EEG-based Emotion Recognition with Music: A Model and Application

Zakariyya Scavotto

Abstract

With the growth of music streaming, both for pleasure and other applications, such as music therapy, being able to understand how music makes someone feel has increased in importance. The goal of this study was twofold: first, create a machine learning model to predict a subject's emotional response to music; then integrate this trained model into an application that can predict someone's emotional response based on live data. Using support vector machines (SVMs) as the basis of the machine learning model, a model was trained to recognize the correct emotional response with 64% accuracy, and the model was successfully implemented into a demonstration web application.

Keywords: EEG, Music, SVM, Web Application

Introduction

Listening to music is something that many if not all people will do throughout their lives. With the growth of music streaming platforms, such as Spotify and iTunes, it has become easier than ever before to listen to whatever music we want, all with the press of a button. This music can elicit a strong emotional response in a person (Schrer & Zenter, 2001). Some songs can elicit more positive emotions like happiness, while others can bring about more negative emotions like anger. An important note is that the same song may cause a different emotional response in different people. For example, one song may make a person happy while it makes another person sad. Given this fact, it would be useful to interpret how a particular song makes someone feel.

Accurately interpreting someone's emotional response to music can be useful in a variety of applications. One example is improving the recommendations that music streaming platforms give to their users for suggested songs to listen to. The platform may have a basis of songs and music genres to start with based on the user's taste in music and prior songs they listened to, but they may not want to listen to a particular song that they generally like at a particular moment. For example, if their significant other just broke up with them, they may not want to listen to the upbeat songs that they usually listen to, they may want to listen to a more sedate song. Being able to interpret how a person is responding to the music they are currently listening to would enable music platforms to improve their recommendations since they could not only take into account a person's taste in music but also how they are feeling.

Another potential application of knowing someone's emotional response to music is in the field of music therapy. Music therapy is the utilization of music as a tool to help people who are having a hard time dealing with mental health problems, such as stress and anxiety (American Music Therapy Association, 2005). If a music therapist could see how a particular song or genre of music is affecting a patient's emotions in real-time, they could better tailor their treatment to that patient.

To further our understanding and technical capacity, I developed a machine learning model based on support vector machines (SVMs) that predicts a person's emotional response to music based on the entropy and energy of electroencephalography (EEG) signals. I then created an application that can demonstrate the model working in near real-time.

Background

EEG Technology

EEG allows for observation of a subject's brain activity by measuring the strength of the electrical signals that are being fired in the person's brain's neurons at the moment of recording. These signals are recorded utilizing a series of EEG electrodes placed on the subject's scalp, which are held by an EEG headset.

These signals can subsequently be filtered and processed to find the different frequency bands of the signal. These frequency bands are Delta, Theta, Alpha, Beta, and Gamma. Each of these frequency bands can reveal a certain behavior or feeling that the subject was feeling at that time. The ranges of these frequency bands and an example behavior associated with each band can be found in Table 1.

Table 1

EEG Frequency Band Ranges (Hz) and Example Behaviors

Name	Frequency Range (Hz)	Example Behavior
Delta	0-4	Deep sleep
Theta	4-7	Drowsiness
Alpha	7-13	Relaxation
Beta	13-30	Concentration
Gamma	30 and above	Attention

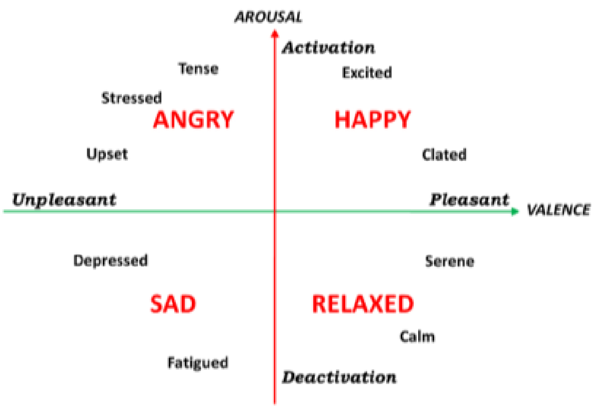
Note. Information in Table 1 from EMOTIV (n.d.)

Modelling Emotions

Many models for analyzing emotions have been developed over the years. One such model is the valence-arousal model proposed by Russell in 1980.

Figure 1

The Valence Arosual Model (Valenza et al., 2014)



In this model (Figure 1), emotions are modeled across two axes: the horizontal valence axis and the vertical arousal axis (Russell 1980). The valence of the emotion is a measure of positive or negative emotions, while arousal is a measure of the strength of an emotion. The valence-arousal model allows for the quantification of emotions.

