Emotion Detection from EEG Signals

Mukhilshankar Umashankar, Lyuxing He

Final Project

CS 545: Machine Learning for Signal Processing

GitRepo: https://github.com/lyuxinghe/M2S.git

Introduction

Electroencephalogram (EEG) signals are rich in information concerning the mental state of an individual. The signals are obtained by attaching electrodes to the brain and recording the electrical impulses to depict the brain state. This data can be analyzed to decipher the emotional state of the individual. Current HCI systems are deficient in understanding the user's emotions, limiting their ability to provide personalized services.

A thorough understanding of the correlation between a user's emotional response and musical dimensions, like genre, beat, and lyrics, can facilitate personalized music suggestions, aid in music therapy, and potentially help tackle sleep disorders. Our ultimate goal is to establish such a correlation for the applications mentioned. The first step towards that goal is to build a classifier that would detect emotions from EEG signals.

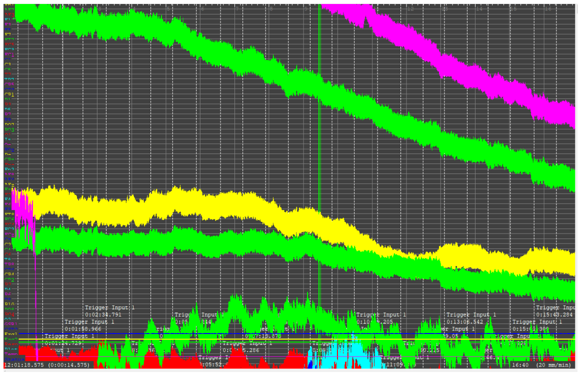

Figure 1: Sample snippet of EEG signal with selected channel measurements.

Details of the Approach

Dataset description

We used an EEG dataset that was developed as part of the eNTERFACE 2006 workshop[1] . The dataset used images from the International Affective Picture System (IAPS) to evoke positive, negative, or calm emotional responses in the participants. There were a total of 420 EEG signal recordings with its corresponding invoked emotions labeled. Each EEG signal recording in this dataset contained the signal captured by 72 electrodes (channels) attached to the scalp of the participant. The sampling rate varied across the recordings, so we downsamples the relevant signals to obtain a uniform rate of 256 Hz.

Preprocessing

The raw EEG datasets in nature are contaminated by various sources of noise: environmental noise, muscle activity, eye blinking, etc. Our specific dataset was further contaminated by fNIRS noise as the dataset collected EEG and fNIRS signals in parallel. The fNIRS probe obstructed the smooth functioning of the EEG data capture. Figure 2 shows the raw signal of F3 channel from the EEG signal recording of Participant 2 Session 3.

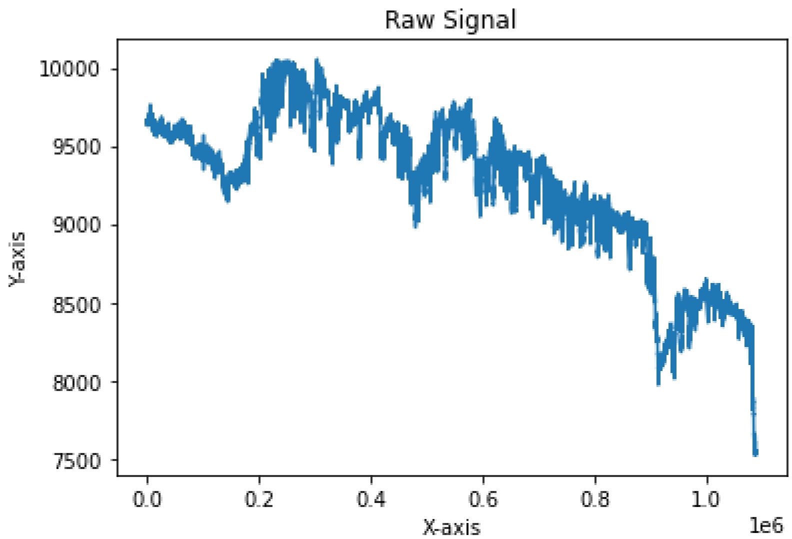

Figure 2: Raw signal of Participant 2 Session 3 EEG signal recording F3 channel

To remove the various sources of noise, the following techniques were adopted:

