One Stop Recommendation

Shloka Ramesh Daga[1], Bhavesh Bholanath Maurya[2], Chiragkumar Shalendra Maheto[3], [4]Kaajal Sharma

[1]*MCT Rajiv Gandhi Institute of Technology, University of Mumbai.*
[2] *MCT Rajiv Gandhi Institute of Technology, University of Mumbai.*
[3]*MCT Rajiv Gandhi Institute of Technology, University of Mumbai.*
[4]*MCT Rajiv Gandhi Institute of Technology, University of Mumbai.*

---***---

Abstract –

Currently, recommender systems are one of the most promising strategies for online businesses that specialize in services and goods related to the Internet. The business models of Google, YouTube, Facebook, LinkedIn, and Amazon are typical instances of how these recommender systems are fundamental to their operations. A category of information filtering systems are recommender systems. These systems are specialized software parts that are typically included in bigger software systems but may also be used independently.

A category of information filtering systems are recommender systems. These systems are specialized software parts that are typically included in bigger software systems but may also be used independently. The primary objective of a recommender system is to offer the user software ideas for things that might be useful. The advice relates to many approaches for making decisions, such as which product to buy, which movie to watch, or which vacation to book. The initiative primarily aids in suggesting titles of films and television shows from various OTT platforms including Netflix, Amazon Prime Video, and Hotstar. The project essentially unifies many recommendation systems onto a single platform.

***Key Words*:**

Recommendation Model, OTT Platform, Netflix, Amazon Prime Video, Hotstar, Streaming Platform, Dashboard .

1.INTRODUCTION

The filtering procedure becomes essential when there are a lot of users on a system and a lot of stuff to provide for them. Nobody can reasonably expect a user to manually shift through tens of thousands or even millions of different items—whether they are movies, goods, or news—in order to discover what he is looking for. Without recommendations, consumers would only be exposed to direct search results, which in cases when there are a lot of items, would only return a few tens or even hundreds of items if the user looked through several pages. Even on smaller news or e-commerce websites with well-organized product categories, there may be too many items for a user to sort through in order to discover what they are looking for. In terms of design, recommender systems typically only focus on a single sort of item, such as movies or music, and both its primary suggestion mechanism and graphical user interface are specialized to that particular type of item. Because recommendations are typically created by taking into account the distinctive qualities of the users, different people or groups of users receive different suggestions. It is simpler to make non-personalized recommendations, which are primarily found in publications or newspapers.

Users may find some particular goods that a system has to offer intriguing, but if the system offers too many items, they may never learn about them. The recommender's objective is to present the user with a fresh selection of options that they might not have discovered on their own.

1. Content Based Filtering

Content-based filtering's fundamental tenet is that every item has certain characteristics. A model or user interest profile of the user's interests is created by recommender systems using a content-based recommendation strategy to analyse a set of documents and/or descriptions of products that the user has previously evaluated. Users have a set of preferences connected to the contents of the items. A profile might be specifically constructed by the user or built automatically based on his or her prior behaviours. The user's interests are organized into a profile, which is then utilized to suggest new, intriguing stuff. Comparing the features of the user profile with the attributes of the item is how the suggestion process works. The result is a relevance score, which reflects the user's level of interest in the particular item. A user profile's ability to accurately predict a user's interests is extremely beneficial to the efficiency of an information retrieval operation.

2. Cosine Similarity Based Recommendation

Here, we use the Cosine Similarity function to calculate the cosine similarity. Regardless of size, the cosine similarity metric can be used to assess how similar two papers are. It mathematically calculates the cosine of the angle formed by two vectors cast in a multidimensional space. Even if two comparable documents are separated by the Distance measure (taking into account the size of such documents), they are likely to be orientated closer together because of the cosine similarity. The angle is narrower the higher the cosine similarity. Applications for text matching, information retrieval, and data mining all make use of the measure. A typical approach in information retrieval is to use weighted TF-IDF and cosine similarity to quickly locate documents that are similar to a search query.

TF-IDF Vectorizer Text vectorization is the process of converting text into a quantitative feature. It compares a phrase's "relative frequency" in a document to the consistency of that term across all papers. The TF-IDF weight shows a phrase's relative importance in the document and throughout the corpus. Phrase Frequency (TF) is a measure that displays how frequently a phrase appears in a document. Due to document size disparities, a term may appear more frequently in a large document than in a short one. As a result, the document's length is usually separated by term frequency. TF-IDF is among the most extensively used text vectorizers, and the computation is straightforward. It distinguishes between the uncommon word heavier weight and the more frequent term reduced weight.

2.RELATED WORKS

The paper presents the development and the comparison of multiple recommendation systems[1], capable of making item suggestions, based on user, item, and user-item interaction data, using different machine learning algorithms[4]. Also, the paper deals with finding different ways of using machine learning models to create recommendation systems, training, evaluating, and comparing the different methods to provide a general but accurate solution for ranking prediction.

