MOVIE RECOMMENDATION SYSTEM

A project report submitted in the partial fulfillment of the requirements for the

Award of the degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE & ENGINEERING

Submitted By

Prakash Nagamani

Reg.no: 121910312048

Under the Guidance of Mr.Arun Muralidharan

24 June 2022

DECLARATION BY THE CANDIDATE

I the undersigned solemnly declare that the project report "MOVIE RECOMMENDATION SYSTEM" is based on my own work carried out during the course of our study under the supervision of Mr.Arun Muralidharan. I assert the statements made and conclusions drawn are an outcome of my research work. I further certify that

I. The work contained in the report is original and has been done by me under the general supervision of my supervisor.

II. The work has not been submitted to any other institution for any other degree/diploma/certificate in this university or any other University of India or abroad.

III. We have followed the guidelines provided by the university in writing the report.

IV. Whenever we have used materials (data, theoretical analysis, and text) from other sources, we have given due credit to them in the text of the report and giving their details in the references.
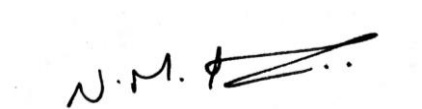
Prakash Nagamani

121910312048

PG-2223-ETSD-864

CERTIFICATE

This is to certify that this project work entitled

"MOVIE RECOMMENDATION SYSTEM"

is the Bonafide work carried out by Prakash Nagamani, Reg.No: 121910312048 submitted in Partial fulfillment of the requirement for the Award of Degree of Bachelor of Technology in Computer Science and Engineering, during May-June 2022.

The results submitted in this project have been verified and are found to be satisfactory. The results embodied in this thesis have not been submitted to any other university for the award of the any other degree/diploma.

Signature of project supervisor

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encouragement crowned the efforts with success. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

The first person I would like to thank my project guide Mr. Arun Muralidharan, who had given continuous critical suggestions and extension of proper working atmosphere, abiding interest has finally evolved into this research work.

It is indeed with a great sense of pleasure and immense sense of guidance that I acknowledge the help and I am highly indebted to Prof. M.Saratchandra Babu, Principal, and School Of Technology, for his support during the tenure of the internship.

I would like to express my sincere thanks to Prof. Sireesha Rodda, Head of the Department of Computer Science and Engineering for providing the opportunity to Undertake this internship and encouragement in the completion of the project.

I am also thankful to all the staff members of the Computer Science and Engineering Department for their valuable suggestions. I would like to thank my team mates and parents who extended their help, encouragement and moral support either directly or indirectly in this project.

Prakash Nagamani

121910312048

CONTENTS

ABOUT

Phoenix Global is a skill-development company that helps students acquire and master professional and soft skills as per the requirements of the industry benchmarked to world's top firms, trained by top class industry professionals.

Phoenix Global is a platform having Industry professionals with esteemed alma mater including the IITs and IIMs to mentor and train students on cutting-edge skills, critical to the emerging industries while also giving them an opportunity to intern on a project under the mentorship of industry professionals from the IITs /IIMs.

Our vision is to be a national leader in skill development and industry readiness training by providing differentiated training from top-class industry experts. The mission is to be a go-to skill development platform for students, imparting skills benchmarked at global standards that help them realize their dream careers profitably

Our core values, the 4Ps – Professionalism, Punctuality, Passion, Perseverance stand for who and what we are as an organization.

SCHEDULE OF INTERNSHIP

Day	Activity Plan
1	Induction Program
2	Pre-Readings/Material Distribution
3	Training Session - 1
4	Training Session - 2
5	Training Session - 3
6	Training Session - 4
7	Training Session - 5
8	Teams formation for Project
9	Weekend Off
10	Training Session - 6
11	Training Session - 7
12	Training Session - 8
13	Training Session - 9
14	Training Session - 10
15	Project Title Allocation
16	Weekend Off
17	Project Session - 1
18	Project Session - 2
19	Project Session - 3
20	Project Session - 4
21	Project Session - 5
22	Project Mid Review
23	Weekend Off
24	Project Session - 6
25	Project Session - 7
26	Project Session - 8
27	Project Session - 9
28	Project Session - 10
29-44	Project Working Sessions
45	Project Final Presentation and Thesis Defense

ABSTRACT

The importance of a movie recommendation system is acknowledged by the users as it helps them enhance their entertainment. It can suggest a set of movies to the users based on their interests and the popularity of the movies. This movie recommendation system that can help users find the best movies. The system can collect various information about the movies, such as their popularity, genre, and plot. It can then generate movie swarms to help the producers plan a new movie.

Problem Statement

The rise of recommendation systems has also been attributed to the increasing number of people constantly looking for the best possible products and services. To help them make informed decisions and avoid spending their cognitive resources on something that they don't need. Create a "Movie Recommendation System" to help the users with the problem.

Description of the Model

The goal of the movie-based filtering Model is to analyze the similarities between different movies based on their metadata. It will then suggest those that are most similar to the one that a user searched.