Machine Learning Modelling using SVMs

Over the past several years, many new machine learning models have been developed, becoming more sophisticated and accurate. One such model is support vector machines (SVMs). SVMs are a linear binary classification model, meaning that they try to separate the input data into two groups by creating a linear separation (a line or plane) between the two groups in such a way that maximizes the distance between the two groups. Now, there are times when the data we are working with, such as EEG data, is not linearly separable. The SVM solution for this data is to utilize a kernel function to map the data into a higher dimensional space, where the data in this new space is then linearly separable.

The Database for Emotion Analysis Using Physiological Signals (DEAP)

When conducting a study on emotion classification based on response to a stimulus, having a good dataset to work with is important. This is where DEAP, presented in Koelstra et al.

(2012), has become an important tool in research on emotion analysis using physiological signals, such as EEGs. To create this dataset, they had 32 participants do 40 trials of watching a one-minute music video clip while having various physiological signals be recorded. Then, they answered a self-assessment for how the music video clip they just watched made them feel.

Prior EEG-based Emotion Recognition Studies

Many studies have been conducted over the years to evaluate combining the power of EEG technology with machine learning to detect emotional responses in subjects. I want to highlight 3 studies in particular since they all utilized music or DEAP for their datasets and utilized SVMs as their machine learning model of choice. The first study was conducted by Lin et al. in 2010. In this study, they created their own dataset from 26 subjects listening to Oscar-winning film soundtracks. They trained an SVM to correctly classify four different emotions (joy, anger, sadness, and pleasure) with approximately 82% accuracy. A subsequent study conducted by Wang et al. in 2014, utilizing a dataset of 6 subjects watching Oscar-winning film clips, was able to train an SVM to classify between two different emotions (positive or negative) with around 87% accuracy on average. Finally, a study conducted by Bazgir et al. in 2018, utilizing DEAP as their dataset, was able to train an SVM to classify the appropriate emotional response with about 91% accuracy.

Methods

Data Acquisition and Data Processing

In my study, I utilized DEAP for training a machine learning model. To filter and process the frequency bands and desired features from the EEG signal, I created a Python program that followed a similar procedure to the one utilized by Bazgir et al. (2018). First, I utilized the average mean reference method to remove some of the extraneous noise that can come with electrical signals. Following this, I then normalized the range of the data to be in the interval of [0,1]. These two steps transform the raw EEG data (Figure 2) into a normalized version of the data across all channels (Figure 3). Then, I windowed the EEG signals into 2-second windows with a 50% overlap. The frequency bands of the EEG signals were then extracted using a discrete wavelet transform using the db4 wavelet function to get the approximate and detailed coefficients (CA and CD). These coefficients show how the signal was changing in a given frequency band over time. The result of the discrete wavelet transform being applied to one window of the EEG signals is shown in Figure 4. Finally, the entropy and energy of each frequency band were calculated as measures of information in the signal. These measurements were input into the SVMs in the machine learning model.

Figure 2

Part of one trial's raw EEG data (amplitude v. sample point)

