

Netflix Contents “Movies and TV Shows”

Project Abstract

Netflix is one of the most popular media and video streaming platforms. They have over 8000 movies or tv shows available on their platform, as of mid-2021, they have over 200M Subscribers globally. This tabular dataset consists of listings of all the movies and tv shows available on Netflix, along with details such as - cast, directors, ratings, release year, duration, etc.

The dataset is provided in .csv format. It contains 8807 TV shows and movies, each of one has 12 features.

This project aimed to understanding what content is available in different countries, Identifying similar content by matching text-based features and Does Netflix has more focus on TV Shows than movies in recent years

Tools

- Pandas for data manipulation
- Scikit-learn for modeling
- re for clean data
- nltk for natural language processing
- Matplotlib for plotting
- streamlit for interactive visualizations

TO DO:

- Explore the data and come up with EDA phases then use a model to fit the data.
- **NOTE:** the used features may be increased or changed and the model as well.