See emotion, Play Music

Play Music by analyzing your mood

By Swetha Jayapathy

Master’s Project Report

Submitted to
Oregon State University

In partial fulfillment of
the requirements for the degree of
Master of Science in Computer Science

Presented March 25th, 2022

COMMITTEE, 2022

Will Braynen
Mike Bailey
Raffaele De Amicis
# TABLE OF CONTENTS

1. Abstract ................................................................................. 1

2. Introduction ........................................................................... 2

3. Related Work .......................................................................... 5

4. Pre-Implementation user Study .................................................. 7
   4.1 Pre-Implementation user study design ...................... 7
   4.2 Pre-Implementation user study results ................. 9

5. Solution .................................................................................. 15
   5.1 Solution Architecture ............................................. 16
   5.2 Implementation of Emotion Detection using Face API 17
   5.3 Connecting to Spotify Web API ....................... 18
   5.4 Generating Customized Playlist from Spotify .......... 19
   5.5 Sample of Application Screenshot ................... 23

6. Post Implementation Survey .................................................... 28

7. Results .................................................................................. 32

8. Limitations ............................................................................ 36

9. Conclusion and Future Work .................................................. 37

10. References ......................................................................... 38
Abstract

Sound is said to have healing power over the human body by inducing certain vibrations into the cells. Music is one such element in our everyday lives that can help reduce stress by increasing the release of pleasure-causing substances in the brain. Researchers have observed that when music is listened to repeatedly, the human body produces a connection between certain emotions by responding to the release of stress-causing hormones.

SEPM (See Emotion, Play Music) is a software that captures the current mood of a person by analyzing their facial expressions and recommending songs that help uplift their state of mind. SEPM connects to the Spotify API and retrieves the user's data. SEPM then generates a customized playlist depending on the user's current mood, which can then be saved and retrieved. SEPM is a game-changer in many people's lives, uplifting their moods and carrying them forward energetically through the day.
Music is said to be the most powerful tool to change the world. Music of any kind is said to have an effect on the human body. It is also said that music is the best companion at all times for an empathic friend. Researchers from the journal of positive psychology [1] have proved that people who listen to upbeat music would improve their moods and boost their happiness in a shorter span of time when compared to others. Whereas people tend to prefer sad music when they are experiencing some interpersonal loss, when in depression, or in contrasting situations. Although this succeeds in soothing one's mood momentarily, it worsens the situation by leading them to listen to songs of the same type ending up sadder.

Music helps a person to remember certain occasions in life which would either make them feel happy or sad. For example, a person can remember his happiest moments with his friends upon listening to a particular song. Also, one can be haunted by past memories while he was struggling in his life as people often tend to relate music to situations. Music, in this scenario, would be the best friend to change the state of mind. More often one tends to be ramping up in the same mood by listening to songs that resonate so much with them. But in a depressing situation, a person tends to become more and more upset by overthinking it. Music, in this scenario, would be the best friend to change the state of mind.

There are various ways to recognize the mood of a person such as by analyzing the songs which he would listen to most often. Most of the software requires the user to categorize their songs into different moods or it would require the user to key in the input, which would end up being so tedious in cases of the huge amount of data. In fact, most of the music applications are determining
the mood from the user’s music data, limiting to the songs which are only present in the device. Also, most of the time this would not be reliable when the user randomly shuffles through his playlist.

One of the most effective and reliable ways of capturing the mood of a person is by capturing facial expressions. Understanding emotions through facial expressions would ease life. Facial recognition technology has found its way into our day-to-day life such as mobile unlock feature, online shopping, identifying people on social media, tracking attendance, etc. Facial expressions are important in facilitating collaborations between humans and moreover facial image-based mood detection techniques may provide a faster approach to mood detection.

Considering the above points in mind, from a user's point of view SEPM was developed. This software captures the current mood of a person by analyzing his facial expression and recommends a song from Spotify depending upon the mood that was captured to lift his state of mind. With Spotify being one of the top cloud-based streaming audio providers, this software would be able to satisfy most of the users across the world.

