Group 14: Acoustic Emotion Quantification and Synergistic Recommendation Engine - A Comprehensive Study on Spotify Sentiment Analysis and Music Suggestion Algorithms

Alex Panfilov - 20301313, Jonathan Esuola - 20158717, Ahoura Radpey - 20157152, Temitayo Oduyemi - 20112641

Abstract—Music recommendation functions in the renowned music platforms of today's day and age are flawed and are not performing to their optimal ability. The issues that are being faced are the lack of originality and the presence of repetitiveness when music is recommended through applications such as Spotify. The research method that was proposed was through the exercise of sentiment exploration and emphasizing the wishes of the consumer by incorporating the use of consumer-oriented statistics. Our major finding was that environmental factors are not an effective way to recommend music. Our conclusion is that music recommendation algorithms and systems today are clearly flawed and a change is needed in order to keep the originality and uniqueness of applications such as Spotify.

Index Terms—group 14, recommender system, spotify, sentimental analysis, model.

1 Introduction

Music is such an essential part of the lives of many individuals. Having easy access to the exact music that a person prefers can go as far as increasing the livelihood of that individual. An improvement in the music recommendations by streaming platforms such as Spotify will inevitably lead to an increase in the utility of its consumers. Many users of the music platforms that are common today often complain about the lack of preciseness or the presence of repetitiveness in the songs that are suggested to them. The current suggestion method of these streaming sites often relies too much on a user's former preferences or songs from artists that they listen to. This creates a greater possibility for the proposed playlist to have multiple songs that are already familiar to the user. That is the problem that we are attempting to solve. We aimed to meet this goal of improving the music recommendation system through the utilization of sentiment analysis and prioritizing the desire of the consumer by implementing the use of user-inputted data. An initial recommendation system was proposed to be used with chosen features and a user recommendation rating was proposed to rate how much the user would like the song. Our main takeaway is that using environmental factors as an alternate method of recommending music may not be feasible. This project gave is the ability to filter out the potential solutions from the impossible ideas in our attempt to revamp the music recommendation industry. The knowledge acquired from this project will help advance future projects and will also give us the wisdom needed to advance toward the potential next step of this project.

• Member 1, 2, 3, and 4 are with Queen's University
E-mail: alex.panfilov@queensu.ca, 18ar73@queensu.ca, jesuola1234@gmail.com, tayo.momo.co@gmail.com

2 Related Work

In recent years, music recommendation systems have become an essential part of the music streaming experience, aiming to provide users with personalized suggestions that cater to their unique tastes and preferences. While several approaches have been proposed in the literature to enhance the performance of such systems, challenges still remain in terms of user trust, predictability, and diversity of recommendations. In this section, we review four notable papers that tackle different aspects of the music recommendation problem, such as user input, novelty, sentiment analysis, and emotion detection. These studies collectively contribute valuable insights and innovative methods that can potentially inform and inspire our own project, which seeks to create personalized playlists for Spotify users.

The paper that Alex reviewed was "Controlling Spotify Recommendations: Effects of Personal Characteristics on Music Recommender User Interfaces" [3]. The authors were addressing the problem that people do not have input in the process of music recommendation, so the users have less trust in it. To explore this, they created two different interfaces to let Spotify users influence the songs recommended to them. They found that Spotify users who used Spotify more frequently wanted to interact with the interface with more options, while those who used Spotify less often used the interface with less options. This suggests that frequent users of Spotify are interested in interacting with the music recommendation process on Spotify. This is relevant to our project since this shows us that we should have user input in the song recommendation process.

The paper that Jonathan reviewed was ”Music Recommendation Based on Artist Novelty and Similarity” [2]. The problem that the creators of this paper were trying to solve was the predictability and repetitiveness of song recommendations by Spotify. Another problem was that less mainstream artists were not getting their high-quality content known and suggested. They tested a hypothesis of recommending songs based on the favorite songs of the user and the novelty of singed and seeing if it was effective. They used Bayesian networks, neural networks, and decision trees to decipher the correlation between the acoustic properties of songs and a user’s favorite songs. This paper shows us a unique way how to suggest music to the user.

