

CS 410 Project Proposal

Team Name: Model Behavior

Team Members:

Names	NetID's
Gautam Putcha	gputcha2@illinois.edu
Ashna Anil (Captain)	aanil2@illinois.edu
Kunal Samant	ksamant2@illinois.edu
Parth Maheshkumar Patel	pp32@illinois.edu

We will be building a search engine that ranks events/activities in your location based on events on websites such as Eventbrite, tweets or Facebook events based on a combination of user preferences/history. This is a text retrieval and ranking task after parsing information from different sources. This uses content based filtering while ranking the events to match the users preferences.

We found this to be an important and interesting task and in today's world there are a lot of events occurring and it is hard to keep track or try to find the right events on social media. Searching on any one source alone typically does not allow users to find the events they are looking for. This search engine will combine and rank the data and make it easier for users to find all of the events they are interested in from a single source of truth.

Our approach is to pull data from a variety of sources containing data about events near any given user's location. This would include Eventbrite, AllEvents.in, Facebook events, and even tweets related to nearby events. We would be using BM25 or similar ranking systems from the Metapy Library that we have used during the 410 Coursework.

Our expected outcome is to develop a centralized system for finding events in a users area by searching for relevant events from several sources of data. We expect to have a well-functioning frontend to accompany our text retrieval and ranking system. We hope to leverage the text retrieval techniques learnt in this course to optimize our ranking process and find the most relevant results for users.

We will get user feedback through relevant or not relevant ratings and improve on our performance with higher success rates. Lastly, we intend to use Python for the Backend and React/Flask for the Frontend of the project.

Workload is at least 80 hours as the tasks involved and approximate times include:

1. Building the framework ~ 5hrs
2. Collecting Data Sets ~ 5hrs
3. Preprocessing Data ~ 10hrs
4. Building the Ranking Model ~ 20hrs
5. Building Frontend/Connecting to Backend ~ 20 hrs
6. Testing/Evaluation ~ 15 hrs
7. Documentation ~ 5 hrs
8. Conclusion and Future Work ~ 5 hrs