Music Recommendation based on Sentiment Analysis using NLP techniques

Annajiat Alim Rasel
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
annajiat@gmail.com

Farah Binta Haque
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
farah.binta.haque@g.bracu.ac.bd

Md. Shakibul Alam
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
shakib20000605@gmail.com

GM Mohaiminuzzaman Apurbo
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
apurbogm1@gmail.com

Md Rishat Sheakh
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
rishatsheakh@gmail.com

***Abstract*—In a world where the consumption of digital music is highly prevalent, personalized music recommendation systems are pivotal for enhancing user experience. This paper presents an approach using Natural Language Processing (NLP) techniques to include sentiment analysis into music recommendations, presenting an innovative methodology. The research entails the analysis of an extensive dataset that includes lyrics from many genres. Sentiment analysis is used to extract subtle emotional indicators from the lyrical material, capturing moods that range from pop's enthusiasm to the blues' melancholy. Furthermore, natural language processing (NLP) methods are utilized to examine the semantic organization and stylistic components included in the lyrics, offering a more profound comprehension of the lyrical composition in diverse genres. We create a model that combines sentiment and linguistic data to precisely categorize songs into different genres using machine learning and NLP methods. We compare our approach's efficacy with conventional genre classification techniques, showing how sentiment-aware lyric analysis can enhance classification accuracy and reveal nuanced aspects of a genre. This study adds to the growing body of research on music classification by emphasizing the role that lyrical content has in identifying genres. The results open the door to more complex and emotionally sensitive genre categorization systems in addition to demonstrating the value of sentiment analysis and natural language processing in enhancing music classification models. The consequences encompass a broader comprehension of the emotional terrain within musical genres, individualized playlist building, and music.**

***Index Terms*—NLP, Vader, TextBlob, Naive Bayes, Music, Sentiment Analysis**

I. Introduction

Human emotions play a role in shaping perceptions of various aspects including experiences, topics, people, and events. And all these aspects show some nature that varies from one another. Sentiment analysis is one way to determine whether one of those aspects gives rise to positive or negative feelings. Music is one of the important elements that influence the mood of its listeners. In the field of the music industry, it is important to understand the tone of a lyric or a piece of text and deliver it to the right audience. Since the industry is growing rapidly and the taste of music among the audience is changing gradually, the music industry also needs to keep pace with it. So it became important to analyze music streaming to see whether it is reaching the target audience or whether the market is growing or not. Sentiment analysis is a valuable tool of machine learning that evaluates the music by looking at its lyrics and coming up with insights that give ideas between positive and negative. It takes the customer's opinion and feelings into consideration and gives recommendations accordingly. Thus, sentiment analysis determines the target audience by looking at the lyrics and reviews given by the audience and recommending music according to them. This work aims to use many Natural Language Processing (NLP) methods to music lyrics in order to determine the effects that various genres of music have on listeners. The main goal is to forecast the sentiment correlated with the song lyrics by applying the naive Bayes model. The study intends to shed light on the complex emotional and thematic aspects hidden in music lyrics and their potential impact on listeners by utilizing a variety of NLP approaches. The last objective is to determine how well the naive Bayes model captures and interprets emotions in the context of music. By means of this investigation, the study hopes to provide significant understanding into the relationship between linguistic elements in song lyrics and the feelings they arouse in listeners.

II. Related Works

Widiyaningtyas et al. [1] performed studies that analyzed hotel reviews' sentiment using Naive Bayes and N-gram algorithms. They utilized the word-weighting approach TF-IDF because to the textual nature of the dataset. They used the Naive Bayes technique, which is commonly used to analyze the sentiment of hotel reviews, for the classification. Following the assessment of the unigram, bigram, and trigram, it was determined that the unigram yielded the highest accuracy value, standing at 100%. Equally 100% was the unigram's precision value. Bigram and Trigram, the other two models, produced accuracy, recall, and precision rates of 94%, 94%, and 95%, respectively, as well as 61%, 83%, and 67%.

Another study conducted by Mantoro et al. [2] From 2018 to September 30, 2021, several keywords like "Papua Merdeka," "Papua bagian Indonesia," and "Otsus Papua" yielded varied results when analyzing the sentiment of the Papuan movement on Twitter using a naïve Bayes algorithm. Opinion mining of Papua is the subject of this study. Using Twitter, hashtags, and filters, they salvaged data in the form of tweets. They discovered that 93% of positive tweets and 96% of negative tweets about the hashtag "Papua Merdeka" on Twitter were generated by the Naive Bayes multinomial algorithm, with a detection accuracy of 95%. With regard to an additional term, "Papua Bagian Indonesia," 94% of tweets were favorable, 88% were negative, and 94% were accurate. And with a degree of accuracy of 97%, the keyword "Otsus Papua" generated 90% unfavorable opinions and 98% good feelings. In the end, the model determined that a 94% accuracy rate was satisfactory for this sentiment analysis.

