DSC 680 -Project 1 Milestone 3

Abhijit Mandal

# Answers to Q&A

**Questions:**

1. **New User:** A newly released movie cannot be recommended to the user until it gets some ratings. A new user or item added based problem is difficult to handle as it is impossible to obtain a similar user without knowing previous interest or preferences. How to handle this scenario?

**Answer:** For any new user we can take the user base with similar geography, age group, interests etc. which the new user might have provided while signing up, we can then recommend movies based on those.

1. **Synonymy** arises when a single item is represented with two or more different names or listings of items having similar meanings, in such condition, the recommendation system can’t recognize whether the terms show various items or the same item. How can we address this issue?

**Answer:** For this we will do some data clean up and validation activities before building our model.

1. Scalability of the model ?

**Answer**: This model is scalable to handle any size of data, however the performance will slow down as the data grows up, we can also look to include other models like XGBoost.

1. Drawbacks and limitations of Collaborative Filtering?

**Answer**: The prediction of the model for a given (user, item) pair is the dot product of the corresponding embeddings. So, if an item is not seen during training, the system can't create an embedding for it and can't query the model with this item. This issue is often called the **cold-start problem**.

1. Drawbacks and limitations of Content Based Filtering?

Answer: The model can only make recommendations based on existing interests of the user. In other words, the model has limited ability to expand on the users' existing interests.

1. Drawbacks and limitations of Demographic Filtering?

**Answer**: This is a very basic filtering technique where the system recommends the same movies to users with similar demographic features. So the recommendations may not be accurate as per the users preferences.

1. Which is the best Algorithm for Recommendation?

**Answer**: Hybrid movie recommendation systems can take advantage of content- based and collaborative filtering as the two approaches are proved to be the best approach.

1. What are the factors that are taken into consideration while recommending a movie to the user, for eg. Age, demography, ethnicity, interests, language etc.

**Answer**: Yes, All these factors are taken into consideration, along with that the 3 top actors, the director, related genres and the movie plot keywords.

1. How much data is enough to predict recommendations for a user or does this changes with the amount of data we process?

**Answer**: For prediction, I would say the more data the better, the recommendation do change with the amount and variety of data we process.

1. How soon this recommender system needs to be re-trained ?

**Answer**: In a real world scenario, I would say this needs to be trained if there is any major changes in the user selection, for this we can have an anomaly detection model to identity any changes in the users preferences and then we can rebuild our recommendation system model for that user.