Adaptive Music Recommendation System: Leveraging Session, Temporal, and Contextual Insights with the Million Songs Dataset

1st Srujan Putta
Masters in AI
San Jose State University
srujan.putta@sjsu.edu

2nd Prithvi Kapilavai
Masters in AI
San Jose State University
prithvi.kaspilavai@sjsu.edu

3rd Tejbhushan Kasibhotla
Masters in Engineering
San Jose State University
tejbhushan.kasibhotla@sjsu.edu

***Abstract*—The novel Adaptive Music Recommendation System presented in this paper provides individualized music experiences by utilizing the extensive Million Songs Dataset. The system customizes music recommendations for each user by combining temporal, contextual, and session-based insights. The system dynamically modifies its recommendations to fit the user's past preferences and current context by examining listening patterns, contextual information, and temporal changes in music preferences. By combining advanced machine learning algorithms with a thorough examination of the Million Songs Dataset, the implementation shows notable gains in user engagement and recommendation relevancy. This method constitutes a noteworthy development in the field of personalized music recommendation technologies.**

***Index Terms*—Adaptive Music Recommendation, Personalization, Session-Based Recommendation, Temporal Dynamics, Context-Aware Systems, User Engagement**

I. Introduction

By utilizing the Million Songs Dataset, the "Adaptive Music Recommendation System" initiative seeks to completely transform how people find music. The growing need for customized music experiences in the digital era is the driving force behind our endeavor. Building a sophisticated recommendation system that combines temporal recommendations, context-aware recommendation systems, and session-based recommender systems is the main goal. This project is unique as it emphasizes temporal dynamics and contextual information, both of which are essential for comprehending and satisfying users' changing musical preferences. Through the utilization of the Million Songs Dataset's rich metadata and technical qualities, this system aims to improve user engagement by providing personalized music recommendations that adjust to each listener's unique listening habits and environment.

A. Motivation & objective

Millions of songs are at our fingertips thanks to the tremendous rise of the digital music business in recent years. But the paradox of choice presents a serious obstacle in the face of this abundance. Users frequently struggle to choose music that suits their mood and particular taste because of the overwhelming number of options available. Even while they have their uses, traditional music recommendation systems frequently rely on static user profiles and collaborative filtering, which can result in repetitive and perhaps irrelevant recommendations. A system that is able to adjust to the user's shifting moods and circumstances while also comprehending their past preferences is becoming increasingly necessary.

This project's main goal is to create an adaptive music recommendation system that makes use of contextual knowledge and session temporal data. With the help of real-time analysis of user behavior throughout a session and a variety of contextual parameters, this system seeks to dynamically modify its recommendations. The system will deliver more precise, tailored, and contextually appropriate music recommendations by including these components. The intention is to improve the user's experience finding music by making it more interesting, simple to use, and fulfilling. In order to achieve a high degree of personalization and adaptability in music suggestions, the project investigates machine learning and statistical algorithms and data processing approaches using the Million Songs Dataset.

II. Literature Survey

The domain of music recommendation systems has experienced swift evolution, including many methodologies to augment consumer satisfaction. Session-Based Recommender Systems, which have attracted a lot of interest, are essential to this development. By using session-specific user interaction data, these systems are able to anticipate future preferences and provide timely, pertinent recommendations. These systems are now much more effective with Transformer models and Recurrent Neural Networks (RNNs). Understanding Temporal Dynamics in Music Preferences is a crucial component in the creation of recommendation systems. This method acknowledges that user preferences change over time due to a variety of circumstances, including trends, seasons, and events in the

user's personal life. By incorporating temporal elements into the recommendation logic, systems can provide more accurate and personalized suggestions, reflecting the dynamic nature of music consumption.
Conversely, Context-Aware Music Recommendation is a big step forward in customisation. This approach considers a number of contextual variables, including loudness, mode, liveliness and tempo. The idea is that user preferences are significantly influenced by the setting in which music is consumed. Recommendation systems can provide more sophisticated and situation-appropriate music recommendations by incorporating contextual data, improving the user experience overall.