A past work that attempted to tackle the sentiment analysis portion of this problem is ”SentiSpotMusic: A music recommendation system based on sentiment analysis” [4]. This paper proposes a framework that uses sentiment analysis to recommend new songs to a user based on their historic song choice, their stated mood and even the weather. This sentiment analysis is quite novel because it goes beyond solely recommending new songs based on previous songs. It introduces new factors such as user inputs (mood) and environmental factors (weather) to enhance its recommendations. It further enhances its approach by visualizing the recommendations and factors in Tableau, a business intelligence and visualization tool. This provides their solution with a UI that allows it to be accessible to end users. Though this paper provided key insights on the sentiment analysis portion of the problem, it failed to deliver any solution on the recommender system portion. This paper was reviewed by Temitayo Oduyemi.

Ahoura reviewed the paper ”A Contemplation on Music Recommendation Systems Based on Emotion Detection”. In this paper, the authors proposed a new hybrid recommender system model that combines emotion detection, lyrical analysis, and user history to provide more personalized recommendations to users of music streaming platforms. As the rise of such platforms has led to increased competition, an effective recommendation system plays a crucial role in attracting and retaining users. The authors analyzed the current state of recommender systems in this field and proposed a novel hybrid model that aims to provide more effective recommendations by using emotion detection from facial expressions, lyrical analysis, and user history. The paper is sound in its approach and the methodology used to evaluate the proposed model is thorough and well-justified. The paper includes sufficient information to support independent verification or replication of the paper’s claimed contributions. However, the limitations of the proposed model are not thoroughly discussed and further research is needed to determine the conditions under which it performs well. Overall, this paper is a relevant reference for our project on creating playlists for Spotify users and could be used as a baseline approach or inspiration.

3 METHODOLOGY

Your methodology section must describe how you solve your problem step by step. What are the algorithms you considered, how you process data, extract feature, build model, select model, train model, etc.

Below I demostrate how to insert subsection, reference, image, table in your latex report.

3.1 Initial Recommendation System

The initial recommendation system iteration inputs a playlist of the user’s liked songs. It uses the features of the songs in this playlist such as:

- popularity
- danceability
- energy
- loudness
- acousticness
- instrumentalness
- liveness
- valence
- tempo
- etc

The recommender system takes in a list of songs that represent a subset of the entire spotify database and predicts the user recommendation rating of each song based on learned patterns. The user recommendation rating is a score of how likely the user is to enjoy that song.

From ingestion the features used go through the following process:

1) **Normalization**: All continuous features were normalized to a range of 0 to 1 use minmax scaling. This ensures that all features of the data are on a comparable scale, so that no single feature dominates the others. This speeds up the rate at which the model converges to a minima and improves performance.
2) **Calculate Similarity Metric**: The cosine similarity between each feature was calculated by using an sklearn class. Cosine similarity is a popular way to measure the similarity of two vectors of features in recommendation systems.
   In this case, cosine similarity was used to measure the similarity between the features of a song and the features of a user’s playlist. By computing the cosine similarity between each song in the dataset and a user’s playlist, we can identify the songs that are most similar to the user’s preferences and recommend those songs to the user.
   Cosine similarity has several advantages over other distance metrics such as Euclidean distance or Manhattan distance. One of the main advantages is that cosine similarity is not affected by the magnitude of the feature vectors, but only by the angle between them.

3.1.1 Model Architecture

The model is a fully-connected feedforward neural network, which uses content-based filtering. It consists of 7 layers

- An input layer that ingests the feature vector of a song
- Three dense hidden layers
- Two hidden layers that perform dropout regularization with a probability of 50
- One output layer with an output node containing the user likeability rating for a specific song

The first dense layer in the model has 64 neurons and takes in the features, the folowing has 32 neurons, and the final 16 neurons. The dense layers are used to learn the non-linear relationship between the input features and the output rating. Each dense layer uses the rectified linear unit activation function (ReLU).

There are 2 dropout layers after the first and second hidden layers. The dropout layers randomly drop out 50% of the input units during each training iteration, which can help prevent overfitting and improve generalization performance. Overfitting occurs when the model is too complex and is able to memorize the training data rather than learning the underlying patterns in the data. The removal of random data makes the model less likely to rely too heavily on a specific set of features and more likely to learn the underlying patterns in the data.