1. Bandpass filter
   Bandpass filter is a simple yet extremely useful technique for EEG signals preprocessing. The frequency intervals of interest in EEG are the sub bands, θ (4-8Hz), α (8-12Hz), β (12-30 Hz), and γ (30-45Hz). Therefore a passband range of 4-45Hz is sufficient for our bandpass filter to initially preprocess the signals. This frequency range also removed the notorious line noise that peaks at 50Hz [2]. The line noise is a result of ambient electrical current in the room, or from electrical devices in the vicinity of the sensors. Across the channels in our dataset, the line noise was significantly present. The frequency range of interest, 4-45Hz, was further bandpass filtered to obtain the sub band signals used for extracting features. Figure 3 shows the bandpass filtered sample signal. As we can see, the signal has been processed with detrending, but there are still strong peaks at random time stamps representing noises residing between 4-45 HZ.

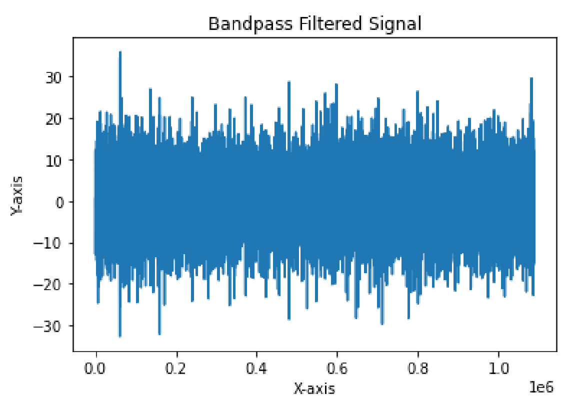

Figure 3: Bandpass filtered sample signal

2. Common Average Referencing[3]
   The EEG signals obtained measure the electrical potential difference between two points. The common electrode used in reference is the ground electrode, which picks up electrical noise. The noise shows up in the voltage difference as it does not get passed on to the channel measurements. A solution to eliminate this noise is to re-reference once the data is obtained. The new reference could be any one of the channels, resulting in the other channels' magnitude being computed in relation to it. However, this leads to channels spatially close to the reference channel to have very low amplitudes as the electrical activity between them will be similar, while channels further away have much larger electrical potential differences. Hence, a better reference point is the mean signal across all the channels as it will be equidistant from all electrical activity in the scalp. This was the reference point used in this project.
3. ICA
   Other artifacts like muscle activity and fNIRS noise require more sophisticated approaches to remove. ICA is the fundamental yet effective approach to take on, since the clean signals are relatively independent to noises like random fNIRS noises and muscle activities. For illustration purposes, the sample signal processed by the bandpass filter illustrated above is further decomposed using ICA, and the major components are shown in Figure 4-7. It is clear that component 1 and 2 are the fNIRS noises originated from light source instability or shot noise that demonstrates spike patterns, whereas component 0 contains muscle activities including large jaw movements or continuous eye ball movements when tracking the visual stimuli. We manually zero out the noise components, and the reconstruction gives ICA component 3 as the cleaned signal shown in Figure 7.