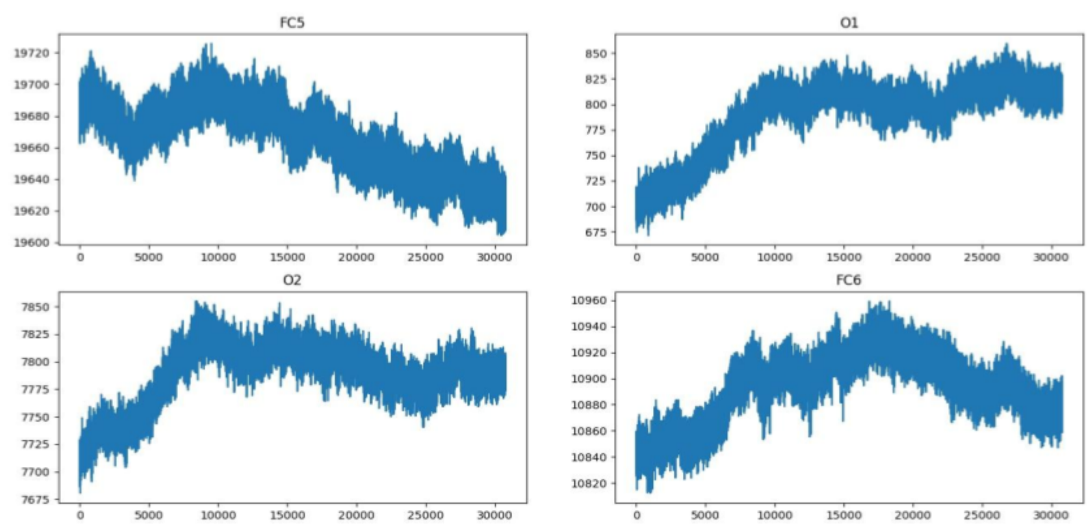


Figure 3

Part of one trial's normalized EEG data (amplitude v. sample point)

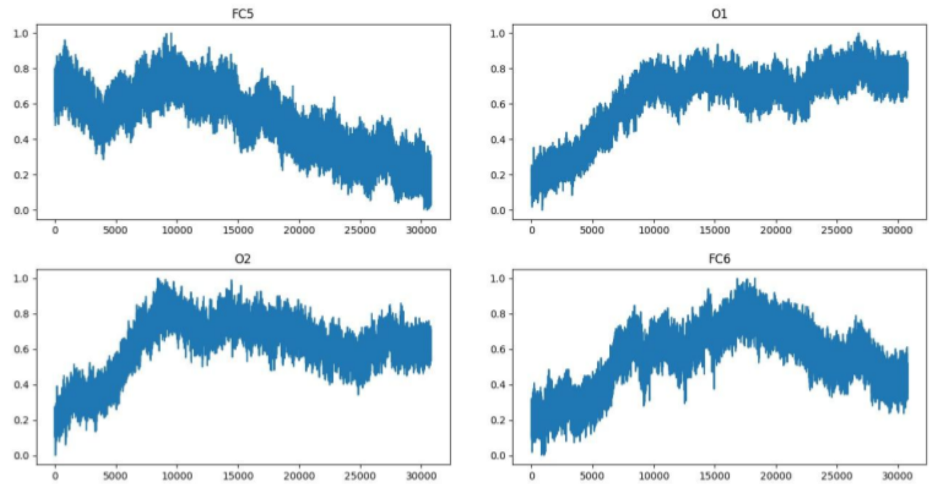
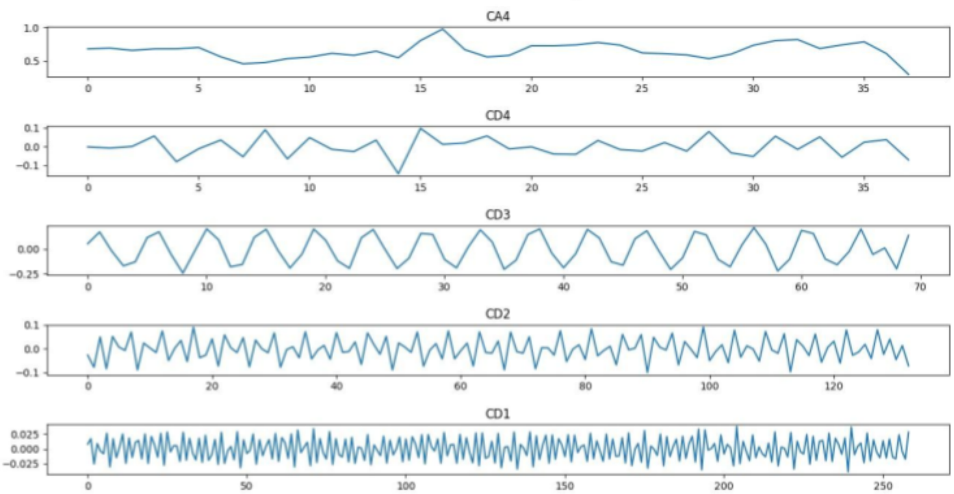


Figure 4

One Window of Discrete Wavelet Transform Applied on O1 Channel (amplitude v. sample point)



Machine Learning Model Training - Model Setup

For my machine learning model, I decided on what I will call a dual voting classification system. As previously mentioned, the valence arousal model allows for emotions to be classified across two different axes: the valence axis and the arousal axis. Additionally, SVMs are binary classifiers, so an SVM could be used to classify which side of the valence axis or arousal axis the emotion the subject is feeling is, allowing for the classification of the emotion in whichever of the four quadrants in the valence arousal model. For both valence and arousal, I trained one SVM for every EEG channel to predict what side of the valence or arousal axis the emotion fell under. The predictions across all the SVMs would then be evaluated, where the prediction for what side of the valence/arousal axis had the majority of votes would be the overall prediction. The result of the vote for the valence SVMs and the arousal SVMs would be the overall classification for which quadrant of the valence arousal model the subject's emotion was in.

