Final Project Proposal

Group Members:

Daniyar Abdrakhmanov, Doszhan Konysuly, Sunkar Kurmet

Problem or Idea Description:

Our project focuses on utilizing Machine Learning (ML) and Deep Learning (DL) techniques to tackle the challenge of recommending movies with unique characters and storylines, rather than just suggesting films from the same franchise. Traditional recommendation systems often prioritize recommending sequels or movies from the same series, which may limit users' exposure to diverse content.

Therefore, our goal is to develop an advanced recommendation system that suggests movies with entirely different characters and plotlines but share similarities in themes, character archetypes, and narrative structures. This approach aims to enrich users' movie-watching experiences by introducing them to a broader range of cinematic storytelling and genres, ultimately enhancing user engagement and satisfaction with the recommendation platform.

Background Information on the Problem or Idea:

Movie recommendation systems play a crucial role in helping users explore diverse content aligned with their interests and preferences. However, conventional recommendation approaches primarily focus on suggesting movies within the same franchise or genre, which may limit the scope of exploration for users seeking novel narratives and characters.

These traditional systems typically employ collaborative filtering or content-based methods, which analyze user behavior or item attributes to generate recommendations. While effective to some extent, they may overlook the intricacies embedded within the textual descriptions of movies, failing to capture the essence of unique storylines and characters.

As our project aims to recommend movies with distinct protagonists and plots, it addresses a gap in existing recommendation systems by leveraging advanced machine learning and deep learning techniques. By delving into the semantic richness of movie overviews and keywords, our system seeks to offer users a more expansive and enriching cinematic experience beyond franchise-centric suggestions.

Available Solutions with Links:

- Collaborative Filtering: https://developers.google.com/machine-learning/recommendation/collaborative/basics?hl=ru

- Content-Based Filtering: https://www.turing.com/kb/content-based-filtering-in-recommender-systems

- Hybrid Recommendation Systems: https://analyticsindiamag.com/a-guide-to-building-hybrid-recommendation-systems-for-beginners/

How to Get the Data?

The dataset for this project will be obtained from a publicly available source or an API that provides information about movies, including their titles, overviews, and keywords. We will explore platforms like TMDB or IMDb for relevant datasets.

Brief Description of Our Solution:

Our solution involves preprocessing movie overviews and keywords, vectorizing them using TF-IDF (Term Frequency-Inverse Document Frequency) vectors, and combining them into a unified feature space. We will then employ BERT (Bidirectional Encoder Representations from Transformers), a state-of-the-art deep learning model, to extract contextualized embeddings from the textual data. Finally, we will compute cosine similarity scores between movies to generate recommendations.

Tech Stack That Will Be Used:

- Python: For programming and data manipulation

- Scikit-learn: For TF-IDF vectorization and cosine similarity computation

- Transformers library: For utilizing BERT model

- Pandas: For data manipulation and analysis

- PyTorch: For implementing deep learning models

- Google Colab: For collaborative coding and running experiments

By implementing this recommendation system, we aim to enhance the movie-watching experience for users by providing personalized and relevant movie suggestions based on their preferences and interests.