Classification of EEG Signals Using Self Reported Stress Values

Bereket Guta

January 18, 2022

Introduction

The classification of human emotion is important for emotional interactions between humans and computer. Due to the increased availability of electroencephalography (EEG) recording devices, EEG-based emotion recognition has received considerable interest [1]. Previous work has shown that EEG signals correlate with arousal and valence scores based on user self assessment [6]. Much of the the focus has been on classification of tasks such as motor functions [8], motor imagery [5] and classification of simple emotional states [9]. Studies generally use two main feature generation techniques, manual feature engineering ([8], [6], [5], [9]) or deep learning based methods ([7], [4]). Based on the success of deep learning in natural language processing, new methods have created models based on Convolutional Neural Network (CNN) Transformer architectures [4]. The goal of this project is to use deep learning based techniques to train a classifier that accurately predicts stress levels from a given EEG input.

Dataset

The data used in the set of experiments was gathered from the work of [2]. The dataset consisted of 64-channel EEG data of 15 undergrad students as they played a video game. In a post game interview, the participants were asked to rate their stress levels during gameplay as they rewatched their gameplay footage. The dataset was prepared for the classification task by normalizing the EEG and stress labels between [0,1].

Feature Generation

The task of classifying the emotional state of a subject based solely on their EEG signal requires the generation of features with emotional insight. While the raw unfiltered features could be used, they are often noisy and lead to poor classification performance [3]. In this project, the latent features from the autoencoder were used as features for classification since it was hypothesized that an autoencoder that could generate accurate reconstruction would be able to create information rich features in its latent space.

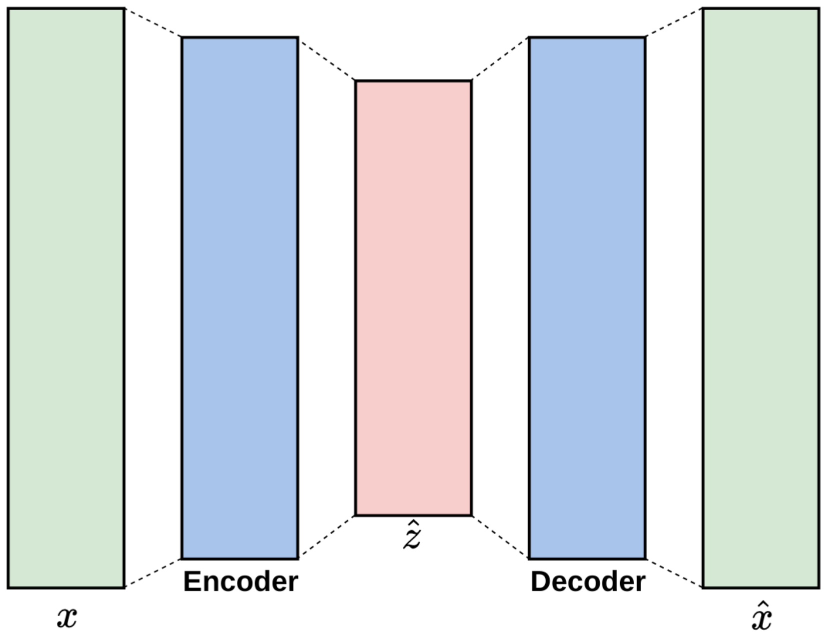

Figure 1: Autoencoder structure with fully connected encoder and decoder layers.

To test this hypothesis, features were generated through the use of an autoencoder. The model generated a latent representation through a bottleneck training scheme. The input to the model was a single time slice of the EEG signal which consisted of 64 channels. The model structure shown in figure 1. The number of features in the latent space $\hat{z}$ and other parameters were varied and the findings are shown in table 1. Several other models were tested including a CNN and LSTM based autoencoder. The CNN model treated the given time window as an image with a width of 64 and generated an encoding by passing this image through the network. The LSTM on the other hand, used the time window and generated an encoded version. Both of these models performed poorly during initial trials as they were unable to recreate the original signal, as a result the focus became the classical model.

