HCID 520 A

A personalized emotion based music recommendation system

SAYENA MAJLESEIN

MARCH 2019

TABLE OF CONTENTS

INTRODUCTION

Audio streaming platforms have become popular as digital music services provide access to millions of songs from record labels and media companies around the world.
A key challenge of many of the audio streaming platforms, including Spotify, is to help users listen to the music that match with their emotional and personal preferences.

Content-based recommendation is a technology in response to this challenge. Based on a profile of user emotional and personal preferences, this technology recommends music that may be of interest or value to the user. Content-based methods play a central role in recommender systems by accommodating the individual differences between users [1, 2].

An accurate profile of users' current interests and emotional preferences is critical for the success of content-based recommendation systems. Some systems [3, 4] require users to manually create and update profiles. This approach places an extra burden on users. Instead, systems can construct profiles automatically from users' interaction with the system.

Our goal is to enable audio streaming platforms to predict users' emotional preferences, as they're listening to music, to provide them with accurate recommendations that reflect users' personal and emotional preferences.

OVERALL CONCEPT

We present our research on developing a personalized emotion based music recommendation system for Spotify.

Spotify, available at ***www.spotify.com/us/download/***, is one of the most popular digital music services in the world.
For users who are logged in and have explicitly enabled music history, our recommendation system builds profiles of users' emotional interests based on their feedback in regards to the emotional category of a track, predicts their emotional interests as they are listening to a track, and recommends music that reflect their current emotional and personal preferences. Here, we focus on the task of continuing tracks that are being played by the user.

That is, given a few tracks, we wish to suggest tracks that follow the same emotional category of the previous tracks that have been played by the user.

We use a Bayesian framework for predicting the user's current emotional and personal preferences. We combine the content-based recommendation mechanism with an existing collaborative filtering mechanism [5] to generate personalized emotion based music recommendations. This hybrid recommender system will be deployed in Spotify.

DESIGN DETAILS

Spotify predicts music preferences using deep learning based on audio signals [6]. Studies have shown that providing more user control is an effective approach that enables users to help steer the recommendation process with additional input [6]. Therefore, our proposed design solution includes a new means of communication for the users to define and input their emotional preferences when listening to music in Spotify.

As a result, we created the table of emotions, demonstrated in image 1. The table of emotion includes eight predefined emotions that are conveyed by different types of music. The emotion categories were inspired by THE PANAS-X Manual for the Positive and Negative Affect Schedule [7]. Considering that different users may associate different emotions with a song when listening to it [7], we allow users to add their own emotions

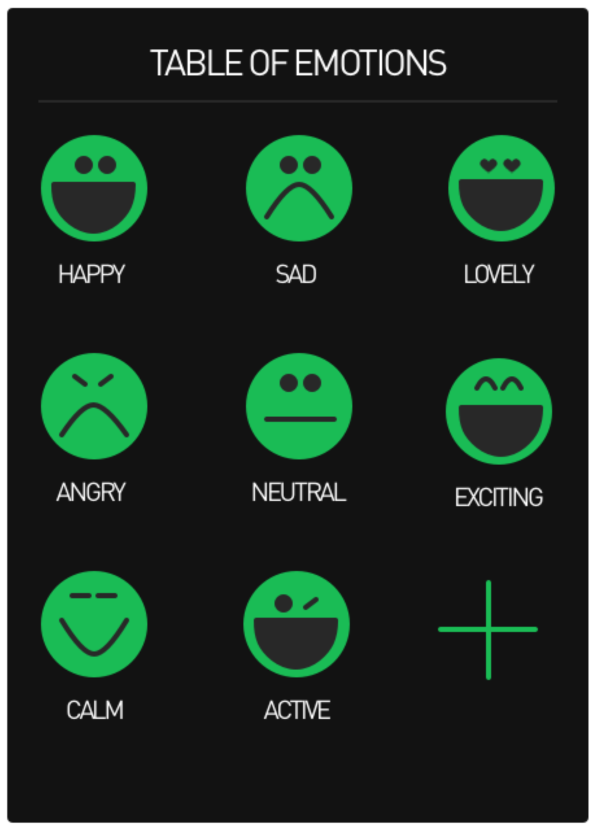

Image 1: Table of emotions

to the table of emotions. Images 2, 3, and 4 demonstrate the process of adding a new emotion to the table. To enable users to reflect their personal preferences in an engaging manner and to use that data to define more standard emotions in the future, the user can define the emotion using an emoji. Then, the user will name the emotion and add it to the table. Image 5 shows the new emotion that has been added by the user.

