Recommendation system based on Netflix dataset

Rana hame khird

Abstract

The goal of this project was to build recommendation system content base using Netflix dataset from Kaggle Netflix Movies and TV Shows About this Dataset: Netflix is one of the most popular media and video streaming platforms., it had many 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.

I use TfidfVectorizer to count the similarity between the to description of the movie and give the user the heist similar movie

Design

Data

Netflix Movies and TV Shows About this Dataset: Netflix is one of the most popular media and video streaming platforms., it had many 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.

With 8807 row there is missing value I cleaned using filling

Algorithms

Counting the similar word in description

use supervised classification algorithms on a multi-label dataset

Naive bayes model

Tools

- Numpy and Pandas for data manipulation
- Scikit-learn for modeling
- Matplotlib and Seaborn for plotting

Communication

```
get_recommendations('Blood & Water')
```

```
5344    Message from the King
1884      Walk Away from Love
4285                    Lilli
4271              Lion Pride
4209              Next Enti?
613                Voiceless
108                Dive Club
1905            Cold Harbour
6289                 Bewafaa
5485               Ram Jaane
Name: title, dtype: object
```

```
[179]
get_recommendations('Mortel')
```

```
256                       Nneka The Pretty Serpent
3674                 PILI Fantasy: War of Dragons
800                                 Mosquita y Mari
6749                             Figures of Speech
4511        Edgar Rice Burroughs' Tarzan and Jane
7535    My Entire High School Sinking Into the Sea
6760                                     FirstBorn
4918                                  Psychokinesis
5481                                      Chamatkar
2190                          The Umbrella Academy
Name: title, dtype: object
```

```
[261] from sklearn.metrics import accuracy_score
      score1 = accuracy_score(y_test, nb_y_pred)
      print("---- Score ----")
      print("Accuracy score is: {}%".format(round(score1*100,2)))

      ---- Score ----
      Accuracy score is: 65.7%
```