The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 64– 79

DQN Model for Music Recommendation Based on Singer Style

M SNV Jitendra[1], Dasari Sai Srikanth[2], Sneha Davarasingi[3], Shrenik Nethi[4], Dronamraju Varun[5]
[1,2,3,4,5]School of Technology, Dept of CSE, GITAM University, Visakhapatnam-India

Abstract

Recommendation systems play an integral role in improving user experience on various digital platforms, including music streaming service. These recommendation systems learn user preferences and then recommend the proper content to foster interaction and satisfaction. Traditional recommendation methods have been applied extensively on music platforms, including collaborative filtering and content-based filtering. With that said, models for traditional recommendation systems all have certain limitations, such as the cold-start problem, sparsity, and the inability for the model to adapt dynamically to changes in user preference. To address these limitations, this paper explores a new music recommendation system, which is the reinforcement learning with Deep Q-Networks (DQN)-based recommendation system, which models music recommendation as a Markov Decision Process (MDP), where the decision-making agent interacts with the environment by selecting songs (recommended to the agent), and receiving a reward or punishment for that action. Each song suggestion incurred a potential reward or punishment, maximizing the agent's expected cumulative reward over time. The model is learned by iteratively adjusting its behavior so that the song recommendations reflect ongoing changes in user behavior, based upon the collected data. The proposed working system will be trained on a data set aggregated from multiple users, consisting of song metadata, listening history and preference to music genre. The reinforcement learning agent learns to optimize song recommendations by maximizing expectations of cumulative reward, thereby develop more personally tailored and interactive recommendations The empirical results validated that the model outperformed traditional methods indicating a recommendation accuracy of 98.33%. Accordingly, the resulting data supports that reinforcement learning is a feasible way to improve music recommendation systems, generating many possibilities for other adaptive and intelligent recommendation models for use in online music platforms.

Keywords: Music Recommendation– Machine Learning– Personalized System– Singer Style Analysis.

1. Introduction

Today, the world has become an increasingly connected place and music has embedded itself into the fabric of daily life. With music being available more conveniently and visibly, and more importantly, with countless exposure to music more than ever, numerous audiences are switching to online consumption of musical entertainment, relaxation, and even productivity. A recent report highlights that there are millions of users around the world, if not billions, using online music platforms like Spotify, Apple Music, and YouTube Music, to explore a diverse and endless stream of songs, playlists, and collections. The ability to listen with ease while subscription pricing is relatively low, has made these platforms an almost permanent companion for many users. Alongside the increase in subscribers seeking out streaming music services, there has become a strong demand in delivering intelligent recommendation engines. With such an enormous and infinite catalogue of songs, trying to manually identify some potentially favorite tracks could be exhausting and overwhelming. The recommendation systems are helpful as a means to curate music consumption for individuals and deliver customized playlists, or song recommendations to provide the best experience. The recommendation systems typically evaluate and observe listening habits, historical listening choices, and demographics of the user in order to provide observers with content that matched a listener's musical preferences, and thus increases engagement and satisfaction of the user. Music recommendation systems have certainly evolved and matured over time, deliberating upon many approaches to provide both user and accurate and precise recommendations. Traditional models of Music recommendation systems are:

Collaborative Filtering: Here the recommendation for songs is based on the songs that the user with similar tastes listens to. For example, if two users have preferences for similar songs, the system recommends songs liked by one user to the other user. Collaboration filtering, however, has been plagued with the cold-start problem, in which new users do not receive recommendations because there is very little historical data to build the recommendations upon.

Content-Based Filtering: This algorithm checks for the attributes of songs, such as tempo, genre, artist, and instrumental elements, to recommend the tracks that share similar characteristics. While this can be helpful, content-based filtering will result in repetitive recommendations, as preferences tend to bias toward songs very similar to previously chosen titles.

