**SI507 Final Project Checkpoint**

Name: Yuru Gong

**Data Collection:** I scraped from multiple pages, the initial URL is https://www.imdb.com/search/title/?groups=top\_1000&view=simple&sort=user\_rating,desc&count=100 Then I changed "&start=101" to "&start=201", ""&start=301" etc. I use movie class to store the information. Finally, I write the data in to a csv file.

The next step: store the data into Binary Search Tree class.

**User interaction:** prompt user to search a number in range of (7.5-9.5) or enter a range (n, n) to return a list of movies starting with the highest rating. If user entered range (n, n), prompt user graphs options list, after user enter a choice number, display graphs using matplotlib or other libraries. The tool can also open the URLs for the user.

**Graphs options:** only available when the user search with a rating range, and will use the returned movie list to generate:

1. Bar chart of movie\_year/movie\_rating/movie\_votes/title\_word\_count
2. Line chart of years/movie\_votes/title\_word\_count vs rating