Classification of musical preference in generation Z through EEG signal processing and machine learning

Classification of Musical Preference in Generation Z through EEG Signal Processing and Machine Learning

Billy Ward[1], Chandresh Pravin[1], Alec Chetcuti[2], Yoshikatsu Hayashi[2], and Varun Ojha[1]

[1] Department of Computer Science, University of Reading, Reading, UK
[2] Biomedical Science & Engineering, University of Reading, Reading, UK
{b.ward,c.pravin,a.chetcuti,y.hayashi,v.k.ojha}@reading.ac.uk

Abstract. This paper proposes a methodology for investigating musical preferences of the age group between 18 and 24. We conducted an electroencephalogram (EEG) experiment to collect individual's responses to audio stimuli along with a measure of like or dislike for a piece of music. Machine learning (multilayer perceptron and support vector machine) classifiers and signal processing [independent component analysis (ICA)] techniques were applied on the pre-processed dataset of 10 participant's EEG signals and preference ratings. Our classification model classified song preference with high accuracy. The ICA based EEG signal processing enabled the identification of perceptual patterns via analysis of the spectral peaks which suggest that the recorded brain activities were dependent on the respective song's rating.

Keywords: EEG signal · Music Stimuli · Signal Processing · Classification · Machine Learning · Feature Analysis

1 Introduction

Music has become an essential part of human population and plays an integral role in popular culture. Music, like many forms of modern media, has the power to elicit strong emotional responses in our brains [15]. This research was motivated to investigate the way in which music affects and is related to human brain activity, especially among the age group between 18 and 24 (so-called *Generation Z*). The motivation behind investigating the mentioned age group particularly, stems from the accessibility of generation Z to available music, which in turn is due to the availability of technology in more recent times [18].

Numerous experiments have looked at the impact music can have on a person's brain activity, some used to treat mental conditions, such as a study by Ramirez et al. [15] which aimed to help treat depression in elderly people. Other studies explore more commercial approaches and investigate the potential uses of real-time brain response functionality [11]. The majority of research classifies emotional states of the participant depending on how their brain reacts to both audio and visual stimuli.

In past experiments, emotional states were classified using the electroencephalography (EEG) signals of a given participant. An EEG signal is a recording of brain activity acquired by attaching sensors to a participant's scalp to pick up the electrical signals produced from different parts of the brain [19]. The participant is usually shown or exposed to a stimulus to which their neurological responses are recorded. Depending on the emotional classification model being used, varying levels of activity in parts of the brain can directly correlate to a person feeling a certain emotion.

Emotional states are complex and often prove hard to determine, so a bipolar model is used in most cases. The most used dimensions to quantify an emotional state are arousal and valence, which measure pleasure or displeasure. For this study, a single-dimensional model is implemented to identify a participant's response, this being a participant's preference to given stimuli.

This research presents an experiment to investigate the responses of 10 participating humans to audio stimuli (songs). An original dataset was obtained where each participant was given 12 audio snippets from a playlist of popular songs to listen to and rate based on their "like" or "dislike" of each song. EEG recordings were taken for each response to the audio snippets using an EEG device[3] and then classified by song rating using machine learning algorithms.

For this experiment, we have collated a dataset of 10 participants, aged between 18 and 24 years with a gender ratio of 1:1 (5 men and 5 women), containing 8 channels of EEG data recorded for 12 different musical clips of length 30 seconds, time period was chosen in order to regulate the lengths of each song clip. The following are the main contributions of this study:

1. An original experiment was designed to collect data related to individuals (gender ratio of 1:1; 5 men and 5 women) musical preference.
2. The musical preference problem was formulated using machine learning algorithms for preference classification. To enhance the learning model's ability to classify musical preference, feature extraction techniques were used.
3. Signal processing techniques were applied to investigate the visual pattern of brain activity related to varied songs and the user's preference ratings.

Sec. 2 reviews related works; Sec. 3 outlines the methods and explains how different classifiers were setup. Results from the filtering, feature extraction, and classification are discusses in Sec. 4 and, Sec. 5 concludes our findings.