Hybrid models: To improve recommendation accuracy, these models merge content-based and collaborative filtering. Hybrid models combine multiple methods to mitigate some of the disadvantages of a single model. Traditional recommendation algorithms are valuable, but they struggle to adapt to changing consumer habits and preferences. Since musical preferences change over time, static models are often unable to generate balanced and engaging suggestions. Users could also become frustrated with the algorithm, particularly if they fail to present new music opportunities. This article presents a reinforcement learning-based method that uses Deep Q-Networks (DQN) to resolve the issues outlined in this section. Reinforcement learning, unlike traditional methods, will continuously be updated according to user interactions. The proposed method will also treat music recommendation as a sequential decision-making problem where an RL agent is tasked with selecting songs according to user preferences and receives information about its selected action via some reward function. Optimization will take place over time as the agent adapts its policy to maximize all future rewards, which would lead to relevant and engaging recommendations. The use of reinforcement learning allows the model to be responsive to changes in user preferences while improving the variety of recommended options it makes available. While collaborative filtering relies on previous data (and may have a hard time recommending songs users have not previously listened to, and content-based filtering relies on similarities in the features of songs, reinforcement learning offers a more flexible and adaptable approach. This is the system guarantees that users continue to experience a range of songs, while representing the user's core preference at the same time, resulting in a more pleasurable and fulfilling listening experience.

2. Literature Survey

Music recommendations in recent years have attracted considerable attention because of the rapid expansion of digital music libraries. Different approaches have been proposed to increase recommended accuracy, from traditional partner filtration techniques to advanced deep learning models. This section provides observation of larger research contributions to this domain. A hybrid music recommendation system that integrates collaborative filtration with a music -based approach was proposed to address boundaries such as computer sparsity and cold start problems [1]. The study says that the music generations, such as melody, rhythm, and style, improve the user's preference to match matching. Although specific accuracy was not given matrix, the results indicate the general improvement in the recommended quality. Another study introduced an automated music recommendation system that uses user -based associated filtration, element -based associated filtration and popularity -based ranking to increase privatization. Experimental results showed that user -based college filtration achieved the highest accuracy by 92% accuracy and 88% recall, demonstrated its effectiveness on other methods.[2]A recommendation system specifically designed for postpartum mothers was created using Support Vector Machine (SVM) classification. The research utilized One-vs-One (OVO) and One-vs-Rest (OVR) classification methods to sort music based on its emotional and therapeutic benefits. The SVM-OVR model with a polynomial kernel yielded the best results, achieving 80% accuracy, 82% precision, and 80% recall, demonstrating that genre classification can significantly improve user satisfaction. Additionally, deep learning methods have been investigated for music recommendations [3]. A study that used Bidirectional Long Short-Term Memory (Bi-LSTM) networks found that incorporating Mel-Frequency Cepstral Coefficients (MFCCs) as input features boosts recommendation accuracy. When compared to the earlier methods, the Bi-LSTM model's observed F1-score was 3% higher, indicating that sequential deep learning models may have applications in music recommendation [4]. A database of 431 user profiles was used in a study comparing Random Forest with Multi-Layer Perceptron (MLP) classifiers for predicting music genres. The results revealed that Random Forest performed better than MLP, with an accuracy of 95.47% compared to 53.07% for MLP. This study showed how effective ensemble learning algorithms are at recommending music depending on genre [5]. Convolutional Recurrent Neural Network Architecture (CRNNA) was assessed for music genre categorization and recommendation in another deep learning-based contribution. The authors utilized Mel spectrograms as input features and achieved very good results with a high accuracy in the genres of hip-hop (90.5%) as well as jazz (84%), and precision = 71.5% and recall = 78.1%. The paper suggests to use the hybrid CNN-RNN design consisting of both architecture for music recommendation because the hybrid architecture captures both the temporal and spectral dimensions of the data stream in music content. Another article presents a review of music recommendation approaches, including content-based filtering, collaborative filtering, and deep learning models. The paper discusses challenges in music recommendation, including data sparsity and cold start problems. This paper reviews studies that look at sentiment, contextual, and hybrid filtering. The discussion continues with studies regarding hybrid models, sentiment-based recommendations, and context-aware filtering. The suggested system successfully integrates machine learning approaches with collaborative filtering to boost personalization. Evaluation results showed high similarity scores with a peak value of 0.9375, implying an effective recommendation performance [7].

i

3. Proposed Model

3.1 Proposed Algorithm

The music recommendation system based on reinforcement learning employs a well-structured algorithm that allows it to learn and enhance its recommendations over time.