Key Claims

- **Usability:** SEPM makes it easy for the user to access the songs upon their mood just by one click upon logging in. The software instantly starts the web cam and captures the mood of the user, which then leads to a page which lists songs from Spotify. User interaction is made easy with capabilities which includes playing music, adding the generated playlist to the library, etc.
• **Effective song recommendation:** SEPM recommends the most relatable tracks to the user depending upon their mood. SEPM creates a playlist of around 20 songs in the same mood range, which the user can listen at the moment. When the mood is detected as sad or depressed, a track related to a slightly higher range is listed to the user so that he would be able to shift his state of mind.

• **Portability:** SEPM is portable, as it is developed with Spring framework. Spring supports challenges of supporting complex environments and this can be used at enterprise level. SEPM can be packaged as a whole and deployed across different platforms without any concerns regarding the dependencies as the system as a whole consist of all that is needed.
3. Related Work

a) Habu

Habu is one of the most popular applications with respect to playing songs depending upon the mood. Habu is an app that generates a mood-based playlist from the music library of the user. It helps to analyze and navigate through a personalized mood map. It produces a tailored playlist of songs from the most prolific music collections. This software however works only on the songs which are present in the user’s music library and therefore it does not pick other songs which are not present on the device. Also, the algorithm groups the songs from the most highly liked music by the user. This however only can recognize the mood of the user when he shuffles through the songs and may fail to detect the actual mood. When the user is randomly shuffling through the songs, it doesn’t satisfy the actual need.

b) Moody:

This software generates mood-based playlist from the songs already present in the music library. It requires the user to tag the songs to the mood axes by mapping them to a color chart by clicking an appropriate color for the song. One advantage of this software is that the mood tags is set by the user, which reflects the individual’s taste. However, it would be difficult for the user to tag all the songs in his music library. In this case it would be a time-consuming one. Moreover, the drawbacks mentioned for Habu would apply to this app too.
c) Stereomood:

Stereomood is an open tagging model-based software. It is an online platform for streaming music where users could hear and create their own playlist depending on their mood. This software is built on Pandora’s model of utilizing listener feedback. It relies on the user to assign a mood to each song in his playlist. The engine then groups the songs into a relevant playlist with appropriate moods such as calm, dreamy, afrodesiac etc. This again becomes tedious when seen from the user’s perspective to tag each and every song in his playlist.
4. Pre-Implementation user study

4.1 Pre-Implementation user study design

A controlled user study was conducted to better understand the target audience. The aim was to get an idea about the user's background, interests, and perspectives. This helped to understand the usability of the application. A set of 20 users were identified randomly for this user study, among which 15 of them responded. These 15 participants consisted of 10 graduate students and 5 IT professionals and were aged between 20 to 30 years. The participants were asked to fill a survey form to understand their background and perspectives on a mood-based playlist.

User study survey consisted of the following questions:

1. What do you do in your leisure time everyday?
   Options:
   - Listen Songs
   - Netflix, Amazon Prime, Binge-watching shows, etc.
   - Outdoor Activities (Sports, driving, etc.)
   - Indoor Activities (Indoor Games, Cooking, Reading books, etc)

2. How often would you listen to music in a day?
   Options:
   - 3+ times a day
   - Once a day
   - Rarely
   - Never
3. Which is your favorite app to listen to music?
   Options:
   
   Spotify
   Apple Music
   Amazon Music
   Other

4. Songs help me to uplift my mood
   Options:
   
   Strongly agree
   Somewhat agree
   Neither agree nor disagree
   Somewhat disagree
   Strongly disagree

5. How satisfied do you feel by shuffling through your own playlist to identify the song that suits your mood?
   Options:
   
   Extremely dissatisfied
   Somewhat dissatisfied
   Neither satisfied nor dissatisfied
   Somewhat satisfied
   Extremely satisfied

6. Would you prefer an app which can capture your current mood and suggest songs to uplift your state of mind?
Options:

Yes
Maybe
No

7. Are you a Spotify user?
Options:

Yes
No

8. What do you do to uplift your mood?
Options:

Listen to songs
Spend time outdoors
Sleep
Exercise
Others

4.2 Pre-Implementation User Study Results

The following points were observed from the Pre-Implementation survey.