Types Of Filtering:

1. Content-based Filtering System:

The concept of content-based filtering is a domain-dependent method that focuses more on analyzing the attributes of web pages and publications to generate predictions. It is very effective when articles and publications are recommended. In content-based filtering technique, recommendation is made based on the user profiles using features extracted from the content of the items the user has evaluated in the past.

2. Collaborative Filtering System:

A collaborative filtering technique is a domain-independent method that can help predict the content that cannot be easily described by metadata such as music and movies. It uses a database to build a list of users' preferences for items, and it then matches those preferences with those of users based on similarities between them.

3. Hybrid Filtering System:

The goal of hybrid filtering is to combine different methods in order to improve the system's efficiency and avoid some of the drawbacks of a single algorithm. It's also beneficial to have a combination of algorithms that can provide better recommendations.

For this Model, we use **collaborative filtering system.**

The goal of a **collaborative filtering system** is to analyze and predict the behaviors and preferences of its users. It then uses this information to recommend the most appropriate movies based on their specific search. Unlike machine-based filtering techniques, this method does not rely on analyzable content. Its ability to recommend complex products without requiring an understanding of the item itself is also a key advantage of this method.

We will create a Model using K-Nearest Neighbour(KNN) Algorithm.

We will build a Collaborative Based Recommender using Movie Genres, Cast, Crew, and Keywords.

COLLABORATIVE FILTERING

Read by both users

Similar users

Read by her,
recommended to him!
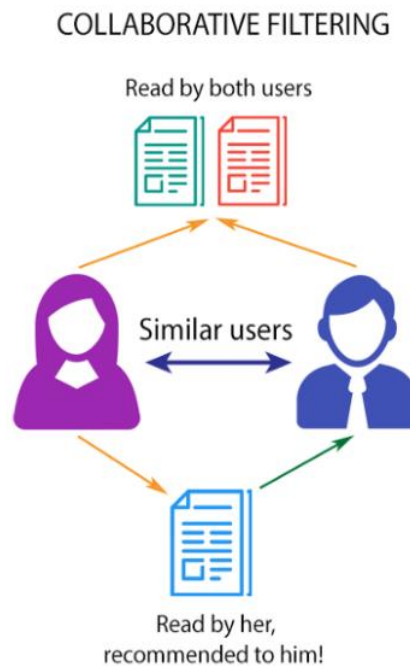
STATISTICAL ANALYSIS

OBJECTIVE: To analyze the most Popular movies according to their genres.
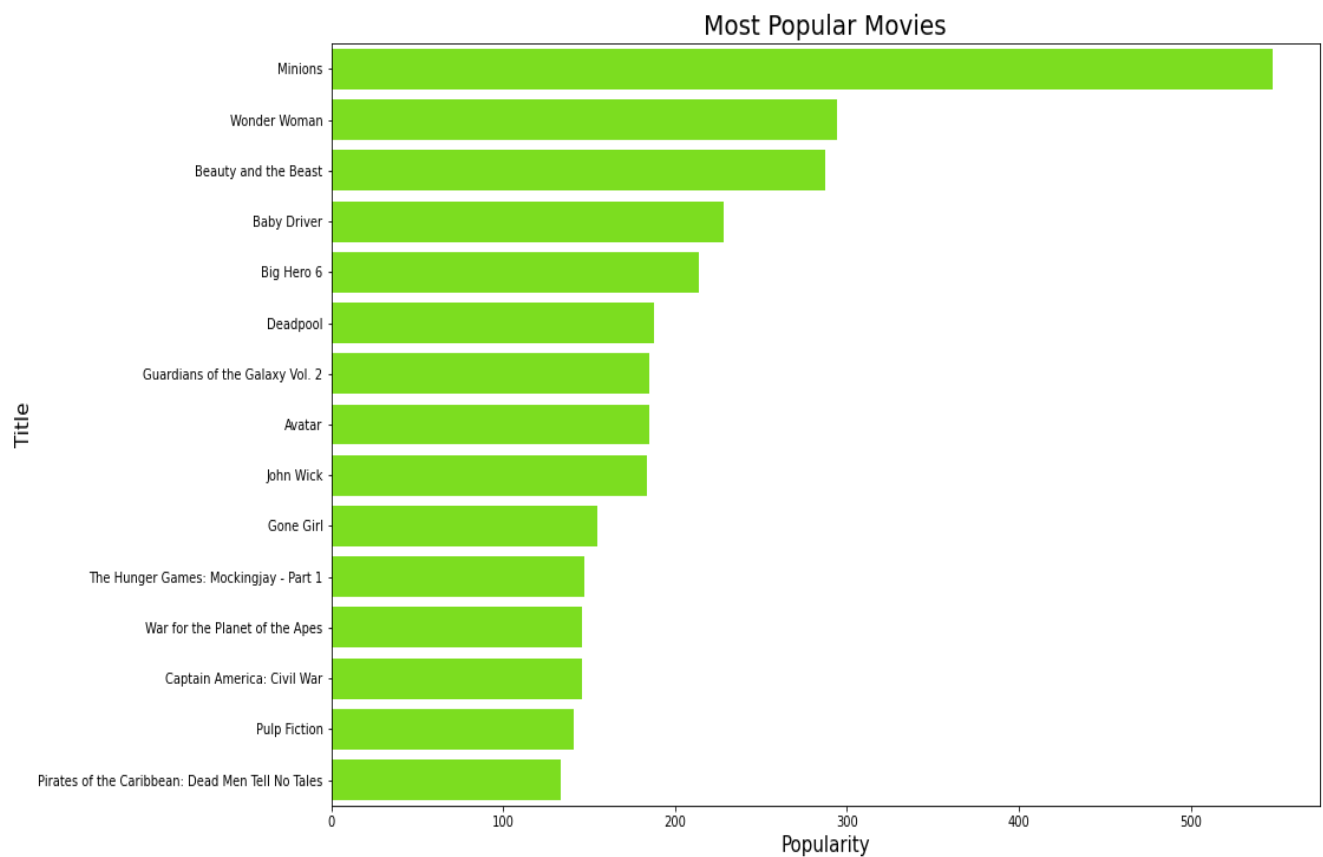
To see full sized image,