The best reconstruction performance was achieved when a bidirectional low pass filter set at 4 Hz was applied to to the input signal. Due to the noisy nature of the original signal, the autoencoder was learned a better representation when the input signal was smoothed, yet table 1 shows that the model can still learn the unfiltered signal. The frequency spectrum of the EEG input was analyzed and it was found to have the highest power in the low frequency range and a noise floor that decreased linearly with frequency as seen in figure 2. The filtering technique is non-causal, as a result would be difficult to apply to real time applications.

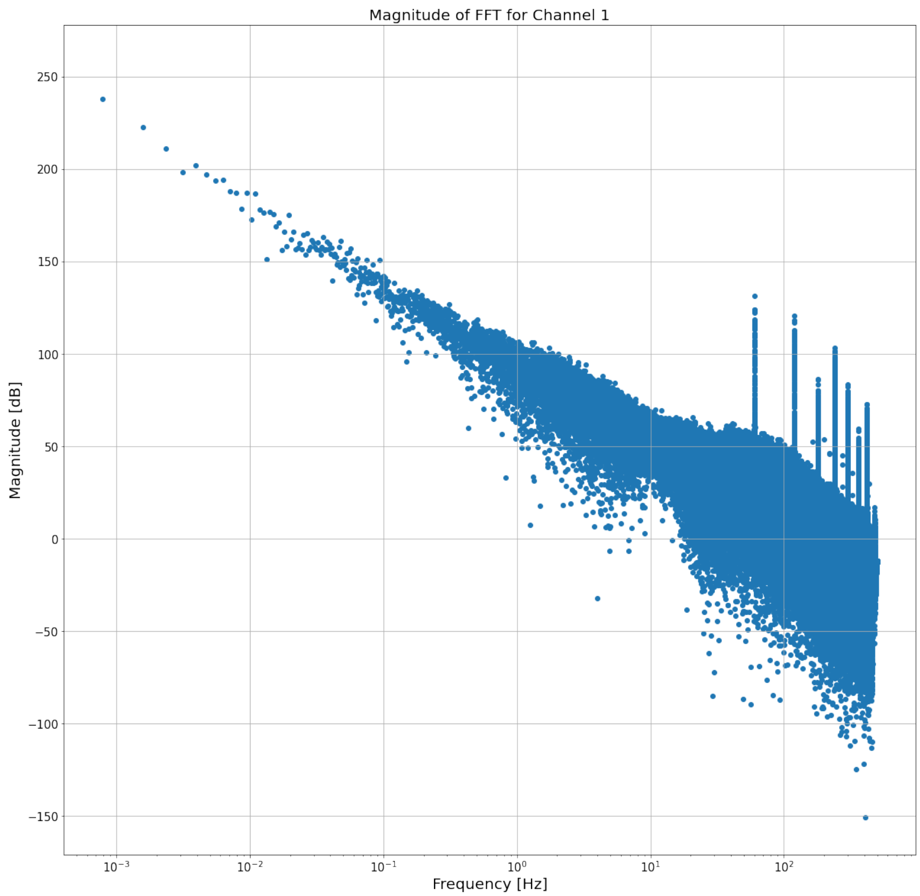

Figure 2: Fast fourier transform of the EEG signal in channel 1 for participant number 9. The spectral power decreases with frequency.