2 Literature Review

There are many ways of classifying emotions that have been explored in previous literature. Work carried out by Ekman and Paul [4] explores the concept of basic emotions, where they explore the relationship between facial expressions and emotions, claiming that there are 6 emotions associated with facial expressions. These are anger, disgust, fear, happiness, sadness, and surprise. They later

[3] https://www.unicorn-bi.com/brain-interface-technology/

added amusement, contempt, contentment, embarrassment, excitement, guilt, pride in achievement, relief, satisfaction, sensory pleasure, and shame. Plutchik and Robert [14] built on work [4] and suggested 8 basic emotions, consisting of the same original 6 emotions expressed in [4], but adding anticipation and acceptance. Both examples show challenges in classifying human emotion.

The general approach of trying to represent these emotions tends to be using a dimensional model [10, 7]. For a dimensional approach, two fundamental dimensions valence and arousal are needed to measure a participant's emotional state. The valence is a measure of pleasure or displeasure. The arousal represents how stimulated the participant feels. As an example, anger could be categorized as a measure of negative valence and high arousal. By using this two-dimensional model of valence and arousal, most basic emotions can be accurately mapped from continuous signals to discrete emotions [4]. Liu et al. [10] claim the most used model for dimensional emotion classification is a bipolar one. However, there are cases of more dimensions being added for more accurate classification. Authors in [9] agree that the bipolar model is most widely used. Moreover, a three-dimensional model is also possible. In this model, third dimension "dominance" is added and quantifies how in control the participant is of their current emotional state. The addition of this extra dimension is justified given that emotions such as anger, a state of high arousal and negative valence, could also be interpreted as fear, another emotion of similar dimensional values.

Most studies that attempt to perform the classification of emotional states use machine learning algorithms with features extracted from EEG signals [7, 2, 6]. Commonly used algorithms are the multilayer perceptron (MLP), support vector machine (SVM) and K-nearest neighbor, all explored and compared in a study by [3]. Asif et al. [1] explored the classification of stress levels in participants with three target classes using machine learning algorithms. In this study, Cohen's kappa is the measure used to evaluate the efficiency of classification. It represents the agreement of classification and considers the possibility that correct classifications are made unintentionally. This makes it a good statistic for verifying the reliability of accuracies obtained.

The spectral powers of frequency bands are often used in feature extraction. In the study performed by [7], power spectra were examined in these distinctive bands to investigate how they correlate with a person's emotional state. One commonly used indicator of a given emotional state is alpha-power asymmetry which was also used in this study. Comparing symmetrical channel spectral powers is also used as a feature for classification in the study by [16].

Other data transformation methods, such as the short-time Fourier transform (STFT) were commonly used for feature extraction [17, 8]. In the study by [8], it is shown that spectral powers are the most common feature used for EEG signals analysis. Schultz et al. [17] suggest calculating the relative power of each frequency band by taking the target frequency component and dividing it by the sum of all frequency components to predict human emotion from data.

3 Music Preference Classification and Analysis

3.1 Study Design

A two or more-dimensional model is used in many previous studies to classify emotional states, such as happiness, sadness, or anger. Emotions, however, are not always so simply defined. Audio and visual stimuli have the potential to induce emotion within a person, but to classify a given emotion felt by a person through listening to a song may not necessarily mean the person has a negative experience. In some cases, a person may want to feel a negative emotion, such as sadness, as this can be therapeutic and, in other instances, may even evoke pleasurable emotions.

A study by [5] investigates this seemingly paradoxical concept, concluding that there is a highly positive correlation between sadness and enjoyment, which comes from a feeling of "being moved" by a stimulus. Reducing the dimensions used to classify an emotional state allows this to be explored. Instead of explicitly defining the emotion felt, the participant's preference rating disregards the underlying emotion as positive or negative and simply measures the pleasure or displeasure that person is feeling. A person may enjoy the way a song makes them feel, even if that feeling is a negative one.

The single-dimensional model used for this study quantifies a participant's liking or disliking of a given song, shown in Table 1. In this study, the participants were asked to wear the EEG headset (Unicorn Hybrid Black). We played each song snippets one-by-one to the participants with an interval of a few seconds relax time in between. The participants were asked to close their eyes during snippets were played to them.