Fig-1: The graph shows "Top 15 Most Popular Movies".

We can see that from the fig-1, **Minions** was the most popular movie with a popularity rating of over 500 followed by Wonder woman, Beauty and the beast, etc.
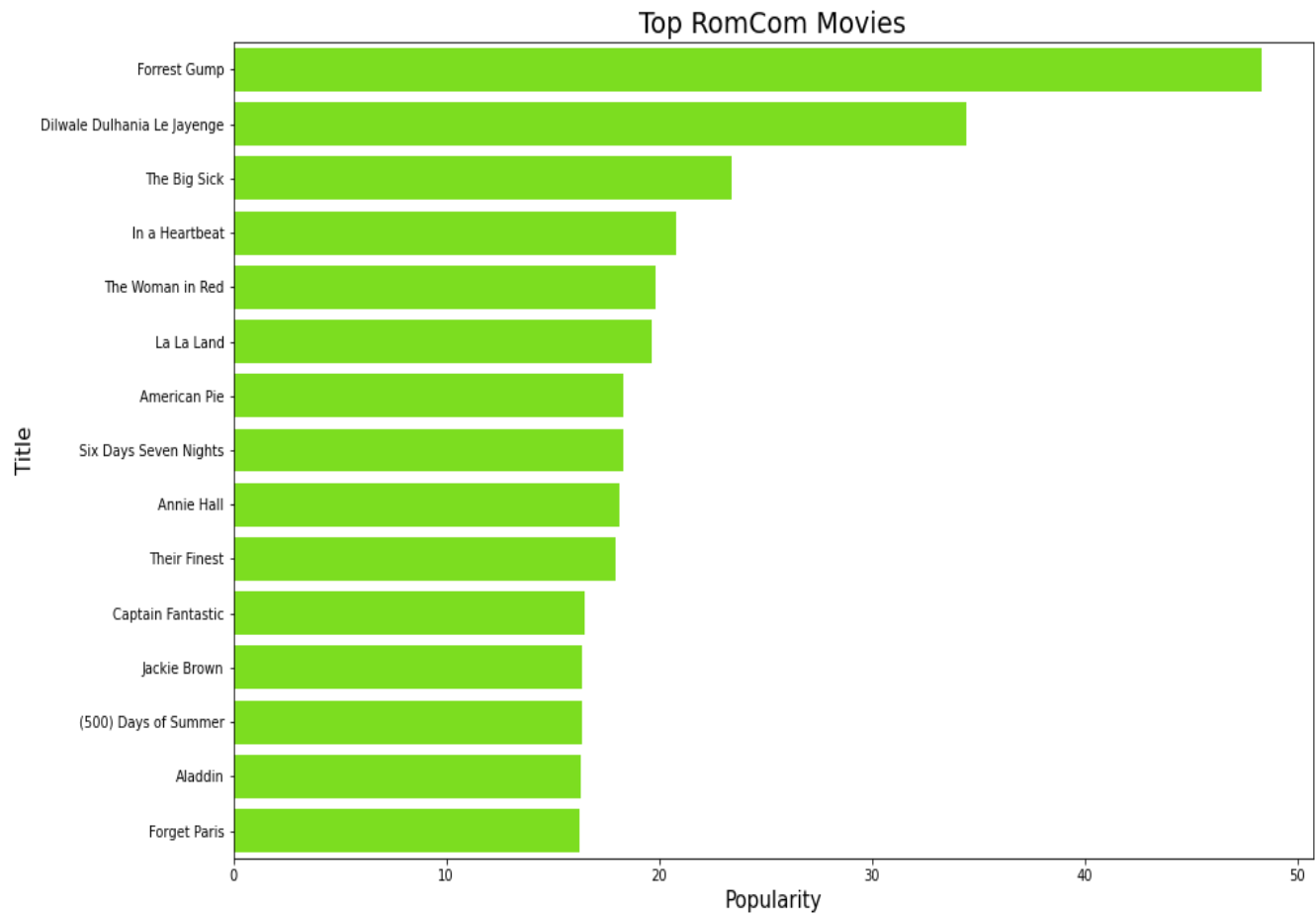
To see full sized image, Click Here

Fig-2: The graph shows "Top 15 Most Popular Rom Com Movies".