Classification

A Long short-term memory (LSTM) based classifier was used to classify the output of the encoder. A novel stacked implementation shown in figure 10 was used. The structure of the classifier was experimented with and results are presented in table 2 for different classification tasks. The normalized stress labels were partitioned into equivalent intervals to produce the classes for classification. Results for this classification are shown in table 2. In addition to stress value classification, stress transitions were classified using a novel labeling technique that partitioned the labels into decreasing,

Time Window	Encoded Features	Subject numbers	EEG Input	Mean Square Error
100 ms	[8]	[9]	Low Pass at 4 Hz	0.00023
100 ms	[8]	[2,9,12]	Low Pass at 4 Hz	0.00037
100 ms	[8]	[2,4,9,12,14]	Low Pass at 4 Hz	0.00054
100 ms	[8]	[9]	Raw signal	0.00023
100 ms	[8]	[2,4,9,12,14]	Raw signal	0.00060
100 ms	[16]	[2,4,9,12,14]	Raw signal	0.00056
100 ms	[12]	[2,3,5,6,7,8,10,11,13,14]	Low Pass at 4 Hz	0.00668

Table 1: Results for autoencoder performance using different parameters.

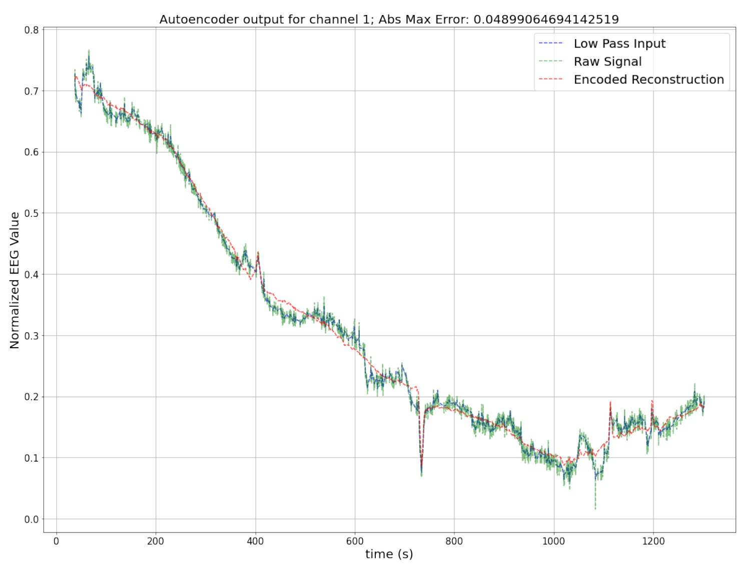

Figure 3: Trained output for subject 9 with 8 features in the encoder and a filtered input.

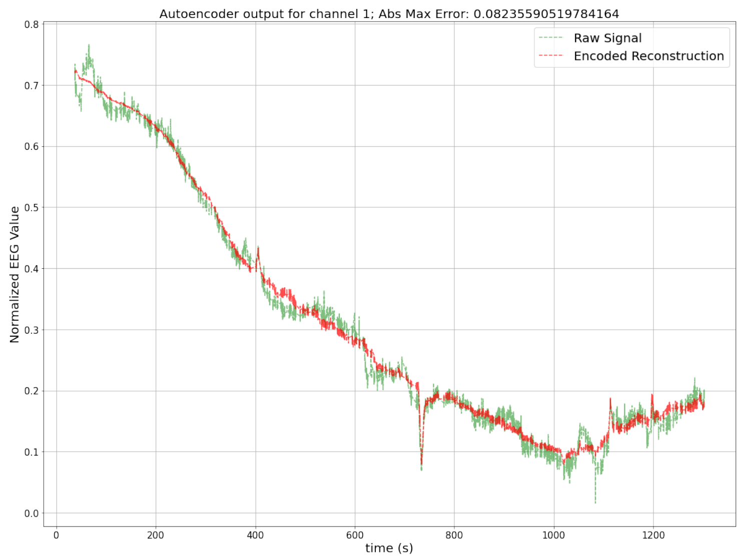

Figure 4: Trained output for subject 9 with 8 features in the encoder and an unfiltered input.