In this paper, they propose a deep learning approach based on autoencoders to produce a collaborative filtering system that predicts movie ratings for a user based on a large database of ratings from other users, using the MovieLens dataset. We explore the use of deep learning to predict users' ratings on new movies, is difficult to patch through and requires high end systems to perform the tasks[3].

3. DATASET

The Netflix, Prime Video, and Hotstar movie data were acquired via Kaggle, each of which contained more than 5000 entries, while the data for the analysis was extracted using Google forms and the majority of the analysis was performed on the data that was collected using our own Google form. The main goal of using a google form to gather information is to learn people's preferences for streaming services, which will help us find the answer that our project is trying to provide. Additionally, the public dashboard for the project includes a link to the google form so that anyone using the dashboard can voluntarily fill out the form and contribute to attaining a better accuracy in the results provided.

Netflix Dataset

The Netflix Dataset comprises of more than 9000 data entries.

Amazon Prime Dataset

The Amazon Prime Video dataset consists of more than 9000 data entries.

Hotstar Dataset

The hotstar dataset consists of more than 6000 unique data.

Google Form Dataset

The Google form dataset is a live dataset which is included in the dashboard, so whenever new entries are included in the google sheet, it gets automatically updated on the dashboard.

4.PROPOSED SYSTEM

The current recommendation system suggests only movies, but in this project, we would suggest movies and television series for several OTT platforms including Netflix, Prime Video, and Hotstar. We would be creating a dashboard with many screens, each of which would feature a different OTT platform. We would gather data using Google Forms as well as different websites, such as Kaggle. In essence, it serves as a single platform to obtain movie suggestions for various OTT services. With its highly dynamic graphs and charts that can be hovered as needed, the project's suggested solution would also assist users in deriving practical insights from the data.

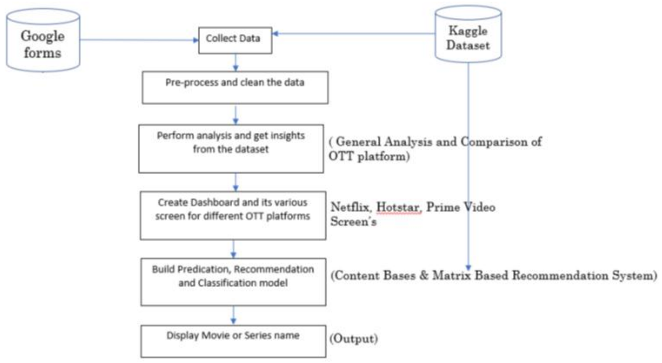

Chart-1: Workflow Diagram

We will obtain the project's dataset from Kaggle, and we'll also use Google Forms to gather user information. The Kaggle dataset will be utilized for content-based and matrix-based recommendations, while the data from Google Forms will be used for collaborative recommendation systems.

Data pre-processing will be carried out after data collection. Pre-processing would be followed by OTT platform-specific analysis on each screen in order to glean some actionable insights from the interactive charts. For each OTT platform, such as Netflix, Prime Video, and Hotstar, there would be a separate screen on the dashboard.

The customer will have the choice to receive several types of recommendations for each OTT platform depending on their needs. The models will calculate the input when the user offers it, and depending on the results, they will recommend movies as input.

A variety of recommendation models are included in the project. 1. Content-Depending Recommendation: The user will receive recommendations for a list of movies based on their favourite genre. Using the genre name as input, the model will output the top films according to their IMDb scores. 2. Matrix-Based Recommendations: In this method, the model compares the user-provided movie description to the descriptions of other films to determine which five films most closely fit the user's description. 3. Collaborative Recommendation: In this case, the model will suggest or recommend a movie or TV show by giving its title based on the preferences of other individuals who prefer a similar genre.

5.ANALYSIS

The dashboard Analysis Page's information was gathered through Google Forms, and it includes graphs that demonstrate how the number of OTT subscribers increased significantly during the COVID19 outbreak. The analysis unmistakably demonstrates that COVID19 revolutionized the use of streaming platforms, and that this use has grown to the point where more people choose to see new movies on streaming platforms than in actual cinemas. Even with its higher prices, Netflix continues to be the most popular streaming platform. Those who do not like Netflix may not like it because it has more expensive subscription costs than other streaming services. The dashboard can also be helpful for a director who wants to choose what kind of film to make based on the audience's age or gender, as the projects help to deal with this issue as well.

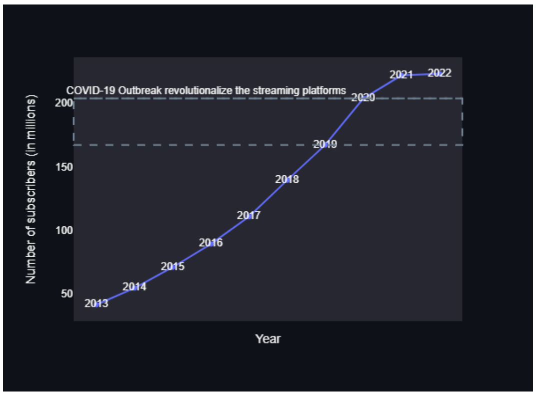

Chart -2: Year wise Subscription Chart

This finding is corroborated by another graph that demonstrates how people now prefer to watch newly released films on OTT platforms rather than going to theatres. The project's dashboard also displays the most popular streaming services, including Netflix, Amazon Prime Video, and Hotstar. Other significant questions are included in the google form and are used in the dashboard to help the audience understand the information and insights. The primary goal of utilizing Google Forms to collect data is to enable more accurate results and conclusions by taking into account the preferences that individuals currently hold.

