Lab 8 – SECTION B, BATCH 3 Date: 20th Sept. 2022

Exer 1: Collaborative Filtering

1. Read about the movielens dataset and write down a summary of metadata.

User-Based Similarity

2. Read the "ratings.csv" file and create a pivot table with index='userId', columns='movieId', values = "rating.
3. sklearn.metrics.pairwise_distances can be used to compute distance between all pairs of users. pairwise_distances() takes a metric parameter for what distance measure to use. Use cosine similarity for finding similarity among users. Use the following packages.

```
4. from sklearn.metrics import pairwise_distances
5. from scipy.spatial.distance import cosine, correlation
```

6. Find the 5 most similar user for user with user Id 25.
7. Use the "movies" dataset to find out the names of movies, user 1 and user 338 have watched in common and how they have rated each one of them.
8. Use the movies dataset to find out the common movie names between user 2 and user 338 with least rating of 4.0

Item-Based Similarity

9. Create a pivot table for representing the similarity among movies using correlation.
10. Find the top 5 movies which are similar to the movie "Godfather".