We can see that from the fig-2, **Forrest Gump** was the most popular romance comedy movie with a popularity rating nearly 50.
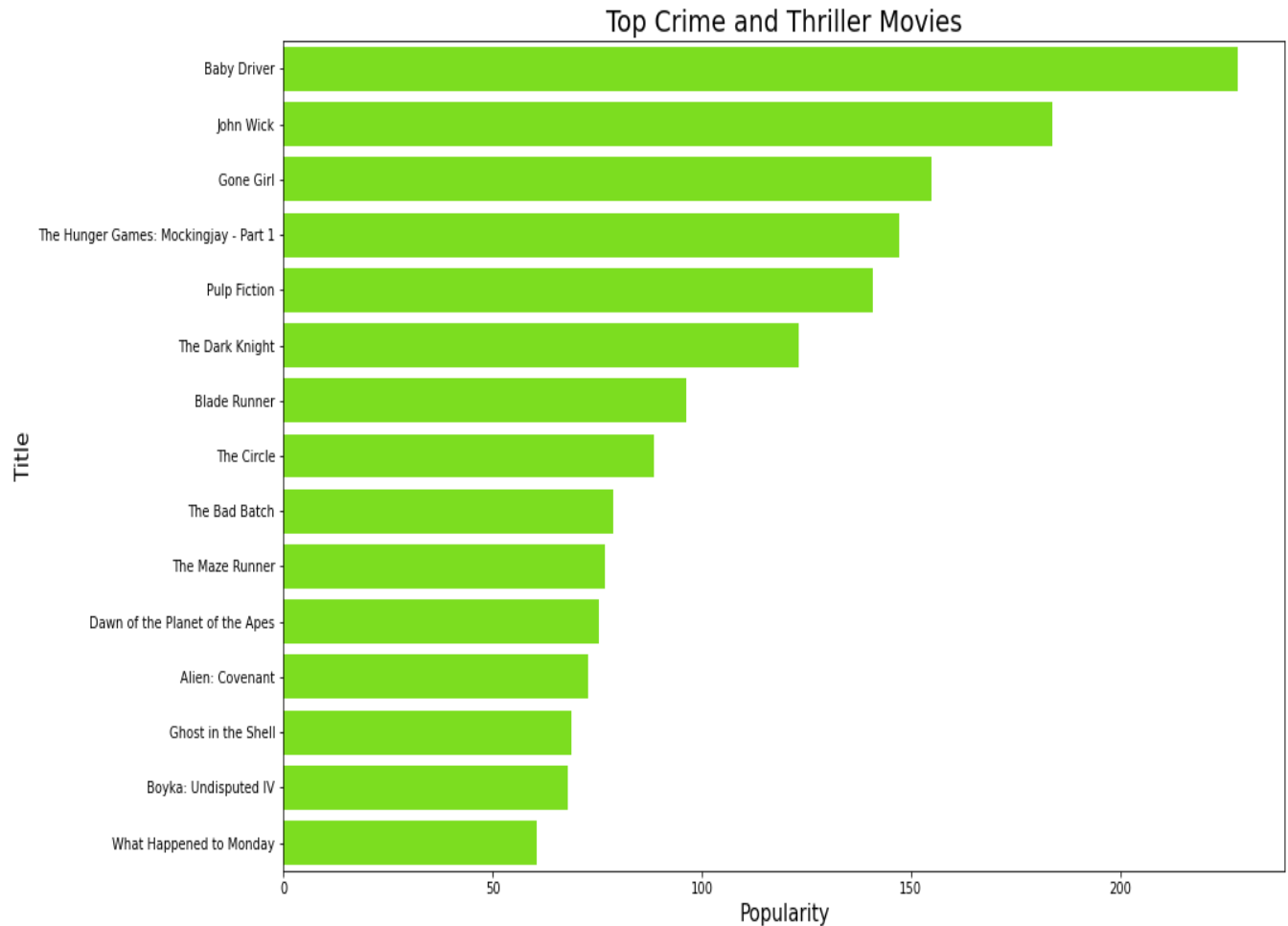
To see full sized image, <u>Click Here</u>

Fig-3: The graph shows "Top 15 Most Popular Crime & Thriller Movies".

