Music Recommender System

Jüri Kivimäe

PAT23

University of Tartu

November 2023

Introduction

In an era where digital music consumption has become the norm, platforms like Spotify have revolutionized how we explore and enjoy music. With vast libraries at our fingertips, the challenge is no longer finding music but navigating this ocean of choices to discover tracks that resonate with our personal tastes. This project was born out of a passion for music and the desire to harness the power of machine learning to enhance the music listening experience.

As a music enthusiast, I've always been fascinated by the potential of technology to curate personalized music experiences. Observing the widespread use of Spotify and similar platforms highlighted a ubiquitous need: a robust system to recommend music that aligns with individual preferences. While these platforms offer recommendation features, the quest for a more personalized and accurate music recommender system remains a compelling pursuit. This inspired me to delve into the realm of data science and machine learning, aiming to contribute a solution tailored to both my interests and the needs of fellow music lovers.

The primary goal of this project is to develop a music recommender system that not only understands users' listening habits but also uncovers hidden gems in their music libraries. By leveraging a hybrid model combining item-based collaborative filtering and content-based filtering, the system aims to provide recommendations that are both diverse and precise. The collaborative filtering component analyzes user playcounts to identify patterns and preferences, while the content-based aspect delves into the intrinsic qualities of songs, such as genre, tempo, and other audio features. This dual approach is designed to capture the essence of what makes a song appealing to an individual, transcending the traditional one-size-fits-all recommendation model. Ultimately, the project seeks to create a more engaging and satisfying listening experience, making music discovery a journey of delight and surprise.

Data Sources and Description

In the contemporary landscape of digital music, the confluence of various data sources presents an unprecedented opportunity to explore and understand musical preferences and patterns. This project leverages a rich dataset amalgamated from three primary sources: the Million Song Dataset, Spotify, and Last.fm, as sourced from Kaggle. The dataset provides a comprehensive view of user interactions and detailed song attributes, serving as a foundational pillar for the development of an advanced music recommender system.

User Playcounts Dataset:

The user playcounts dataset, a crucial component of this analysis, boasts over 9 million observations, providing a window into listener habits on an extensive scale. This dataset is characterized by a left-skewed distribution, where a majority of the playcounts cluster under 20-30 plays, although some tracks reach the heights of over 2400 plays. This skewness reflects a typical user behavior pattern in digital music consumption, where certain tracks enjoy repeated plays, while a vast majority have fewer listens.

To address this skew and enhance the model's efficacy, a decision was made to focus on playcounts under 20. This thresholding not only streamlines the data but also ensures a more balanced representation of user preferences, thereby avoiding bias towards a handful of overly popular tracks. Such a choice aids in refining the recommendation process, ensuring it captures a wide array of user interests.

Song Metadata Dataset:

Complementing the user playcounts is the song metadata dataset, initially comprising 50,000 tracks. This rich dataset includes an array of features for each track: `track_id`, `name`, `artist`, `spotify_preview_url`, `spotify_id`, `tags`, `genre`, `year`, `duration_ms`, and a suite of audio features such as `danceability`, `energy`, `key`, `loudness`, `mode`, `speechiness`, `acousticness`, `instrumentalness`, `liveness`, `valence`, `tempo`, and `time_signature`.

These features provide a multidimensional view of each song, encapsulating not only basic information but also nuanced audio characteristics that are instrumental in understanding the essence of each track. For the numeric features, scaling was employed to standardize the data, bringing every attribute to an equitable level for analysis. The categorical features, particularly `tags` and `genres`, underwent a process of multi-label encoding. This approach efficiently transformed these attributes into a format conducive to computational analysis while preserving their categorical richness.

Data Consistency and Reduction:

A pivotal challenge encountered was the inconsistency in track representation between the playcounts and the metadata datasets. For the integrity and coherence of the analysis, it was paramount to ensure that each track in the playcounts dataset had corresponding metadata. This alignment necessitated the pruning of the song metadata dataset to around 30,000 tracks, aligning it with the playcounts dataset.

Although this reduction marginally impacted the diversity of the song dataset, it had a minimal effect on the playcounts dataset. This refinement was a necessary step to uphold the consistency and reliability of the recommender system. The resulting datasets, post-alignment, provided a robust and harmonious foundation for the subsequent stages of model development and analysis.

Methods and Results

The pursuit of an optimal music recommendation system led to an exploration of various machine learning models, each offering distinct advantages and challenges. Central to this exploration were two models from the `scikit-surprise` library: `SVD` (Singular Value Decomposition) and `KNNWithZscore`. The process involved rigorous hyperparameter tuning and meticulous methodologies to synthesize insights from both user playcounts and song metadata, ultimately culminating in a sophisticated, hybrid recommendation engine.

