- **CSE 523 Machine Learning Project Report**

Movie Recommender System

Weekly Report

-Kashish Jivani(AU1940161)

-Neel Popat(AU1940165)

-Yashvi Navadia(AU1940123)

Content based Recommender System:

Task Completed:

This week, we performed Content based filtering. It is a step ahead of the Simple recommender system. Simple recommendation system used to provide generalized recommendations. Content based filtering will be able to provide recommendations based not only on the ratings and popularity but also on the metadata of the movies such as overview, tagline, cast, crew, keyword, genre. We will be making two recommenders:

1. Based on Movie Overview and Tagline
2. Based on Movie Cast, Crew, Keyword, Genre

We will be using the small dataset for content based filtering as processing time increases for the big dataset. As we do not have a quantitative metric to judge our machine's performance, we have done it qualitatively. Therefore we have used TF-IDF Vectorizer to transform the text data into a vector. We performed cosine similarity functions on all the movies. The following is the formula to calculate Cosine similarity:

$$cosine(x, y) = \frac{x.yT}{||x||.||y||}$$

Outcome of the task performed:

Cosine similarity calculates a numeric quantity that denotes the similarity between two movies. We will now have a pairwise cosine similarity matrix for all the movies in our dataset and can list down similar movies by entering any movie of our choice. Please find the results below:

1) Overview, tagline

   get_recommendations('The Dark Knight').head(10)

Sr.No	Movie
7931	The Dark Knight Rises
132	Batman Forever
1113	Batman Returns
8227	Batman: The Dark Knight Returns, Part 2
7565	Batman: Under the Red Hood
524	Batman
7901	Batman: Year One
2579	Batman: Mask of the Phantasm
2696	JFK
8165	Batman: The Dark Knight Returns, Part 1

2) Cast, crew, genre

get_recommendations('The Dark Knight').head(10)

Sr.No	Movie
8031	The Dark Knight Rises
6218	Batman Begins
6623	The Prestige
2085	Following
7648	Inception
4145	Insomnia
3381	Memento
8613	Interstellar
7659	Batman: Under the Red Hood
1134	Batman Returns

Task to be performed in upcoming week:

Implement Collaborative based filtering which provides personalized recommendations based on the users' preferences.

References:

- https://rdrr.io/cran/bcv/man/cv.svd.html
- https://stackoverflow.com/questions/56273351/how-to-validate-test-set-on-trained-svd-model