

Initial Plan- Musical Sentiment Analysis

By Timothy Tismo-Capili

Supervisor: Federico Liberatore

CM3203 - One Semester Individual Project - 40

CARDIFF
UNIVERSITY

PRIFYSGOL
CAERDYDD

Project Description

Music has become an essential part of many people's lives. People's feelings are often expressed through the art of song and those that listen to their feelings are often connected with them mentally and spiritually. The songs that a person listens to quite often reflects how that person has been feeling lately which is why looking at what songs, artists, and albums a person has been listening to is a great way to identify how a person has been feeling recently. Sometimes, a person's mood could change drastically within a day or a week and this would also mean a change in music. For my project, I want to develop an application that will analyse how they are feeling and represent this data in the form of graph and statistics.

Spotify being one of, if not, the most popular music streaming service with more than 286 million monthly users as of 19 February 2020. Other popular music streaming platforms like Apple Music (over 60 million monthly subscribers) and Amazon Music (over 55 million monthly subscribers) pale in comparison to Spotify in terms of the number of monthly active users.^[1] One of the biggest reasons for this is because of Spotify's music discovery and user recommendation system. Since Spotify is the most popular of the music streaming platforms, I plan to use Spotify to retrieve the type of songs that a user listens to and then analyse those songs to determine how they are feeling.

In a way, Spotify already somewhat achieves sentiment analysis by recommending the user songs and playlists that are similar to what the user already listens to. For example, if a user were listening to a lot of sad and slow songs recently, it would recommend songs and playlists that are also 'sad-sounding'. This, however, does not explicitly show the user how much they have listened to sad songs recently should they ever be curious.

Project Aims and Objectives

I intend to develop my project as a web application instead of a mobile application. The reasons for this are because I wanted it to be accessible to on any device that supports HTML5 and it is generally a lot less time consuming developing a web application.

The Spotify API allows developers to get track info and user info. To analyse how the person is currently feeling, the most accurate piece of data to use for this would be the songs that the user has played most recently. Fortunately, Spotify allows us to retrieve this information as well as the user's top tracks/artists and their most recently liked songs which could all be used to further analyse that user's mood.

For the analysis, the track info may be used. Each song on Spotify is assigned an attribute and given value to represent that attribute. From Spotify's API documentation, some of the attributes that I deem relevant to my project are:

- *Danceability: Danceability describes how suitable a track is for dancing*
- *Energy: Energy represents a perceptual measure of intensity and activity.*
- *Instrumentalness: Predicts whether a track contains no vocals.*
- *Loudness: the overall loudness of a track in decibels (dB).*
- *Valence: A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track.*

- *Tempo*: The overall estimated tempo of a track in beats per minute (BPM).

The audio feature that is of most relevant for the analysis is valence since it measures how positive or negative a song is. The other features may be considered for use when analysing the sentiment of a song to enriching.^[2]