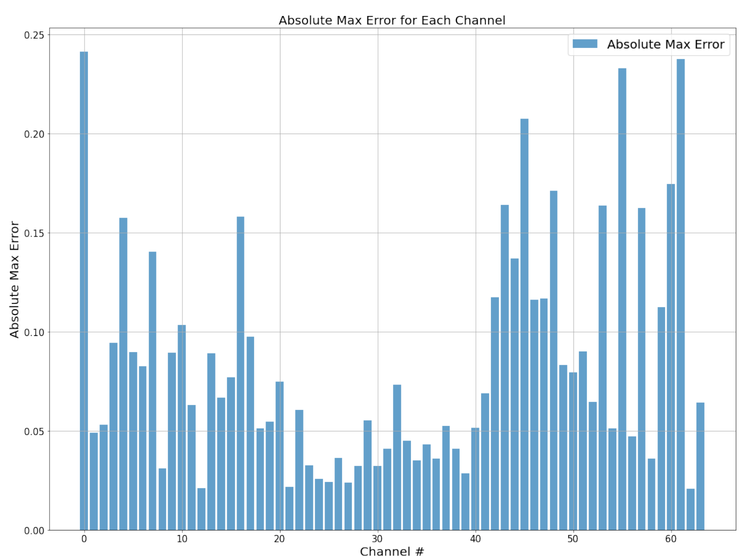

Figure 5: Absolute max error for each channel for subject 9 with 8 features in the encoder and a filtered input.

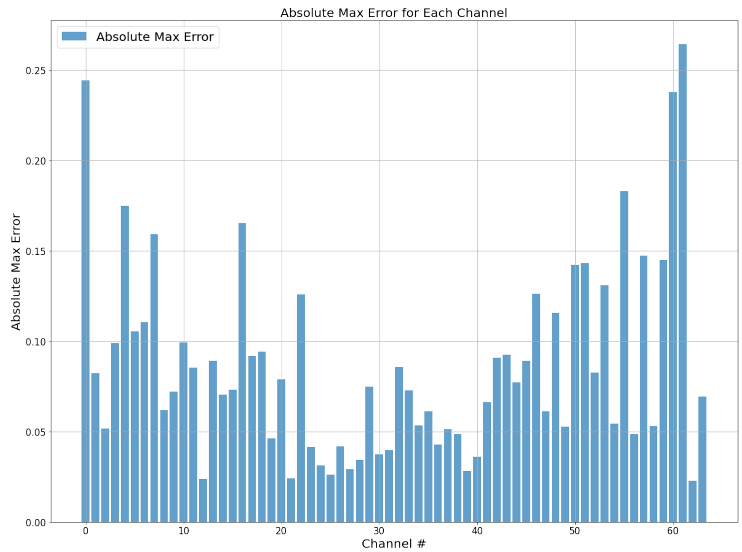

Figure 6: Absolute max error for each channel for subject 9 with 8 features in the encoder and an unfiltered input.

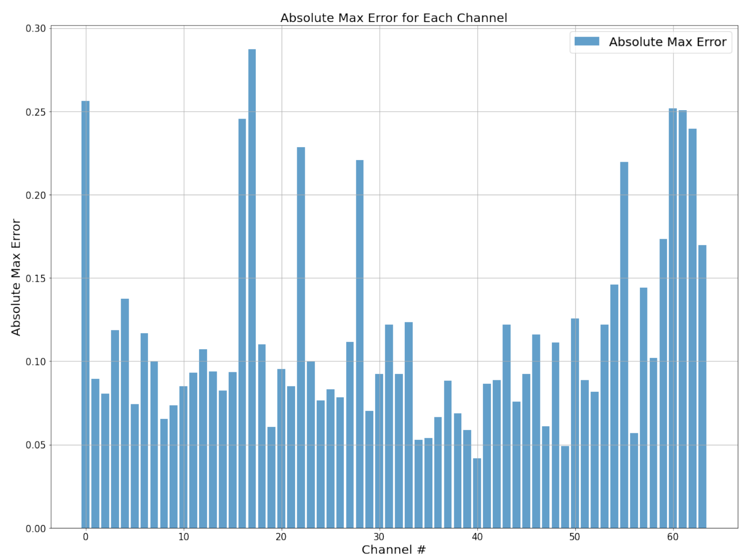

Figure 7: Absolute max error for each channel for combined subjects [2,9,12] with 8 features in the encoder and a filtered input.