When the participants were what would they do in their leisure time, binge-watching shows and listening to songs were the most frequent ones in everyone’s life. We can see from the below result chart that among the 15 participants, 10 responded that they listen to songs in their leisure time.
When participants were asked how frequently they would listen to music, 38% of the participants responded that they listen to music at least once a day and 25% of the participants have responded that they listen at least 3 plus times a day. It is also observed from the below graph that listening to songs is the top-rated one when it comes to uplifting one's mood.
When the users were asked their most often used app for listening music, there was a significant response for Spotify. 63% of the participants uses the Spotify app to listen to music whereas 31% of the participants use Apple music. This shows that Spotify is the most used music app in current world.

![Figure 4.3: Result of Q3- Participant’s favorite app to listen to music](image_url)

The below graph shows whether music makes an impact in peoples life and it can be seen that 31% of the participants strongly agree, while 44% somewhat agree and 25% of the participants are unsure of this fact. This indicates that songs does have an impact in uplifting ones mood.
In order to see if the participants are satisfied by shuffling and selecting songs on their own each time they want to listen to a music, when asked, it is observed that 19% of the users were dissatisfied and 56% were neither satisfied nor dissatisfied. On an average 25% of the users felt satisfied by selecting their own song. This shows that users had a mixed reaction for this question.
Figure 4.5: Result of Q5 - Participant’s satisfaction in shuffling through their own playlist

To further analyze this, when they were asked if they would prefer an app that would generate a playlist by capturing the mood, a huge response of yes was observed (Figure 4.6). Almost 75% of the participants were really interested in using such application and the rest of the 25% of them have responded ‘May be’. This response indicated that the participants would be highly interested to use such applications which would minimize their need of search and listen to recommended music.
Figure 4.6: Result of Q6- Participant’s preference of an application like SEPM
5. Solution

SEPM is designed and developed mainly for the users to easily access a playlist depending on their current mood without searching manually. SEPM captures the mood of the user through the user’s facial expression and generates a list of customized songs for the user from Spotify. The tracks are chosen in a way that it lifts the current state of mind of the user. It allows users to select and play songs from the recommended set of tracks that have been generated for the user. The interactive user interface allows the user to navigate through the list of recommendations and save them to their playlist.

Spotify application does not have features wherein the user can retrieve his top tracks or top artists based on a certain time period. Therefore, SEPM has other features as accessing and playing:

- User’s top Tracks
- User’s top Artists
- User’s Library
- User’s most recently played tracks

SEPM allows the user to select and play tracks from their top tracks or top artists based on a certain time period as below:

- All time
- Past 6 months
- Past month

SEPM Customized playlist generation has two main features:
- Detecting the mood of the user by connecting to Face API
- Retrieving Spotify data by connecting to Spotify Web API

**5.1 Solution Architecture:**

The system is implemented using the Spring Model-View-Controller (MVC) framework. The model handles the data part of the application, the view is implemented using the Thymeleaf template engine which supports both web and standalone environments. The dispatcher servlet acts as the front controller and routes the request to the respective controller depending upon the request received. Front end is built using HTML, CSS and JavaScript. The backend is implemented using Java and MySQL database.

![Figure 5.1: SEPM Architecture](image-url)
SEPM is a web-based application where the request from the client side is sent to server. The front-end pages are managed by the templates and the backend code has respective controllers and services to implement each of the above functionality. A request from the front-end is sent to Face API to identify and detect the facial features and expressions. The detected mood is then passed on to the backend to further fetch the songs from Spotify API. The logic to filter and generate the customized playlist is implemented in the backend service and the result is sent to the front-end template where each of the track items are rendered in a viewable format with clear descriptions such as the track name, album, artist name.

5.2 Implementation of Emotion detection using FACE API

Face API is a JavaScript library which helps to build a browser-based face recognition system. It is built over tensorflow.js core API and supports Face detection, Face Recognition and detecting facial expressions. This face-api library has pre-trained models. The following models are used:

- loadSsdMobilenetv1Model – SSD (Single Shot Multibox Detector) object detection model
- loadFaceLandmarkModel – detects facial landmark
- loadFaceRecognitionModel – detects facial recognition
- loadFaceExpressionModel – detects facial expression

SEPM captures the image of the user using web cam and then loads all the models. Then it gets the face detection and face landmarks points which is then passed on to the face expression model to determine the facial expression of the user.
Face API’s Tiny face detector is used to detect the face in the web cam. It is a real time face detector which is faster and less resource consuming. This face detector is trained on an average of 14k images with bounding boxes. This model predicts bounding boxes which entirely cover facial feature points and thus produces best result for face landmark detection. The face landmarks are and then the result is passed on to the expression detecting model. This returns a score corresponding to the various emotions such as happy, sad, neutral and anger. Based on this expression detected, the custom playlist is generated by accessing Spotify Web API.

