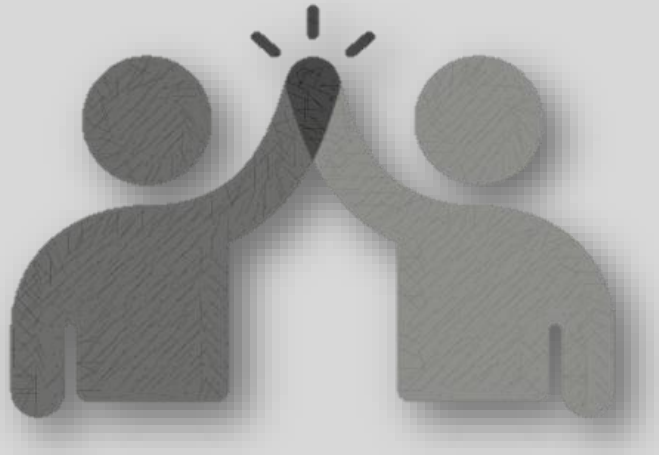
Movie Analysis And Recommendation System

Group No. 41

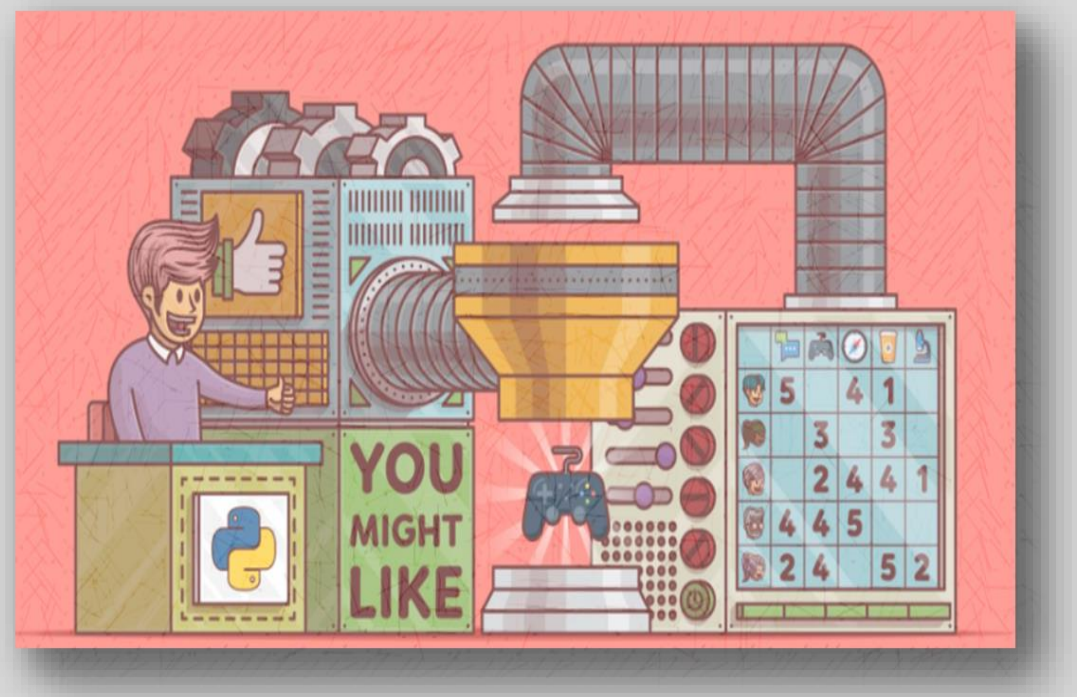
Members Details :

1. Abhinav Dudeja (21111001) (abhinavd21@iitk.ac.in)
2. Malkeet Singh Dhalla (21111035) (malkeetsd21@iitk.ac.in)
3. Manthan Kojage (21111039) (manthank21@iitk.ac.in)
4. Sandeep Saket (21111055) (sandeeps21@iitk.ac.in)
5. Shivam Kharwar (21111058) (skharwar21@iitk.ac.in)



Introduction

The Bollywood industry plays a vital role in India's economy. Every year hundreds of Bollywood movies are released in India with the hope of becoming the next blockbuster, but only one-third of them gets success. The producers, actors, investors, and sponsors are all interested in predicting the movie's box office success. Our project is on analyzing the effect of the genres, the release date around holidays, the release month, actors, directors, ratings on the box office success of a movie.



Movie Analysis And Recommendation System

It can predict best movie based on rise and fall of ratings.

It can also virtually present the analysis for the clear understanding of the user.