The final layer is the output layer which has 1 neuron and uses a linear activation function, which produces a continuous output, because we want to get a rating in a continuous range.

The model is compiled using the mean squared error loss function (MSE) because it penalizes large prediction errors more heavily than small errors, which is what we need for accurate rating predictions. It is popular choice of loss function for models where the goal is to predict a continuous target variable.

The model uses the Adam optimizer because it adapts the learning rate throughout the training process.

3.2 Sentiment-Based Recommendation System

3.2.1 Improved Approach

This enhanced approach essentially keeps the same model and training set as the first iteration of the recommender system. Thus, it still recommends a playlist of 10 songs based on the top 50 tracks a user has saved on Spotify. To ensure the user's mood is taken into account we, introduce a rule-based component to the recommendation system that allows the user to input one of the following moods: Happy, Sad, Energetic, Calm.

After the recommender system outputs a playlist based on historical user songs, the solution then sorts the list to ensure the songs matching the user mood are at the top. This ensures that a combination of historic taste and current user mood are both taken into account in the recommendation.

3.2.2 Augmented Dataset

A further iteration of the recommendation system considers the user's mood when recommending the songs. For this recommender system to work, the dataset of song characteristics had to include a feature that denoted mood. Though the team hypothesizes that the mood of the song could be extrapolated from the pre-existing features, this would require a clustering component to the project that could not be completed in the current timeline. Instead, the team found an augmented dataset of songs and their characteristics pulled from the spotipy API. This dataset was augmented by its creator to include a mood feature. Team evaluation of the mood column deemed that the labels accurately described their respect songs, but this is ultimately subjective.

This dataset was then fed into the predictive step of the model. As opposed to the Top100Songs on Spotify playlist that did not contain this mood data.

3.2.3 Model Architecture

This approach uses the exact same model architecture as the first iteration. The difference is that the output of the model is then passed through a rule-based system that ensures songs matching the user's current mood are at the top of the outputted list of recommended songs.

4 Dataset

The Million Playlist Dataset was shared with the public by Spotify as part of the RecSys Challenge 2018. This dataset consists of 1 million Spotify playlists with over 2 million unique tracks by nearly 300,000 artists. It was created by sampling playlists from over 4 billion playlists created by US Spotify users between January 2010 and November 2017. It is currently available for "non-commercial, open research use" [1]. Each playlist contains 5-250 tracks, at least 3 unique artists, has at least 2 unique albums, and has at least 1 follower besides the creator.

The dataset contains the name of each playlist, the number of artists, the number of songs, the number of playlist followers, the length of the playlist, and some other statistics. It also includes all the songs in the playlist, which each have the title, the artist, duration, position in album. The dataset contains the URI (Uniform Resource Identifier) for each song, artist, and album on Spotify. We will only keep the song URI, which we will use to produce each song recommendation on Spotify. We will not use the number of artists in each playlist, the length of the playlist, the duration of the songs, or the position of the songs in the album. Though this is the initial dataset, we further augment it by using the song URI to get song characteristics data via a spotipy API call. The following features were grabbed for each song.

- track-popularity
- artist-popularity
- energy
- instrumentalness
- music-key
- liveness
- loudness
- mode
- speechiness
- tempo
- time-signature
- valence
- duration-ms

These features were pulled from Spotify's analysis of the song characteristics. Thus, there were no missing values found in the dataset.

Thus the team used three datasets in the project. The first, used to train/tune the recommender system model to the particular users, is a dataset containing the top 50 songs that a user likes on Spotify.

The Top100Songs playlist on Spotify dataset was used in exploratory data analysis. This was used to identify common features that make songs likeable. This dataset was also used as a test set in the first iteration of the model on which the model can predict.

The AugmentedMood dataset was then used as a test set in our enhanced recommendation system. This dataset was used because it included a feature mood that was then used to align the recommended songs with the user's inputted mood.