Model Exploration and Challenges:

Initially, the project embarked on evaluating `SVD` and `KNNWithZscore` as potential candidates for the recommendation system. The `SVD` model, renowned for its effectiveness in matrix factorization and latent feature identification, showed promising results. It outperformed in terms of scoring metrics, indicating a superior ability to predict user preferences based on playcounts. However, a significant hurdle emerged in adapting the `SVD` model for item-to-item recommendations, a key requirement for the intended system.

Shifting focus to `KNNWithZscore`, known for its user-based and item-based collaborative filtering capabilities, I encountered a different set of challenges. The model’s computational intensity, particularly during the creation of the Mean Squared Difference (MSD) similarity matrix, posed substantial memory demands. Such requirements stretched the limits of standard desktop computing resources.

Successful Implementation with KNNWithMeans:

In light of these challenges, the project pivoted to employing `KNNWithMeans`, a variant of the K-Nearest Neighbors algorithm that integrates user and item mean ratings. Despite the model's memoryintensive nature, successful hyperparameter tuning was achieved on a test set, followed by training on the full dataset. This feat was accomplished by maximizing the available computing resources, utilizing 32GB of RAM and a 20GB swap file to their near capacities.

Integration with Song Metadata:

Parallel to the development of the playcount-based model, the project harnessed song metadata to create a cosine similarity matrix. This matrix encapsulated the nuanced relationships between tracks based on their audio features and encoded genres, offering a content-based perspective to the recommendation process.

A crucial step in the methodology was aligning the KNN similarity matrix, inherently ordered by the model's internal IDs, with the ordering of the tracks in the cosine similarity matrix. This realignment ensured that the similarities calculated from both collaborative and content-based approaches corresponded to the same set of songs.

Hybrid Similarity Matrix and Recommendation Functionality:

With both matrices aligned, the next pivotal phase was to amalgamate these distinct insights into a unified recommendation framework. This was achieved by creating a weighted similarity matrix, assigning equal weights (0.5 each) to the KNN and cosine similarity matrices. This hybrid approach effectively balanced the collaborative filtering insights with the content-based attributes, enriching the recommendation system's capability to suggest relevant and diverse tracks.

The culmination of this project was the development of functions that accept an artist and song title as input, locate the corresponding index in the combined matrix, and retrieve the top N recommendations based on the aggregated similarities. This functionality not only provided a seamless interface for generating personalized recommendations but also represented the integration of complex, multi-faceted data processing into a user-friendly application.

Conclusions and Future Directions

The journey of creating a music recommender system has been both challenging and enlightening, drawing parallels to a similar venture I undertook last year with a video game recommender system. While there were similarities in the underlying principles of recommender systems, the music domain presented its unique set of challenges and intricacies, particularly in handling the significantly larger datasets of playcounts and track metadata. These challenges were not just computational but also conceptual, requiring a nuanced approach to data processing and model building.

The most formidable obstacle encountered was the immense memory demand imposed by the expansive datasets. Successfully navigating this issue was a triumph of resourcefulness and meticulous memory management. Through careful optimization and strategic data handling, I was able to mitigate these constraints, underscoring the importance of adaptability and problem-solving in data science.

Future Roadmap:

Looking ahead, the project opens several exciting avenues for enhancement and innovation. A key area of focus will be the implementation of dimensionality reduction techniques on the track metadata dataset. By distilling the data to more relevant features, the model can achieve greater efficiency and potentially uncover deeper insights into song characteristics that influence user preferences.

Exploring Mood as a Feature:

An intriguing aspect to explore further is the mood of a song. Mood, as a quantifiable attribute, could significantly enrich the recommendation process. One approach is to augment the dataset with a mood feature, possibly sourced from external datasets. Alternatively, a mood feature could be ingeniously crafted from existing attributes such as valence, energy, and danceability. This would involve delving into the intricate interplay of these features and how they collectively translate into the perceived mood of a track.

User-Input Mood Feature:

To distinguish this recommender system from mainstream offerings like Spotify, I envision integrating a user-input mood feature. This innovation would allow users to input their current mood, which the system could then use to tailor recommendations. Such a feature would add a layer of personalization and interactivity, making the music discovery experience more dynamic and responsive to the user's emotional state. It aligns with the broader vision of creating a system that not only understands musical preferences but also resonates with the listener's current state of mind.

Conclusion:

This project stands as a testament to the transformative power of machine learning in personalizing entertainment experiences. It underscores the dynamic nature of data science, where challenges become opportunities for growth and innovation. The future roadmap lays out a path not just for technical enhancement but also for creating a more empathetic and responsive system that aligns technology with human emotion. As the project evolves, it holds the promise of transforming how we interact with music, making every recommendation a step closer to the perfect soundtrack for the listener's life.