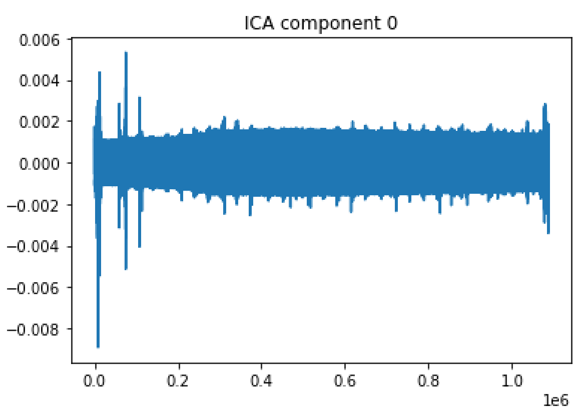

Figure 4: ICA decomposition component 0

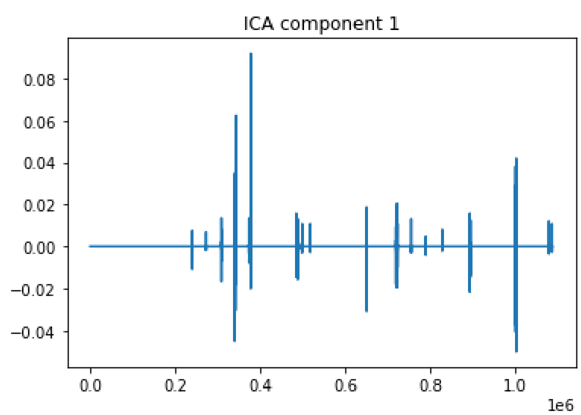

Figure 5: ICA decomposition component 1

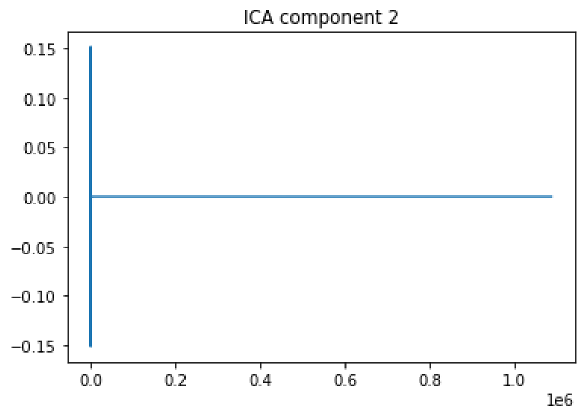

Figure 6: ICA decomposition component 2

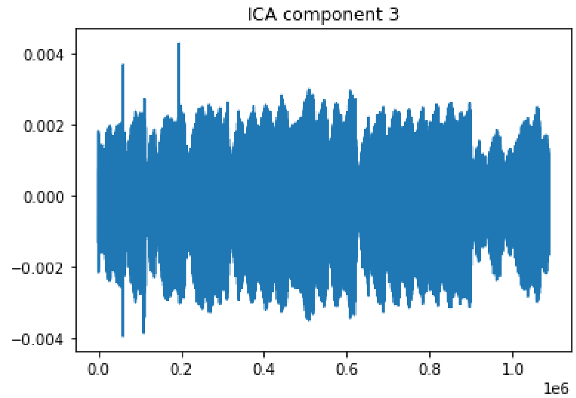

Figure 7: ICA decomposition component 3

Feature extraction

The next step in building the classifier was to extract features from our preprocessed signals. The features were computed per channel of interest. Following previous work we explored several different channels of interest, but finally chose an implementation with 32 channels[4][5]. The features extracted are listed in Table 1.

Table 1: Extracted Features

Feature Type	Features	No. of features
EEG Power Features	Average band power for each frequency band (θ (4-8Hz), α (8-12Hz), β (12-30 Hz), and γ (30-45Hz))	4 x 32 = 128
EEG Power Difference	Difference in average band power, in each frequency band, for 14 channels pairs between right and left scalp	4 x 14 = 56
Statistical Time Domain Features	Mean, Variance, Zero-crossing rate of preprocessed interest channels	3 x 32 = 96
Hjorth Parameters	Hjorth Mobility and Complexity parameters of interest channels	2 x 32 = 64
Fractal Dimension	Higuchi's fractal dimension of interest channels	32

Welch's periodogram was used to obtain the power spectral density function that was then used to compute the average band power by finding the area under the curve for the appropriate frequency band domain[6]. The EEG power difference between symmetrical pairs of electrodes was used as a feature to capture any asymmetrical electrical activity in the scalp corresponding to emotional experiences[4]. The Hjorth mobility parameter is an estimate of the mean frequency, while the Hjorth complexity parameter captures the similarity in shape of the signal with a pure sine wave, thereby giving an estimate of the bandwidth of the signal. The fractal dimension is a measure of the complexity of the signal. Fractal dimensions are tougher to interpret, however, they have witnessed success in classifying EEG signals into emotions in previous works [7], so they were incorporated as a feature.