Algorithm: Reinforcement Learning-Based Music Recommendation

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 64– 79

1. Start by adding the user's listening history and pertinent metadata to the environment.
2. Define the state space based on song metadata and user preferences.
3. Choose a song from the available dataset to represent the action space.
4. Use random weights to start the Q-Network.
5. For every episode:
   a. Use an epsilon greedy policy to select an action (song recommendation).
   b. Receive feedback in the form of a reward (positive for genre aligned selections, negative otherwise).
   c. Store the experience in the replay buffer.
   d. Sample a mini-batch from the replay buffer and update the Q-Network using the Bellman equation.
   e. Adjust the exploration-exploitation balance through epsilon decay.
6. Continue this process until convergence or performance stabilizes.
7. Deploy the trained model to deliver real-time recommendations. This algorithm guarantees that the model consistently improves by learning from user interactions and dynamically adjusting its recommendations.

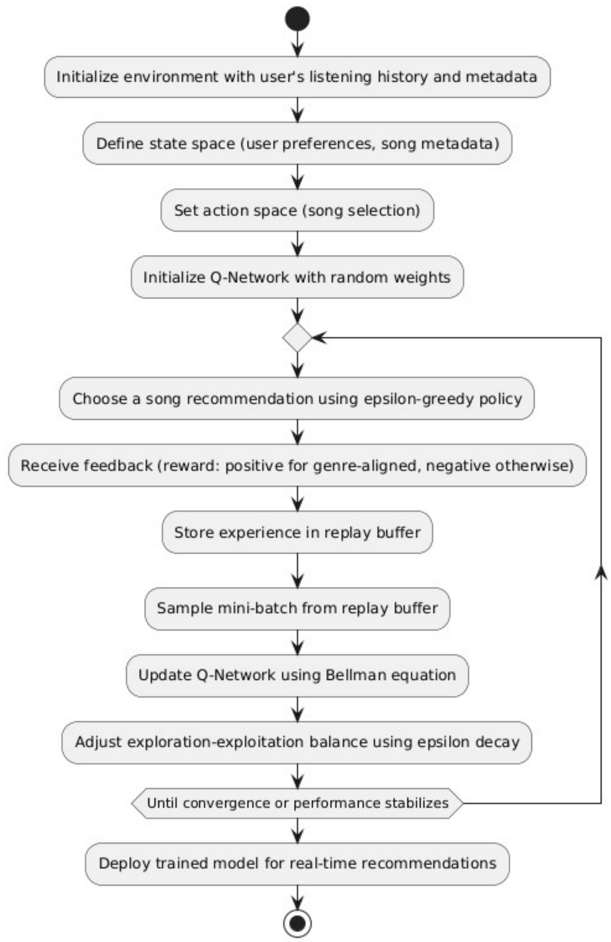


Fig.1 Proposed Algorithm for the model

3.2 Activity Flow

The workflow of the Reinforcement Learning-Based Music Recommendation System includes several essential stages:

User Data Acquisition: Spotify, Apple Music and other music streaming services supply the system with user listening history. The metadata includes the listening session timestamp, track name, artist, genre, and popularity among other things. This data allows us to better understand user preferences and levels of engagement.

Data Preprocessing: To level the playing field, all genre labels are changed to lowercase and removed of excess whitespace. To avoid biases in the data for the model later, duplicates are removed. Furthermore, songs labeled by different genre are split into multiple records, which allows for a further rich categorization for recommendations.

Model Training (DQN): The Deep Q-Network (DQN) is trained using reinforcement learning. By optimizing rewards based on past user preferences and user feedback, the model learns the optimal action. An experience replay buffer is used to improve the stability of the learning process and reduce sample variance.

Recommendation Generation: Each user is given a dynamic song recommendation based on the trained model. The algorithm will use learnt policies based on previous interactions to predict the next best song to recommend. The user will be given a prioritized collection of recommendations based on their listening preferences.

Model Update: Based on user feedback, the reinforcement learning model will be updated on a regular basis. As user preferences evolve over time, the recommendations remain applicable thanks to the ongoing training. The system responds quickly and attempts to update suggestions such that they remain fresh for the user and are neither overly repetitive nor unnecessary.

