CS 410 Project Proposal – Media Match

1. Team Members: Kanjana Chandren

Net Id: kanjana3

1. Project Description: This project will create a chatbot that offers personalized movie and book recommendations based on the user’s mood and/or preferences. The chatbot will utilize open-source large language models and gen ai frameworks for language understanding and text generation. It will engage users with an easy-to-understand UI and conversational format to understand their mood and preferences with natural language processing and provide tailored recommendations for books and movies.

Task and Importance: The task is to build a chatbot that can understand and respond to user input effectively, and then be able to return online book and movie titles in the context of a conversation. This will be interesting because the chatbot will show capabilities of providing personalized recommendations which can be significant in increasing user satisfaction and user engagement with movies and books. This project will also showcase how large language models can be adopted into real world applications to create conversation AI for recommendation systems.

Approach: The initial steps of the project will be to first select an open-source large language model that will be suitable for text generation and natural language understanding. Next will be setting up an environment that can use python libraries, frameworks like PyTorch and the chosen model. I will collect text datasets of books and movies to be used as an input from the model. The implementation will either consist of a Retrieval Augmented Generation approach or fine tuning the model. This model will be integrated with a custom built chatbot interface using stream lit. Lastly, I will test the chatbot with diverse scenarios to make iterative improvements and deploy it on a public to use website.

Tools, Systems, and Datasets: This project will use an open-source large language model as well as Python or PyTorch model implementation. The datasets for books, movies or training datasets for fine tuning will be hosted in an object storage system that will be used by an index search engine system. Lastly the chatbot user interface will be created with Stream Lit and deployed on a website either using AWS CloudFront distribution or as an application to be publicly available.

Expected Outcome: The expected outcome is a fully functional chatbot that can understand user sentiment from their mood or preference and give personalized recommendations for books and movies.

1. I will be using Python.
2. Here is a breakdown of planned time commitment for each step in my approach: Environment Configuration (1hr), Data collection and preparation (4 hours), model selection and processing (RAG or fine tuning) (6 hours) chatbot development (4 hours) chatbot integration (1 hour) recommendation and sentiment analysis tuning for the bot (3 hours) testing and iterating (2 hours) deployment on a website (1 hour). At a minimum this project would roughly take 23 hours without including the time for creating the presentation and any documentation. Additional complexities for the project will increase the time which goes beyond the 20-hour mark allocated for one person.