### 5.3 Connecting to Spotify Web API:

Spotify has turned out to be the most popular music application currently. The advantage of using Spotify is that it stores a user’s data for a particular range of time, and it has a huge range of songs which allows us to consider almost all possibilities.
Spotify Authentication:

Spotify implements the OAuth2.0 authorization framework.

![Figure 5.3 Connecting to Spotify Web API](image)

The end-user here corresponds to the SEPM user, the end-user needs to grant access to the protected resources (e.g. playlist, personal Information, etc). MyApp is the client that requests access to the Spotify web app. The access to the resources is determined by several scopes. Scopes enable the application to access specific functionality (e.g. read playlist, read library). This set of scopes is set while authorizing a user.

Spotify web app provides the following types of authentications:

- Authorization code + PKCE extension
- Client credentials
- Implicit grant

SEPM uses the Authorization code with PKCE extension. The authorization code flow allows the client secret to be safely stored and the user needs to grant permission only once. The PKCE allows to access the Spotify web API using access tokens and refresh tokens. Access tokens are deliberately set to expire in a short time, therefore refresh tokens are needed to get new access.
tokens at certain intervals. New access tokens are granted by supplying the refresh token which is obtained originally during authorization code exchange.

5.4 Generating Customized Playlist from Spotify:

The Customized Playlist generation based on the mood is implemented by first getting the user’s top tracks and his favorite artists from his most recent time (6 months). Depending upon the mood recognized through FaceAPI service, a list of songs is selected which falls within a certain mood range and is suggested to the user. This selection of tracks involves accessing the Spotify API by passing on specific attributes such as Valence, danceability, and energy.

```json
{
    "danceability" : 0.635,
    "energy" : 0.678,
    "key" : 4,
    "loudness" : -10.840,
    "mode" : 0,
    "speechiness" : 0.0461,
    "acousticness" : 0.514,
    "instrumentalness" : 0.0902,
    "liveness" : 0.143,
    "valence" : 0.611,
    "tempo" : 95.002,
    "type" : "audio_features",
    "id" : "06AKEBrZackW0KREERnvQ",
    "url" : "spotify:track:"06AKEBrZackW0KREERnvQ",
    "track_href" : "https://api.spotify.com/v1/tracks/06AKEBrZackW0KREERnvQ",
    "analysis_href" : "https://api.spotify.com/v1/audio-analysis/06AKEBrZackW0KREERnvQ",
    "duration_ms" : 255444,
    "time_signature" : 3
}
```

Figure 5.4: Spotify Track attributes

These attributes are selected for the below reasons:

- Valence measures the positiveness of a song, a high value of valence represents happy, cheerful and euphoric type of songs. While tracks with low valence sounds more negative (sad or angry).
• Energy represents a measure of intensity and activity. Energetic tracks feel fast and loud leading the user to be more energetic.

• Danceability measures how suitable a song is for dancing based on tempo, rhythm, beats. A high value of danceability represents most danceable.

The above attributes vary from a range of 0 to 1. 1 being more positive and a high measure of the feature to 0 being the lowest.

When these attributes are considered to generate the playlist, it gives a great match to the mood range provided and would suggest the best suitable tracks.

We have the following sentiments mapped:

• Happy
• Sad
• Neutral
• Angry

User’s Top Tracks are retrieved by hitting the Spotify end point for user’s top tracks. Then the user’s Top 5 Artists are retrieved and their top tracks are further fetched. The result set of top tracks of user and top tracks of the artists are combined into a single result set. The detected emotion along with this result set is passed on to further generate the customized playlist.