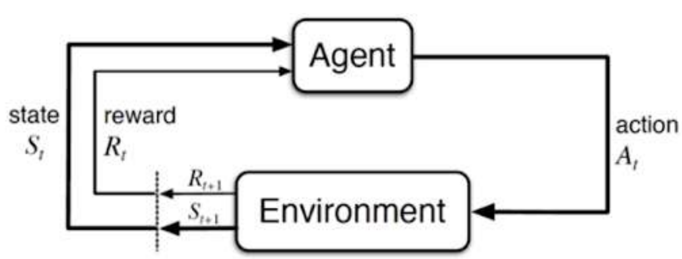
3.3 Markov Decision Process (MDP) Formulation

A Markov decision process (MDP) refers to a stochastic decision-making process that uses a mathematical framework to model the decision-making of a dynamic system. It is used in scenarios where the results are either random or controlled by a decision maker, which makes sequential decisions over time. MDPs evaluate which actions the decision maker should take considering the current state and environment of the system. The policy for the MDP model reveals the agent's following action depending on its current state.

The MDP framework has the following key components:

- S: states ($s \in S$)
- A: Actions ($a \in A$)
- $P(St+1|st.at)$: Transition probabilities
- R (s): Reward

The graphical representation of the MDP model is as follows:



3.4 Reinforcement Learning Agent

The DQN architecture of the reinforcement learning agent is used to approximate the optimal policy using a neural network. To improve the stability of the learning process and decrease variance, an experience replay mechanism is incorporated into the learning agent's training process. The neural network consists of multiple fully-connected layers that ultimately represent the state and produce Q-values for each possible action a recommendation agent might take. With a Q-value estimation for each action available, the recommendation agent can make optimal recommendations. The agent continually improves its policy in response to user feedback, reinforcing successful recommendations, and reducing the likelihood of recommendations that were unrelated to user preferences. This Q-learning approach implemented with a reinforcement learning agent also uses an epsilon-greedy exploration strategy to balance exploration against exploitation, aiming to introduce some new song recommendations while still prioritizing successful songs from the past.

