Project: Context Based Movies

My project is live on Heroku and can be accessed using the below url:
https://context-based-movies-498141f87819.herokuapp.com/

Abstract: Movie posters play a crucial role in movie identification, serving as attention-grabbing visuals that often form the audience's initial connection with a film. These posters visually encapsulate the movie's theme, characters, and style. In our project, our primary objective was to integrate the poster's context into a movie recommendation system. We developed a system that suggests movies based on poster features, including color and composition, as well as extracting textual information describing the image. These recommendations are tailored to user preferences and the context of their current movie preferences.

To achieve this, our application incorporates three main features. Firstly, it provides an initial recommendation based on the poster selected by users. Additionally, users can employ a text search, enabling them to input queries related to the poster or movie, resulting in a list of 10 suggested movies matching their interests. Furthermore, an image search feature allows users to upload an image, with the system returning a list of 10 movies sharing contextual similarities with the input image. In our project, we utilized cosine similarity for calculating similarity, employed BM25 along with DESM for ranking, and implemented a neural network model to extract context from images.

Introduction: My goal is to offer users the finest movie suggestions based on their preferences by leveraging the potent combination of BM25 and Cosine Similarity algorithms. Through this approach, I aim to deliver search results that are finely tuned to individual tastes, ensuring an optimal and enjoyable movie-watching experience. In addition to textual preferences, I extend our recommendation capabilities to image search. By extracting context from images and incorporating it into our model, I strive to recommend movies that closely align with the visual elements and themes captured in the provided images. This innovative approach ensures that our suggestions resonate not only with users' stated preferences but also with the nuanced aspects that images can convey.

Dataset: While the movie recommendation system will be built in a traditional sense while leveraging user ratings the search system will be built based on the context of the poster based on a search query. We will be using existing methodologies such as user-user CF, item-item CF, BPR, MF, and content-based CF to generate movie recommendations. The dataset I am using is this: https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset

"The Movies Dataset" is a comprehensive dataset related to movies. It includes a variety of information such as movie metadata, user ratings, credits, and keywords. Here are some key features of the dataset:

- movies_metadata.csv: Contains information about movies, including details like title, release date, budget, revenue, etc.
- keywords.csv: Provides keywords associated with each movie.
- credits.csv: Contains information about the cast and crew associated with each movie.
- links.csv: Includes information for linking movies to external databases.
- ratings.csv: Consists of user ratings for the movies.

Proposed solution / methods used: My proposed approach for the movie recommendation system involves employing a variety of techniques and models, including:

1. BM25 (Term Frequency-Inverse Document Frequency): I have integrated the BM25 ranking algorithm, a variant of TF-IDF, to assess the relevance of textual information extracted from movie posters. This algorithm facilitates the effective ranking of movies based on user preferences.

2. DESM (Deep Embedding Semantic Model): The DESM technique was implemented to enhance the recommendation system. By utilizing deep neural networks, DESM learns semantic representations of textual content, improving the system's understanding of the movie's context and relationships between different films.

3. TF-IDF (Term Frequency-Inverse Document Frequency): TF-IDF was utilized to analyze and extract meaningful features from textual information associated with movie posters. This method quantifies the importance of each word in the document, aiding in understanding the context and relevance of movies.

