

Movie Recommender System

Author: DEBANJAN DUTTA

Introduction:

Recommender systems simplify decision-making processes and enhance user experiences. They have become integral to diverse industries, including e-commerce and streaming services. Our Movie Recommender System aims to personalize the movie discovery process by understanding individual tastes and predicting preferences.

Dataset:

We used the `tmdb_5000_movie_dataset`, which contains details of around 5000 movies. The dataset includes information such as movie titles, genres, keywords, cast, and crew.

Content-Based Recommender Systems:

Our system is based on a content-based approach, which analyzes the content or features associated with items and compares them to a user's profile. It creates a user profile based on historical interactions and computes similarity scores between the user's profile and each item's features.

Data Preprocessing:

We removed unnecessary columns and handled missing values. We extracted relevant information from columns like genres, keywords, cast, and crew. We also performed text vectorization to convert the movie descriptions into numerical vectors.

Building the Recommendation System:

We calculated similarity scores between pairs of movies using cosine similarity. The similarity scores were stored in a matrix, and recommendations were made based on the highest similarity scores.

Examples:

We tested our system with input movies and obtained relevant recommendations. For example, inputting "Pirates of the Caribbean: At World's End" resulted in recommendations related to pirates or the sea. Inputting "Harry Potter and the Chamber of Secrets" led to recommendations of other Harry Potter movies or movies related to magical fantasy.

Conclusion:

Our content-based movie recommender system combines data-driven insights with the joy of storytelling. It provides personalized recommendations based on user profiles and movie features. While content-based approaches have advantages such as handling the "cold start" problem and offering transparency, they may lack serendipity and over-specialize. Balancing these strengths and weaknesses is crucial for an optimal movie recommendation experience.

Future Work:

We plan to further improve our system by incorporating collaborative filtering techniques and exploring hybrid approaches. Additionally, we aim to develop a user-friendly application using the 'streamlit' library to make the recommender system accessible to users.