Using EEG-validated Music Emotion Recognition Techniques to Classify Multi-Genre Popular Music for Therapeutic Purposes

Dejoy Shastikk Kumaran, NUS High School

Acknowledgements:

Saravanan Yukesh Ragavendar, NUS High School
Aung Aung Phyo Wai, Nanyang Technological University
Liu Xiaoyu, Apple Inc.
Mr. Claude Chua, NUS High School
INTRODUCTION
“…lowering apical heart rates and raising peripheral temperatures were more successful in the relaxation and music therapy groups than in the control group. The incidence of cardiac complications was found to be lower in the intervention groups…”

(Guzetta, 1989)
Increasing numbers of U.S. college freshmen and those in high school are reporting depression and high levels of stress.

This has been coupled with increasing abuse of Xanax and other Benzos, which has caused a teenage health crisis amongst teenagers.
Music Therapy has potential to reduce depression and stress, but…

Most existing studies have a small sample size and subjective response experimental techniques and thus any specific findings on the implementation of music therapeutic programs are not broadly generalizable.
Is there a scientific way to use the features extracted from audio files to predict music emotion?
Machine Recognition of Music Emotion has **low accuracy rates** in existing literature, with accuracies below 66% for studies using only **low-level features**.

Machine Recognition of Music Emotion has not been applied to **popular music**, and has mostly been applied to obscure music which is of **less public relevance**.

Machine Recognition of Music Emotion has not been applied to **popular music**, and has mostly been applied to obscure music which is of **less public relevance**.

Machine Recognition of Music Emotion is a technique that uses **classification** of **extracted audio features** (or language features, if NLP is used) to classify music into various emotions.
Valence – How positive or negative an emotion is

Arousal – How intense an emotion is

We are only studying the circled emotions – (happy, angry, afraid and sad) in this study. (Ask me why later!)
1. Obtain **subject-independent** classification accuracies for EEG and Music Feature data

2. Experimentally validate the viability of **subjective annotation** in determining music emotion

3. Experimentally validate the viability of **machine recognition** of music emotion

4. Analysing **subjective annotations** and **topographic maps** for further use

5. Design methods for **emotional induction** and classify 2500 songs by emotion

6. Implement **validated application** allowing for **genre specific** playlist creation
EXPERIMENTATION
Setting up the EEG setup → User Survey and Instructions → Music Excerpt Playing (26s) → Survey on Played Song/Genre

Repeat for 5 genres, 8 songs each

Experimentation → Segmentation → EEG Feature Extraction

Music Feature Extraction → Cross-Validation of Classification Accuracy

Creation of Topographic Maps → Data Visualization and Organization
Expert based annotation used to annotate music in first stage

The annotations are then crosschecked with internet chatter

If both internet chatter and the initial annotation agree, we will label that song.

Emotion Labelling Process
Selection of Songs

- Hip Hop
- Classical
- Heavy Metal
- Electronic
- Popular
10-20 Electrode Placement Map
Experimental Program

We introduced a fully automated experimental program for experiment participants – we would not interact at all with the participants during the experiment.

Why is this important?
Setting up the EEG setup

User Survey and Instructions

Music Excerpt Playing (26s)

Survey on Played Song/Genre

Repeat for 5 genres, 8 songs each

Experimentation

Segmentation

EEG Feature Extraction

Music Feature Extraction

Cross-Validation of Classification Accuracy

Creation of Topographic Maps

Data Visualization and Organization
Instructions

Several songs will be played to you based on your previous selections.

As each song is being played you are required to keep as still as possible with your eyes closed and your hands on your lap.

After each song you are required to complete a short survey and take a 10s break.

Once you are ready, close your eyes.
Segmentation and Extraction of Features

Only low-level features are used in jAudio to demonstrate higher generalizability, as some high level features are not easily extracted.

