Literature Review

Implementation of Machine Learning Tools and Techniques in Music Recommendation Systems

Introduction

The working topic for this literature review is Music Recommendation Systems. These systems provide users with personalized recommendations, accounting for their preferences and listening habits. They also serve as the primal source for providing access to new music and likewise to new artists (as compared to the horse and buggy days of word of mouth or radio) (Daga, 2023).

About 524 million music streaming subscribers were reported as of June 2022 (Balcombe et al., 2022). A more specific, perhaps clinical example is the usage of music recommendation systems in music therapy, as an alternative therapy for patients with dementia (Nunes et al., 2023). These music recommendation systems are being used all over the world, which raises an interesting research question to serve as the focus of this literature review: How is Machine Learning (ML) vital in producing music recommendations for users with varied musical tastes, across the globe?

What remains a gap in this research question is whether ML alone achieves the heights of relevant music recommendations, or if there are other techniques that can aid/do better. Nonetheless, this is an important topic to review as it provides a crucial focus to machine learning and its charm towards providing music recommendations for a varied group of people.

The sources used in this literature review were found using the search term “music recommendation systems” within the University of Essex online library and Google Scholar. These sources are either academic papers or published, standalone literature reviews.

The key topics that will be discussed in this literature review are ML filtering techniques, and their usage (sometimes combined) in industrial or academic modeling of recommendation systems. There will also be a hunt for other techniques potentially used to conduct music recommendations.

ML Filtering Techniques

There are various types of music recommendation techniques that are used individually, or paired together, to create various kinds of recommendation systems. Let’s dive into the filtering techniques first:

Collaborative Filtering

Based on an assumption that users will continue to rate items (in the past) similarly (in the future), collaborative filtering generates self-directing responses and/or predictions about a user’s interests. It relies on collecting taste or preferences information from users (Paul, 2019).

To provide recommendations, collaborative filtering uses a ML technique called the k-nearest neighbor algorithm. Ratings are divided into two categories, the first being explicit (provided explicitly by the users, such as 1-5 star ratings), and the second being implicit (interpreted by user behavior, such as song repetition will cause the song to have a higher implicit rating) (Paul, 2019).

Naturally, a system that relies on ratings will not perform well without sufficient ratings, and henceforth produce lackluster recommendations (i.e. a cold-start problem). Which, in turn, would offer less motivation for a user to provide ratings to a system that doesn’t offer recommendations commendably (Paul, 2019).

Content-Based Filtering

Contrary to collaborative filtering, this method of filtering provides recommendations based on an internal system comparison between a user’s profile and the content (i.e. descriptors or terms in text) of the items (acoustic features such as loudness, tempo, rhythm, timbre, etc). This technique uses K-means clustering and expectation-maximization with Monte Carlo sampling ML techniques to compute similarity, and avoids the cold start problem by not requiring as much data as Collaborative filtering does (Paul, 2019).

This type of filtering faces the following set of issues that contribute to its glass-ceiling effect. Firstly, the terms as mentioned above can either be assigned manually or automatically (in the latter case, a method needs to be picked that can extract these terms from the items mentioned above). Secondly, the representation of these terms must be purposefully done so that comparison between the user profile and the terms can be done consequentially. Lastly, to learn the profile of the user based on prior viewing activity and make recommendations accordingly, the right learning algorithm must be made available (Paul, 2019).

Apart from the issues, there are several drawbacks to this filtering technique; the biggest handicap being that it relies heavily on the item model for accuracy. Moreover, this technique disappoints in differentiating vital details between more or less similar songs (Paul, 2019).

Metadata-Based Filtering

Diving into the most traditional form of filtering technique, this filtering technique uses editorial metadata of a particular song (such as the artist name, genre, album name, etc.) to provide recommendations. The lack of utilization of any user information can be seen in the uninspiring recommended music (Paul, 2019).

Emotion-Based Filtering

With the innate connection of music and human emotions, so much so that music can activate reward centers in the brain (Arjmand, 2017), this happens to be an ideal, highly user-satisfying form of filtering. A song, with its different acoustic features, can trigger different emotions. Inversely, a user is more likely to select music based on their present mood too (Paul, 2019).