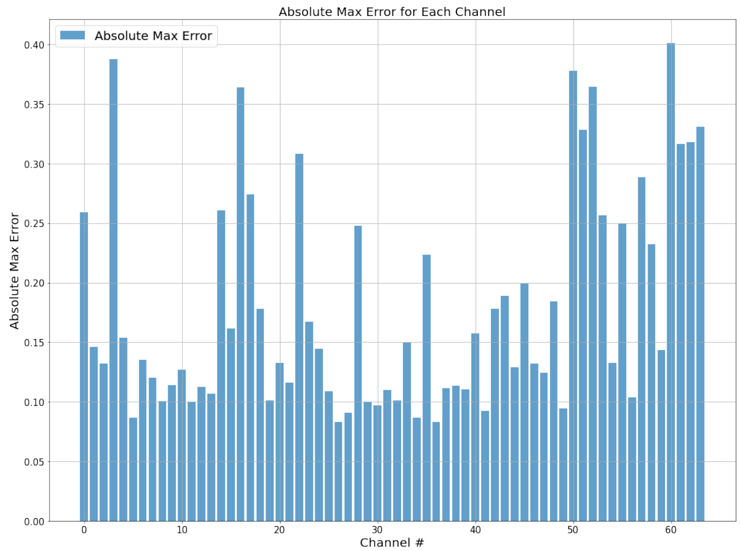

Figure 8: Absolute max error for each channel for combined subjects [2,4,9,12,14] with 8 features in the encoder and a filtered input.

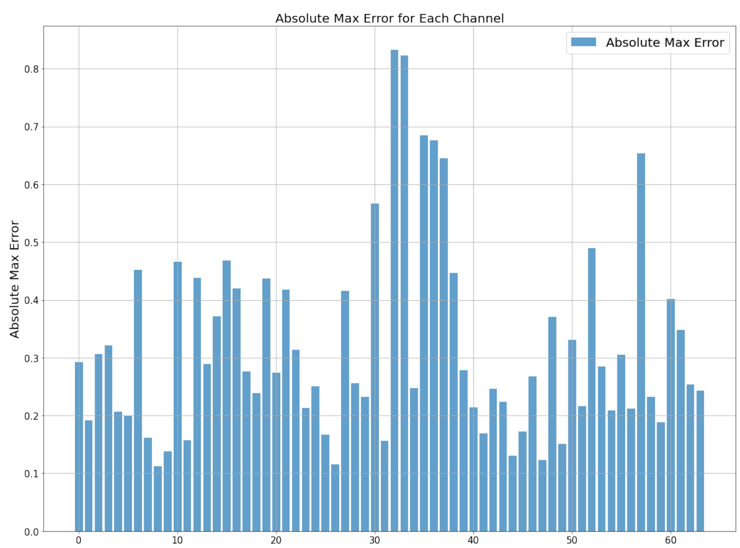

Figure 9: Absolute max error for each channel for combined subjects [2,4,9,12,14] with 8 features in the encoder and an unfiltered input.

increasing and not changing using the slope in a time window. The initial results were poor and heavily depended on the chosen window size.