We can see that from the fig-2, **Baby Driver** was the most popular romance comedy movie with a popularity rating over 200.
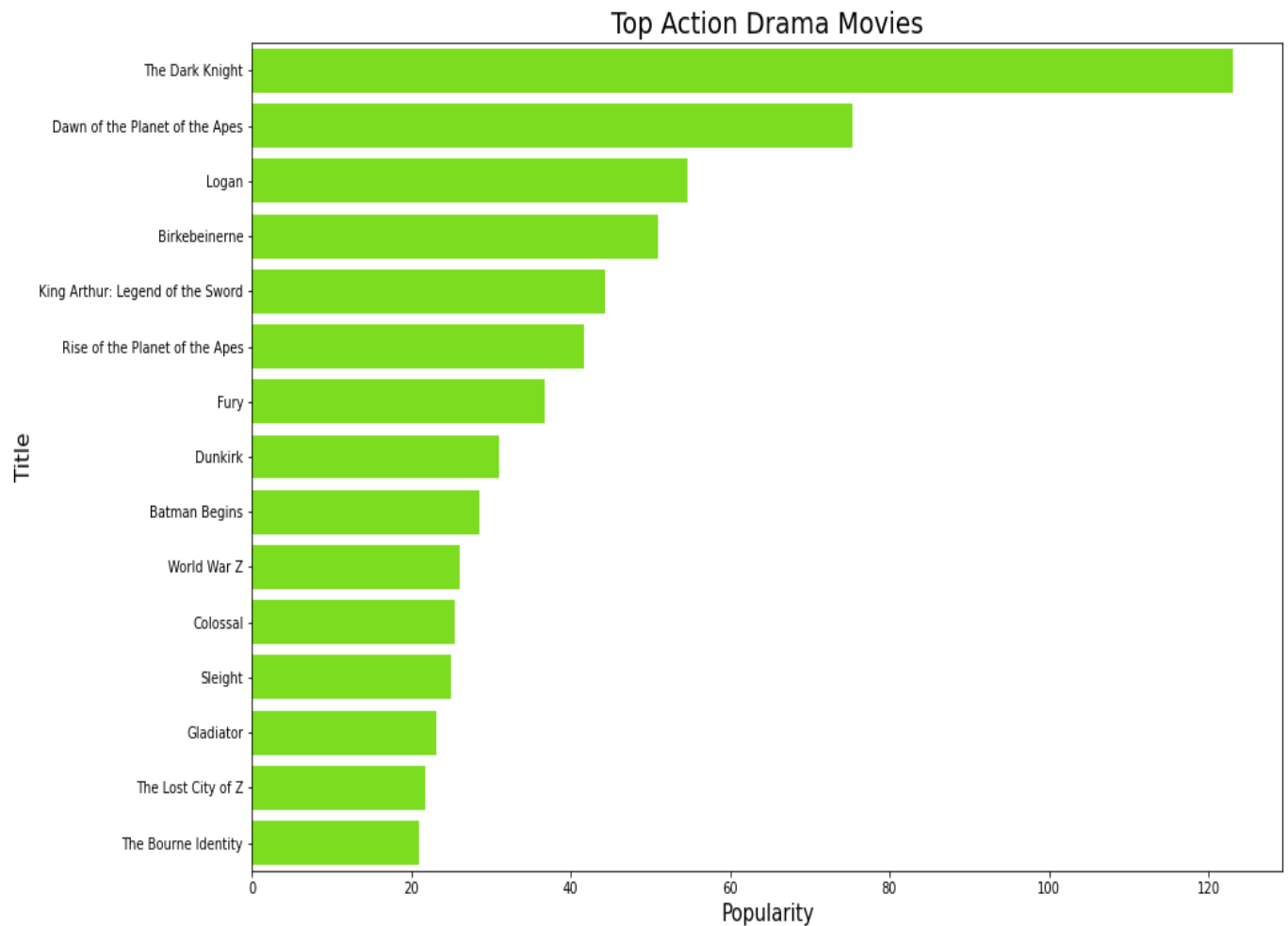
To see full sized image, Click Here

Top Action Drama Movies

Fig-4: The graph shows "Top 15 Most Popular Action Drama Movies".