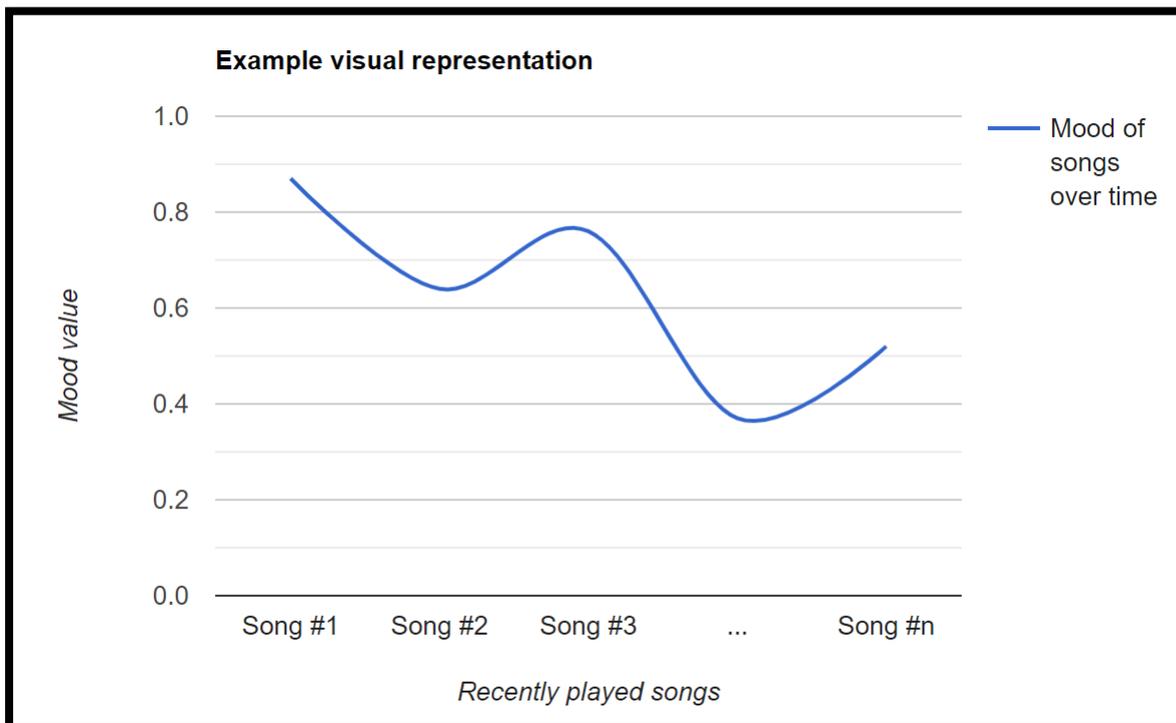


This picture shows the different feelings that can be inferred from the tempos.^[3] As we can see, any type of tempo can infer either positive or negative emotions which is why it is best to use multiple audio features to find the exact emotion that the song is conveying. For example, a higher tempo may imply that a song could mean happiness or excitement, but it could also convey anger as well. This is where we use the values from Energy, Loudness and Valence to determine and specify exactly what emotion or emotions the song is conveying.

Along with the audio features, the sentiment analysis of a song will be further enriched by using natural language processing. My intentions with this are to retrieve the lyrics of a song and then use NLP to examine each lyric of the song and determine each lyrics sentiment in context with the rest of the song. Since songs on Spotify do not come with lyrics, Genius API will be used to gather lyrical data. Genius is a digital media company with a website that houses a large collection of song lyrics. On top of that, the website allows users to provide annotations and interpretations of a song's lyrics. Genius' API allows us to retrieve the lyrics and the annotations that come with them. The annotations of the lyric will be examined to further enrich the analysis of the song's sentiment. I feel that Genius is a better choice over other popular lyric sites such as Musixmatch which also has its API that retrieves lyrics. The API for Musixmatch only lets users view the first 30% of a song's lyrics with the free version and it also does not provide annotations for song lyrics since that is a Genius exclusive feature.^[4]

Part of the project will be devoted to researching which will NLP model will be best suited to achieve the aims and objectives. I would need to search for models that are can be used with Python and are specifically designed for sentiment analysis. This is something I will also need to discuss with my supervisor.

Once I have gathered the data and conducted an analysis on the user's music taste, I will then present the data back to the user through a line graph that shows the 'mood values' of the songs that they have most recently listened to. The song's 'mood value' will be a mixture of the valence value that was provided by Spotify and a value that will be calculated from running a sentimental analysis on the lyrics of the song.



This is an example of what the data will look like when presented to the user. The horizontal axis will have the name of the song up to the n th song which will be the latest song played by the user. The songs will be displayed in chronological order. The vertical axis will have a corresponding mood value. The higher the mood value, the higher the positivity of a song.

Following this, the application will also display a prediction of what the user's mood is. It will calculate the percentages of the different moods that are conveyed from the songs that the user listens to and will show the most prominent mood followed by a list of the other moods that were conveyed.