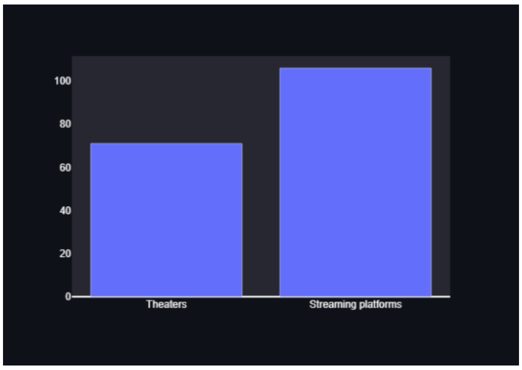

Chart -1: Comparison Chart

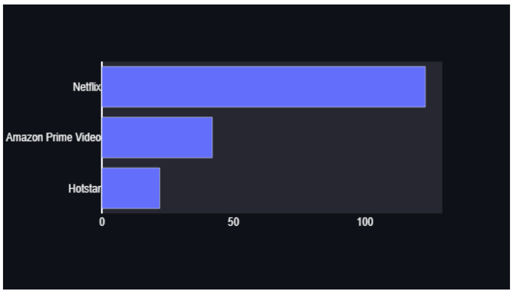

Chart-2: OTT preference chart

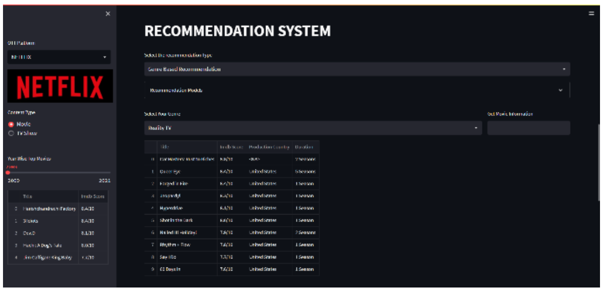

Chart-3: User's Interface

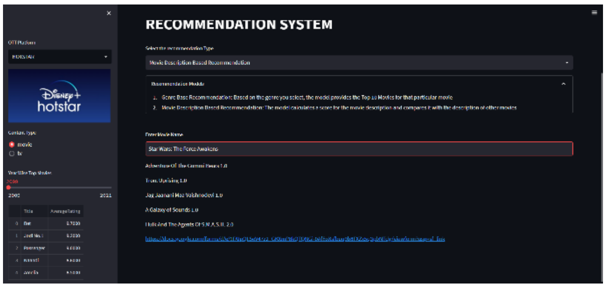

Chart-4: User's Interface

6. CONCLUSIONS

New possibilities for finding personalized information on the Internet are made possible by recommender systems. It also enables users to access goods and services that are not immediately available to users on the system, which helps to relieve the issue of information overload, which is a fairly typical situation with information retrieval systems. We have developed a recommendation algorithm for Netlfix, Prime Video, and Hotstar that can suggest films and television shows based on a user's affinity for a particular genre. The models also make recommendations for films that are comparable to the movie description that users enter in the input area. Additionally, our dashboard offers a thorough picture of how Covid19 contributed to the era of streaming platforms, as well as information on which OTT platform users prefer, how much time they spend there, and their preferred genres.

REFERENCES

- G.C. Capelleveen, C. Amrit, D.M. Yazan, W.H.M. Zijm, "The recommender canvas: A model for developing and documenting recommender system design". Expert systems with applications, pp. 97-117, 2019.
- F. Ricci, L. Rokach, B. Shapira, Introduction to Recommender Systems Handbook. Boston, Massachusetts, United States of America: Springer, 2010.
- Y. Lim, "A Primer to Recommendation Engines", Sep 10, 2019
- By Daniil Korbut, Statsbot, "Recommendation System Algorithms: An Overview", Jan. 2022.
- Richmond Alake, "Understanding Cosine Similarity And Its Application: Understand the basics behind a technology that is used across different fields and domains of Machine Learning" Sep. 15, 2020
- Mukesh Chaudhary, "TF-IDF Vectorizer Scikit-learn: Deep understanding of TF-IDF Calculation by Various examples, why is so Efficiency than other Vectorizer Algorithm, Apr. 24, 2020.
- Plotly, "Introducing Plotly Express", Marc. 20, 2019.
- Andy McDonald, "Getting Started with Streamlit Web Based Application: A gentle Introduction To Creating Streamlit Web apps", May 23, 2022.