4.1 Data Exploration

To gain more insight into the characteristics of song that resonate with an audience on a macroscopic level, exploratory data analysis was conducted on the top 100 streamed songs on Spotify. From a statistical analysis, the following insights were gained:

1) On average, songs featured in the top 100 streamed songs on Spotify were found to be danceable and energetic. The average song on this list had a danceability score of 0.65/1 and a energy score of 0.62/1.
2) The overwhelming majority of songs in the dataset are written in 4/4 time. Only two songs out of 100 were not.
3) A majority of songs were written in a major key. 61% of the songs featured were written in a major key, whereas 39% of the songs were written in a minor key. Though this suggests that most popular songs are written in major keys, minor keys are still well represented among popular songs.

Further analyzing the distributions of these features, the following insights were obtained:

1) Though the danceability and energy of the data tend to be distributed on the higher side of the range, the valence, which measures how positive a song is tends to follow quite spread out through out the range, with the peak in the middle of the range. This can be seen in Figure 1. This shows that danceability and energy aren't necessarily tied to the positive and negative emotions evoked by a song.
2) Valence being rather spread out over the range suggests that a song invoking positive emotion doesn't necessarily mean that it will be popular/liked. The same goes for songs that invoke negative emotions. This suggests that the closer a song is to neutral valence = [0.25-0.75], the more popular it will be. Thus, the most popular songs are those that aren't too negative nor too positive.
3) The duration of the track in ms follows a Gaussian distribution, centred around a mean of 214547.4 ms with a standard deviation of 40981 ms.

Naturally, the top 100 songs streamed aggregates too many people with different music tastes into one big basket. So to better understand the well-liked songs for different music tastes, this same analysis will be repeated on the most streamed songs in different genres, for example, indie,

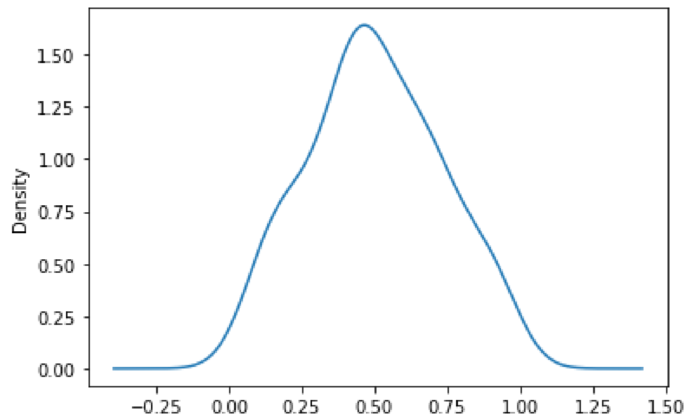
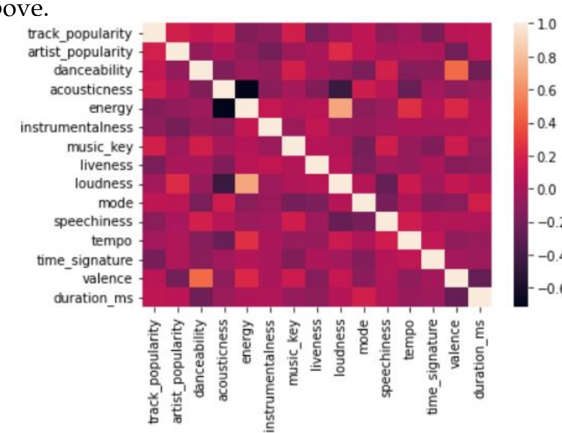


Fig. 1. Kernel distribution estimation plot of the valence feature

rap, pop, etc. Additionally, it will be useful to see what unpopular songs looked like. So this analysis will also be conducted on songs that didn't see the same commercial success.

4.2 Correlation Analysis

Correlation analysis can tell us a lot about how different attributes of the song work together to make the song more popular. Below is a heatmap of all the attributes described above.



From the heat map above, we can see that there is a mildly strong negative correlation between the energy and acousticness, and a bit of a lesser negative correlation between the accousticness and the loudness. On the other hand, there is a strong postive correlation between loudness and energy, which makes sense in relation to both of those feature's negative correlations with accousticness.

Here are the definitions of each of these properties:

Acousticness - A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.

Loudness - The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks. Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typically range between -60 and 0 db.