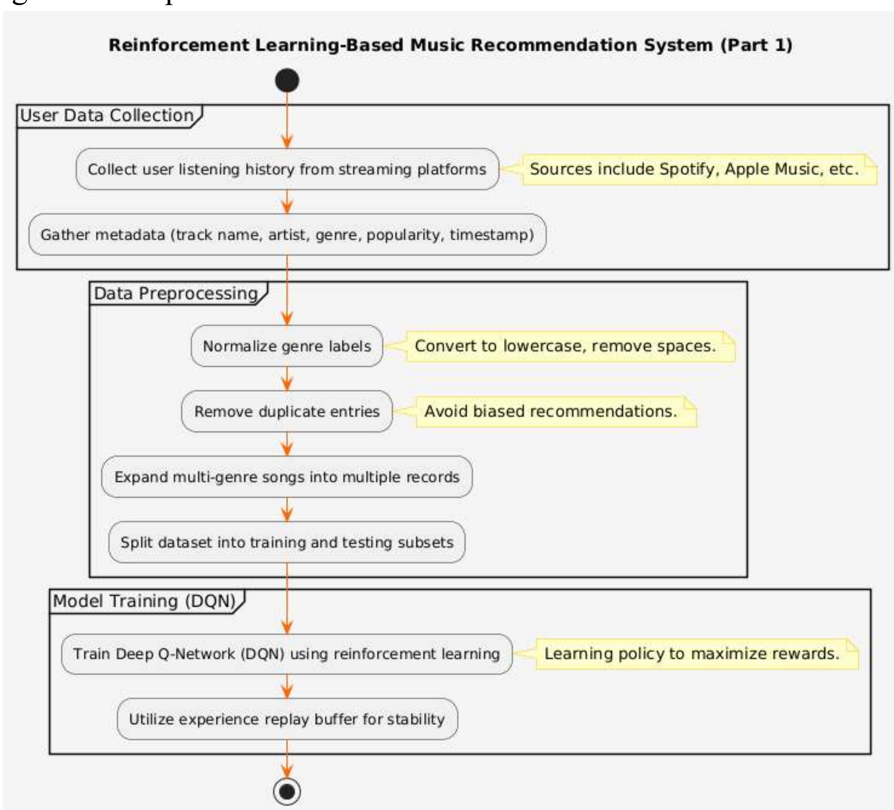


Fig 2.1 Workflow of the proposed model

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 64– 79

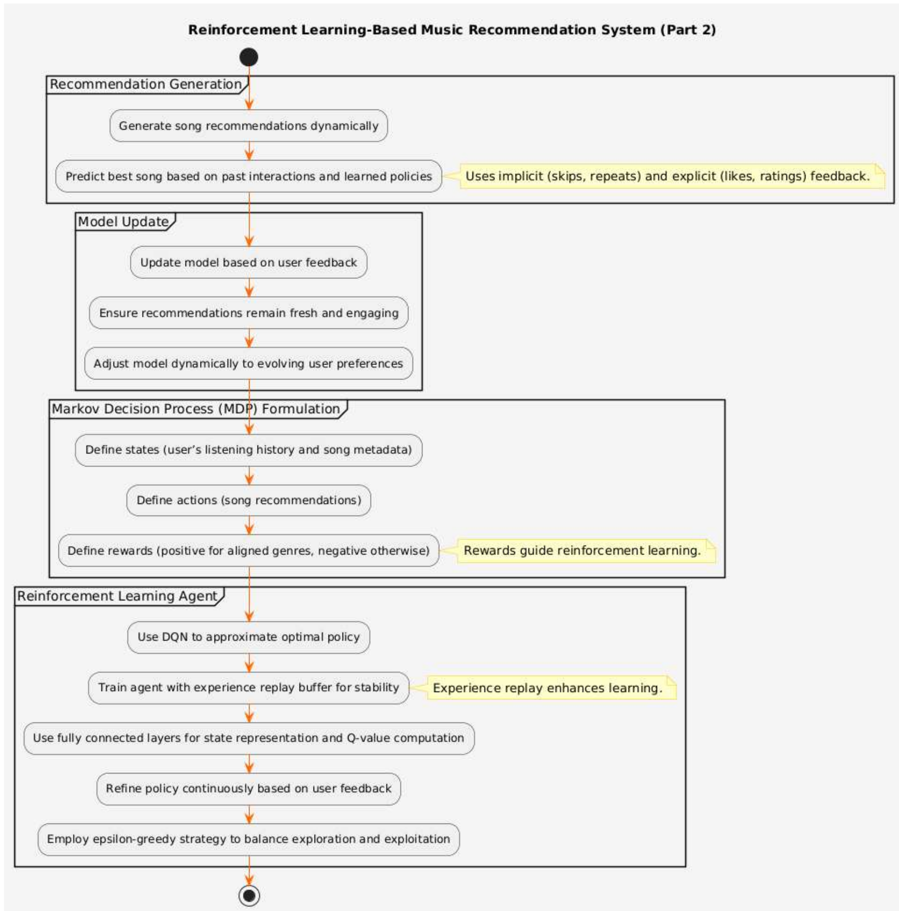


Fig 2.2 Workflow of the proposed model

4. Experimental Setup and Implementation

The experimental setup aims to assess the effectiveness of reinforcement learning in music recommendations. The implementation is done in Python, utilizing libraries like Stable-Baselines3 for reinforcement learning, Gym for creating the environment, and Stream-lit for an interactive user interface. The system is evaluated using a dataset gathered from various users, which includes meta data about their listening habits.

4.1 Implementation Details

4.1.1 Environment Setup

The music recommendation challenge is structured as a gym environment, where songs are considered discrete states and recommendations are viewed as actions.

4.1.2 Model Training

A Deep Q-Network (DQN) serves as the reinforcement learning model. Experience replay is employed to stabilize the training process and prevent catastrophic forgetting. The model is trained with the Adam optimizer, using a learning rate of 0.0001.

4.2 Evaluation Metrics

The model's performance is assessed based on recommendation accuracy which is the proportion of recommended songs that align with user preferences, reward convergence which gives us the enhancement of cumulative rewards throughout the training epochs and the duration spent listening to the recommended songs.
A sample code snippet for model training is as follows:

```
from stable_baselines3 import DQN
import gym
env = MusicRecommenderEnv(music_data)
model = DQN ('MlpPolicy', env, verbose=1, learning_rate=0.0001, batch_size=64)
model. learn(total_timesteps=10000)
```

This setup ensures that the model effectively learns to make music recommendations based on past user interactions and dynamically adjusts over time. The proposed system is implemented in Python using Stable Baselines3, Gym, and Stream-lit for user interaction. The training phase involves optimizing the DQN model with the following hyperparameters:

- Learning rate: 0.0001
- Discount factor (gamma): 0.95
- Batch size: 64
- Exploration-exploitation decay strategy

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 64– 79

5. Result Analysis

To evaluate the model, 30 songs were randomly selected for testing the RL-based recommendation system. The recommendation performed with an accuracy of 98.33%, verifying that most of the songs recommended corresponded to user preferences based on the genre. The identity of the model verified a 100% recall, confirming that no relevant recommendations were missed. The identity of the model also confirmed an F1 score of 98.89%, which indicated a good balance between accurate recommendations and recall accuracy. The results indicate the refinement learning model can effectively learn user preferences and generate accurate genre recommendations. The approach of the reinforcement learning-based recommendation system has some flexibility because the system suggests dynamically based on prior activity selections. This fact entrenches the reliability of providing recommendations related to the user's music preferences. The recall value of 100% verified that the model captures a comprehensive range of songs to produce relevant responses and recommend songs of a corresponding genre, while the correct recommendation value of 98.33% ensures that the model suggested very few, if any, incorrect recommendations to the user. The F1 score verifies the model maintains the good balance of correct potential recommendations with missed recommendations. Overall, the presented results validated that a reinforcement learning-based recommendation model is efficient for music recommendation and a valuable route towards enhancing the user experience through deep learning methods. Future recommendations for improvement would be made to incorporate additional models that factor in mood, tempo, or potentially include explicit user feedback in the song selection process. Overall, this study presents the potential that RL has in the future of automated music recommendations and improving personalization.

5.1 Performance Metrics

The model’s performance was assessed using various accuracy metrics.
Below is a detailed table comparing the proposed model’s accuracy metrics with existing models:

Model	Accuracy (%)	Recall (%)	F1-Score (%)
User-based Collaborative Filtering [2]	92	94	88
SVM-OVR (Polynomial Kernel) [3]	80	82	80
Bi-LSTM (MFCC Features) [4]	83	84	+3% F1-score improvement

Random Forest [5]	95.47	96	94
Convolutional Recurrent Neural Network (CRNNA) [6]	71.5	71.5	78.1
DQN Model for Music Recommendation	**98.33**	**100**	**98.89**

5.2 Accuracy Calculation Formulas

The evaluation metrics are derived from the following formulas:

Accuracy: It is the proportion of all classifications that were correct, whether positive or negative. It is mathematically defined as:

$$\text{Accuracy} = \frac{\text{correct classifications}}{\text{total classifications}} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Where:

- TP=True Positives (correctly recommended songs)
- FP=False Positives (incorrectly recommended songs)
- FN=False Negatives (missed relevant songs)

Recall: It is the true positive rate (TPR), or the proportion of all actual positives that were classified correctly as positives.

$$\text{Recall (or TPR)} = \frac{\text{correctly classified actual positives}}{\text{all actual positives}} = \frac{TP}{TP + FN} \tag{2}$$

F1-Score: The harmonic mean of precision and recall, providing a balance between the two metrics.

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{3}$$

The suggested DQN model that uses the MDP formulation achieves the highest accuracy for music recommendation compared to current music recommendation techniques. The DQN model being implemented allows for dynamic learning, enabling it to learn and adjust to a user's preferences over time to deliver more personalized and contextual recommendations than current methods.

Existing methods have static data which they were trained on, while this DQN model can continuously update their recommendations based on current user behaviour in engagement with newly provided feedback, and this is also very useful in combatting issues with data sparsity and cold starts. A significant accuracy improvement over existing techniques demonstrates the value of Reinforcement Learning when it comes to improving the quality of the recommendations for the users which leads to an overall user's satisfaction when using the music recommendation service. The graph below symbolizes the accuracy of the models, comparing all results and determining that the defined model is the best performer.