Time Window	Hidden Size	Subject numbers	Feature input	F1-Score
100 ms	[32]	[9]	Encoded Low Pass 8 features	0.31 - Emotion Transition - 3 Classes
100 ms	[32]	[9]	Encoded Low Pass 8 features	0.61 - Emotion Labels - 3 Classes
100 ms	[32]	[0,1,2,3,5,7,8,9,13,14]	Raw 64 Channels	0.66 - Emotion Labels - 5 Classes
100 ms	[32]	[0,1,2,5,6,7,9,11,14,15]	Raw 64 Channels	0.682 - Emotion Labels - 5 Classes
100 ms	[32]	[0,1,2,3,4,5,9,10,11,15]	Raw 64 Channels	0.784 - Emotion Labels - 3 Classes
100 ms	[32]	[0,1,2,5,9,11,12,13,14,15]	Encoded no filter 12 features	0.604 - Emotion Labels - 5 Classes
100 ms	[32]	[2,3,5,6,7,8,10,11,13,14]	Encoded Low Pass 12 features	0.469 - Emotion Labels - 5 Classes

Table 2: Results for transition and state classification for various parameters.

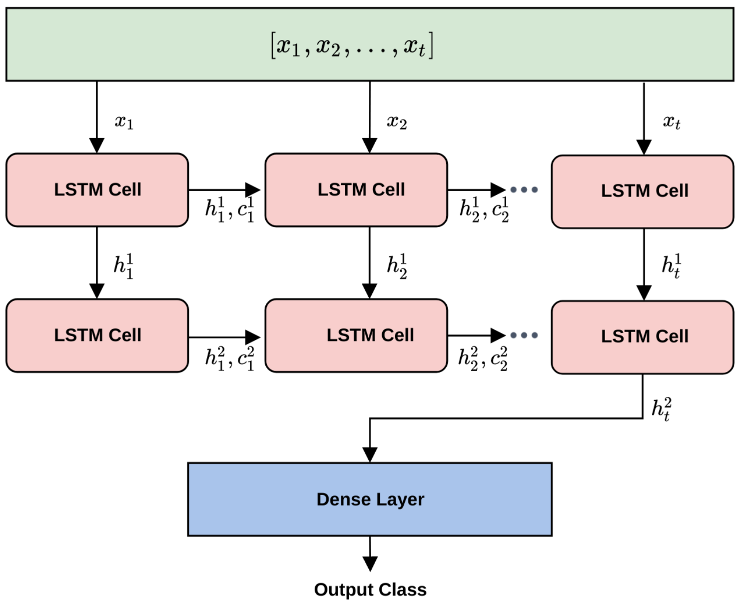

Figure 10: The Long short-term memory classifier model that takes in the EEG features as input and generates the predicted class.

Discussion

The metric used to assess the autoencoder was the mean square error (MSE) of the reconstruction and the original input. This value was on the order of 0.0001 for models trained on a single participant but increased by a factor of 10 when trained on 5 individuals and a factor of 100 when trained on 100 individuals. To fully assess the performance of the autoencoder, the classifier must be taken into account. There was an imbalance in the number of labels as individuals chose the values near 0.5 at a higher rate than the extremes. The dataset was balanced by oversampling the under

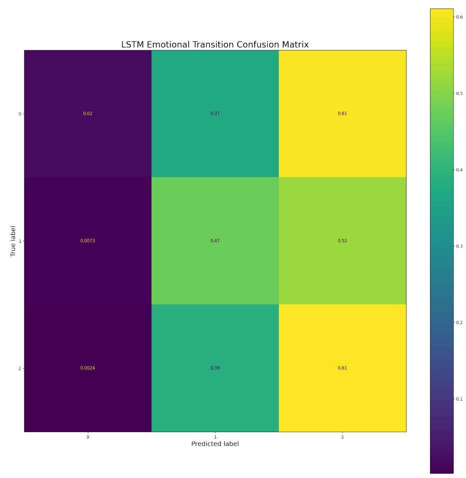

Figure 11: 3 class emotional transitions confusion matrix for participant 9 with 8 features in the encoded input and a low pass filter applied to the EEG signal. Class 0 is decreasing, class 0 is staying constant and class 1 is increasing in each time window of 100 ms.