According to the study, many natural language processing (NLP) and other approaches struggled to identify the sentiment analysis procedure. One study examined the elements impacting customers' perceptions of various music streaming services by analyzing social media data using topic modeling and text regression [3]. Amazon Music, YouTube Music, Spotify, Deezer, and YouTube Music were the five most popular music streaming services in the US that were taken into consideration.

Liu et al. conducted another study [4] in which they proposed an LSTM-based method for optimal music selection. At the time of suggestion, it will take into account the user's emotional state.They created a model to represent the listener's emotional states and a way to analyze the user's mental status in this study. An example of a music-to-vector model would be any piece of music from the past that has been input. The results demonstrated the substantial societal and economic benefits of personalized music recommendation (PMR).

Sholihat et al. [5] performed additional research on the topic of sentiment analysis pertaining to illicit investments by utilizing a Naive Bayes algorithm with tweets as input. This study makes use of an orange data mining dataset retrieved from Twitter's social media platform by use of the Twitter API. The data utilized for training was 55% of the total, whereas the data used for testing was 25% of the whole. When looking for comments, they utilized a variety of keywords, including "illegal investment," "crazy rich," and others. Out of 571 comments, 152 were positive, 244 were negative, and 175 were neutral, according to the sentiment analysis.

Zul et al. [6] offered to present a paper at the conference that used K-means and the Naive Bayes algorithm to examine the sentiment present in social media. Their research relies on data extracted from Twitter's API and Facebook's Power Query, respectively. Using Frequent Itemset Mining (FIM), they preprocessed the data and picked the features. After ten rounds of testing with three different FIM values—2, 3, and 4—the 2 value yielded the best accuracy, at 82,500. Testing revealed that 80.526%-82.523% was the accuracy of the Naive Bayes algorithm without K-means, whereas 80.323%-81.523% was the accuracy of the combination of the two.

III. Methodology

In this section, we are going to explain the processes we have implemented using Natural Language Processing(NLP) throughout our research. We have used some techniques of NLP to train our dataset and to obtain our desired results. Our workflow can be seen below from the flowchart.

Here are the following steps of our working methods:

Step 1: *Data Preprocessing and Cleaning*

Preparing the data for analysis requires a critical and necessary step called data preprocessing. To make our data more usable and of higher quality, we organize and clean it. Inaccurate processing can produce biased outcomes and incorrect conclusions. The data preprocessing has been implemented through steps to enhance the quality and findability of our data. We have imported the required libraries and preprocessing the dataset on Google Colab according to the following processes:

- Dropping rows with null values
- Identifying the language of the lyrics using FastText
- Filtering only English songs
- Removing repeating songs
- Plotting popular artists in Top 100 Songs List
- Removing unnecessary characters from the lyrics
- Removing stop words and musical slangs from the lyrics
- Generating a word cloud from the lyrics

Step 2: *Vader*

VADER, an off-the-shelf sentiment analysis tool, is primarily designed for analyzing textual data. VADER excels in processing text that contains peculiarities such as slang, emoticons, and other casual language commonly seen in social media, unlike alternative methods of sentiment analysis. We have enhanced the TF-IDF vectors by using VADER's sentiment scores, hence incorporating sentiment-related attributes into our model. Through this combination, our model is capable of capturing both the emotional tone and content-based information of the text. The utilization of Vader in sentiment analysis connects lexical attributes to sentiment scores, which quantify the level of emotional intensity. The cumulative intensity of individual words in a text can be calculated to ascertain the sentiment score of the text. By employing this methodology, we generated three additional columns denoted as:

- sent_scores
- comp_score
- Sentiment

After adding together all the positive, negative, and neutral scores, the compound score is adjusted to fall somewhere between -1 (the most negative) and 1 (the most positive). Then, we titled and labeled the bar chart and determined the percentage of each sentiment. Next, we use VADER to get sentiment scores, and then we use scikit-learn's TfidfVectorizer to create TF-IDF vectors, which we then combine with sentiment scores. By integrating sentiment-related data (VADER) with content-based data (TF-IDF), a more complete picture of the text is produced. This combined approach has the potential to be useful in many NLP scenarios where understanding the tone as well as the content is paramount.