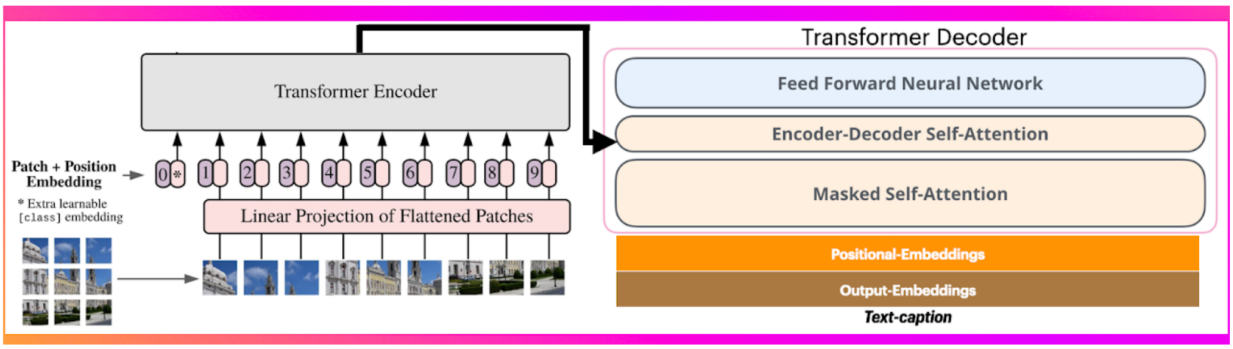
4. Word2Vec: The Word2Vec model, a widely used word embedding technique, was employed to transform textual data into numerical representations. This transformation enhances semantic understanding and facilitates similarity analysis between movie descriptions, contributing to the recommendation process.

5. ViT-GPT2-Image-Captioning: I have utilized Visual Transformer (ViT) models in combination with GPT-2 and image captioning techniques to extract contextual information from movie posters. This approach captures visual cues, providing a deeper understanding of visual content and improving the accuracy of recommendations.

6. CLIPImageProcessor: The CLIPImageProcessor, based on Contrastive Language-Image Pretraining (CLIP), was employed to generate visual embeddings from movie posters. These embeddings capture visual information, enabling similarity comparisons between different movie posters.

In the recommendation process, these techniques and models were leveraged to calculate cosine similarity between textual and visual embeddings. This approach contributes to generating personalized movie recommendations based on user preferences and contextual similarities between movies.

By integrating these diverse methods, our movie recommendation system effectively analyzes both textual and visual content, delivering accurate and context-aware recommendations to users.



Evaluation: The application underwent testing with various examples, and below are screenshots showcasing the testing of three distinct functionalities. You can also explore custom queries at https://context-based-movies-498141f87819.herokuapp.com/

Initial Recommendations: Upon the user's system login, they will be prompted to choose three movies from a movie collection. Subsequently, the system will provide recommendations for top movies, determined internally through the utilization of both BM25 and Cosine Similarity, based on the user's selected films.