Following steps are used to generate the customized playlist depending upon the user’s mood:

1. Get the attributes of each of the tracks.
2. We then filter the tracks based on the below mapping of Sentiments.
Mapping of Sentiments to Spotify Features

SEPM recommends tracks to lift a person's state of mind and therefore the following mapping is used:

<table>
<thead>
<tr>
<th>Emotions</th>
<th>Danceability</th>
<th>Valence</th>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happy</td>
<td>&gt; 0.6</td>
<td>&gt; 0.6</td>
<td>&gt; 0.6</td>
</tr>
<tr>
<td>Neutral</td>
<td>&gt; 0.6</td>
<td>&gt; 0.6</td>
<td>&gt; 0.6</td>
</tr>
<tr>
<td>Sad</td>
<td>0.4 to 0.7</td>
<td>0.4 to 0.7</td>
<td>NA</td>
</tr>
<tr>
<td>Angry</td>
<td>0.3 to 0.6</td>
<td>0.3 to 0.6</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 5.1: Mapping of emotion to track attributes

If a person is happy or neutral, SEPM recommends songs that have more positive features and therefore, we filter songs with danceability, valence, and energy greater than 0.6. If a person is sad, then SEPM would recommend songs that are slightly in the neutral mood range with audio features in the range of 0.4 to 0.7. If a person is angry, SEPM would recommend songs that are close to sad and neutral ranging from 0.3 to 0.6. Energy for sad and angry mood is not considered as it has the same range as the other two attributes.
5.5 Sample of Application screenshot:

Figure 5.5: Login screen for the SEPM application

Figure 5.6: SEPM user registration page
Figure 5.7: Spotify Authorization
Figure 5.5.4: SEPM Home Page

Figure 5.8: SEPM Get top artists of user
Figure 5.9: SEPM Get top tracks of user

Figure 5.10: SEPM User’s saved tracks
Detecting Facial expression

Expression: Neutral

Figure 5.11: SEPM Face expression detection

Customized Tracks

SAVE PLAYLIST

Figure 5.12: SEPM Customized playlist
6. Post-Implementation User Study

After the implementation phase, a controlled informal user study was conducted to get their feedback on the application. The same 15 participants, who participated in the pre-implementation survey were involved in this study to better understand how well this application captures their needs and usage. The participants were allowed to use the software via zoom meetings and in person meetings. At the end of the usage, they were asked a couple of questions.

Study Objectives:

1. Does the user feel it easy to capture his mood through a real time web cam and is interested to listen to the recommended songs?
2. How well a user perceives the recommended songs? Does he feel to continue listening to the songs suggested by the software?
3. Does SEPM have an interactive and user-friendly UI which can be navigated without any guidance?

User Study Design:

The system was evaluated from user surveys, interviews and observations, and each of these followed a certain methodology listed below:

(1) Observations during user meetings when they were using the software which helped in assessing the usability of the application and the challenges which that the user face. Various important factors like the ease of use of the software without any help or explanation, whether the
user was satisfied with the recommendation that he got, most often used features of the software being used (e.g. detecting the user’s mood, adding the songs to playlist) The observations were focused on rich insights into the satisfaction of the user with respect to the look and feel of the software and the recommendations.

(2) Surveys and Semi-structured Interview was conducted:

The interviews primarily focused on getting an in-depth understanding of what are the most used music application and what did the user feel about the music application they were using, what challenges they face and would they prefer SEPM over it.

![Figure 6.1: Logic behind User Study](image-url)
Participants Validation

After the post-implementation user study, in order to make sure that there is a correct understanding of what participants have told in the interviews and that a correct comprehension of what was seen through observations, a survey was arranged. 3 of the participants were asked to fill a questionnaire and let us know if they agree or disagree with each of the conclusions. Also, participants were asked to leave a short description in cases that they disagree with the conclusions. Any generalization or any hidden implication of information in the questionnaire were avoided.

The following survey questions were answered by the participants:

- It is easy to use and navigate through the SEPM application without any guidance
  
  Options:
  
  - Strongly agree
  - Somewhat agree
  - Neither agree nor disagree
  - Somewhat disagree
  - Strongly disagree

- Rate the level of satisfaction using SEPM to detect your mood on a scale of 1 to 5, where 5 being the highest and 1 being the lowest.

- Rate the level of satisfaction using SEPM for song recommendation on a scale of 1 to 5, where 5 being the highest and 1 being the lowest

- How often do you think you would use this SEPM application?
  