III. System Design & Implementation details

A. Algorithms Considered/Selected

Context-Aware Music Recommendation is a big step forward in customisation. This approach considers a number of contextual variables, including the loudness, mode, liveliness and tempo. The idea is that user preferences are significantly influenced by the setting in which music is listened. Recommendation systems can provide more sophisticated and situation-appropriate music recommendations by incorporating contextual data, improving the user experience overall.
A key component of our music recommendation project is the Session-Based Recommender System, which uses individual session listening histories to prioritize in-the-moment music recommendations. Because of their superior ability to capture temporal dependencies in sequential data, recommender systems like this one are ideal for modeling user behavior patterns that change over the course of sessions. However, transformer-based models—which are well-known for their efficacy in sequence modeling—use self-attention mechanisms to manage intricate dependencies among user sessions. A number of factors, including the size of the dataset and the complexity of the user behavior patterns, influence the decision between RNNs and Transformers. The ultimate goal of this approach is to improve users' music-listening experiences by offering session-aware music recommendations that are tailored to their current tastes.
Understanding the temporal dynamics of user preferences is central to the Temporal Recommendation System. It seeks to track changes in listener behavior over time and make music recommendations that reflect changing preferences. To address this, we have determined two main algorithmic methods. First, the efficacy of matrix factorization techniques such as Alternating Least Squares (ALS) and Singular Value Decomposition (SVD) in capturing long-term trends in user preferences is evaluated.
Referred to as the Context-Aware Recommender System, emphasizes the incorporation of contextual elements to personalize music recommendations based on the user's preferences and current context. In this regard, we consider two main paths for algorithmic implementation. The first strategy is called collaborative filtering with context, and it builds on traditional collaborative filtering techniques by adding contextual elements like tempo, dancibility, and liveliness.

B. Technologies Tools used

The programming language and libraries you choose for your music recommendation project will have a significant impact on how well the recommendation system performs in the context-based approach. Python's rich ecosystem of machine learning and data science libraries makes it a popular choice for these kinds of tasks. NumPy and pandas are two essential libraries for preprocessing and manipulating data. While pandas offers flexible data structures for managing and cleaning the Million Songs Dataset, NumPy offers effective numerical operations. When it comes to data visualization, Matplotlib comes into play, enabling you to make intelligent charts and plots that examine the variety of contextual factors affecting music recommendations. Variables like user location, device type, and mood may be included in this diversity, and Matplotlib facilitates the visual exploration of patterns and trends.
Cosine similarity is a basic metric used in modeling and recommendation that assesses how similar items or users are based on their feature vectors. Cosine similarity can be used in your context-aware recommender system to measure how similar user profiles and songs are, allowing for more customized recommendations.Developed with TensorFlow, an advanced deep learning framework, you can efficiently create, train, and implement RNN-based recommendation models. When combined, these technologies allow the project to take advantage of the data's richness, produce recommendations that are sensitive to context, and improve the user experience when it comes to music recommendations.

C. System design and architecture

Three different recommendation methods—Session-Based, Temporal, and Context-Aware—are seamlessly integrated into our music recommendation system's architecture. Contextual data like location, type of device, and mood, as well as user listening histories and the Million Songs Dataset are some of the sources of data that feed into the system. To provide session-aware music recommendations, the Session-Based Recommender System analyzes user sessions in real-time using LSTM and Cosine Similarity. While the Context-Aware Recommender System customizes music recommendations based on user context, the Temporal Recommendation System uses time series analysis to capture changing user preferences over time.

IV. Experiments / Proof of concept evaluation

A. Dataset

The Million Songs dataset is a collection of music-related information, and the whole dataset is a whopping 300GB in size, but there's a smaller subset that's more manageable, at 1.8GB, for smaller projects. The data is organized in JSON format. In this subset of the Million Songs Dataset with a wealth of technical attributes and metadata. The dataset includes information such as the song's title, artist name, release year, tempo, danceability, energy level, and more. Notably, this subset does not contain audio data but focuses on the track descriptions and associated metadata. In addition to the general attributes mentioned above, the dataset provides more in-depth information about each track. For example, it includes details like the song's key, mode (major or minor), loudness, time signature, and various confidence measures associated with features like beats, bars, energy, sections, and segments. These features can be valuable for building music recommendation systems and conducting music-related analyses.