It reduces the search and prediction time thus it is fast as small dataset needed for a user.

Accounted Datasets

Our Recommendation and Analysis System analyze and make use of the following datasets from which some are found online and sone are scrapped using data mining and web scrapping techniques. The datasets which are present online can be found at: Dataset1,Dataset2,Dataset3,Dataset4,Dataset5,Dataset6,Dataset7

Some datasets are scrapped from BoxOfficeIndia using python library BeautifulSoup (WEB SCRAPPED)

This datasets contains Various attributes related to movies which will be used by us for the analysis part.

Few of those important attributes from the datasets found includes:

1.The title of the movie

2.IMDB Rating of the movie

3.The Release date of the Movie

4.The total nett collection of the movie in india

5.Lead actors from the movie

6.The worlwide total gross figure for the movie

7.The budget of the movie.

8.Language in which the movie is made

9.Director of the movie

10.Producer of the movie

11.Genre of the movie

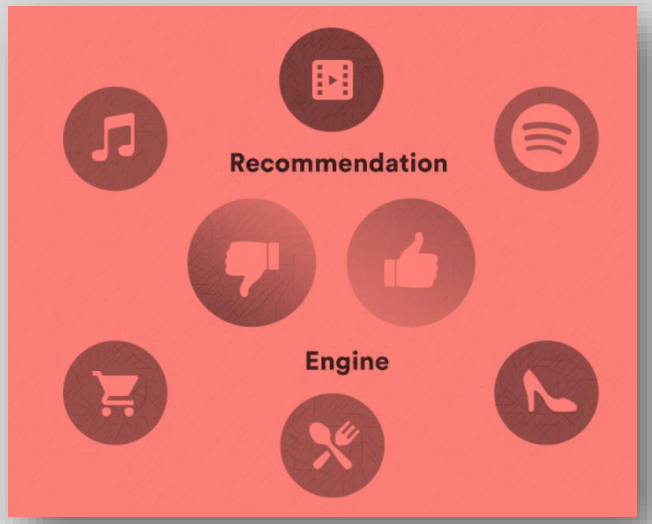
12.The overseas nett collection of the movie

13.The tickets sold for the Movie

14.Verdict of the Movie (Hit,Flop etc.)

Project Analysis

- Movies Recommendation System is a program or set of programs that predicts or filters the movies by using various Machine Learning Algorithm techniques based on the taste of users and gives movie or list of movies that a user will prefer to watch on above of all other movies.
- Recommending Movies is a two-step process that consists of the data management for the recommendation system and second step that consists of using that extracted data to fit an efficient ML-model to make the prediction or recommendation system.
- Our Project's focus was made on the first step that consists of data handling and management step. We also made a demo recommending system without any ML algorithm just to give view how recommendation system uses these dataset to predict movies.



Project Focus

- Our project focuses on the information extraction from web scrapped data, then using that information frame to solve the queries regarding box-office movies considering only Bollywood movies and developing a demo recommending system.
- We used the datasets and for some queries we web scrapped datasets to implement in our recommendation and analysis systems.
- Data Handling and Extraction comes under Data Mining part which is necessary to model data into a structure which is efficient to be fit for the model for Machine Learning.
- We used various Data Mining and web scrapping techniques and software to extract data frames for our project which is mentioned in next slide.

Project Results and Findings

Our project solved various queries which will help in prediction various rise and falls in Bollywood trend accordingly.

We have made following findings for the Movies like Gross-collection, Profit-made, Genres, Average collections for the movies, different rating for the different genres and their average ratings, popular directors and actors over time along with their hits and average ratings.

We have also made a Demo Recommendation system that works on content filtering and recommends movies based on genres and actors.

Conclusion

We fractionated our project into two parts namely : Analysis System and Recommendation System. Analysis System was the main theme of our project in which we use several data scrapping techniques to extract Bollywood movies data from internet web sites namely Bollywood Hungama, IMDB, Box Office.

Analysis System is the part where we extracted information from scrapped data modelled them into the structures and used them to solve the queries.

Recommendation System is the part where we used a subset of datasets extracted and made a demo recommended system just on the attributes of the data frame.

Benediction

We would like to thank our instructor, ‘Mr. Arnab Bhattacharya’ Sir for approving and being a constant source of help and guidance for the project which helped us in exploring the insights of the Data Mining. We are very much clear with the basics of the subject because of the completion of the project. The challenges and the obstacles we encountered in the path of this project helped us to gain a concrete knowledge of Data Mining and Extraction.

We would also like to express our gratitude towards our friends and internet resources that helps in completing our project.

Thank You