Abstract

Personalized recommendation systems are essential for improving user experience in the current digital content era by providing customized recommendations. In order to increase the recommendations' relevance, this study proposes a sentiment analysis-enhanced content-based movie recommendation system. The main goal is to create a system that recommends films based on content similarity, including cast, genre, and storyline keywords, while making sure that the suggestions support favorable audience sentiment.

The suggested method refines movie descriptions using TF-IDF (Term frequency-inverse document frequency) transformation after extracting features from them using CountVectorizer. Cosine similarity is used to quantify how closely related films are based on their content. The system uses a sentiment analyzer tool to calculate sentiment scores after dynamically retrieving user reviews over the application programming interface in order to account for user sentiment. To guarantee that only well-received films are recommended, only those with an average sentiment score above a predetermined level are added to the list. By removing films with similar substance but negative reviews, this dual-filtering method increases the precision and applicability of movie recommendations.

Key libraries like scikit-learn for text vectorization, nltk (Natural language tool kit) for sentiment analysis, and Streamlit for the interactive user interface are used in the Python implementation of the system. This study shows how sentiment analysis and content-based filtering increase user experience and suggestion quality while offering a flexible and scalable framework for further advancements.

Keywords: Content-based, TF-IDF, CountVectorizer, VARDER, nltk, Streamlit.