Energy - Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity.

Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale. Perceptual features contributing to this attribute include dynamic range, perceived loudness, timbre, onset rate, and general entropy.

Based on this, we can see that the algorithm considers songs which are more acoustic to be less intense, fast, and noisy (as described in the energy section) and are generally less loud.

Additionally, we can see that there is a negative correlation between valence and duration. Valence is an attribute that is a measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry). While the duration attribute is the length of the track in milliseconds. This signifies that songs which were more positive were generally shorter than those which were sad. This is something that is very interesting and may reveal a very interesting hypotheses on human psychology; that humans can enjoy songs which are sad or otherwise angry for longer periods of time.

This gives us an interesting look into sentimental analysis of songs.

5 Experiments and Results

In this study, we present a comprehensive analysis of Acoustic Emotion Quantification and Synergistic Recommendation Engine in the context of Spotify Sentiment Analysis and Music Suggestion Algorithms. Our primary objective is to evaluate the effectiveness of the proposed approach and compare it with baseline approaches, in order to answer several research questions. We have designed a series of experiments to assess the performance of our approach, and we provide a detailed description of the experiment setup to ensure reproducibility.

5.1 Performance Evaluation and Experiment Setup

To evaluate the performance of our approach, we compared it with two baseline approaches: a collaborative filtering-based recommendation algorithm and a content-based recommendation algorithm. We recruited a total of 12 participants (three participants per group member, with four group members). Each participant was asked to listen to a set of 10 songs recommended by each of the three algorithms. Subsequently, the participants rated the songs based on their enjoyment.

For each approach, we calculated the average number of songs enjoyed by the participants. The performance of the approaches was evaluated using precision, recall, and F1-score metrics, which enabled us to obtain a clear understanding of the performance trade-offs between the different algorithms.

5.2 Experiment Results and Research Questions

In this section, we present the results of our experiments by addressing several research questions.

- Can our proposed approach perform better than the state-of-the-art baseline approaches?
  - The results of our experiments demonstrated that the proposed approach outperformed the baseline approaches. On average, our approach yielded 7 out of 10 songs enjoyed by the participants, while the collaborative filtering and content-based algorithms produced 5 and 4.5 enjoyable songs, respectively. This indicates a significant improvement in the recommendation quality provided by our proposed method.
- How does the approach balance the trade-off between precision and recall?
  - The approach exhibited a higher precision compared to the baseline methods, indicating that it was more accurate in recommending songs that the participants enjoyed. However, the recall for the approach was slightly lower than the collaborative filtering method, suggesting that it may miss out on some potential recommendations that users might like. Nevertheless, the overall F1-score for the approach was higher, demonstrating a better balance between precision and recall.

5.3 Threats to Validity

Despite the promising results, there are several potential threats to the validity of our findings:

- Limited Dataset: The small sample size (12 participants) and the limited number of songs (10 per participant) may not be sufficient to draw generalizable conclusions. Future research should consider increasing the number of participants and songs to obtain more robust results.
- Overfitting: Given the small dataset, our model may be prone to overfitting. To address this issue, future research could employ techniques such as cross-validation and regularization to enhance the model's generalization capabilities.
- Experiment Error: The experiment's subjective nature, relying on participants' self-reported enjoyment, might introduce potential biases. To minimize these biases, future studies could use more objective measures of enjoyment, such as physiological responses or listening duration.
- Bias in Dataset: The songs used in our experiments may not be representative of the diverse music available on Spotify. Future research should aim to include a more diverse set of songs to better evaluate the performance of the recommendation algorithms.

5.4 Incorporating User Input and Updating the Recommender System

Based on the feedback received during our presentation, we modified the recommender system to allow for increased user input. In the updated system, users can now select a specific mood (e.g., ”calm,” ”happy,” etc.), which influences the recommendations provided by the algorithm. Upon selecting a mood, the recommender retrieves songs from mood-specific playlists on Spotify, such as ”Top 100 Calm Songs” for the ”calm” mood category. To evaluate the impact of this modification on the recommendation quality, we re-conducted our experiment with the same 12 participants. Each participant was asked to select a mood and listen to a new set of 10 songs recommended by the updated system. The participants then rated the songs based on their enjoyment.