Ethical Considerations

The ethical implications and issues that may come from the development of this application will need to be considered.

Firstly, to protect the user's privacy, their credentials and data will not be stored. All their data is stored on Spotify and we should only be retrieving the data that is needed and should only be retrieved when it is needed. The application should discard their data once it is no longer needed i.e. when the user logs out of the application or when the user stops using the application.

I will need to make sure that any data that is collected for my projects needs to be collected ethically. As I am collecting data from users, I would need to make sure that the users I am collecting from stay anonymous.

Work Plan

During the development process, I plan to keep a developer diary that will be updated frequently to keep track of where I am with the development.

With both the Spotify API and the Genius API, I would need to conduct a bit of research and read into the documentation of both APIs to learn how to work with the two. I will also need to register my application with both of them so that I have access to their respective APIs.

I would also need to research further the attributes that Spotify provides for each song. To be more specific, how are the values of the attributes determined and how reliable they are. To do this, I would need to examine the Spotify API documentation and I may also have to look at articles and other projects that talk about these attributes.

Another thing that I would need to research is machine learning in general, the core principles of machine learning and how it would work with my application. This is because sentiment analysis is a sub-category of machine learning partially due to the use of natural language processing. Doing so I would allow me to fully understand the implementation of sentiment analysis in my project.

To test the sentiment analysis of my application, I would need to do some research on some datasets that I could use on websites like Kaggle and more. I would use this along using my data on Spotify.

I also plan to set up fortnightly meetings with my supervisor to ensure that I remain on track and to discuss certain aspects of my project.

Week 1

- Read up on the documentation of Spotify and Genius API to familiarise myself with its syntax and to help me plan the methodology
- Register my application with both Spotify and Genius for access to their respective APIs
- Look into what web application framework I want to use
- Start learning about the key concepts of machine learning and how exactly it will be integrated into my project

Week 2

- Continue to research the key concepts of machine learning
- Continue to read up on the documentation for Spotify and Genius API whilst also experimenting with them in Python. At this point, I will use my data to experiment with first
- Start to build the web application with the selected framework
- Read articles about how a person's emotions and the songs that they are listening to correlate

Week 3

- Look into the different natural language processing models that can be used for sentiment analysis and discuss this with my supervisor
- Continue to build the web application. The basic structure of the web application should be ready at this point
- Start collecting some datasets and experimenting with them
- Complete reading up on machine learning

Week 4-5

- Analyse my own music dataset and try to determine what emotions each song conveys using the Spotify attributes. Some further research may need to be done to help with this
- Start to experiment with the chosen NLP model in Python

Week 6-7

- Work with the other datasets that were collected
- Use the chosen NLP model to conduct a sentiment analysis on the lyrics of each song and the annotations of those lyrics. This should enrich the analysis further.

Week 8-9

- Begin integrating the analysis functionality into the web application itself.
- Login functionality for application should be implemented
- Use the datasets to test the accuracy of the analysis function and I will also get other users to test it

Week 10-11

- Work on the front-end of the application. The aim of this is to make the application more aesthetically pleasing
- Implement how the results will be displayed to the user
- Start to write the final report

Week 12

- Finish writing the final report

References

1. Musically.com. 2021. How many users do Spotify, Apple Music and streaming services have?. [online] Available at: <<https://musically.com/2020/02/19/spotify-apple-how-many-users-big-music-streaming-services>> [Accessed 8 February 2021].
2. Developer.spotify.com. 2021. Web API Reference | Spotify for Developers. [online] Available at: <<https://developer.spotify.com/documentation/web-api/reference/#objects-index>> [Accessed 8 February 2021].
3. Flat - Blog. 2021. How to project emotions through music: setting the tempo. [online] Available at: <<https://blog.flat.io/setting-the-tempo/>> [Accessed 8 February 2021].
4. Developer.musixmatch.com. 2021. FAQ. [online] Available at: <<https://developer.musixmatch.com/faq>> [Accessed 8 February 2021].