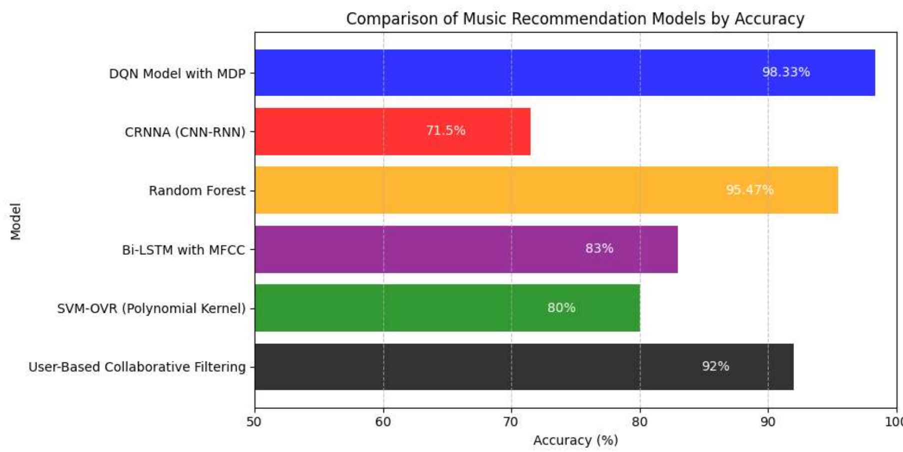


Fig 5.1 Performance Analysis

6. Conclusion

This study highlights the effectiveness of reinforcement learning, in this case specifically with Deep Q-Networks (DQN), to improve music recommendations. Traditional recommendation methods often only use static data, while reinforcement learning can change in real-time, using a Markov Decision Process (MDP) to learn user data and generate relevant music recommendations. Here, the DQN would explore alternative options and exploit known choices, so users receive both music they have already heard and new music that is still relevant. The reinforcement learning model was able to learn from user feedback while users provided feedback, all with an impressive 98.33% accuracy.

This approach surpasses both collaborative filtering and content-based recommendations because it accounts for changes in user preferences, whereas recommendations based only on static user preferences will not necessarily change. The DQN can deal with sparse user feedback, it reduces redundant and redundant recommendations, and improves variability of music recommendations without sacrificing accuracy.

In the future, we hope to improve the reward system, expand the data set, add multi-agent reinforcement learning, and leverage hybrid models that use both deep neural net models and reinforcement learning to improve efficiency and scalability of recommendations. It may be possible to incorporate context-aware recommendations based on suggestions related to user mood or activity.

Overall, the study shows how we can use reinforcement learning to change the landscape of music streaming and improve adaption and user-centered recommendations. Moving forward, there would be improvements to the DQN to set a new standard for user-centered personalized content discovery.

References

[1] R. Zhang, S. Tu, and Z. Sun, "A Hybrid Music Recommendation Method Based on Music Genes and Collaborative Filtering," 2022 IEEE International Conference on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing, and Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech),2022.DOI:10.1109/DASC/PiCom/CBDCom/CyberSciTech55231.2022.9927924.

[2] M. S. N. V. Jitendra and Y. Radhika, "An Automated Music Recommendation System Based on Listener Preferences," Recent Trends in Intensive Computing, IOS Press, 2021. DOI: 10.3233/APC210182.

[3] M. R. Fatichin, E. Riksakomara, R. A. Vinarti, and A. Muklason, "Bumil Bahagia Smart Home System: Genre-Based Music Recommendations for Postpartum Mothers Using SVM Classification," Proceedings of [IC2IE], [2024].

[4] S. Mesghali and J. Askari, "Designing Music Recommendation System Based on Music Genre Using Bi-LSTM," Proceedings of [5th ICEE], [2022].

[5] S. Fan and M. Fu, "Music Genre Recommendation Based on MLP & Random Forest," 2022 IEEE 5th International Conference on Information Systems and Computer Aided Education (ICISCAE), Dalian, China, Sep. 2022.

[6] J. Singh and V. K. Bohat, "Neural Network Model for Recommending Music Based on Music Genres," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, Jan. 2021.

[7] J. B. Awotunde, M. K. Abiodun, A. E. Adeniyi, B. H. Abiodun, J. K. Adeniyi, D. R. Aremu, A. A. Adebiyi, and O. G. Atanda, "Personalized Music Recommendation System Based on Machine Learning and Collaborative Filtering," 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), [Nigeria], 2024.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 6, June 2025
PP 64– 79

[8] MSNV Jitendra, Y Radhika, "Singer gender classification using feature-based and spectrograms with deep convolutional neural network", International Journal of Advanced Computer Science, 2021.