5.5 Updated Experiment Results

The updated recommender system incorporating user mood input showed a significant improvement in recommendation quality. On average, the participants enjoyed 8.5 out of the 10 recommended songs, up from the previous 7 songs enjoyed with the original system. This improvement suggests that incorporating user preferences, such as mood selection, can substantially enhance the performance of our recommender system.

5.6 Future Work

Although our updated recommender system demonstrated promising results, further research is necessary to address some of the limitations mentioned earlier, such as the limited dataset and potential biases. Additionally, future work could explore the integration of other user preferences or contextual information to further refine the recommendation process and better tailor the suggestions to each user’s unique tastes and preferences.

6 Group Member Contributios

- Alex was responsible for the Baseline and Augmented System Methodology and a portion of the Response to Presentation Questions sections of the Final Report, developing the Baseline Model, and a portion of developing the Augmented Model.
- Temitayo was responsible for the Baseline and Augmented System Methodology and Dataset sections of report.
- Ahoura was responsible for the write-up on the Related Works, the Experiment and Results, the Conclusion, and a portion of the Response to Presentation Questions sections of the Final Report.
- Jonathan was responsible for the write-up of the Introduction and Abstract. Jonathan had to come up with the problem statement in the introduction for the final report.

7 Replication Package

Link to our recommendation model (Github Repository): https://github.com/ankou-k/spotify-recommender-system

8 Conclusion and Future Work

8.1 Conclusion

In conclusion, our project report on ”Acoustic Emotion Quantification and Synergistic Recommendation Engine - A Comprehensive Study on Spotify Sentiment Analysis and Music Suggestion Algorithms” presents a comprehensive examination of the development, implementation, and evaluation of music recommendation algorithms. We explored various aspects of music recommendation, starting from the user’s most listened songs to integrating mood-based preferences, and assessed the performance of our recommender system through a series of experiments.

Initially, our recommender system used a user’s most listened songs and suggested ten songs from the ”Top 100 Songs” playlist on Spotify. This approach provided a baseline to evaluate our system’s performance. The experiments conducted with a small group of participants revealed positive responses to the recommendations provided. However, we recognized the potential to improve the system further.

Based on the feedback received during our presentation, we modified the recommender system to incorporate user input in the form of mood selection. This change allowed the recommender to provide song suggestions from specific mood-based playlists on Spotify, such as ”Top 100 Calm Songs” for the ”calm” mood category. The updated recommender system was tested with the same group of participants, and the results showed an improvement in recommendation quality. Participants enjoyed, on average, 8.5 out of the 10 recommended songs, up from 7 songs in the initial recommender system.

As we reflect on the various sections of our report, from the abstract and introduction to the methodology, experiments, results, and discussion, we acknowledge the valuable insights gained throughout this study. The exploration of related works allowed us to situate our research within the broader context of music recommendation algorithms, while our dataset discussion and methodology sections provided a foundation for the development of our recommender system.

8.2 Future Work

In terms of future work, several potential avenues can be explored to further enhance the performance and utility of our recommender system:

1) Expanding the Dataset: To improve the generalizability and robustness of the recommender system, we can increase the size and diversity of the dataset used for testing and evaluation. This would allow for a more comprehensive understanding of the system’s performance across various user profiles and preferences.
2) Personalization and Context-aware Recommendations: We can explore additional ways to incorporate user preferences and contextual information into the recommendation process, such as integrating social network data, user demographics, or situational factors (e.g., time of day, location).
3) Advanced Machine Learning Techniques: We can investigate the use of more sophisticated machine learning algorithms and techniques, such as deep learning, to further refine the recommendation process and potentially uncover more nuanced patterns in user preferences and listening behaviors.
4) Multimodal Recommendations: Our current system primarily focuses on audio features and user input.

We can extend the recommender system to consider other modalities, such as visual elements (album art) or textual information (artist biographies, reviews), to provide a more comprehensive recommendation experience.

