

# Song Recommendation Generator

MUTMA ADEBAYO, NANDANA DILEEP, SMRUTHI SRINIVASAN, and SELINA TANG

## ACM Reference Format:

Mutma Adebayo, Nandana Dileep, Smruthi Srinivasan, and Selina Tang. 2024. Song Recommendation Generator. 1, 1 (March 2024), 1 page. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

## 1 SUMMARY

We plan to analyze song lyrics across various artists in order to create a song recommendation generator. At a high level, we want to detect sentiment in text in order to give recommendations of other songs based on the lyrics of a song that the user likes. A user will input a song name as the query and our generator will return the the most relevant songs to the query based on lyrics.

## 2 INTRODUCTION: RELEVANT LITERATURE

Content based music recommendation system:

Olvera, K. Sood, C. Reyes and W. Tu, "Content-Based Music Recommendation System using Supervised Learning," 2023 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), Malang, Indonesia, 2023, pp. 120-127, doi: 10.1109/COMNETSAT59769.2023.10420756.

Music recommendation system using machine learning:

Gundabatini, Dr. Sanjay Gandhi and Bindu Sneha, R and Vamsi, M. and Yagna Yaswanth, N. and Krishna Vamsi, K., Music Recommendation System Using Machine Learning (May 13, 2023). International Journal for Innovative Engineering & Management Research, Vol. 12, No. 04, 2023, Available at SSRN: <https://ssrn.com/abstract=4447197>

Recent surveys on music recommendation systems:

Hidasi, B., & Tikk, D. (2013). Survey of music recommendation systems. In I. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), Recommender Systems Handbook (pp. 589-628). Springer US.

Liem, T., Vu, T. Q., & Nguyen, L. P. (2020). A literature review on music recommendation systems: Foundations, algorithms, and challenges. Artificial Intelligence Review, 53(5), 2593-2634.

Using unsupervised algorithms to generate song recommendations:

<https://www.enjoyalgorithms.com/blog/music-recommendation-system-using-ml>

Using python to create a music recommendation system:

<https://towardsdatascience.com/create-music-recommendation-system-using-python-ce5401317159>

Top recommendation algorithms: <https://itnext.io/what-are-the-top-recommendation-engine-algorithms-used-nowadays-646f588ce639>

Guide to recommendations systems:

<https://towardsdatascience.com/recommender-systems-a-complete-guide-to-machine-learning-models-96d3f94ea748>

## 3 DESCRIPTION: ACTIVITIES

For our song recommendation generator, our intended goal is that we will use our database of songs and song lyrics, take in an input from the user, and generate a playlist of about 10 songs that are similar to the lyrics, keywords, genre that the user inputs.

To achieve this goal, we plan to first gather our data from online datasets on Kaggle that contain song lyrics and from lyric APIs such as Genius and Musixmatch. We will need to preprocess the data possibly using the following:

- Removing stop words, such as "the", "in", "of" because they are not relevant.
- Removing punctuation and making all the words lowercase.
- Accounting for numbers: either removing numbers entirely or changing numbers to words to be included in the vocabulary.
- Stemming with online library/ packages.

Next we will split up the words by the space delimiter and create the vocabulary and calculate the frequency of each word in each song.

From the user's input of a song, we will fetch the song lyrics for that song and use that as the query. If the user gives keywords, then those keywords will be the query. We can use a text retrieval model to score and rank songs that match the genre of user's request. Then, we can obtain the top 10 most similar songs.

the activities you are planning to do

## 4 DATA

The data that we will be using will be a song lyrics dataset. The data will provide several song lyrics from albums from multiple artists. We plan to process and parse through this dataset before training our recommendation model on it, in order to ensure that all the fields are valid. We have researched multiple datasets and are currently planning to use <https://lyrics.github.io/> for our dataset based on its size, and collection of song lyrics.

Authors' address: Mutma Adebayo; Nandana Dileep; Smruthi Srinivasan; Selina Tang.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM XXXX-XXXX/2024/3-ART

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>