

Movie Recommendation Chatbot Project Report

By: Niruthiha Selvanayagam

1. Introduction

The Movie Recommendation Chatbot project aims to provide personalized movie recommendations based on user queries using advanced natural language processing (NLP) techniques and vector search capabilities.

2. Methodology

Project Setup

I utilized Streamlit for creating an interactive user interface and integrated Pinecone for efficient vector similarity search. Additionally, I employed a large language model (LLM) from the LangChain community for generating text prompts.

Data Handling

Movie data was sourced from a `movies.txt` file containing titles, genres, and descriptions. This data was parsed and structured into a suitable format for querying and recommendation generation.

Functionality

- **Pinecone Integration:** Pinecone was used to store and query movie vectors, allowing me to retrieve top-k nearest neighbors based on user input vectors.
- **LLM Integration:** The LLM generated prompts based on user queries to refine the search and recommendation process, enhancing the relevance of movie suggestions.

3. Challenges Faced

Challenge 1: Integrating Pinecone

Solution: I faced initial challenges in setting up Pinecone and configuring API keys. Through thorough documentation and support resources, I successfully integrated Pinecone for efficient vector search operations.

Challenge 2: LLM Fine-tuning

Solution: Fine-tuning the LLM for movie recommendations required experimentation with prompt structures and response formats. I refined the model by iterating over different configurations until achieving satisfactory results.

Challenge 3: Streamlit App Optimization

Solution: Optimizing the Streamlit app for responsiveness and user experience involved caching mechanisms and layout optimizations. This ensured smooth interaction and quick response times for users.

4. Results and Discussion

Performance Evaluation

I evaluated the chatbot's performance by measuring the accuracy of movie recommendations against user preferences. Initial user feedback indicated a high degree of satisfaction with the relevancy and diversity of recommended movies.

Discussion

The integration of Pinecone and LLM proved effective in enhancing the recommendation accuracy. Future iterations could explore incorporating user feedback loops to further personalize recommendations based on individual preferences.

5. Conclusion

In conclusion, the Movie Recommendation Chatbot successfully leverages advanced NLP and vector search techniques to deliver personalized movie recommendations. The project highlights the potential of combining machine learning models with efficient data retrieval systems for interactive applications.

6. Future Work

Future enhancements may include:

- Implementing sentiment analysis for refining recommendations based on user mood.
- Integrating user profiling to tailor recommendations to individual preferences more accurately.
- Scaling the application to handle larger datasets and support more diverse movie genres.