We can see that from the fig-4, **Dark Knight** was the most popular Action Drama movie with a popularity rating over 120.
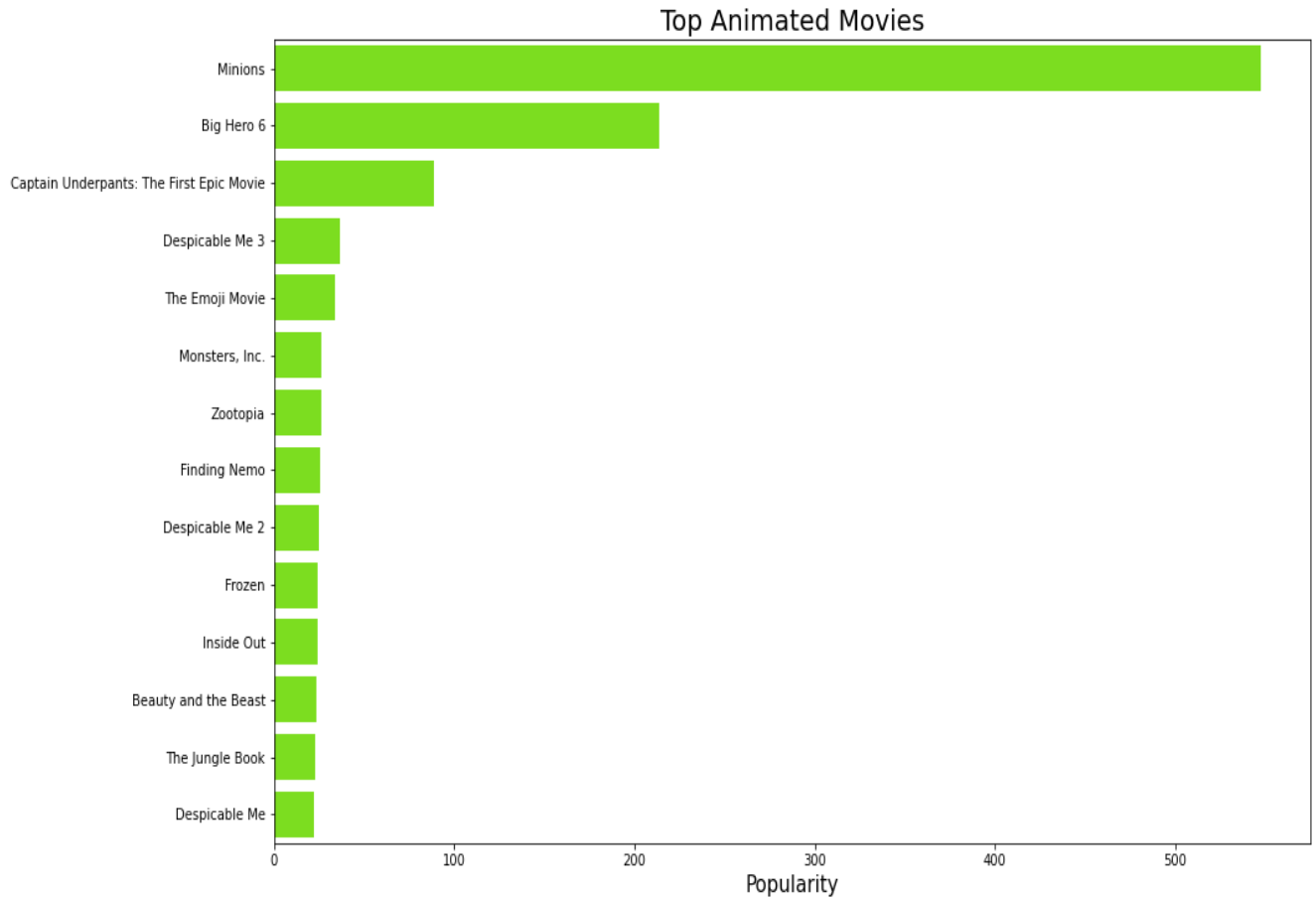
To see full sized image, <u>Click Here</u>

Fig-5: The graph shows "Top 15 Most Popular Animated Movies".