B. Data preprocessing

To aid in the creation of a successful music recommendation system, a significant amount of data preprocessing was done to convert the initially difficult Million Songs Dataset from H5 format into a more readable CSV format by means of a specialized Python script. This conversion made data manipulation much simpler and laid the foundation for later project stages. Additionally, the dataset—the "music" dataset and the "user history" dataset—was carefully chosen into two separate subsets. The'music' dataset is the primary source of data for song analysis and recommendation. It includes essential track information such as track ID, name, artist, genre, year, duration, and audio features. In addition, the 'user history' dataset records user interactions, playcounts, and user IDs, offering insightful information about user behavior that can be used to tailor recommendations. The foundation for building a strong music recommendation system that takes into account unique user preferences and engagement patterns has been laid to prepare the data.

C. Evaluation

For every recommendation system variant, a different technique was used in the evaluation methodology used for the music recommendation project. An 80-20 split was applied to the data for the Session-Based Recommender System, meaning that 80 percent of the data was used for training and 20 percent for testing. Nevertheless, this method required close examination of user sessions, with each user's session acting as a benchmark for recommendations. Heuristics were used to organize the user sessions because timestamps and missing data were present. Two main techniques were applied: the first involved creating sessions at random, and the second method used a set of tracks to strategically create sessions.
The Temporal Recommendation System, on the other hand, used an alternative strategy by taking into account the entirety of a user's listening history. Every user's past information was considered in order to make recommendations. In a similar vein, the Context-Aware Recommender System adopted the same methodology, taking into account all accessible user context data. These differences in evaluation techniques were made to accommodate the distinct features and data accessibility of every recommendation system, guaranteeing that the assessment procedure is in line with the particular demands and difficulties presented by every approach.

D. Comparison

We experimented with the Million Songs dataset in our thorough assessment of the Adaptive Music Recommendation System, carefully preprocessing the data in order to make it ready for analysis. For the Temporal Recommendation System, Context-Aware Recommender System, and Session-Based Recommender System, we used different evaluation approaches. Our findings showed that, in comparison to the original LSTM models, the cosine similarity-based approach greatly increased the accuracy and relevance of recommendations in the Session-Based Recommender System. Adaptive recommendations were made possible by the Temporal Recommendation System, which successfully recorded shifting user preferences over time. Although the Context-Aware Recommender System demonstrated potential in tailoring music recommendations according to user context, certain contextual attributes were restricted by data limitations.

E. Analysis of results

Our examination of the findings emphasizes how important it is for the Adaptive Music Recommendation System to use a variety of recommendation techniques. Users were guaranteed to receive music recommendations that were more closely aligned with their immediate preferences and context, as the Session-Based Recommender System proved when it used cosine similarity-based techniques to improve recommendation accuracy. We found that taking into account the complete listening history of the user produced recommendations in the Temporal Recommendation System that dynamically adjusted to the user's evolving preferences over time. The availability of contextual attributes in the dataset limited the potential of the Context-Aware Recommender System, despite its incorporation of user context data. These results highlight the value of a multimodal strategy for music recommendation that takes user preferences and contextual relevance into account. The contextual based approach suggests that the more attributes are taken into consideration, better is the recommendation. This was supported by diversity factor using cosine similarity and when all the attributed were considered shouwed significat imporvement by 53.67 percent

V. Discussion Conclusions

A. Decisions made/Things that worked

A thorough understanding of the distinctive qualities and user expectations connected with each recommendation method served as a basis for a number of strategic decisions made throughout the development of the music recommendation system. The deliberate choice made in the Session-Based Recommender System was to randomly select 10 songs from each user's listening history in order to create user sessions. By giving priority to songs that aligned with the user's recent preferences, this method made sure that recommendations were precisely tailored to their current context.

For instance, the system would suggest folk songs that fit the user's mood after 30 minutes of listening to folk music, so improving their listening experience. On the other hand, the Temporal Recommendation System chose a more all-encompassing approach. By considering the complete history of a user's listening choices, the system was able to offer more comprehensive suggestions. This method took into account the fact that users' musical preferences change over time and captured both transient and permanent preferences. Two unique attributes were taken into account in the Context-Aware Recommendation System: name, artist, and genre in addition to a wide range of attributes like track ID, tags, and audio features. It was a wise decision to include cosine similarity in the diversity factor calculations for suggested songs because it gave users access to a wide range of musical qualities and made sure their music recommendations were varied and well-rounded.

