MOVIE RECOMMENDER SYSTEM

Certification Project

edureka!

edureka!

Domain – OTT Platform

Overview
Congratulations!!!! for making it so far. This is the Certification Project for Data Science and Machine Learning Internship Program and here we will use all the concepts learned so far in this course.

Context
Over the past two decades, there has been a monumental shift in how people access and consume video content. With the universal access to broadband internet, numerous platforms like YouTube, Netflix, and HBO Go emerged and steadily grew to prominence.

Although not a household name in itself, OTT is the exact technology that made the streaming revolution possible.

OTT stands for "Over The Top" which refers to any video streaming service delivering content to the users over the internet, however, there are subscription charges associated with the usage of such platforms such as PrimeVideo, Netflix, HotStart, Zee5, SonyLiv, etc.

But choosing your next movie to watch can still be a daunting task, even if you have access to all the platforms.

Business Requirement:
"MyNextMovie" is a budding startup in the space of recommendations on top of various OTT platforms providing suggestions to its customer base regarding their next movie.

Their major business is to create a recommendation layer on top of these OTT platforms so that they can make suitable recommendations to their customers, however, since they are in research mode right now, they would want to experiment with open-source data first to understand the depth of the models which can be delivered by them.

The data for this exercise is open-source data that has been collected and made available from the MovieLens website (http://movielens.org), a part of GroupLens

Research The data sets were collected over various periods of time, depending on the size of the set.

You have recently joined as a Data Scientist at "MyNextMovie" and plan to help the existing team to set up a recommendation platform.

https://www.kaggle.com/code/ayushimishra2809/movie-recommendation-system/data ?select=ratings.csv

Objective:

1. Create a popularity-based recommender system at a genre level. The user will input a genre (g), minimum rating threshold (t) for a movie, and no. of recommendations(N) for which it should be recommended top N movies which are most popular within that genre (g) ordered by ratings in descending order where each movie has at least (t) reviews.

   Example:

   Input:
   - Genre (g) : Comedy
   - Minimum reviews threshold (t) : 100
   - Num recommendations (N) : 5

   Output:

S.No	Movie Title	Average Movie Rating	Num Reviews
1	A	4.5	422
2	B	4.4	495
3	C	4.3	342
4	D	4.22	531
5	E	4.21	454

2. Create a content-based recommender system that recommends top N movies based on similar movie(m) genres.

   Example:

   Input:
   - Movie Title (t): Toy Story
   - Num recommendations (N): 5

   Output:

Sl.No	Movie Title
1	A
2	B
3	C
4	D
5	E

3. Create a collaborative based recommender system which recommends top N movies based on "K" similar users for a target user "u"
   Example:
   Input:
   - UserID: 1
   - Num recommendations(N): 5
   - Threshold for similar users (k: 100

   Output:

S.No	Movie Title
1	A
2	B
3	C
4	D
5	E

Data Description

The data consists of 105339 ratings applied over 10329 movies. The average rating and minimum and maximum rating are 0.5 and 5 respectively. There are 668 users who have given their ratings for 149532 movies.

There are two data files which are provided:

Movies.csv
- movieId: ID assigned to a movie
- title: Title of a movie
- genres: pipe-separated list of movie genres.

Ratings.csv
- userId: ID assigned to a user
- movieId: ID assigned to a movie
- rating: rating by a user to a movie
- Timestamp: time at which the rating was provided.

Steps and Tasks

- Import libraries and load dataset

- Exploratory Data Analysis including:

    o   Understanding of distribution of the features available

    o   Finding unique users and movies

    o   Average rating and Total movies at genre level.

    o   Unique genres considered..

- Design the 3 different types of recommendation modules as mentioned in the objectives.

- Additional/Optional: Create a GUI interface using Python libraries (ipywidgets etc.) to play around with the recommendation modules.

edureka!