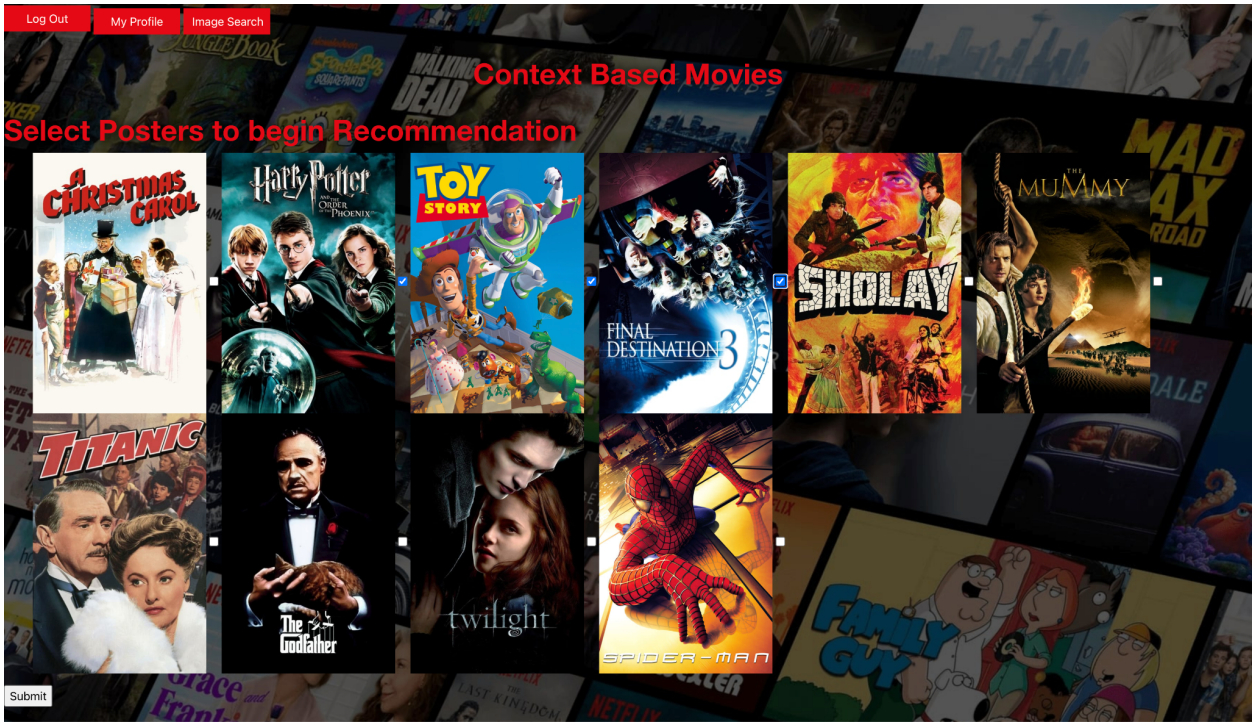


Fig: Home Page

As evident, I have the option to pick three films - “Harry Potter”, “Toy Story”, and “Final Destination 3”, and subsequently click on "Submit." The recommendations for the chosen movies are provided below:

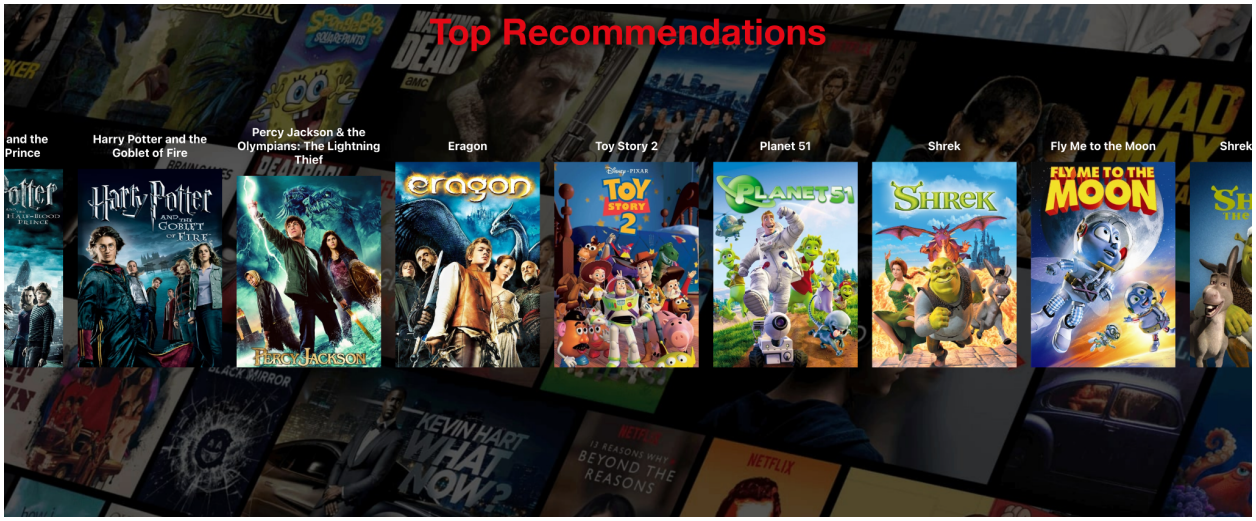


Fig: Top Recommendations.