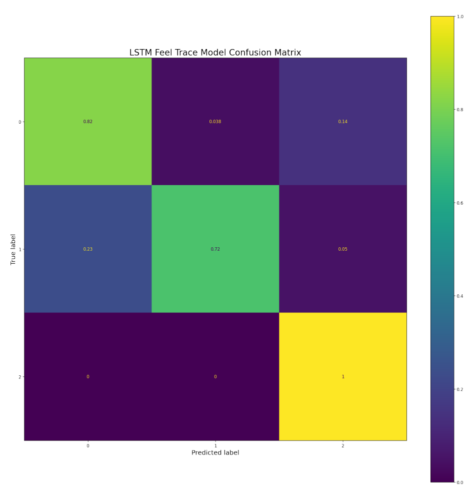

Figure 12: 3 class emotional label confusion matrix for participant 9 with 8 features in the encoded input and a low pass filter applied to the EEG signal.

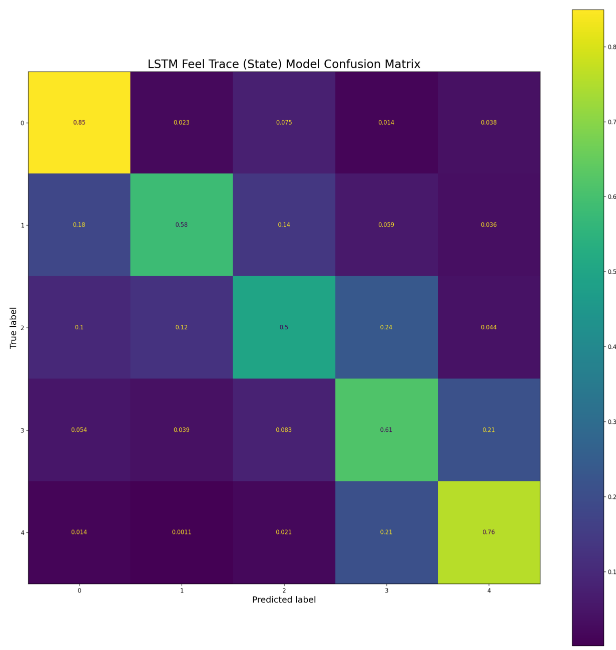

Figure 13: 5 class emotional label confusion matrix for participant 10 with 8 features in the encoded input and 64 channel raw feature input.

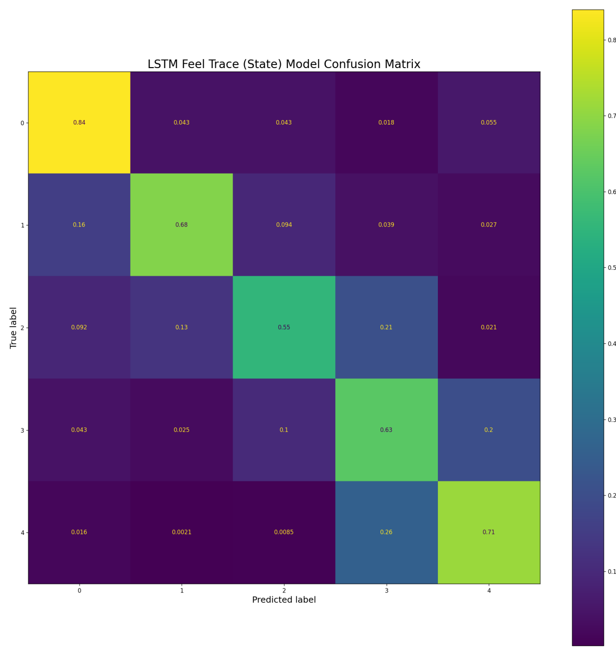

Figure 14: 5 class emotional label confusion matrix for participant 10 with 8 features in the encoded input and 64 channel raw feature input.