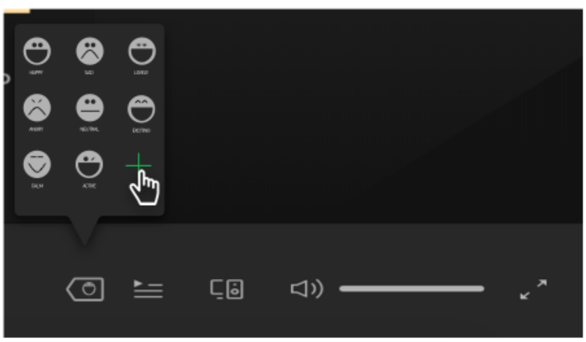

Image 2: Adding a new emotion

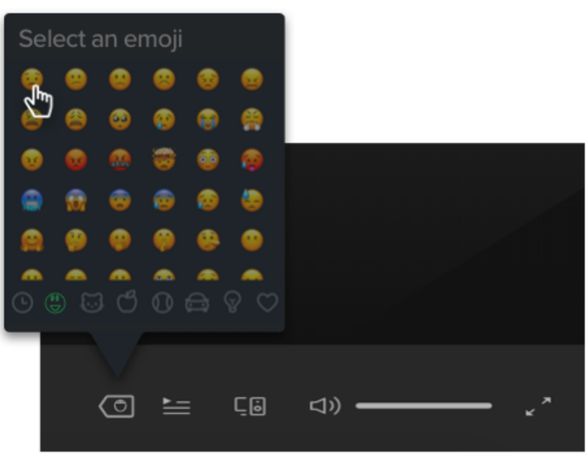

Image 3: Selecting an emoji for the new emotion

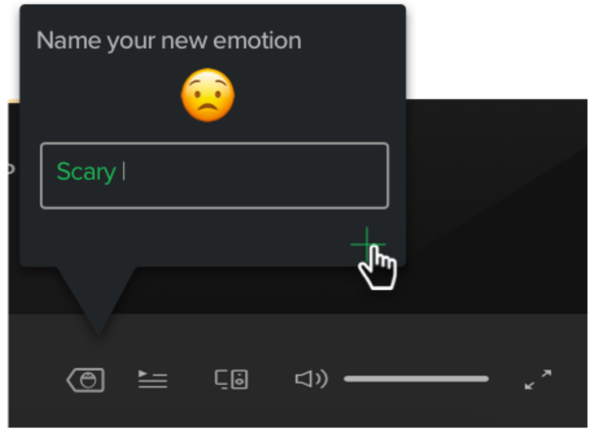

Image 4: Naming the new emotion

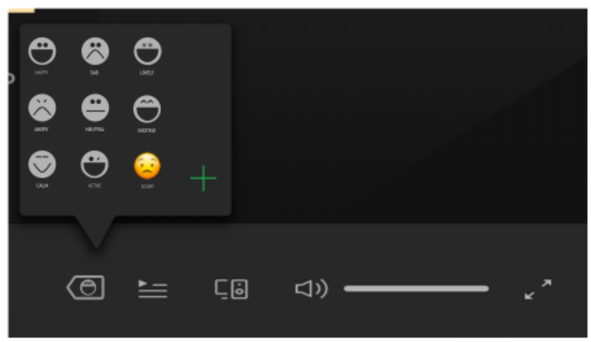

Image 5: New emotion has been added

To provide a flexible and reliable system for users, the emotions that have been defined by the user can be deleted from the system by performing right click on the emotion in the desktop version or tapping on the emotion and holding it in the mobile application version of Spotify. Image 6 demonstrates the delete icon that appears on top-left corner of the emotion that has been defined by the user as user right clicks on it.