Machine Learning Model Training - Model Implementation and Creation

I implemented the SVMs utilizing the scikit learn machine learning library. To find the optimal hyperparameters, I utilized a grid search over a set of different parameters for *c* (the value that controls the margin of error), gamma (the value that controls the curvature/flexibility of the decision boundary), and kernel (which kernel function to use). The SVMs were trained on a laptop with an NVIDIA GPU, utilizing 80% of the data in training and 20% for accuracy testing.

Application Creation

For my demonstration application, I decided to make an application that would allow me to stream data from my EMOTIV EPOC+ EEG headset (Figure 5) to the application, then process that data and feed it through the already-trained machine learning model to have the prediction shown on the screen.

Figure 5

The EMOTIV EPOC+ EEG Headset

I decided on a web application as my application of choice because it would serve as a test to see if an application like this could become easily usable by a wider audience. If this sort of technology were used for medical purposes, I would not want it to be hard to access. If someone has access to an EEG headset, I would like for this application to be a tool available to them. Having the application online would make it easier to access than shipping it as a software package they would have to download.

The first decision I needed to make was what framework I would use to help create my website. I decided to make it with a Python-based framework since that would allow me to more easily work with my SVMs and data processing code since that code was also written in Python. I chose Flask as the framework I would use for the site thanks to it being a fairly straightforward framework to work with.

In my demonstration, the first thing I needed to figure out was streaming the data from my EMOTIV EPOC+ EEG headset to my application. EMOTIV had a feature built-in to their EmotivPRO software called the Lab Streaming Layer (LSL), which allows for different pieces of data from the headset, most importantly the main EEG data, to be streamed to an outside application. They also made a library that they published on Github that makes it easy to work with the LSL in an outside application, such as a Python application (EMOTIV, 2022). With these tools, streaming EEG data to the application and processing it using the same code I used for the machine learning model training became a fairly straightforward process.

An additional consideration for the site was avoiding the page from reloading with every new data collection since you would then not be able to see the previous emotional response on the page. I found that the solution to this issue was using a tool called AJAX that aids in asynchronous page updates (MDN). By using AJAX, when a user clicks the “Connect Headset” button to stream the EEG data from the headset to the application. The SVMs then predict the emotional response, which can be handled asynchronously, allowing for the prior prediction to stay displayed until the new one has been computed and shown on screen.

Results

Machine Learning Model Accuracy

After training, the grid search determined that for every SVM, the best hyperparameters were c = 0.1 and gamma = 1, and the best kernel function for the SVMs was the RBF kernel. Across all EEG channels, the SVMs trained were 62.34% accurate across the arousal axis and 66.88% accurate across the valence axis.

Figure 6

Arousal SVM Accuracy

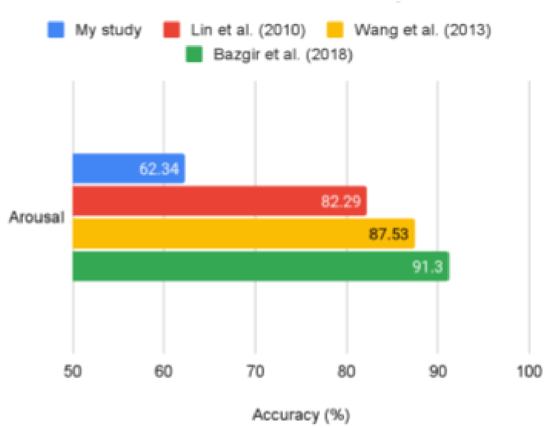
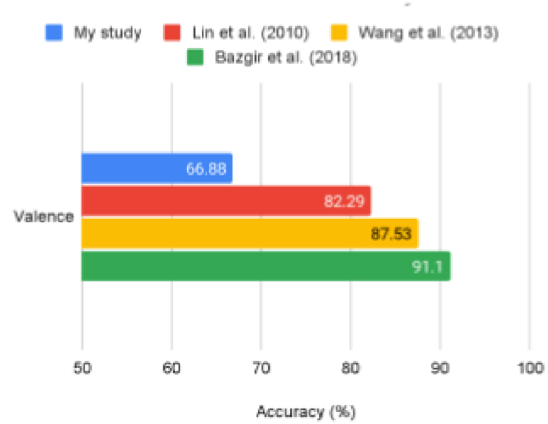


Figure 7

Valence SVM Accuracy



For reference, I have compared the accuracy of the SVMs in my study to those in the aforementioned studies in the literature review (Bazgir et al. 2018; Lin et al. 2010; Wang et al. 2013) in Figures 6 and 7.