Step 3: *TEXTBlob*

A straightforward API for typical natural language processing (NLP) operations including parts-of-speech tagging, noun phrase extraction, sentiment analysis, and more is provided by the extBlob python module, which streamlines text processing and NLP tasks. Although there is no direct relationship between textblob and TF-IDF vectorization, textblob does assist in enhancing the TF-IDF model by incorporating extra analytical layers. For this reason, we have used TEXTBlob to extract key points from the texts. By integrating sentiment and content analysis, this set of techniques allows for a more comprehensive understanding of the text data (TF-IDF). Although a new column called sentiment score has been added, we will be focusing on polarity score for our work. The sentiment label and comparison of positive, negative, and neutral results are dependent on the polarity score.

Step 4: *Splitting the Dataset*

After separating the dataset into a training set and a test set, we vectorized the sets using TF and IDF. To make our work more reusable, we've included two libraries: sklearn naive bayes and sklearn.metrics. These libraries provide a straightforward way to convert raw text data into a format that our approaches can handle.

TF:

$$TF(t,d) = \frac{\textit{Total no. of terms in document d}}{\textit{No. of times term t appears in document d}}$$

IDF:

$$IDF(t,d) = \frac{\textit{log(no of documents containing term t+1}}{\textit{Total no of documents in the corpus Ns}}$$

The resulting vector denotes the significance of each term in a document relative to the entire corpus:

$$TF - IDF(t,d,D) = TF(t,d) \times IDF(t,d)$$

Step 5: *Naive Bayes*

Through the utilization of the training data, we have trained the Naive Bayes model. Particularly useful applications of Naive Bayes include text classification and spam filtering. As an appropriate evaluation method, Naive Bayes has been utilized to compare the actual and predicted sentiments in our labeled dataset. A number of common evaluation criteria are recall, accuracy, precision, and F1 score. Thus, we have ascertained the precision of the methods employed. Plotting the proportion of accurately predicted sentiments to the total number of sentiments in the dataset yields the accuracy.

Accuracy = Number of Correct Predictions/Total Number of Predictions

IV. Dataset

A. Dataset Description

The provided dataset contains information for the top 100 songs for the year 2012 to 2022. Every input in the dataset follows a regulated and structured format that complies with the accuracy requirements of professional data analysis procedures. The data is arranged tabularly to provide an easily navigable and effective framework for thorough investigation and well-informed decision-making in the field of music analytics. The dataset contained the scrapped lyrics of the Top 100 songs on Billboard Charts for the past 10 years. The dataset contains a null value. After pre-processing the dataset is reduced to 907 tuples. Important variables in this dataset include the song title, artist, official release year, recording or release year, lyrics status, full lyrics, and musical genre.

References

[1] Triyanna Widiyaningtyas, Ilham Ari Elbaith Zaeni, and Riswanda Al Farisi. Sentiment analysis of hotel review using n-gram and naive bayes methods. In *2019 Fourth International Conference on Informatics and Computing (ICIC)*, pages 1–5, 2019.

[2] Teddy Mantoro, Meita Merdianti, and Media Anugerah Ayu. Sentiment analysis of the papuan movement on twitter using naïve bayes algorithm. In *2021 IEEE 7th International Conference on Computing, Engineering and Design (ICCED)*, pages 1–5, 2021.

[3] Jaemin Chung, Jiho Lee, and Janghyeok Yoon. Understanding music streaming services via text mining of online customer reviews. volume 53, page 101145, 2022.

[4] Zhiyuan Liu, Wei Xu, Wenping Zhang, and Qiqi Jiang. An emotion-based personalized music recommendation framework for emotion improvement. volume 60, page 103256, 2023.

[5] Anisa Sholihat, Fahurachman Bei, Regina Ainaya, Falentino Sembiring, and Amy Lattu. Twitter tweet: Sentiment analysis on illegal investment using naive bayes algorithm. In *2022 IEEE 8th International Conference on Computing, Engineering and Design (ICCED)*, pages 1–5, 2022.

[6] Muhammad Ihsan Zul, Feoni Yulia, and Dini Nurmalasari. Social media sentiment analysis using k-means and naïve bayes algorithm. In *2018 2nd International Conference on Electrical Engineering and Informatics (ICon EEI)*, pages 24–29, 2018.