Table 1. Musical preference questionnaire used to determine the level of musical preference for each song played during the experiments. Participants were asked to select their feelings towards the song using this questionnaire.

-2 Hate it	-1 Dislike it	0 Neutral	+1 Like it	+2 Love it

For this research, we define a *binary classification* task ($r = 2$), being the participant liking or disliking the music being played, the method of which is explained further in Sec. 3.2. There are 8 EEG channel measurements used for predicting emotions induced in participants for a total of 12 songs being played for each experiment. Each recording is of length 20 seconds (with a 5-second blank break before and after the recording), sampled at 250 Hz by the measurement device, thus resulting in 5000 data points ($d = 5000$) for each EEG channel. The recorded dataset was composed of 120 EEG recording sets per person and a total of 960 EEG channel waveform ($N = 960$) in the dataset.

3.2 Experiment Setup and Measurements

To conduct this experiment, a new data set was obtained from 10 participants, 5 men and 5 women, from Generation Z, which are people between the ages of 18 and 24. This seemed an appropriate number of participants as previous literature shows this is enough to obtain well-classified data [10]. Participants were asked to confirm they did not have a history of mental illness, as this has been shown to produce unreliable EEG recordings.

The experiment setup was influenced by previous experiments in this area. For the stimuli, 12 song clips of 30 seconds each were chosen, all of which had received awards in recognition of their quality. To attempt to diversify the music, 3 different awards were chosen: The 2019 BRIT awards, the Mercury Prize ranging from years between 2010 to 2019 and Rolling Stone's Most Influential Albums of all time. In the case of the Mercury and Rolling Stone awards, a whole album is given the award, so the most played song from that album was chosen according to Spotify statistics. The music chosen is shown in Table 2, along with the award given to each song.

Table 2. Songs used for the musical preference classification task. Index ID representing the song label used within the model, song name shows the original name of the song, as released by the artist(s), followed by the artist(s) name and award code, which lists the musical award that the song and artist has won. The 'BRIT' code represents the BRIT Awards (2019), 'MERC' represents the Mercury Prize (2010 - 2019) and the 'ROLS' code represents the most influential albums (of all time) Rolling stone.

Index	Song Name	Artist	Award Code
S1	One Kiss	Calvin Harris and Dua Lipa	BRIT
S2	Don't Delete the Kisses	Wolf Alice	MERC
S3	Money	Pink Floyd	ROLS
S4	Shotgun	George Ezra	BRIT
S5	Location	Dave	MERC
S6	Smells Like Teen Spirit	Nirvana	ROLS
S7	God's Plan	Drake	BRIT
S8	Breezeblocks	alt-J	MERC
S9	Lucy In The Sky With Diamonds	The Beatles	ROLS
S10	Thank U, Next	Ariana Grande	BRIT
S11	Shutdown	Skepta	MERC
S12	Billie Jean	Micheal Jackson	ROLS

The participants were given an answer sheet with a rating table for each song, ranging from -2, -1, 0, $+1$ and $+2$, shown in Table 1. They were also asked to indicate whether they recognized the song played to them to identify prior knowledge of songs.

For the data acquisition process, a Unicorn Hybrid Black device was used to record the participant's brain activity during experimentation. The device uses electrodes positioned according to the 10–20 system at Fz, C3, Cz, C4, Pz, PO7, Oz and PO8. These 8 channels were recorded, processed and used to develop the learning models.

3.3 Signal Preprocessing and Feature Extraction

The EEG dataset was processed using a bandpass filter, with a lower cutoff frequency of 2Hz and a high cutoff of 30Hz, accompanied by a notch filter at 50Hz. The notch filter was used to remove any environmental noise, such as electrical current wiring near the EEG device or eye movement. The type of filtering method chosen to reduce noise was the Chebyshev Type 1 filter of order 6 and had a peak-to-peak passband rippled 10 decibels and a passband edge frequency of 0.6. The filter was developed based on prior testing of randomly chosen data and evaluated using the signal-to-noise-ratio calculation shown as per:

$$\mathbf{SNR} = \frac{P_{signal}}{P_{noise}}, \tag{1}$$

where P_{signal} is the power of the signal and P_{noise} represents the power of the noise within the signal. Further details on the results of the filtering and classification model accuracies can be found in Sec. 4.