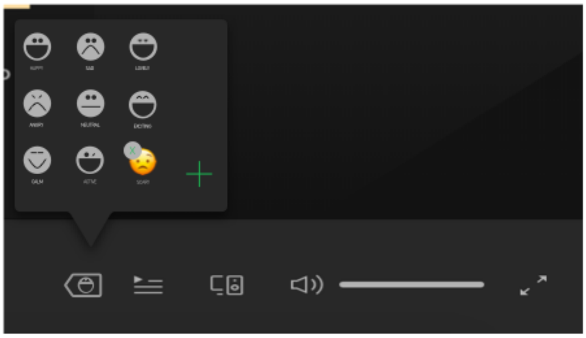

Image 6: Deleting an emotion

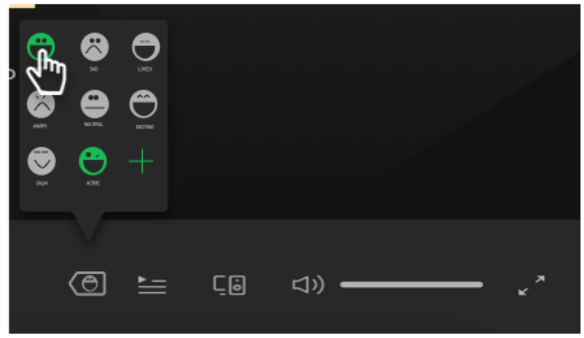

Image 7: Selecting multiple of emotions

Doris Baum's study on classifying music according to emotion [8] states that different users may associate different feelings with a given piece of music, however, the majority of users agree on the connoted emotions for quite a lot of songs. As a result, we provide users with another level of flexibility by enabling them to associate more than one emotion with a song.

As an example, image 7 demonstrates that two emotions have been associated with a piece of music.

To create a cohesive and consistent design structure, the emotion button is placed next to other available features when playing a song. Image 8 shows how the emotion button follows the same visual design of other components as the user hovers on it.

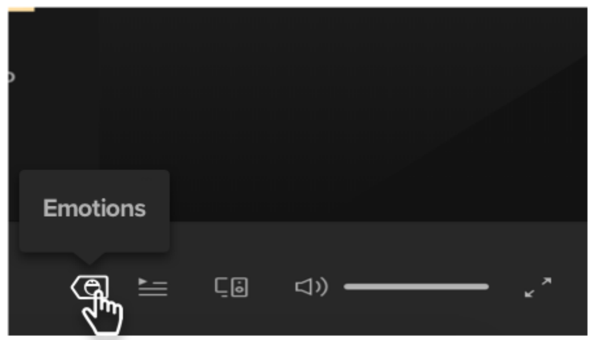

Image 8: The toolkit is shown as user hovers on the button

The emotion button, shown in image 9, has been designed to allow users to associate emotions with tracks.

Image 9: Emotion button

Finally,the Personal Emotions section, shown in image 10, is implemented using our recommendation system to support automatic creation of the user's personalized emotion based playlists.

The playlists will include songs that reflect the user's emotional and personal preferences.