Feature selection is done to select important features.
3

VALIDATION
Machine Learning Algorithms Featured

Random Forest

Can avoid overfitting, and can thus be highly accurate as it does not recognize noise. Averages multiple decision trees (this can make predictions harder to interpret)

Instance-based

Can adapt to unseen data – thus, highly useful for a forever expanding database of new data and new music. Potentially allows for individual user customization.
How We Validate Subjective Annotation and Machine Recognition of Music Emotion

Subjective Annotation, Music Features, EEG Data

Expert Annotations on Music Pieces
Frequency Bands

Useful for Emotion Analysis:

- Alpha
- Low-Beta
- Beta

Not Useful for Emotion Analysis:

- Delta (sleep analysis)
- Theta (semi-sleep states/ drowsiness)
- Gamma (sensory processing in the visual cortex)
10-fold Cross Validation Accuracies

<table>
<thead>
<tr>
<th></th>
<th>Random Forest</th>
<th>Instance-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electroencephalography Data</td>
<td>98.2%</td>
<td>86.8%</td>
</tr>
<tr>
<td>Low-Level Music Features</td>
<td>95.0%</td>
<td>84.0%</td>
</tr>
</tbody>
</table>

Most Predictive Frequency Bands

<table>
<thead>
<tr>
<th></th>
<th>Alpha (α: 8–13 Hz)</th>
<th>Low beta (β: 14–17 Hz)</th>
<th>Beta (β: 18–31 Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90.4%</td>
<td>91.2%</td>
<td>88.0%</td>
</tr>
</tbody>
</table>
Why is the accuracy so high?

Instance-based classification allows for more precise calculation of emotion in each segment (less variation).

2/3 of all instances in a song have the same predicted music emotion.

That song will be predicted to possess that said emotion.
If two-thirds of the instances within a song agree on the predicted emotion, that emotion will be listed as predicted for that song and compared to the expert-annotated data.

For subjective annotation data, user responses are normalized and classified into emotions, after which they are compared to the expert-annotated data.

<table>
<thead>
<tr>
<th>EEG Classification</th>
<th>Music Feature Classification</th>
<th>Subjective Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% (40)</td>
<td>100% (40)</td>
<td>74.2%</td>
</tr>
</tbody>
</table>

Accuracy of individual music classification methods (as compared to expert-annotated set)
Integrated Model for Prediction of Music Emotion

- Prediction of emotion through low-level feature classifier
- Use of expert-based annotation to annotate music (blinding used)
- If both agree, song is classified as such. **2500 songs classified as such**
4
SECONDARY ANALYSIS
Genre preference appears to be linked to perceived positivity in music – which correlates to positive emotions as per the valence-arousal model.

Preferred – Two most favorite genres among the five

Disliked – Two least favorite genres among the five
Is there a third axis to the emotion model?

We have shown from the **normalised average arousal and valence ratings** that for at least 2 of the genres there is indeed a **perceived link between valence and genre preference**.

However, we can **only propose that there is a third axis to the emotion model**, as it may itself be a confounding factor and we have **limited data to validate our work here as it was not our main objective**.

We suggest an **extension to this project** with a **diverse collection of songs per genre** to solidify that conclusion. The number of subjects also has to be increased to increase the reliability of results.
Valence
Alpha
Beta
Arousal
Intensity

Topographic Maps
5
APPLICATION & VALIDATION
Integrated Model for Prediction of Music Emotion

- Prediction of emotion through low-level feature classifier
- Use of expert-based annotation to annotate music (blinding used)
- If both agree, song is classified as such. **2500 songs classified as such**
Bruno Mars
24K Magic

XXXTentacion
Moonlight
Problems with Existing Work

• Methodologies exist to go from music at one end of the valence-arousal scale to another

• Current work does not consider psychological findings to create automatic playlists for users to use

• High complexity on user’s end – this is undesirable!

• Users may not understand how to effect meaningful emotional change, and this leads to worse results

(Mr Emo., National Taiwan University)
Lack of application for general public to **induce emotional change** for therapeutic benefits

Music must **transition** from original to final emotion to **keep the user engaged**.