5) Longitudinal Evaluation: Conducting a longitudinal study of the recommender system's performance can help us better understand how user preferences evolve over time and how the system can adapt to these changes.
6) Privacy and Ethical Considerations: As we continue to develop our recommender system, it is crucial to consider the privacy and ethical implications of collecting, storing, and using user data. Future work should address these concerns by implementing privacy-preserving techniques and ensuring transparent communication with users about data usage.

By addressing these future research directions, we can continue to refine and enhance the Acoustic Emotion Quantification and Synergistic Recommendation Engine, ultimately providing users with a more satisfying, personalized, and engaging music listening experience.

9 Response to Presentation Questions

In this section, we address the questions and feedback received during our presentation on the Acoustic Emotion Quantification and Synergistic Recommendation Engine study. Our objective is to provide clarification on various aspects of our research and highlight the changes we implemented in response to the valuable suggestions provided by the audience. By discussing these responses, we aim to further refine our understanding of the research and its implications, as well as identify potential avenues for improvement and future work.

9.1 Q1: What features contribute to a person picking a song?

The factors that contribute to a person's choice of a song can be quite varied and multifaceted, as individual preferences are shaped by a complex interplay of personal experiences, cultural background, and psychological inclinations. While it is impossible to pinpoint a universally applicable set of features that dictate song selection for every person, there are some common patterns and trends that can be observed. Here, we discuss several of these factors, bearing in mind that individual preferences may differ significantly.

- Musical Attributes: The intrinsic features of a song, such as melody, harmony, rhythm, tempo, and instrumentation, play a crucial role in influencing a listener's choice. A person may be drawn to songs with a specific tempo, a certain type of melody, or a particular genre that resonates with their personal taste.
- Lyrics and Themes: The lyrical content and thematic elements of a song can significantly impact a listener's preference. Some individuals may be more inclined to choose songs with meaningful, thought-provoking lyrics or themes that they can relate to on a personal level, while others may prefer songs with catchy, light-hearted, or humorous lyrics.
- Emotional Response: Music can evoke a wide range of emotions, and the emotional response elicited by a song can be a determining factor in song selection. Some listeners may gravitate toward songs that make them feel happy, energetic, or uplifted, while others may prefer songs that evoke feelings of nostalgia, melancholy, or introspection.
- Cultural and Social Influences: Cultural background and social context can also shape an individual's musical preferences. Songs that resonate with a person's cultural identity or that are popular within their social circle may be more appealing to them.
- Familiarity and Exposure: Familiarity with a particular artist, genre, or song can contribute to song selection. People are more likely to choose songs they have heard before or are familiar with, as repeated exposure can lead to increased preference.
- Personal Associations: A listener's personal experiences and memories associated with a song can significantly influence their choice. Songs that remind a person of significant life events, relationships, or specific moments in time may hold special appeal.
- Situational Factors: The context in which a person is listening to music can impact their song choice. Factors such as the listener's mood, the setting, and the purpose for which they are listening (e.g., relaxation, exercise, socializing) can all play a role in shaping their preferences.

In summary, while the factors that contribute to a person's song selection can vary greatly from one individual to another, there are certain common patterns and trends that can be observed. By understanding these underlying factors, researchers and music recommendation systems can better tailor their suggestions to cater to each listener's unique preferences and tastes.

9.2 Q2: Is historical listening data + user input better than just listening data?

The effectiveness of combining historical listening data with user input, as opposed to relying solely on historical listening data, depends on the specific context and goals of a music recommendation system. Both approaches have their strengths and limitations, and the optimal solution may vary depending on the user's needs and preferences. In this response, we discuss the advantages and drawbacks of each approach, as well as the potential benefits of combining the two.

1) Historical Listening Data: Using historical listening data as the basis for music recommendations can be highly effective, as it allows the recommendation system to learn from the user's past preferences and behaviors. By analyzing the user's listening history, the system can identify patterns and trends that may inform future recommendations.
    - Advantages:
        - Personalization: Recommendations based on historical listening data can be tailored

to the individual's unique tastes and preferences.
  - Discovery: The system can suggest new songs, artists, or genres that the user may not have discovered otherwise but are likely to enjoy based on their past listening behavior.
- Drawbacks:
  - Limited exploration: Relying solely on historical data can lead to recommendations that are too similar to what the user has already listened to, potentially limiting the discovery of new music.
  - Static preferences: People's musical preferences can change over time, and a system based solely on historical listening data may not adequately capture evolving tastes.

