**NCF Recommendation System for Music Consumption Applications**

**Executive Summary**

**Problem and Hypothesis**

Modern business models rely on online consumption to generate profit and are therefore constantly seeking novel methods of increasing customer interaction. E-commerce platforms such as Amazon and Netflix (Chui et al., 2018) have designed complex recommendation systems to motivate their users to spend more on their platforms. This study implements a state-of-the-art recommendation system – Neural Collaborative Filtering – to recommend curated music artists to users. The proposed hypothesis of the project and the study outcome are summarized as follows:

Since the final model reaches a validation recall of 85%, the null hypothesis which states that an NCF model cannot be designed to reach 90 % recall accuracy, cannot be rejected.

**Data Analysis**

The dataset, downloaded from Kaggle (Larxel, 2021) contains data from real Spotify users whose identities have been anonymized for security. For each user, the dataset includes the songs that the user has consumed and the corresponding artist of the song. To prepare the dataset for the analysis, each user id and artist had to be tokenized. Users who listened to less than 6 songs were removed because the model would not be able to recognize the consumption patterns correctly. Once removed, there were +14,000 users and +289,000 artists remaining for the model. To create the training and testing datasets, a “leave-one-out” technique was used – for each unique user one item was removed from the original dataset to build the testing dataset. The removed item is known as the holdout item and would be used to test the model accuracy. The remaining items stayed in the training dataset which is used to fit the model.

Once data was processed, an NCF model, composed of General Matrix Factorization and a Multilayer Perceptron, was constructed (Figure 1). Upon training the model, the accuracy and the loss function of the training and the testing data were able to be assessed (Figure 2,3). Generating the accuracy and loss of the training and testing data was particularly useful because it revealed the optimal number of epochs that the model should be trained with to prevent overfitting and underfitting. After numerous runs of fitting the model, the optimal number of epochs was established at 25. Upon building and fitting the model, it had the ability to generate personalized music artist suggestions for each user. On average, generating predictions for a single user took 0.4 seconds – for users with more consumed items generating predictions would be slower and for users with less consumed items the generation would be quicker. The process of creating recommendations for all +14,000 users took 1 hour and 40 minutes. Since the output layer of the NCF network applied a sigmoid activation function the generated recommendation came in the form of a probability ranging from 0 to 1. In this study, predictions with a probability of greater than 0.5 were given a positive label signifying that the user would be interested in consuming that item.

**Figure 1**

*Model Structure*

**Diagram

Description automatically generated**

**Figure 2 Figure 3**

Chart, line chart, histogram

Description automatically generated

**Findings**

Tuning the training and model parameters allowed the final model to reach a recall accuracy of 85%. In other words, the model was able to correctly give the holdout item a higher than 0.5 probability of being consumed in 85% of records. Although the outcome of the study does not permit rejecting the null hypothesis, it highlights the power of Neural Collaborative Filtering as a recommendation system. To demonstrate the model performance further, exploratory data analysis was performed to analyze the relationship between the recommendation recall accuracy and the number of consumed items by user (Figure 7). As the graph shows, the model’s predictive capabilities are relatively consistent when faced with varying pool sizes of consumed items. Performing the exploratory analysis reveals that the model is generalizing well when faced with varying input sizes.

**Figure 7**

**Chart, bar chart

Description automatically generated**

**Limitations**

The analysis faced certain limitations. Primarily, the accuracy of the model was not high enough to reject the null hypothesis. Additionally, the data analysis was computationally expensive. Training the model lasted for 2 hours for 8 epochs only. Generating predictions for each user added an extra 1 hour and 40 minutes. By decreasing the learning rate or batch size, or increasing the model complexity, the training and validation computations further increased, making it increasingly more time and computation costly to perform the analysis. The third noteworthy limitation was that the model’s performance did not change as the neural network was made to be deeper and wider – adding additional units to existing layers or adding more layers did not impact the accuracy and only made the model more complex.

**Proposed Actions**

Before the model can be deployed for the business use case its accuracy needs to be increased. There are three main approaches which have the ability to improve the study :

1. Increasing the number of users could improve the model by giving it more data to learn from. Specifically, increasing the number of users would allow the model to gain a deeper understanding of music consumption behavior that listeners exhibit.
2. Creating a supplementary machine learning model which will be trained on item metadata such as the artists’ genre, mood, tags, and language. By incorporating these features, it is possible to perform a supplementary classification model (such as KNN) to identify similar artists to serve as recommendations.
3. Designing an NLP algorithm to reduce the dimensionality of items. As seen in Figure 8, an artist named Akon is categorized as a separate item for each new featuring artist he has collaborated with. In reality, a user who has listened to Akon featuring Eminem is likely to listen to Akon featuring Snoop Dog. Therefore, an NLP algorithm should be devised to categorize all of the below artists as Akon.

**Figure 8**

*Example of High Dimensionality*

**A picture containing graphical user interface

Description automatically generated**

**Study Benefits**

The created model and the proposed actions have the potential to create a industry-standard music recommendation system. The study is beneficial because it demonstrates the power of Neural Collaborative Filtering as a recommendation generator while also identifying its limitations and outlining potential data analytics techniques that can serve as remedies. Since the model gave relevant results for 85% of users, The NCF could be used as a beta feature on the music platform to help users discover new artists.

**References**

Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P., & Malhotra, S. (2018, April 17). *Notes from the AI frontier: Applications and value of Deep Learning*. McKinsey & Company. Retrieved April 14, 2023, from https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning#

Larxel. (2021, November 15). *Spotify playlists*. Kaggle. Retrieved April 13, 2023, from h https://www.kaggle.com/datasets/andrewmvd/spotify-playlists

Pashtukov, D. (2022). *How Spotify's algorithm works? A Complete Guide to spotify recommendation system [2022]: Music tomorrow blog*. How Spotify's Algorithm Works? A Complete Guide to Spotify Recommendation System [2022] | Music Tomorrow Blog. Retrieved April 4, 2023, from <https://www.music-tomorrow.com/blog/how-spotify-recommendation-system-works-a-complete-guide-2022>