Develop application to create musical playlists which will **induce desired mood** in users

**Music must be preferred** by listener to **emotionally engage** the listener
Validation of our Work

There were 3 groups for the experiment – a control group, a group which used only 1 genre and a group which used 2 preferred genres.

There were 31 participants, with an age range of 36 and a median age of 18. Participants did not have more than a year of music education.

Custom playlists were created on popular streaming platforms.

A total of 31 users were asked to listen to the playlist after providing a mood rating (upon 10).

After listening to the playlist, they were asked to rate their mood again.
<table>
<thead>
<tr>
<th>TITLE</th>
<th>ARTIST</th>
<th>ALBUM</th>
<th>DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aesthetics Of Hate - Explicit...</td>
<td>Machine Head</td>
<td>The Blackening</td>
<td>2018-03-21</td>
</tr>
<tr>
<td>The art of dying</td>
<td>Gojira</td>
<td>The way of all-fi...</td>
<td>2018-03-21</td>
</tr>
<tr>
<td>Disciple</td>
<td>Slayer</td>
<td>God Hates Us All</td>
<td>2018-03-21</td>
</tr>
<tr>
<td>No Smoka</td>
<td>YoungBoy Neve...</td>
<td>Al YoungBoy</td>
<td>2018-03-21</td>
</tr>
<tr>
<td>Mask Off</td>
<td>Future</td>
<td>FUTURE</td>
<td>2018-03-21</td>
</tr>
<tr>
<td>Inked in Blood</td>
<td>Obituary</td>
<td>Inked in Blood (...</td>
<td>2018-03-21</td>
</tr>
</tbody>
</table>
## Results of Validation

<table>
<thead>
<tr>
<th></th>
<th>Happy Songs Only</th>
<th>Custom (2 Genres)</th>
<th>Custom (1 Genre)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean Change</strong></td>
<td>1.13</td>
<td>2.58</td>
<td>4.18</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td>1.64</td>
<td>1.16</td>
<td>1.25</td>
</tr>
<tr>
<td><strong>P-Value</strong></td>
<td>-</td>
<td>0.016</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

*Validated* that our custom methodology provides more positive mood change for users (1 genre preferred)
5

CONCLUSION, FUTURE WORK AND REFERENCES
Obtained very high classification accuracies with EEG and Music Feature Data

Validated subjective annotation and low-level features for machine recognition uses

Analysed links between genre preference and perceived valence and arousal

Created topographic maps to analyse links between frequency bands and emotion

Designed music playlist creation technique to induce desired emotion

Classified 2500 songs using an integrated approach in a validated application

Commercialization of technology alongside extension to more genres and affects

Development of single-image emotion recognition technology to remove need for manual emotional input
References


References


References


References


References


Thanks!

Connect with me at
shastikk@gmail.com or shastikk@ieee.org
5

FURTHER QUESTIONS
Why did you pick the machine learning algorithms selected?

- Existing papers often use Support Vector Machines (SVMs) so we wanted to explore a wider range of classification algorithms.

- High accuracy achieved for these classifiers in similar prediction applications and research.

- The advantages of each machine learning algorithm as earlier stated – instance-based are able to adapt to unseen data and random forests avoid overfitting.
What is the research standard for classification accuracy in your experiment?

• We have achieved very strong results, with accuracy of >98% for EEG data as compared to the research standard of ~94% with SVM.

• We have also achieved strong results for music feature classification at the low level, achieving 95% in comparison to the 66-70% research standard, perhaps due to the splitting of each song into instances.
Why does the initial emotion state’s songs need to be played if we are intending the user to move to the final emotion?

• As Bailey states, alongside multiple similar research papers, users are more emotionally engaged when the music they are listening to is emotionally similar to them.

• Important to engage users in order to get them to listen to the music for longer with higher effectiveness in engagement.