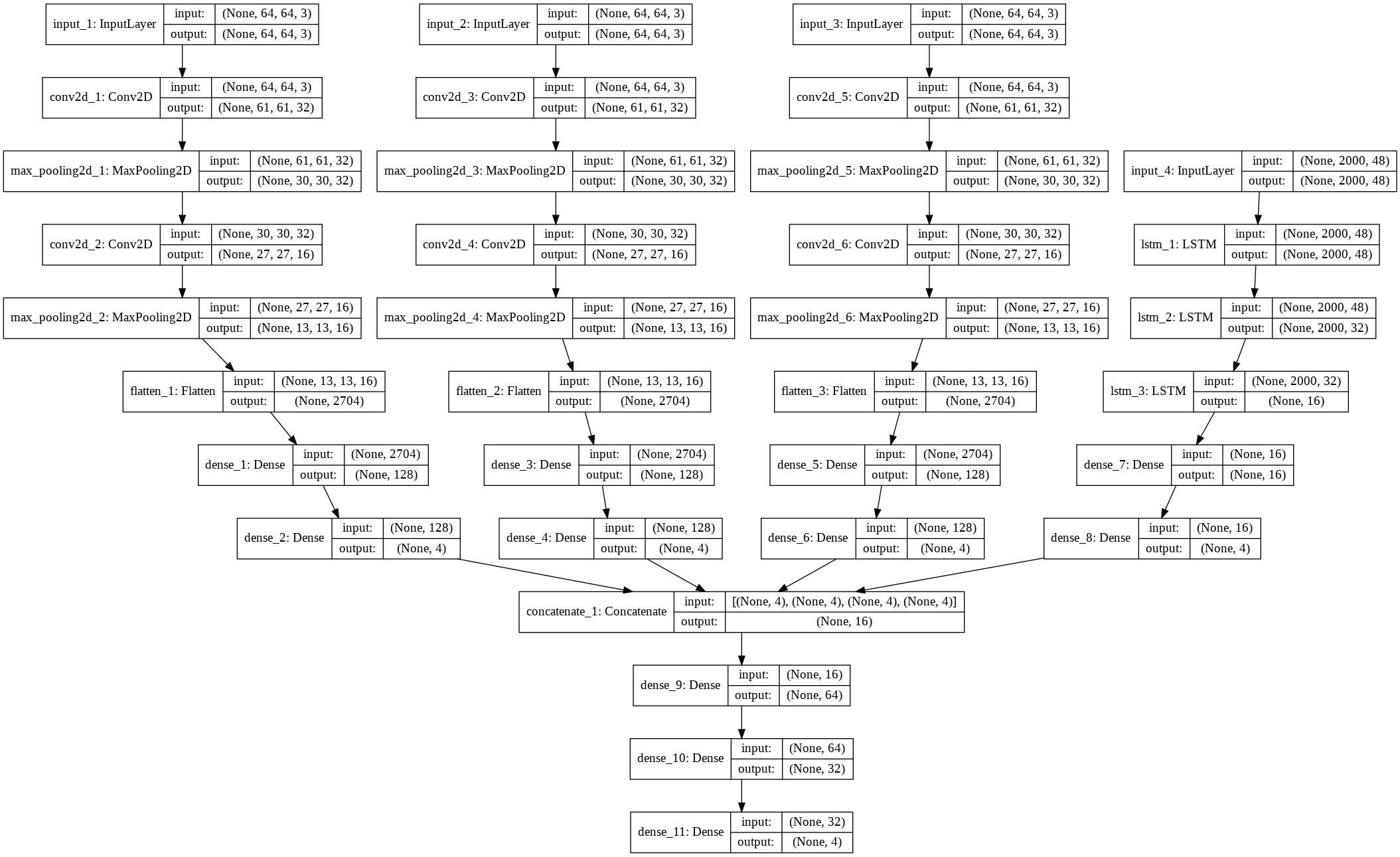


Figure: Hybrid deep learning model.