Text Search and Image Search: Examining the provided example, it is evident that both image search and text queries yield some shared recommendations. In the image search, where an

image of an animal is uploaded, 10 distinct recommendations are generated. Similarly, in the text search, inputting "bride in wedding," which aligns with the image, results in another set of 10 recommendations. Notably, 6 recommendations overlap between these two functionalities. This observation underscores the application's capability to comprehend the context behind both images and textual queries effectively.

Text Search:

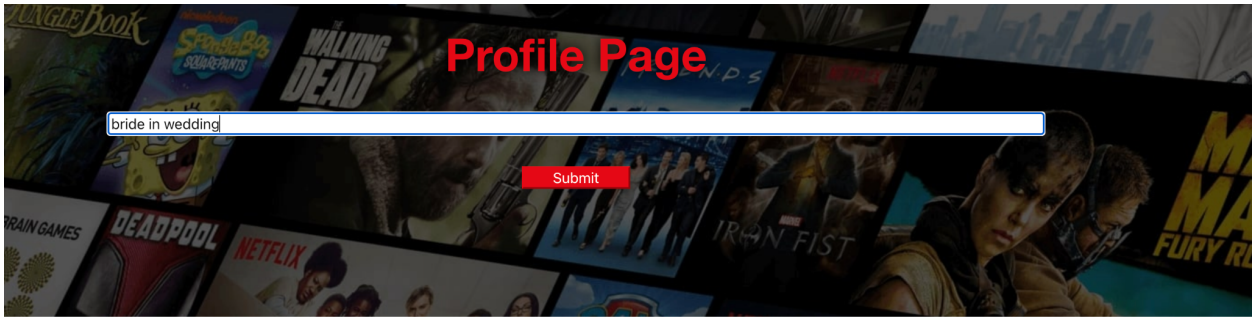


Fig: Text Search

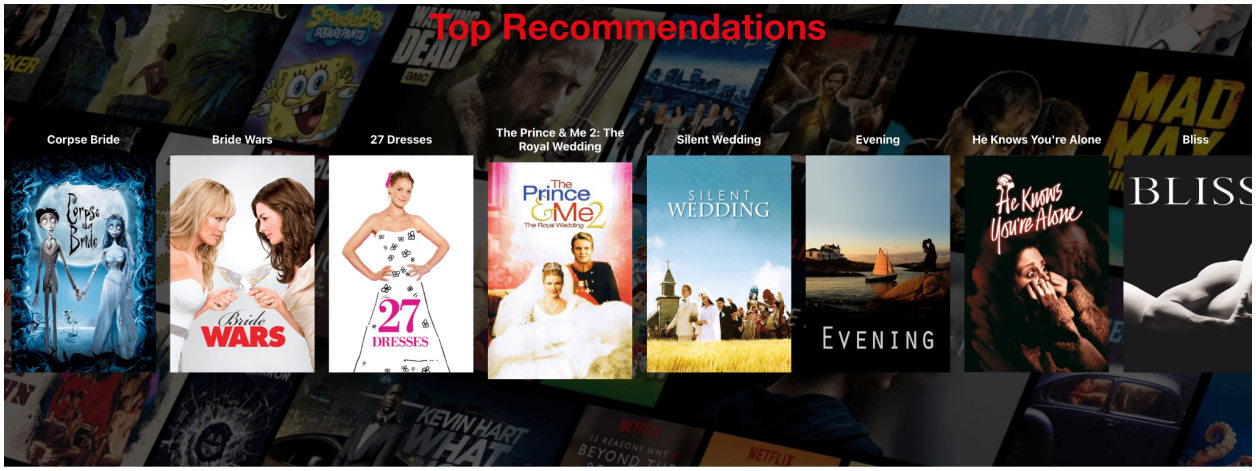


Fig: Text Search Output

Image Search:

Fig: Input Image

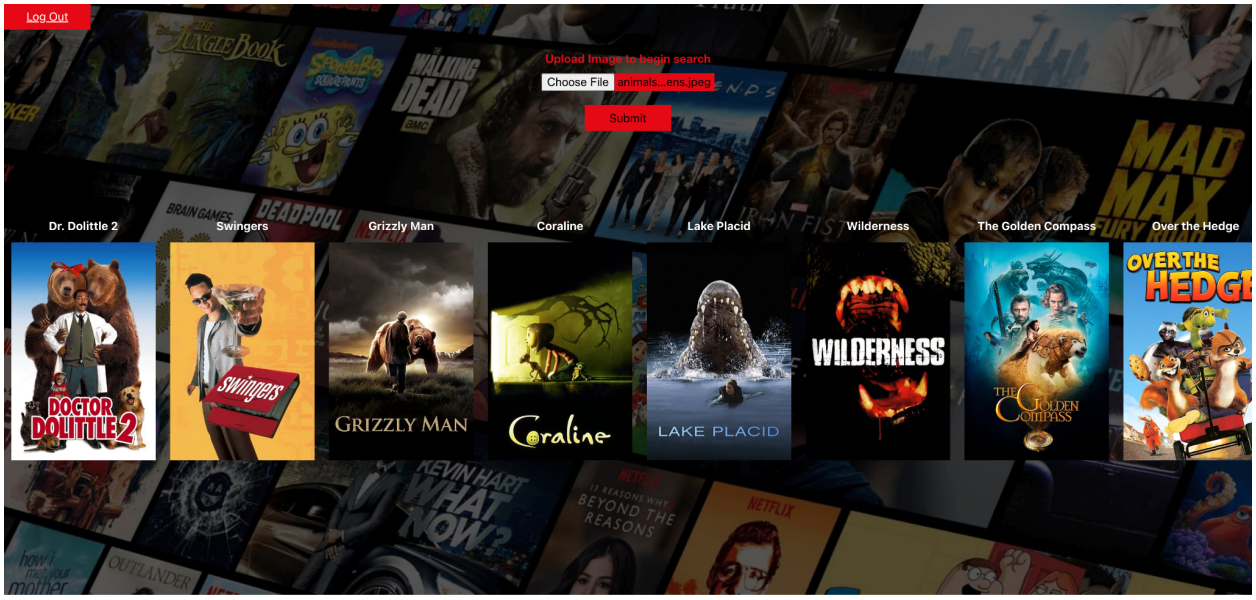


Fig: Recommendations for the above image using image search

Conclusions: This system has demonstrated a high degree of success in delivering movie results through both image and text queries, along with furnishing personalized recommendations tailored to user preferences. This underscores the effectiveness of the utilized methods, including BM25, DESM, TF-IDF, Word2Vec, ViT-GPT2-Image-Captioning, and CLIPImageProcessor.

Furthermore, I acknowledge the considerable potential for advancing and expanding this project. For instance, broadening the model's scope to analyze movie clips sourced from various platforms, including social media applications, could empower the system to curate a list of potential movies associated with those clips.

By capitalizing on the strengths of the implemented methods and exploring additional avenues for application, we anticipate ongoing enhancements and refinements to the movie recommendation system. This will not only enrich the user experience but also provide valuable movie recommendations based on a diverse range of inputs.

Future enhancements:

1. Improving the overall user interface for a more seamless and intuitive experience.
2. Currently, the weightage for BM25 and Cosine Similarity is evenly distributed (50-50). In future iterations, considering implicit feedback from users and dynamically adjusting this ratio can be explored for more personalized recommendations.
3. The current Image Search process experiences delays in loading results for a given input. To address this, exploring the option of hosting the model locally instead of relying on the Hugging Face URL is suggested. This adjustment can significantly reduce network calls and enhance the speed of the image search.
4. Improving the accuracy of movie recommendations can be achieved by incorporating implicit feedback from users. This includes considering the movies they selected and the posters they clicked on, providing a more nuanced understanding of their preferences.