The feature extraction process results in a 376 length feature vector for each EEG recording that is then passed into classifiers for learning.

Deep Learning Approach
We also explored the prospect of using deep learning to build our classifier. Given our very limited knowledge of the domain, we figured we might need deep learning to identify features by itself to reach a decent accuracy. To this end, we implemented a Multilayer Perceptron, an LSTM and an artificial neural network from scratch. We trained them on the FFT of 90 EEG signal recordings preprocessed in the time domain.

Results and Discussion

Evaluation of feature extraction
K-nearest neighbors, Random Forests and an artificial neural network were the classifiers used to classify the feature vector into emotions. The surprising result was that each of the classifiers performed very well on training data, reaching accuracies as high as 95%, however, they performed abysmally on test data: accuracies between 33-35% (random classification amongst the three classes). The results did not change with differing literature-backed sets of interest channels, differing classifier parameters (epochs, learning rate, number of neighbors, hidden nodes, activation layers, tree depth etc.), or a different set of features chosen. These results suggest that the models were overfitting and did not learn a generic representation for the emotions.

This indicates a couple of things. The features chosen may not have good enough discriminative power in differentiating between positive, negative, and calm emotional responses in EEG signals. This is a bit likely given that previous works have used a combination of these features to reach accuracies much higher than random classification. A second takeaway could be that the dataset we used was largely confounded with noise, requiring even more deliberate efforts in preprocessing or using another dataset to develop the classifier. This reason is more likely as on top of the various sources of noise present in any EEG signal, this dataset also had disturbances from the fNIRS probe. Moreover, the dataset was collected as early as 2006 with no significant literature work using it.

Evaluation of deep learning approach

For our MLP implementation, the best performance we achieved is using a hidden size of 128 with a 2-layer architecture. For our implementation of LSTM, the best performance we achieved is using a hidden size of 512 and a stack of 4 LSTMs. For our customized neural networks, we used a 6-layer pyramid neural network architecture, with the increasing hidden size to 1024 to help create increasingly abstract representations of the data, dropout layers with dropout rate of 0.5 to avoid overfitting, and activation layers using leakyRELU to get around with dead gradients. We chose cross entropy loss during the optimization step, as it performs well for classification tasks. All learning rates are set to 1e-5. It turns out that the deep networks were able to achieve better performance than the classical machine learning techniques with training accuracy breaking 90%, but evaluation accuracy for MLP and customized neural network ranges between 20% to 60%, which differs drastically from the training accuracy and is highly unstable. The LSTM model is relatively more reliable, with an averaged evaluation accuracy of around 52% with relatively smaller variation around the mean. The confusion matrix of trained LSTM is shown in Figure 8.