Feature extraction methods, such as *principal component analysis* (PCA) and *independent component analysis* (ICA) were used in this study. The PCA algorithm was used with the highest 2 and 3 components prior to the data being applied to the classification model and aimed to improve the accuracy of the model. The ICA algorithm was used to interpret and understand the acquired results for *spectral analysis* of the EEG signal.

3.4 Musical Preference Classification

There are two main supervised machine learning techniques used for classifying emotional states that appear to be used most frequently. These are SVM and MLP. The data for classifiers were organized as follows: For the formulation of the musical preference problem, the collected EEG based extracted feature of each participant ($P1$ to $P10$), each song was labeled with related like and dislike rating into binary form (0 indicates “Hate it” and “Dislike it” rating and 1 indicates “Neutral,” “Like it,” and “Love it” rating). The dataset was further processed for these classification tasks, which consisted of the relative bandpowers calculated during feature extraction from EEG signals. The original and pre-processed dataset is available at the following repository[4].

The SVM algorithm was developed using two and three dimensions, resulting in varying model accuracy dependent on whether the data was normalized or non-normalized. Results for SVM algorithm are shown in Sec. 4. As with the implementation of the SVM, the MLP model was used to classify binary song preference and song recognition.

[4] doi: http://doi.org/10.5281/zenodo.4071944

4 Results

The two methods of feature extraction chosen: relative bandpowers and PCA. This created two data sets that would also be trained using a single layer MLP. As well as for these two methods, a comparison of their performances when using normalized and non-normalized data is also made. The data used for these tests was from Participant 2 and all song recordings were included in the classification. According to Table 3 normalized bandpowers feature dataset was used for final modeling for all participants.

Table 3. Showing difference in classification accuracy for regular and normalized data, using the bandpower, PCA using the top 25% of components and PCA using the highest 27% of the components.

Feature Extraction (Regular)	Test 1	Test 2	Test 3	Average Accuracy (%)
Bandpowers	51.562	48.438	56.25	52.083
PCA (2 components)	57.562	57.494	57.007	57.354
PCA (3-dimensions)	57.362	56.988	57.889	57.413
Feature Extraction (Normalized)	**Test 1**	**Test 2**	**Test 3**	**Average Accuracy (%)**
Bandpowers	67.188	60.938	64.062	64.063
PCA (2-dimensions)	49.871	50.313	49.384	49.856
PCA (3-dimensions)	50.061	50.159	49.761	49.994

4.1 MLP and SVM classifiers

The data was classified based on each participant's rating values to the 12 songs listened to in the experiment. The data set was treated as a binary classification problem, where a user's preference was converted into either 0 or 1, representing a participant disliking or liking the song, respectively. Table 4 shows the classification results for all 10 participants using the MLP and SVM models. The results show relatively low accuracy for some participants and as much as 76.7% accuacry in some others of association between their implicit response against explicit response.

4.2 Spectral Analysis and ICA Components

Song 10, "Shutdown" by Skepta, had varying participant responses to it, and was chosen to compare topographic plots of ICA components. Fig. 1 shows the difference in brain activity between positive and negative ratings, with the positive responses in the left column, Participants 3, 4 and 7, and the negative response in the right column, Participants 2, 5 and 9. Participant 4 had 5 of their 8 components rejected as they mainly consisted of noise, however, the 3 components displayed depict reliable brain activity.

Table 4. Showing the classification accuracies of 10 people (P1 - P10) for musical preference. Models evaluated is the SVM algorithm and MLP model. The calculated values of Cohen's Kappa are also shown in the table.