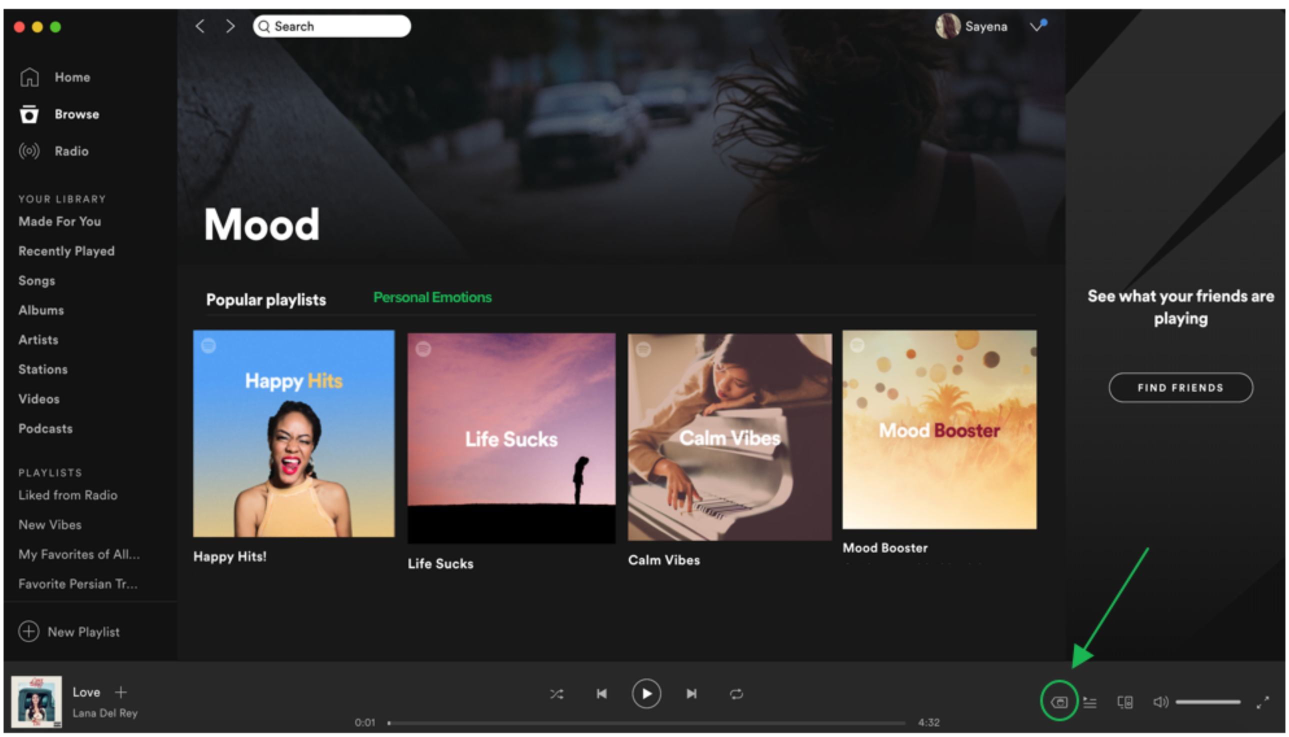

Image 10: Spotify desktop version

ENGINEERING DETAILS

As mentioned before, the main goal of our recommendation system is to recommend tracks that reflect the user's personal and emotional preferences.

Spotify already supports the personalized recommendation by applying the following mechanisms: content-based recommendation, knowledge-based recommendation, and collaborative filtering. We modified and reused the combination of these mechanisms to integrate the personalized emotional based recommendation system.

Considering the feedback that is received through the emotion button, the number of times that a track gets played can be treated as positive votes for the emotional category of that track.

Image 11 shows the attributes associated with a track in Spotify. To support our recommendation system, a new attribute called "emotion" should be defined for each track to represent the emotional categories associated with a track.


```
{
  "danceability" : 0.560,
  "energy" : 0.527,
  "key" : 2,
  "loudness" : -9.783,
  "mode" : 1,
  "speechiness" : 0.0374,
  "acousticness" : 0.516,
  "instrumentalness" : 0.0000240,
  "liveness" : 0.156,
  "valence" : 0.336,
  "tempo" : 93.441,
  "type" : "audio_features",
  "id" : "2z7D7kbpRcTvEdT71tdiNQ",
  "uri" : "spotify:track:2z7D7kbpRcTvEdT71tdiNQ",
  "track_href" : "https://api.spotify.com/v1/track
  "analysis_url" : "http://echonest-analysis.s3.am
  "duration_ms" : 168720,
  "time_signature" : 4
}
```


Image 11: Spotify song attributes [9]

The combination of two different mechanisms are used in this recommendation system: content-based recommendation and collaborative filtering.