The boon also happens to be the curse, with massive data collection and managing excessive datasets being required in the model, and human effort accordingly. Keeping song interpretations in mind for different people with different backgrounds, ambiguity is also an issue with the datasets (Paul, 2019).

Context-Based Model

This is an interesting form of filtering as it recommends music partly utilizing public perception and social engagement of a song (such as using social media sites such as Facebook, Twitter, Reddit, and Youtube). Additionally, to recommend music, this technique utilizes the user's listening history (similar to what content filtering does), a social media twist (i.e. the new word of mouth), and user's location (adding on the cultural impact of musical taste).

It is an efficient and well performing model (perhaps even better than the other recommendation models) to work with minimal amounts of data, for e.g. music streaming platforms (such as Spotify and Apple Music) generating top charts based on the music that is heard the most from their user database (Paul, 2019).

ML Music Recommendation Systems

The above filtering techniques, in combination with various parameters, can help create the following music recommendation systems.

Hybrid System

The name is self-explanatory herein. A hybrid model can outperform recommendation efficiency as it can use a combination of recommendation systems and does not have to rely on a single recommendation system, An example of this could be a model utilizing both content and collaborative techniques, defeating individual frailties and instead succeeding on their combined virtues (Paul, 2019).

Listening History

It was noted in the context-based modeling discussion above that a user's listening history is pivotal in making useful song recommendations, especially because it is very likely that a user will listen to the same song repeatedly (or at least similar songs, frequently) (Paul, 2019).

Breaking down the entire listening history in sessions (i.e. utilizing a profile for each session instead of making a profile for each user) is one of the many techniques available to extract essential information from the user's listening history (to be able to recommend songs) (Paul, 2019). An example of this usage would be a session-based collaborative filtering model (SSCF) that "*finds similar sessions with an active session from given partial information and reference items in the similar sessions*" (Park et al., 2011). This is an interesting example of the collaborative filtering technique used in a proposed academic model, and how it appreciates referencing items of users with similar musical taste, and henceforth increasing recommender system performance.

Furthermore, this breakdown can help the recommendation model understand the user's short term vs long-term performance, as the model gets more information on what songs are listened to in succession (Paul, 2019). An exciting revelation on this forefront is - while the next-track can be easily governed by the short-term listening history, the actual personalization quality actually increases via knowing long-term preferences (Kamehkhosh, n.d.). To serve as an anecdote, a parallel to human connections can easily be made here.

User-Centric Experience

Music industry is massive, and a simple explanation to why we need the variety is because no one person is cognitively alike, and neither is their personal appeal to music. This is where personalization, preference learning approaches to machine learning, and information retrieval techniques enter to be able to design user-aware music applications (Schedl et al., 2013). The modern streaming websites are therefore observing the music-listening habits of users and creating a user-specific experience, such as a personalized "for you" playlist (Paul, 2019).

Main Findings From Literature Review

While the ML filtering techniques are distinct from each other, there are also similarities. Furthermore, these similarities are being leveraged by commercial and academic models that are being designed to help fill the gaps.

The strength of this literature review lies in its wholesome, systematic, and methodological structure; covering most of the overall methods used within this field. The limitations of this literature review lie in the exploration of the use of this field in combination with other fields (music paired with visual media, for example).

In conducting the literature review, it was found that there are in fact other comparators that can perform in combination (or solo) to recommend music, such as deep learning (Schedl, 2019) and sound technology (Reddy, 2024). There is also a music recommendation system that utilizes the user's current emotions, and physical data such as body temperature (Jazi, 2021). This addresses the research gap in the original research question stated above, but is out of scope of this literature review for further analysis.

Conclusion

Reiterating the main research question here, this literature review was conducted to review how machine learning is vital in producing music recommendations for a varied group of users with varied musical preferences and location. It can be seen how specific filtering techniques, rooted in machine learning, can recommend music based on a user's location. There is also a keen discussion on how user's listening data can help make those recommendation systems better.

