# **NETFLIX MOVIES AND TV MOVIES AND TV SHOWS Recommendation System**

**ABSTRACT**

The project will be on predicting Netflix Movies and TV Shows based on personal history viewing using recommendation system. In order to help understand people preference and how Netflix shape similar catalog of movies and TV shows using genres, title, and release year as features. Netflix Main business is its subscription-based streaming service, which provide online streaming of a collection of films and television series, including those produced in-house. The project is using TV Shows and Movies listed on the Netflix dataset from Kaggle. First, Exploratory Data Analysis (EDA) will be presented to understand and describe the data in a better way through interactive graphs and visualizations. After that, a recommendation engine for Movies and TV shows on Netflix is demonstrated.

**DESIGN**

Netflix use predictions to make personal movie and tv shows recommendations based on each customer’s unique tastes.The project is using TV Shows and Movies listed on the Netflix dataset from Kaggle. Visualization would enable a better understanding of the most content type on Netflix movies or TV shows and the variety of Content by Rating. Also, the highest amount of content produced by country. Another aspect is to identify the most popular director on Netflix with the most titles. On another hand, a Content-Based Recommendation model, is proposed to find Movies and TV Shows users might like based in history.

**DATA**

The dataset contains contains 12 columns for exploratory analysis with 8807 records

FEATURES:

|  |  |
| --- | --- |
| **Indicator** | **Description** |
| show\_id | unique id of each show |
| type | The category of a show, can be either a Movie or a TV Show |
| title | Name of the show |
| director | Name of the director(s) of the show |
| cast | Name of actors and other cast of the show |
| country | Name of countries the show is available to watch on Netflix |
| date\_added | Date when the show was added on Netflix |
| release\_year | Release year of the show |
| rating | Show rating on netflix |
| duration | Time duration of the show |
| listed\_in | Genre of the show |
| description | Some text describing the show |

**ALGORITHMS**

***CLEANING THE DATASET***

* I used dropna() to drop any row with null value.There are a total of 4,307 null values across the entire dataset with 2,634 missing points under "director", 825 under "cast", 831 under "country", 10 under "date\_added", and 4 under "rating".
* Imputation is used as a treatment method for missing value in this project by filling it In this module,I used the fillna function from Pandas for this imputation. Drop rows containing missing values. Can use the dropna function from Pandas
* For missing values, since "director", "cast", and "country" contain the majority of null values, we chose to treat each missing value is unavailable. The other two label "date\_added" and "rating" contain an insignificant portion of the data so it drops from the dataset. Finally, there are no more missing values in the data frame.

***MODELS***:

* For this recommender system I used Content-based Filtering technique to find the similarity between content that the user has already consumed and the recommended.
* I appled the The nearest neighbour algorithm model to calculate the “distance” between the target tv show and every other tv show in the database (nearest 5), then it ranks its distances and returns the top K nearest neighbor tv show as the most similar tv show recommendations.
* I normalize the data by scaling each feature to a given range between zero and one min-max scaling.

**TOOLS**

* Numpy and Pandas for data manipulation
* Scikit-learn for modeling
* Matplotlib and Seaborn for plotting
* Tableau for interactive visualizations
* Matplotlib
* Seaborn

**Communication**

To get better result more algorithm models will be applied in the second stage like

The collaborative filtering algorithm where it will uses “User Behavior” for recommending.

**REFERENCES**

Kaggle.com. 2021. *Netflix Movies and TV Shows*. [online] Available at: <https://www.kaggle.com/shivamb/netflix-shows> [Accessed 25 November 2021].