The content-based approach recommends tracks based on profiles. These profiles are created by analyzing the attributes of tracks that the user played and/or favored in the past. In contrast, the collaborative filtering approach does not consider the attributes, but uses the opinions of peer users to generate recommendations.

We combined the content-based method presented in [5] with the collaborative filtering method previously developed for Spotify [10] to generate personalized emotional based music recommendations.

The basic assumption of personalization is that users have reasonably consistent emotional interests. A user's history will only be useful if the history help us predict her future actions [5].

PREFERENCE DISTRIBUTION [5]

To integrate the personalized emotional based recommendation system, we classify emotions into a predefined set of emotional categories, **$\{C_1, C_2, \ldots, C_n\}$**, including "happy", "sad", and "angry". Then, we compute the distribution of the number of tracks that have been played over the set of emotional categories for individual users as well as the group of users in a country . We divide the time period into 6 months. For each user **u** , we compute the distribution of the number of tracks that she listens to in every month **t** , **D(u,t)** , represented as a vector over the set of emotional categories:

$$D(u,t) = \left(\frac{N_1}{N_{total}}, \frac{N_2}{N_{total}}, \ldots, \frac{N_n}{N_{total}}\right), \quad \text{where } N_{total} = \sum_i N_i$$

N_i is the number of tracks classified into the category **C_i** that were played to by the user **u** in the month **t**. This number includes the number of repeated tracks.

N_{total} is the total number of tracks that have been played by the user in the time period. Similar to **N_i**, the total number includes the number of repeated tracks.

BAYESIAN FRAMEWORK

We use a Bayesian framework [5] to predict users' current emotional interest based on the patterns that have been created using the emotional categories of the tracks played by the individual users. The predicted emotional interests are used in the recommendation system.

The approach works as follows: first, the system predicts user's genuine emotional interests regardless of the music trend, using the history of the tracks that have been played by the user and the emotions that have been associated with them in each past time period; second, the predictions made with data in a series of past time periods are combined to gain an accurate prediction of the user's genuine emotional interests; finally, the system predicts the user's current emotional interests by combining her genuine emotional interests and the current emotional preferences of other users in her country.

PREDICTING USER'S GENUINE EMOTIONAL INTEREST

The genuine interest of a user in emotional category **Ci** is modeled as p^t (number of times that the tracks associated with a specific emotion have been played **|** emotional category = **Ci**), and can be computed using a Bayesian rule [5].

p^t (category = **Ci** | Number of times that the tracks associated with a specific emotion have been played) is the probability that the tracks that will be played belong to the emotional category **Ci** . It can be estimated by the preference distribution **D(u,t)** observed in time period **t**. As more tracks are associated with a given topic category, more new tracks will be analyzed and placed in that category.

Normalization

To normalize the feedback entered by the user using the emotion button, we apply the Batch normalization method. Due to the fact that multiple of terms may refer to the same emotion, we use Batch to find the most common term that refers to a specific emotion triggered by a track. Moreover, to analyze the emotional category of tracks, we use Robert Thayer's traditional model of mood [11] and use it in a method of music mood classification.

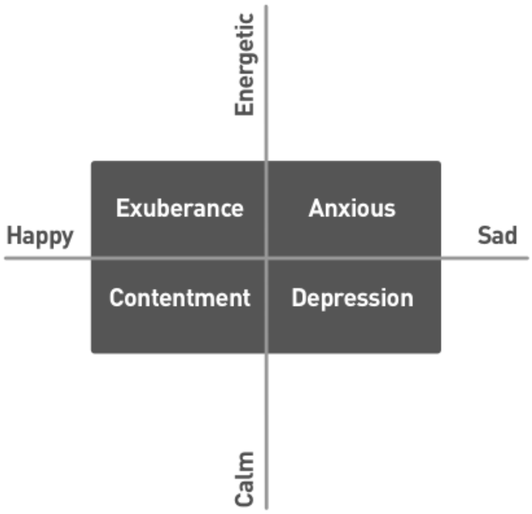

Image 12: Thayer's mood model [11]

The eight categories created by Thayer's model include the extremes of the two lines as well as each of the possible intersections of the lines. The model can be seen in image 12. Using the combination of Robert Thayer's model and the PANAS-X emotion descriptors as mood labels for music [7], the system can automatically organise the tracks according to emotion, generate mood playlists, and recommend tracks that reflect user's emotional preferences.