We can see that from the fig-5, **Minions** was the most popular Animated movie with a popularity rating over 500.
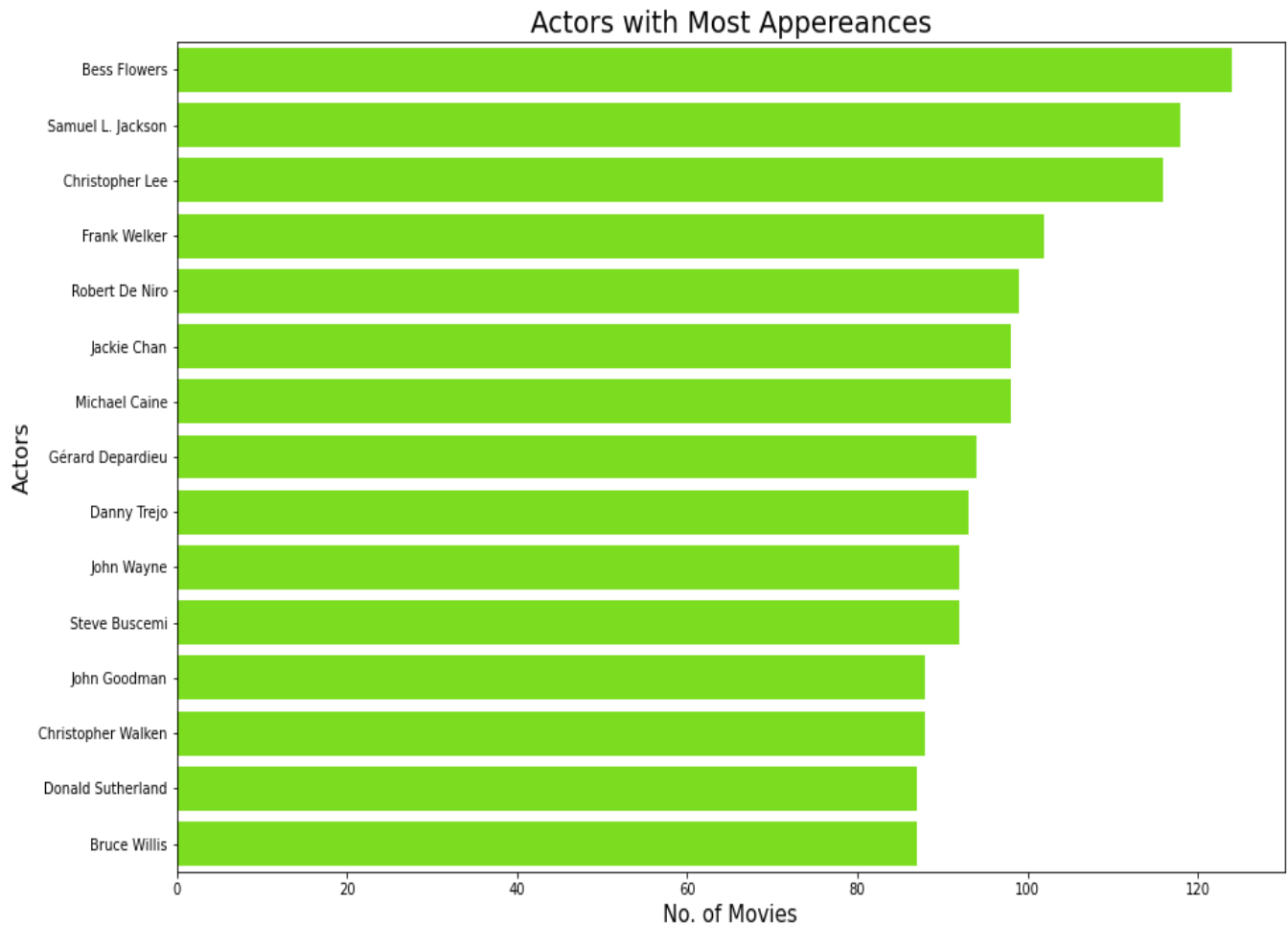
To see full sized image, <u>Click Here</u>

Fig-6: The graph shows "Top 15 Actors with Most Appearances".

We can see that from the fig-6, **Bess Flowers** was the Actor with the Most Appearances in over 120 films(i.e., 124 Movies). followed by Samuel L. Jackson, Christopher lee with 118 and 116 Movies respectively.
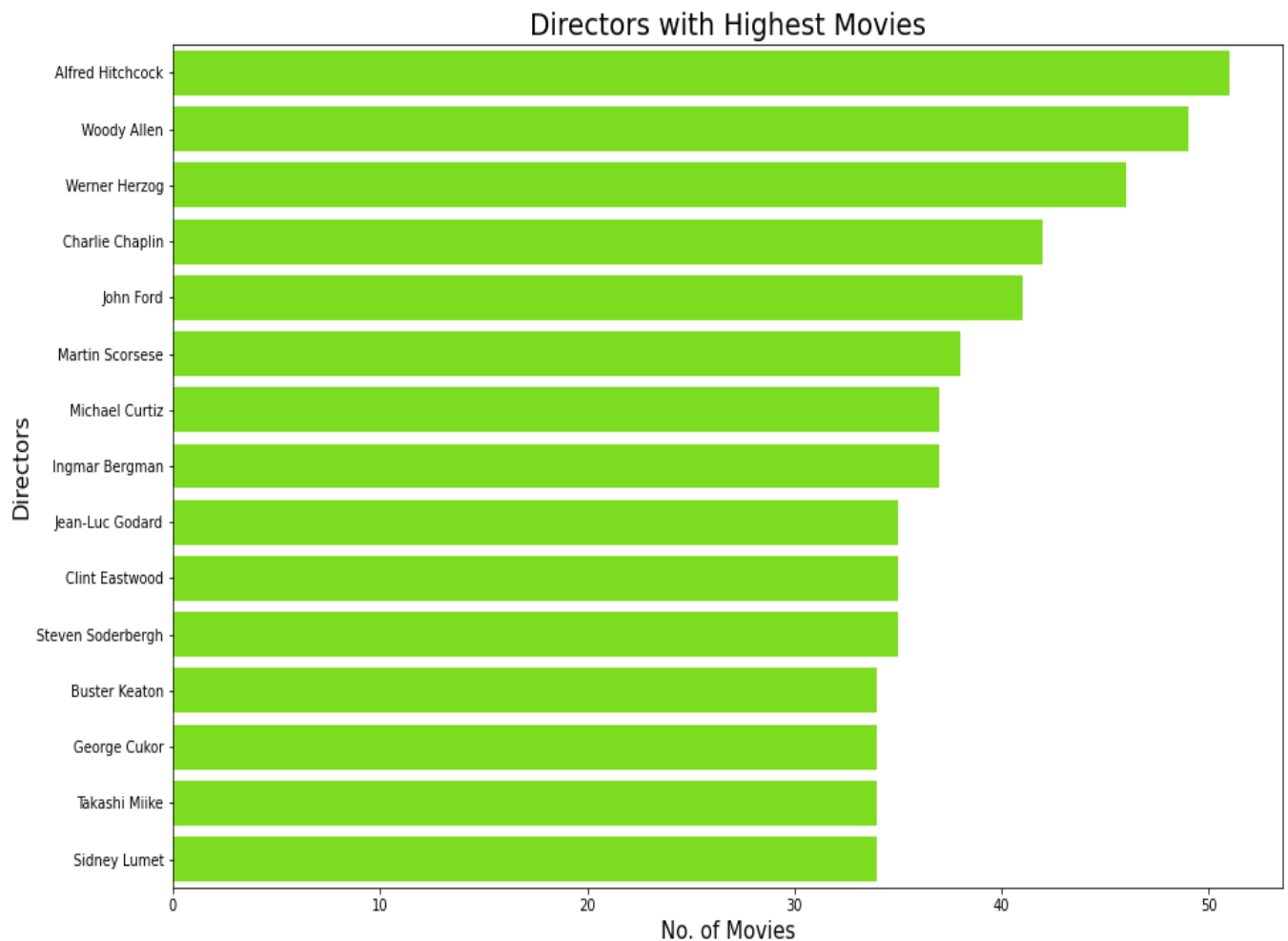
To see full sized image, Click Here