Application

In the end, the application was able to successfully take in and process the data from the EMOTIV EPOC+ EEG headset and feed it into the pre-trained SVMs to determine the appropriate emotional response. A small demo of the application working successfully can be found in this video I recorded: https://youtu.be/5QBksX91XfA.

Discussion

The goal of this study was to develop a machine learning model that can predict in real time a subject's emotional response while listening to music and create an application that utilizes the pre-trained model to predict a subject's response from live EEG data. Utilizing the DEAP dataset and following a data processing procedure from a prior study on EEG-based emotion recognition, a series of SVMs were trained to recognize the correct emotional response with approximately 64.6% accuracy. The trained model was then implemented into an application that can take in data from an EMOTIV EEG headset using their LSL software. While successful, this study is just the beginning of work that can be done regarding both the machine learning and live application aspects of a music emotional response study.

First, the SVMs could be more accurate. It is not clear why my SVMs are not nearly as accurate compared to those in other studies. This is especially the case comparing my results to Bazgir et al. (2018). I followed a very similar procedure to their study, yet my SVMs are around 25-30% less accurate than theirs. While there could be some slight difference or error in my procedure compared to their study, I don't think there is anything so significant that would create such a large difference in accuracy.

Second, there would have to be a lot more work done to not only deploy and host the website but also make it work for different types of EEG headsets. Right now, for the application to work, a user would need an EMOTIV headset and an EmotivPRO paid license to utilize their LSL software. This makes the application very restrictive to use, since not only would a user need a headset from a specific company, but they would also need to have an additional subscription to use the software. Transforming the application to make it usable with more types of EEG headsets is a crucial next step for making the application more widely available.

Further research and model development would not only benefit music listeners personally but could have wide-ranging applications in the medical community.

Acknowledgments

I would like to thank Dr. Nathalia Peixoto of George Mason University and Dr. Peter Gabor of the Thomas Jefferson High School for Science and Technology for their guidance as I worked on this project.

References

American Music Therapy Association. (2005). *What is Music Therapy?* American Music Therapy Association. https://www.musictherapy.org/about/musictherapy/

Bazgir, O., Mohammadi, Z., & Habibi, S. A. H. (2018). Emotion Recognition with Machine Learning Using EEG Signals. *2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME)*, 1–5. https://doi.org/10.1109/ICBME.2018.8703559

EMOTIV. (n.d.). The Introductory Guide to EEG (Electroencephalography). *EMOTIV*. Retrieved August 20, 2022, from https://www.emotiv.com/eeg-guide/

EMOTIV. (2022). *Emotiv Lab Streaming Layer Interface*. EMOTIV. https://github.com/Emotiv/labstreaminglayer (Original work published 2020)

Italia, P. (2018, November 22). Validation of EMOTIV EPOC+ for extracting ERP correlates of emotional face processing. *EMOTIV*. https://www.emotiv.com/independent-studies/validation-of-emotiv-epoc-for-extracting-erp-correlates-of-emotional-face-processing/

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. https://doi.org/10.1109/T-AFFC.2011.15

Lin, Y.-P., Wang, C.-H., Wu, T.-L., Jeng, S.-K., & Chen, J.-H. (2009). EEG-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine. *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 489–492. https://doi.org/10.1109/ICASSP.2009.4959627

MDN. (2022, August 12). *AJAX Getting Started*. https://developer.mozilla.org/en-US/docs/Web/Guide/AJAX/Getting_Started

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In *Music and emotion: Theory and research* (pp. 361–392). Oxford University Press.

Valenza, G., Citi, L., Lanatá, A., Scilingo, E. P., & Barbieri, R. (2014). Revealing Real-Time Emotional Responses: A Personalized Assessment based on Heartbeat Dynamics. *Scientific Reports*, *4*(1), 4998. https://doi.org/10.1038/srep04998

Wang, X.-W., Nie, D., & Lu, B.-L. (2014). Emotional state classification from EEG data using machine learning approach. *Neurocomputing*, *129*, 94–106. https://doi.org/10.1016/j.neucom.2013.06.046