COMBINING PREDICTIONS OF PAST TIME PERIODS

To accurately gauge the user's genuine emotional interests, we combine the predictions made over multiple time periods. The more emotions we have associated with tracks, the better the prediction is going to be [5].

$$interest(category = c_i) = \frac{\sum_t \left(N^t \times interest^t(category = c_i)\right)}{\sum_t N^t}$$

$$= \frac{\sum_t \left(N^t \times \frac{p^t(category = c_i \mid \#tracks)\, p^t(\#tracks)}{p^t(category = c_i)} \right)}{\sum_t N^t}$$

PREDICTING USER'S CURRENT EMOTIONAL INTEREST

The user's emotional interest is decomposed into two parts: the genuine emotional interest and the influence of the country's emotional preferences.

We use the emotional preferences of the general public in a short current time period. The emotional preferences have been associated with tracks by other users that belong to the same country. Because of the large number of users, there are enough clicks in the short current time period to accurately recommend new songs that belong to the same emotional categories [5].

TRACK RECOMMENDATION

In order to rank the list of candidate tracks to be recommended, the system generates a content-based recommendation score and a collaborative filtering score for each track [5].

The content-based recommendation score of a track is based on the emotional category of that track and the predicted user's emotional interest.

The collaborative method implemented in [5] computes the collaborative filtering score. The two scores are combined in ranking the candidates for track recommendation using the approach described in [5]. Combining the content-based method and the collaborative method offers the advantages of both methods.

CITATIONS

1. Billsus, D. and Pazzani, M. J. A hybrid user model for news story classification. In Proceedings of the Seventh International Conference on User Modeling. 1999.

2. Billsus, D. and Pazzani, M. J., User Modeling for Adaptive News Access, User Modeling and UserAdapted Interaction, v.10 n.2-3, p.147-180, 2000.

3. Carreira, R., Crato, J. M., Gonçalves, D., Jorge, J. A. Evaluating adaptive user profiles for news classification, Proceedings of the 9th international conference on Intelligent user interfaces, 2004.

4. Chen, C. C., Chen, M. C., Sun, Y. PVA: a self-adaptive personal view agent system, Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, 2001.

5. Jiahui Liu, Peter Dolan, and Elin Rønby Pedersen. 2010. Personalized news recommendation based on click behavior. In Proceedings of the 15th international conference on Intelligent user interfaces (IUI '10). ACM, New York, NY, USA, 31-40. DOI: https://doi.org/10.1145/1719970.1719976.

6. Martijn Millecamp, Nyi Nyi Htun, Yucheng Jin, and Katrien Verbert. 2018. Controlling Spotify Recommendations: Effects of Personal Characteristics on Music Recommender User Interfaces. In Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization (UMAP '18). ACM, New York, NY, USA, 101-109. DOI: https://doi.org/10.1145/3209219.3209223

7. Watson, D., & Clark, L. A. (1994). The PANAS-X: Manual for the Positive and Negative Affect Schedule-Expanded Form. Ames: The University of Iowa.

8. Baum, Doris. "EmoMusic – Classifying Music According to Emotion." (2006).

9. Spotify for Developers at https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

10. Spotify's Collaborative Filtering by Soren Nelson available at http://www.math.utah.edu/~gustafso/s2018/2270/projects-2018/submittedprojects/sorenNelson Spotify's%20Collaborative%20Filtering.pdf

11. Thayer, Robert E. (1998). "The Biopsychology of Mood and Arousal". New York, NY: Oxford University Press