Arguably, the next step in this field of research would be to find a way to improve these recommendation systems, perhaps by furthering the parallelism of audio and artificial intelligence techniques. For example, combining a technique that suffers in its recommendations (due to lack of data) with a technique that can work with minimal data, can prove to be advantageous. Another arena of research would be to recommend music to pair with visual media, which can benefit from the already explored field of music recommendations.

It is important to note that this literature review has resulted in an observation about other techniques, whether used solo or in combination with ML, that are also vital in developing the music recommendation systems. They work best when they work together, and ML alone cannot achieve what a combination of these techniques can.

References

Arjmand, H. A., Hohagen, J., Paton, B. & Rickard, NS (2017) Emotional Responses to Music: Shifts in Frontal Brain Asymmetry Mark Periods of Musical Change. *Front. Psychol.* 8:2044. Available from: doi: 10.3389/fpsyg.2017.02044 [Accessed 16 September 2024].

Balcombe, L. & De Leo, D. (2022) Linking music streaming platform advertisements with a digital mental health assessment and interventions. *Frontiers in Digital Health*, 4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9677233/#:~:text=As%20of%20June%202022%2C%20there,with%20422%20million%20unique%20users. [Accessed 16 September 2024].

Daga, B., Kadam, H., Shrungare, S., Fernandes, S. & Johsnon, A. (2023). Music Recommendation System. *International Journal of Computer Trends and Technology* 71(5): 26–36. Available from: https://www.ijcttjournal.org/2023/Volume-71%20Issue-5/IJCTT-V71I5P105.pdf [Accessed 16 September 2024].

Healey, M., Matthews, K., & Cook-Sather, A. (2020) *Writing about learning and teaching in higher education: Creating and contributing to scholarly conversations across a range of genres*. Center for Engaged Learning Open-Access Books, Elon University. 142-152. Available from: https://www.centerforengagedlearning.org/books/writing-about-learning/ [Accessed 16 September 2024].

Ingrid B. N., Santana, J., Arianne, T., Charron, N. C., Machado, G., Gusmão, C. & Santos (2023) Music recommendation systems to support music therapy in patients with dementia: an exploratory study. *Research on Biomedical Engineering*, 39(3): 777–787. Available from: https://link-springer-com.uniessexlib.idm.oclc.org/article/10.1007/s42600-023-00295-7 [Accessed 16 September 2024].

Kamehkhosh, I., Jannach, D. & Lerche, L. (n.d.). *Personalized Next-Track Music Recommendation with Multi-dimensional Long-Term Preference Signals*. Available at: https://web-ainf.aau.at/pub/jannach/files/Workshop_IFUP_2016.pdf [Accessed 16 September 2024].

Park, S. E., Lee, S. & Lee, S-g (2011) Session-Based Collaborative Filtering for Predicting the Next Song. *2011 First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering, Jeju, Korea (South)*. 353-358, Available from: https://ieeexplore.ieee.org/abstract/document/5954341 [Accessed 16 September 2024].

Paul, D. & Kundu, S. (2019). A Survey of Music Recommendation Systems with a Proposed Music Recommendation System. *Advances in Intelligent Systems and Computing*, 279–285. Available from: https://link-springer-com.uniessexlib.idm.oclc.org/chapter/10.1007/978-981-13-7403-6_26 [Accessed 13 September 2024].

Purdue Writing Lab (2020). Writing a Literature Review. Purdue Writing Lab. Available from: https://owl.purdue.edu/owl/research_and_citation/conducting_research/writing_a_literature_review.html [Accessed 26 August 2024].

Schedl, M. (2019). Deep Learning in Music Recommendation Systems. *Frontiers in Applied Mathematics and Statistics*, 5. Available from: https://www.frontiersin.org/journals/applied-mathematics-and-statistics/articles/10.3389/fams.2019.00044/full [Accessed 16 September 2024].

Schedl, M., Flexer, A. & Urbano, J. (2013). The neglected user in music information retrieval research. *Journal of Intelligent Information Systems*, 41(3):523–539. Available from: https://link-springer-com.uniessexlib.idm.oclc.org/chapter/10.1007/978-981-13-7403-6_26#ref-CR11 [Accessed 16 September 2024].