Ratings	SVM	MLP	Cohen's Kappa (SVM)	Cohen's Kappa (MLP)
P1	0.677	0.656	0.354	0.313
P2	0.672	0.594	0.344	0.188
P3	0.703	0.552	0.406	0.104
P4	0.672	0.615	0.344	0.229
P5	0.766	0.615	0.531	0.229
P6	0.693	0.625	0.385	0.250
P7	0.641	0.641	0.281	0.281
P8	0.714	0.594	0.427	0.188
P9	0.635	0.583	0.271	0.167
P10	0.656	0.609	0.313	0.219
Average	0.683	0.608	0.366	0.217

In the positive responses, a peak on the Pz electrode is shown, suggesting use of the parietal lobe. This suggests that the parietal lobe is activated when a person conducts episodic memory retrieval. Episodic memories consist of memories of a specific time and place that they can recall. A study by [20] concludes that there are multiple areas of the parietal lobe that are activated during episodic retrieval. In all cases, when a positive response is given, the song heard was recognized by the participant. The spectral peaks at the Pz nodes indicating positive responses may instead be signaling the participant remembering the song or a memory associated with the song. In this sense, recognizing a song can be determined as an important factor in the response experience. This detail is reinforced by the results shown in the classification of signals, where classification by recognition of song resulted in the highest classification accuracy.

The model parameters could have been a limiting factor on the quality of classification. Since each participant's EEG data is unique to them, using a generalized model for all data could have limited performance. Despite this, accuracies between 63.5% and 76.6% were obtained when classifying by user rating using the SVM model, with fair to moderate agreement for all results from the Cohen's kappa coefficient. The MLP did not perform as well, achieving accuracies between 55.2% and 65.6% with slight to fair agreement.

5 Conclusions

In this paper, we aimed to implement a novel methodology to classify participant preference to songs using a single-dimension emotional classification model. The data used for the study was from an original data set recorded from 10 participants using an EEG device with 8 channels. The classification was performed using two machine learning algorithms, a multilayer perception and support-vector machine, where the target classes were the participants preference and recogni-

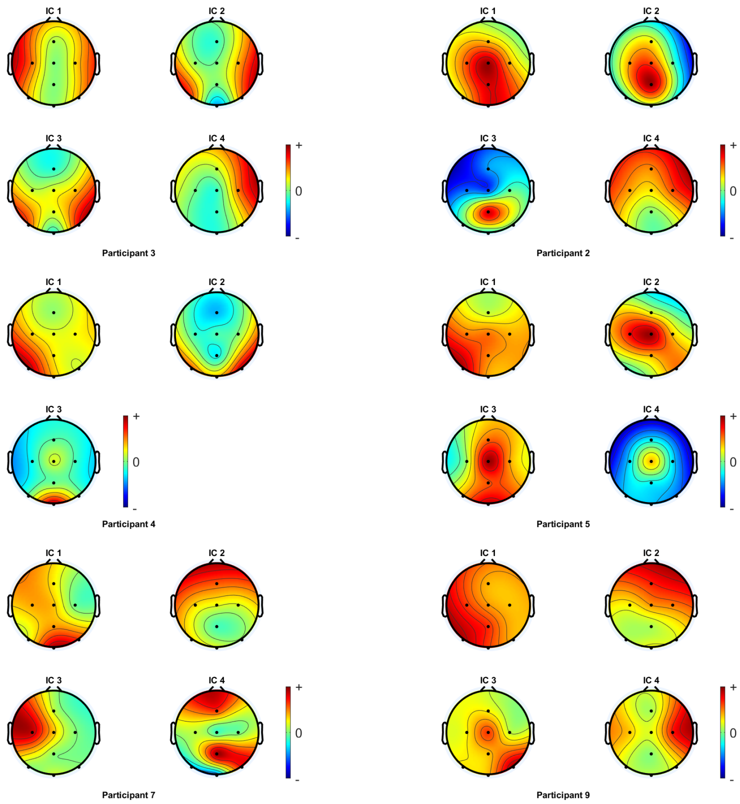

Fig. 1. Shutdown by Skepta ICA Components

tion of the songs. ICA components were created to perform visual analysis of spectral peaks for participants. The accuracy obtained from the classification performed shows that binary separation of EEG signals based on like and dislike is possible. Given a larger quantity of more contrasting data, model performances can be improved. Visual analysis of ICA components suggests spatial patterns of brain activity dependent on song rating. It highlighted the significance of song recognition and the potential relationship between the activation of the parietal lobe with song recognition.

