

# CSE 523 Machine Learning Project Report

## Movie Recommender System

### weekly Report

-Kashish Jivani(AU1940161)

-Neel Popat(AU1940165)

-Yashvi Navadia(AU1940123)

#### Task completed:

The task of the Simple Recommender System is to offer generalized recommendations to every user based on popularity of movie and sometimes genre. We sorted our movies based on the ratings and popularity and displayed the top movies of our list. We have used weighted mean for the same.

Previously we saw that the top 3 movies were from the same director. The output indicates a strong bias of TMDb users towards particular genres and directors. Now as an added step, we have also passed in a genre argument to get the top movies of a particular genre. We have reduced the cutoff for being in the top to 85% from 95%. And we kept the genre as “Romance”. We use IMDB’s weighted mean formula:

$$\text{Weighted Rating (WR)} = \left( \frac{v}{v+m} \cdot R \right) + \left( \frac{m}{v+m} \cdot C \right)$$

where,

- $v$  is the number of votes for the movie
- $m$  is the minimum votes required to be listed in the chart
- $R$  is the average rating of the movie
- $C$  is the mean vote across the whole report

```
s = md.apply(lambda x:
pd.Series(x['genres']),axis=1).stack().reset_index(level=1, drop=True)
s.name = 'genre'
gen_md = md.drop('genres', axis=1).join(s)
def build_chart(genre, percentile=0.85):
```

```
qualified['wr'] = qualified.apply(lambda x:
(x['vote_count']/(x['vote_count']+m) * x['vote_average']) +
(m/(m+x['vote_count']) * C), axis=1)
```

```
build_chart('Romance').head(15)
```

Outcome of the task performed:

	title	year	vote_count	vote_average	popularity	wr
10309	Dilwale Dulhania Le Jayenge	1995	661	9	34.457	8.565285
351	Forrest Gump	1994	8147	8	48.3072	7.971357
876	Vertigo	1958	1162	8	18.2082	7.811667
40251	Your Name.	2016	1030	8	34.461252	7.789489
883	Some Like It Hot	1959	835	8	11.8451	7.745154
1132	Cinema Paradiso	1988	834	8	14.177	7.744878
19901	Paperman	2012	734	8	7.19863	7.713951
37863	Sing Street	2016	669	8	10.672862	7.689483
882	The Apartment	1960	498	8	11.9943	7.599317
38718	The Handmaiden	2016	453	8	16.727405	7.566166
3189	City Lights	1931	444	8	10.8915	7.558867
24886	The Way He Looks	2014	262	8	5.71127	7.331363
45437	In a Heartbeat	2017	146	8	20.82178	7.003959
1639	Titanic	1997	7770	7	26.8891	6.981546

The top 15 romantic movies from the list showed up which got more than 85% of votes and is rated high.

Task to be performed in the upcoming week:

We will look upon our next stage of recommender system which is content based filtering which is a personalized recommendation system. It gives personalized recommendations based on the past views and votes of the user.

References:

<https://www.kaggle.com/code/rounakbanik/movie-recommender-systems/notebook>

<https://www.analyticsvidhya.com/blog/2020/11/create-your-own-movie-movie-recommendation-system/>