  Options:
  
  - More than once a day
  - Once a day
A few times a week
A few times a month
Not at all

- Would you prefer to use the SEPM application for Song recommendations in the future?
  
  Options:
  
  Yes
  No
  Maybe
7. Results

The results are derived from both the post-implementation survey and semi structured interviews. When the users where asked how often they would listen to music, an average of 60% of the participants said that they would listen more than once a day. 34% responded that they listen at least once a day. While only 6% responded that they listen to songs very rarely. This shows how significant music is in everyone’s daily life.

More than 90% of the users on an average agree that it is easy to use the SEPM software without any guidance and said that the UI is user friendly and interactive. Also, from the observation made from the user study meetings, it was observed that they had no difficulty in accessing the application and navigating through it.

![Figure 7.1: User’s thoughts on the SEPM usability](image)
When the users were asked to give a score from 1 to 5 for how satisfied were they with the face detection model to capture their mood, it is observed from the below figure 7.2 that almost all of them on an average were very satisfied. Although we could see one of them had given a score of 3, the rest of the observation looks like the users were satisfied.

![Figure 7.2: User's level of satisfaction for the Face expression detection](image)

The level of satisfaction that the user perceives from the list of recommended song is showed on a scale of 1 to 5, with 5 being the highest and 1 being the lowest. It can be seen from the graph that 9 users have agreed on a score of 4 and 4 users have agreed on a score of 5. With 4 being a more than average effectiveness, it can be said that the users were satisfied upto a certain level with the recommended songs.
Q6 - Rate the level of satisfaction using SEPM for song recommendation on a scale of 1 to 5, where 5 being the highest and 1 being the lowest

Figure 7.3: User’s level of satisfaction for the song recommendations

Q8 - How often do you think you would use this SEPM application?

- More than once a day: 43.75%
- Once a day: 25.00%
- A few times a week: 31.25%

Figure 7.4: User’s perspective on using the SEPM application in daily life
The participants were asked would they prefer to use SEPM in day-to-day life for listening to music in the future and their response was significantly positive. 73% of the users said that they would prefer and 25% of the users said may be.
8. Limitations

SEPM has a few limitations as listed below:

- Although facial expression is reliably associated with certain emotional states, not all the time, the mood can be recognized through a person’s facial expression. There would be certain people who would not react or show their mood outside, whereas they tend to have a different feeling inside. Also, in the cases where people have mixed emotions it would be difficult for the software to correctly suggest the songs depending upon one’s facial expression.

- Not an ideal solution when the user is extremely upset. When the user is extremely upset, the recommendation of the songs at times fail to satisfy the purpose. This might be because of the fact that songs alone cannot uplift the mood of a person depending upon the situation that they are in. For example, when a person goes through a heart break, he would often tend to listen to sad songs and in this case SEPM would help him in suggesting a track with slightly higher range value. Whereas in situations where there is an extreme loss, it is difficult for a person to listen to songs in that stage and this doesn’t apply in such scenarios.

- There are still few outliers in the recommended song list, like around 20% of the suggested songs were not much likeable by the user. To account this, other attributes of the tracks such as loudness, Beats per minute(BPM) can to be considered.
9. Conclusion and Future Work

SEPM is a unique software which analyzes the facial expression to capture the mood of a person and recommends a list of songs matching the mood. User study has proven the insights and effectiveness of the SEPM functionality. It was also observed from the user study that the users enjoyed playing with the application as it is very interactive and proves to be one of an entertaining way of suggesting songs through real time web cam. It can also be concluded that users are more receptive to this application and would like to start using it in day today life.

However, there are still some areas of improvement for the application. Not all times a person’s mood could be detected through his facial expression. In order to precisely find the mood of a person, we must consider other factors such as his personality, tone of voice, changes in skin tone, the visual context of their background, etc. It is said that a person’s skin tone changes depending upon his emotional state and this would be one of a challenge to work with. To account for the outliers in the recommended song list, other attributes such as loudness, Beats per minute(BPM), and acoustics of the tracks can be considered. When the mood is detected as upset or depressed, along with the set of songs, other uplifting podcasts can also be recommended to the user.
10. References