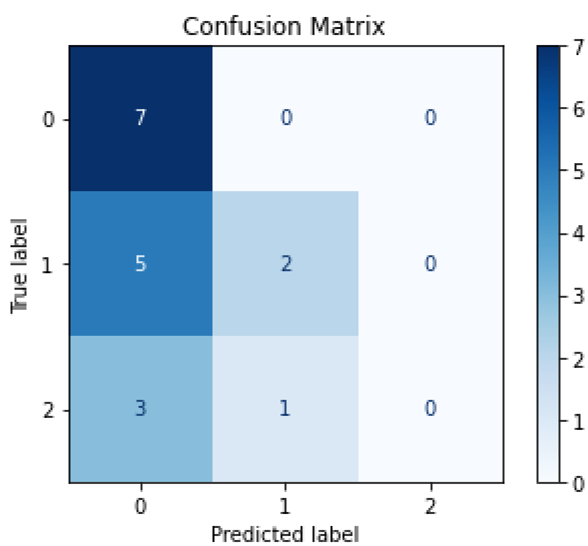

Figure 8: Confusion matrix of LSTM trained on eNTERFACE 2006

Other dataset approach

The eNTERFACE 2006 dataset we've been working with presented unexpected challenges from various sources that we cannot pinpoint, and therefore we've tested our system on another dataset EEG Brainwave Dataset: Feeling Emotions. The dataset was collected from two people (1 male, 1 female) for 3 minutes per state - positive, neutral, negative [8]. We used a Muse EEG headband which recorded the TP9, AF7, AF8 and TP10 EEG placements via dry electrodes. Six minutes of resting neutral data is also recorded after the visual stimuli [8]. This dataset is relatively newer than eNTERFACE 2006 dataset and has been processed with statistical extraction resampled. Using our proposed system, we've achieved a 96.48% accuracy on test data, about 2% below the benchmark.

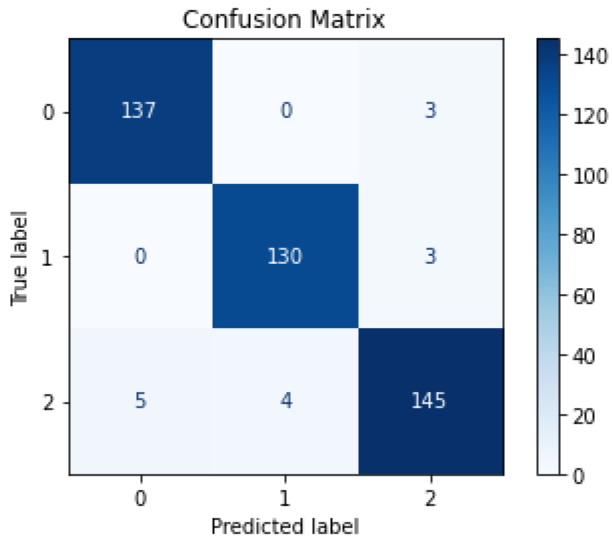

Figure 9: Confusion matrix of LSTM trained on EEG Brainwave Dataset: Feeling Emotions.

References

[1] SAVRAN, Arman et al. Emotion Detection in the Loop from Brain Signals and Facial Images. In: Proceedings of the eNTERFACE 2006 Workshop. Dubrovnik (Croatia). [s.l.] : [s.n.], 2006.

[2] Grummett, T. S., Fitzgibbon, S. P., Lewis, T. W., DeLosAngeles, D., Whitham, E. M., Pope, K. J., & Willoughby, J. O. (2014). Constitutive spectral EEG peaks in the gamma range: Suppressed by sleep, reduced by mental activity and resistant to sensory stimulation. Frontiers in Human Neuroscience, 8.

[3] Products, B. (2022, February 14). Choosing your reference & why it matters - A brainvision analyzer tip. Brain Products Press Release.

[4] Koelstra, S., Muhl, C., Soleymani, M., Jong-Seok Lee, Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2012). DEAP: A database for emotion analysis ;using physiological signals. IEEE Transactions on Affective Computing, 3(1), 18–31.

[5] Kalashami, M. P., Pedram, M. M., & Sadr, H. (2022). EEG feature extraction and data augmentation in emotion recognition. Computational Intelligence and Neuroscience, 2022, 1–16. https://doi.org/10.1155/2022/7028517

[6] Raphael Vallat, P. (n.d.). Raphael Vallat, Phd. Bandpower of an EEG signal. https://raphaelvallat.com/bandpower.html

[7] Ackermann, P., Kohlschein, C., Bitsch, J. A., Wehrle, K., & Jeschke, S. (2016). EEG-based automatic emotion recognition: Feature extraction, selection and classification methods. 2016 IEEE 18th International Conference on E-Health Networking, Applications and Services (Healthcom). https://doi.org/10.1109/healthcom.2016.7749447

[8] Bird, Jordan J. “EEG Brainwave Dataset: Feeling Emotions.” Kaggle, 19 Dec. 2018,

[9] (N.d.). Figure 6: Artifacts Removal of EEG Signals. https://doi.org/10.7717/peerj.13096/fig-6

[10] Artifacts. (n.d.). https://www.learningeeg.com/artifacts