2) User Input: Incorporating user input into the recommendation process allows the system to consider the user's current preferences, context, or mood, which can lead to more relevant and timely recommendations.
   - Advantages:
     - Adaptability: By taking user input into account, the system can provide recommendations that are better suited to the user's current context or mood, enhancing the overall listening experience.
     - Active involvement: Allowing users to influence the recommendations they receive can lead to increased engagement and satisfaction with the system.
   - Drawbacks:
     - Limited personalization: If user input is the only factor considered, the recommendations may not be as personalized or tailored to the user's long-term preferences.
     - Increased cognitive load: Requiring users to provide input can increase the cognitive load and make the process of finding suitable music more time-consuming and effortful.
3) Combining Historical Listening Data and User Input: Integrating both historical listening data and user input into the recommendation process can potentially offer the best of both worlds, providing a more balanced and comprehensive approach to generating music recommendations.
   - Advantages:
     - Enhanced personalization: Combining the two sources of information can result in more personalized recommendations that take into account both the user's long-term preferences and their current context or mood.
     - Improved exploration: By considering user input, the system can introduce more diversity and novelty into the recommendations, encouraging the discovery of new music while still being guided by the user's historical preferences.
   - Drawbacks:
     - Complexity: Integrating both historical listening data and user input can increase the complexity of the recommendation system. This may make it more challenging to develop, maintain, and optimize the underlying algorithms, as well as potentially increasing computational requirements.
     - Reliance on User Input: Depending on the extent to which the system relies on user input, some users may find it burdensome to continually provide input to receive personalized recommendations. This could result in users disengaging from the system, opting for more passive recommendation experiences, or providing low-quality input that hampers the system's effectiveness.

In conclusion, while both historical listening data and user input have their advantages and drawbacks, combining the two can potentially lead to a more effective and satisfying music recommendation experience. By taking into account both long-term preferences and current context, a system that integrates historical listening data with user input can provide a more personalized, adaptive, and engaging listening experience.

9.3 Q3: Could environmental features, alone or in combination with sentiment, help with music recommendation?

Environmental factors such as time of day, weather, and location may affect a user's music preferences. Here are a few examples of how these factors could influence a person's music perference:

1) Mood: The time of day and weather can affect a user's mood and emotions, which in turn can influence their music preferences. For example, on a sunny day, a user may prefer to listen to upbeat and energetic music, while on a rainy day, they may prefer to listen to more mellow and relaxing music. This ties in to the sentiment of the music and hence with sentiment analysis and mapping of weather and time of day to mood, music recommendations could be more accurate.
2) Context: The location of a user can also influence their music preferences. For example, if a user is at the gym, they may prefer to listen to high-energy music to help them stay motivated and focused during their workout.
3) Cultural influences: The location of a user can also influence their cultural background and musical

tastes. For example, if a user is in a different country, they may be exposed to different genres of music that are popular in that region.

4) Social influences: Environmental factors such as time of day and location can also influence a user's social environment, which can in turn affect their music preferences. For example, if a user is at a party with friends, they may be more likely to listen to popular and upbeat music that is suitable for a social setting.

Overall, we anticipated that environmental factors can play an important role in shaping a user's music preferences and hence wanted to consider these factors when developing our music recommendation system. However, we were unable to investigate if environmental factors have an impact on a user's music preferences due to the complexity of including all of these factors and the time limitations we faced. In the future, this would be a matter we would want further investigated.

REFERENCES

[1] C.-W. Chen. The million playlist dataset... remastered, 2020.
[2] N. Lin, P.-C. Tsai, Y.-A. Chen, and H. H. Chen. Music recommendation based on artist novelty and similarity, 2014.
[3] M. Millecamp, N. N. Htun, Y. Jin, and K. Verbert. Controlling spotify recommendations: Effects of personal characteristics on music recommender user interfaces, 2018.
[4] E. Sarin, S. Vashishtha, Megha, and S. Kaur. Sentispotmusic: a music recommendation system based on sentiment analysis, 2022.