Stage 4 - User Interface

- **Section 1: Edit Distance Similarity Search UI** The UI for finding semantic similar tweets based on Edit Distance integrated Word2Vec (EDW) and K-medoids clustering (implemented by Lingfei Zeng).

  The mode of user interaction with the data is text queries and mouse clicks. There are two input boxes. The first one is for text input such as a tweet, some key word or a topic (such as hashtags in tweets). The second one is for an integer to indicate how many (top) similar tweets one would like to display. Then mouse click for submit, the UI will return the output, which are the similar tweets.

  The initial UI screenshot in shown Figure 1. After input a tweet and the number of tweets one like to show, the result is shown in Figure 2 and Figure 3 . One can also input one word as topic or hashtags in the tweets. The search result is shown in Figure 4. Overall, the output is quite accurate, which indicates the good performance and versatility of our algorithm.

  The error massage will pop out if user does not input anything but hit submit button. It's shown in Figure 5.

Figure 1: Initial UI screen shot

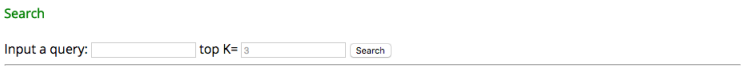


Figure 3: Output after another tweet

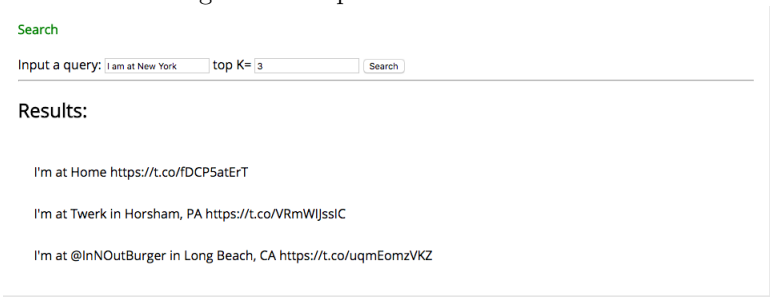


Figure 2: Output after input of a tweet

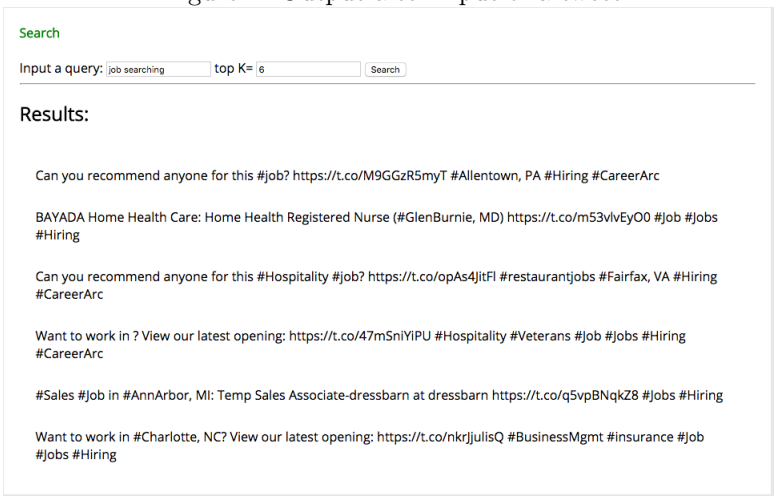


Figure 4: Output after input a hashtag

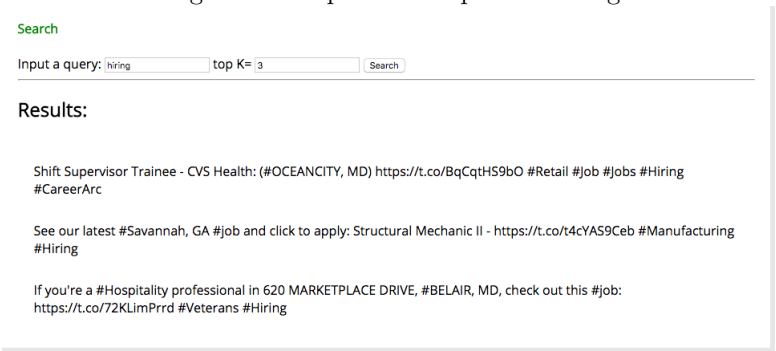
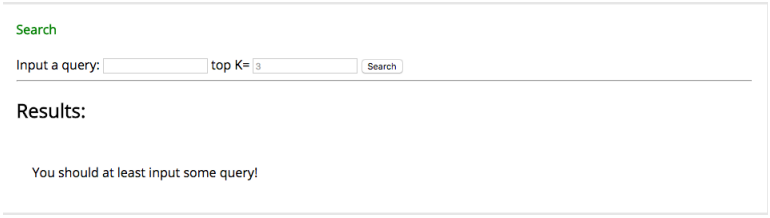


Figure 5: Error message when user doesn't input query but hit the submit button



- **Section 3: Network Similarity Search UI**

The Network Similarity Search UI was written in the Python Tkinter library. When the UI starts, it loads the following from the initial dataset: its graph, tweet and hashtag IDs, and scores.

The UI (started by running: run_UI.py) can be found on:
https://github.com/circlefive05/SimRank-on-Twitter-UI

A 3 minute demo of the user functions and exception handling cases can be found here: https://youtu.be/CHiQ8C7cdtg

User functions:

- Input tweet and search for similar tweets or hashtags
- Input hashtag and search for similar tweets or hashtags (different from the input)
- Input a set of new hashtags, search for new tweets, and return similar tweets or hashtags. Add the newly searched dataset of scores to set of dataset IDs
- Show tweet IDs in current dataset
- Show hashtags in current dataset
- Show current dataset IDs
- Switch datasets by inputting ID

UI displays:

- Tweet IDs and hashtags in current dataset
- If user inputs tweet ID, UI displays both the content of that tweet ID and the similar tweets or hashtags
- Similar hashtags/tweets
- Set of dataset IDs

Exception Handling- Returns error message on UI if user performs:

- If type is integer that is not in current dataset
- If type in hashtag not in current dataset
- If, in switching dataset, input is not a valid dataset number
- If type in hashtag with spaces and don't search 'Gather new tweets'
- If no similar tweets/hashtags found
- If similar tweets/hashtags found but no new tweets/hashtags found