Fig-7: The graph shows "Top 15 Directors with Most Directed Films".

We can see that from the fig-7, **Alfred Hitchcock** was the Director with the Most with Most Directed Films i.e.,51 Movies. Followed by Woody Allen, Werner with 49 and 46 Movies respectively.

MACHINE LEARNING MODEL

We have created 2 machine learning models:

1. Movie Recommendation System using Cosine Similarity.

2. Movie Recommendation System using KNN Algorithm.

1. Movie Recommendation System using Cosine Similarity

Cosine similarity measures the similarity between two vectors of an inner product space.

To find Cosine Similarity :

1.1   First we have converted textual data into feature vectors using **TfidfVectorizer.fit_transform**. The "**TfidfVectorizer**" converts a collection of raw documents to a matrix of TF-IDF features. The "**fit_transform**" learn vocabulary and idf, return document-term matrix.

1.2.   Now, we find similarity between Feature vectors using Cosine Similarity.

1.3   We'll check distance between each movie using output from 1.2 and display the first K lowest difference to the user.

2. Movie Recommendation System using KNN Algorithm

The k-nearest neighbor algorithm is a type of "**Supervised Machine Learning Algorithm**" used to solve classification problems. This Algorithm doesn't perform any training when you supply the training data. Instead, it just stores the data during the training time and doesn't perform any calculations. It doesn't build a model until a query is performed on the dataset.

Steps to build Model using KNN Algorithm:

2.1   Using **Cosine similarity** we will check the euclidean distance between the 2 points(I.e., two Movies).

2.2   Choose the value of K and select K neighbors closet to the new point.

2.3   Count the votes of all the K neighbors / Predicting Values.

For More details **click here** and go to **Creating a "Machine Learning Model using KNN Algorithm"** Header.

1. <u>**Cosine Similarity**</u>:

We took four parameters for measuring distance between 2 points. They are:

    a) Genres

    b) Cast

    c) Director

    d) Keywords

2. In this step, we need to choose the value of K. In our case value of K is equal to 15.

3. After declaring K, we will iterate through the distances, and we'll pick the 15 nearest neighbors. These 15 nearest neighbors are our recommended movies.

CONCLUSION

The importance of a good recommendation system has become more apparent due to the increasing number of information overloads. This is why we are working on a collaborative-based system that aims to improve the accuracy of the movie's suggestions.

Movie Recommendation system helps to clear the problem of information overload with information retrieval systems. We come up with a strategy that focuses on dealing with the user's personal interests, and based on his previous search list, movies are recommended to users.This strategy helps in improving the accuracy of the recommendations.

REFERENCES

Dataset:

The dataset contains data about movies details including cast, crew etc. in form of **CSV** file of nearly 47000 values.

Dataset - <u>click here</u>

Resources:

◆ To Access Dataset - [Click Here](#)

◆ To Access Complete Model using KNN Algo - [Click here](#)

◆ To Access Complete Model using Cosine Similarity - [Click here](#)