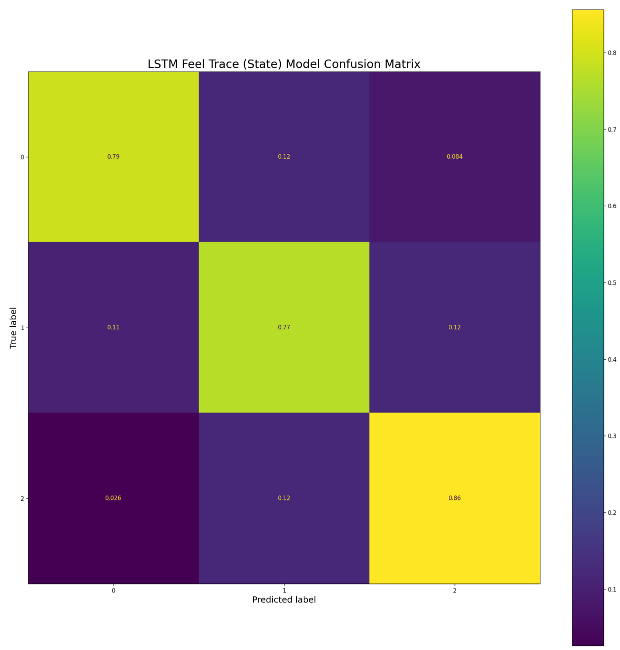

Figure 15: 3 class emotional label confusion matrix for participant 10 with 8 features in the encoded input and 64 channel raw feature input.

represented classes. Training with fixed time windows of 500ms and 5 classes on single individuals, F1-Scores were higher for features that came from the autoencoder then the raw EEG signal. Running the model on multiple individuals generated mixed results. The autoencoder did not perform significantly better than the raw features and when more data resulted in lower F1-Scores. Unexpectedly, the filtered features performed poorly as the number of data points increased. These indicates that an LSTM based classifier may be a suitable candidate for real time classification of EEG signals capable of generalizing to multiple subjects.

Conclusion

The autoencoder and LSTM model structure showed promising performance on an individual subject classification level but the encoded features were unable to generalize when it came to classification of multiple participants. Classifying stress transitions did not provide statistically significant results, however, the LSTM stress state classifier performed best with the raw unfiltered 64 features obtaining an average F1-Score of 0.66. This model has the potential to be deployed for real-time applications as it can generalize to multiple subjects.

References

[1] S. M. Alarcão and M. J. Fonseca. Emotions recognition using eeg signals: A survey. *IEEE Trans. Affect. Comput.*, 10(3):374–393, jul 2019.

[2] X. Cang, Q. Feng, and K. MacLean. What's in a name? issues in triangulating self-reported emotion to establish trustable labels. In *CHI '21 Workshop on Momentary Emotion Elicitation and Capture*.

[3] Z. Iscan, Z. Dokur, and T. Demiralp. Classification of electroencephalogram signals with combined time and frequency features. *Expert Systems with Applications*, 38(8):10499–10505, 2011.

[4] D. Kostas, S. Aroca-Ouellette, and F. Rudzicz. BENDR: using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data. *CoRR*, abs/2101.12037, 2021.

[5] S. Kumar, A. Sharma, and T. Tsunoda. Brain wave classification using long short-term memory network based optical predictor. sci rep 9: 9153, 2019.

[6] L. Martinez-Tejada, A. Puertas-González, N. Yoshimura, and Y. Koike. Exploring eeg characteristics to identify emotional reactions under videogame scenarios. *Brain Sciences*, 11:378, 03 2021.

[7] S. Mousavi, F. Afghah, and U. R. Acharya. Sleepeegnet: Automated sleep stage scoring with sequence to sequence deep learning approach. *PLOS ONE*, 14(5):1–15, 05 2019.

[8] G. Zhang, V. Davoodnia, A. Sepas-Moghaddam, Y. Zhang, and S. A. Etemad. Classification of hand movements from EEG using a deep attention-based LSTM network. *CoRR*, abs/1908.02252, 2019.

[9] W.-L. Zheng and B.-L. Lu. Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development*, 7(3):162–175, 2015.