References

1. Asif, A., Majid, M., Anwar, S.M.: Human stress classification using EEG signals in response to music tracks. Computers in biology and medicine **107**, 182–196 (2019)
2. Basterrech, S., Krömer, P.: A nature-inspired biomarker for mental concentration using a single-channel EEG. Neural Computing and Applications pp. 1–16 (2019)

3. Bazgir, O., Mohammadi, Z., Habibi, S.A.H.: Emotion recognition with machine learning using EEG signals. In: 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME). pp. 1–5. IEEE (2018)
4. Ekman, P.: Basic emotions. In: Handbook of cognition and emotion. vol. 98, p. 16 (1999)
5. Hanich, J., Wagner, V., Shah, M., Jacobsen, T., Menninghaus, W.: Why we like to watch sad films. the pleasure of being moved in aesthetic experiences. Psychology of Aesthetics, Creativity, and the Arts **8**(2), 130 (2014)
6. Li, M., Xu, H., Liu, X., Lu, S.: Emotion recognition from multichannel EEG signals using k-nearest neighbor classification. Technology and Health Care **26**(S1), 509–519 (2018)
7. Lin, Y.P., Wang, C.H., Jung, T.P., Wu, T.L., Jeng, S.K., Duann, J.R., Chen, J.H.: EEG-based emotion recognition in music listening. IEEE Transactions on Biomedical Engineering **57**(7), 1798–1806 (2010)
8. Lin, Y.P., Wang, C.H., Wu, T.L., Jeng, S.K., Chen, J.H.: EEG-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine. In: 2009 IEEE international conference on acoustics, speech and signal processing. pp. 489–492. IEEE (2009)
9. Liu, Y., Sourina, O.: EEG databases for emotion recognition. In: 2013 international conference on cyberworlds. pp. 302–309. IEEE (2013)
10. Liu, Y., Sourina, O., Nguyen, M.K.: Real-time EEG-based human emotion recognition and visualization. In: 2010 international conference on cyberworlds. pp. 262–269. IEEE (2010)
11. Nathan, K.S., Arun, M., Kannan, M.S.: Emosic—an emotion based music player for android. In: 2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). pp. 371–276. IEEE (2017)
12. Ojha, V.K., Abraham, A., Snášel, V.: Metaheuristic design of feedforward neural networks: A review of two decades of research. Engineering Applications of Artificial Intelligence **60**, 97–116 (2017)
13. Ojha, V.K., Griego, D., Kuliga, S., Bielik, M., Buš, P., Schaeben, C., Treyer, L., Standfest, M., Schneider, S., König, R., et al.: Machine learning approaches to understand the influence of urban environments on human's physiological response. Information Sciences **474**, 154–169 (2019)
14. Plutchik, R.: Emotions and life: Perspectives from psychology, biology, and evolution. American Psychological Association (2003)
15. Ramirez, R., Palencia-Lefler, M., Giraldo, S., Vamvakousis, Z.: Musical neurofeedback for treating depression in elderly people. Frontiers in neuroscience **9**, 354 (2015)
16. Rozgić, V., Vitaladevuni, S.N., Prasad, R.: Robust EEG emotion classification using segment level decision fusion. In: 2013 IEEE international conference on acoustics, speech and signal processing. pp. 1286–1290. IEEE (2013)
17. Schultz, T., Schaaff, K., Wand, D.M.M.: EEG-based emotion recognition. Universitat Karlsruhe, Institut fur Algorithmen und Kognitive Systeme, Cognitive Systems Laboratory (2008)
18. Turner, A.: Generation z: Technology and social interest. The journal of individual Psychology **71**(2), 103–113 (2015)
19. Viola, F.C., Debener, S., Thorne, J., Schneider, T.R.: Using ICA for the analysis of multi-channel EEG data. Simultaneous EEG and fMRI: Recording, Analysis, and Application: Recording, Analysis, and Application pp. 121–133 (2010)
20. Wagner, A.D., Shannon, B.J., Kahn, I., Buckner, R.L.: Parietal lobe contributions to episodic memory retrieval. Trends in cognitive sciences **9**(9), 445–453 (2005)