B. Difficulties faced/Things that didn't work well

There were a number of significant obstacles and aspects of the music recommendation system's development that did not produce the best outcomes. First, given the size of the dataset, the conversion of the lengthy.h5 file into the more manageable CSV format presented a major challenge because of computational resource limitations. Second, poor recommendations resulted from the Session-Based Recommender System's use of RNN models due to the unsatisfactory prediction accuracy. This led to the adoption of cosine similarity-based techniques. Last but not least, the Context-Aware Recommender System's ability to make more dynamic and contextually relevant recommendations was hampered by limitations in the dataset attributes that were available. These limitations prevented the system from accounting for important environmental attributes like user location, mood, and the timing of song listening sessions. These challenges highlight how difficult it is to create recommendation systems and how crucial it is to improve model performance and data constraints in order to raise the overall quality of recommendations.

C. Conclusions

One enhancement in the field of personalized music services is the creation of the Adaptive Music Recommendation System. The system efficiently incorporates temporal, contextual, and session-based insights to improve users' music discovery experience by leveraging the Million Songs Dataset. After addressing issues with model accuracy and data preprocessing, a context-aware, adaptive recommendation mechanism was successfully put into place. Subsequent enhancements might concentrate on augmenting the contextual attributes of the dataset and further perfecting the recommendation algorithms. This system establishes a new standard for personalized music recommendations and shows how different data-driven insights can be combined to meet the changing needs of users in the digital age.

VI. Project Plan / Task Distribution

As part of the project's task distribution, Srujan Putta, Prithvi Kapilavai, and Tejbhushan Kasibhotla worked together to tackle the first phase of data preprocessing, which ensured a strong basis for the project's later stages. Subsequently, the team members concentrated on distinct facets of the recommendation system: Srujan Putta worked on the session-based methodology to customize recommendations according to individual user sessions. Prithvi Kapilavai went with the temporal approach, exploring the subtle changes in users' musical tastes over time. The context-aware approach was done by Tejbhushan Kasibhotla, who concentrated on incorporating different contextual elements to improve the relevance and accuracy of the recommendation system.
Prithvi focused his efforts on creating a PowerPoint presentation that summarized the project's key findings and methodologies, while Tejbhushan and Srujan worked together to compile the comprehensive report, which synthesized all aspects of the research and development process.

References

[1] Session-Based Recommendations with Recurrent Neural Networks - Hidasi et al., 2016. Presented at the IEEE International Conference on Learning Representations (ICLR), this paper discusses the application of RNNs in session-aware music recommendation systems.

[2] Time-aware Recommender Systems: A Comprehensive Survey and Analysis of Existing Evaluation Protocols - Campos et al., 2014. This IEEE paper provides an extensive survey of the impact of temporal dynamics on the accuracy and effectiveness of recommendation systems.

[3] Context-aware Music Recommendation: An Empirical Comparison of Pre-filtering, Post-filtering, and Contextual Modeling Approaches - Baltrunas et al., 2011. In this study, presented in an IEEE publication, various strategies for incorporating context into music recommendation systems are analyzed, highlighting the significant role of contextual information.

[4] Adomavicius, G., Sankaranarayanan, R., Sen, S., Tuzhilin, A.: Incorporating contextual information in recommender systems using a multidimensional approach. ACM Trans. Inf. Syst. 23(1), 103–145 (2005)

[5] Adomavicius, G., Tuzhilin, A.: Multidimensional recommender systems: a data warehousing approach. Second International Workshop on Electronic Commerce, pp. 180–192, Heidelberg (2001).

[6] Cao, H., Chen, E., Yang, J., Xiong, H.: Enhancing recommender systems under volatile user interest drifts. Eighteenth ACM conference on Information and knowledge management, pp. 1257–1266, Hong Kong (2009)

[7] Celma, O.: Music recommendation and discovery in the long tail. PhD Dissertation, Universitat Pompeu Fabra, Spain (2010)