CS 410 Project Progress Report

Tyler VanderLey

1.

The main task I've completed so far is setting up the PRAW API to read comments from Reddit. To do so, I first created my client ID, client secret, and user agent so that I can successfully query data from Reddit. After getting the basics set up, I explored and validated my application's ability to take in the URL for any post-game thread from the college football subreddit, read its comments, and filter based on user flairs. Thus, most of the text retrieval component of my project is complete. Another key task that has been completed is the exploration of how to perform sentiment analysis on the comments from these reddit threads. In particular, I found that using VADER, a pre-trained sentiment analysis tool from the NLTK Python library, is an effective method of quickly determining whether any individual comment is mostly positive, negative, or neutral. Since my application will require analyzing many of these comments across multiple threads, I plan to use this method for its ease and efficiency. Hence, I have a plan for how I will analyze sentiment for my project.

2.

A few tasks are still pending for my project. First, I need to implement the pieces of the application that will be user-controlled. For instance, I need to enable the application to take as input what set of reddit threads to analyze and whether to filter on user flairs. Next, I need to curate a subset of pre-selected threads that a user can choose to analyze in case they don't want to manually provide links to each thread. This will entail selecting which teams to use as this pre-chosen subset, finding the URLs of their post-game threads, and then setting these in the application as options for a user to select. Additionally, I still need to configure the entire user interface, which will likely comprise most of the remaining effort. In this UI, I plan to have an input screen where the user can specify the inputs mentioned above, and then an output screen where they can see how sentiment changed over the course of the season based on the threads they chose to analyze. After doing all of these implementation tasks, I will also still need to create documentation in order to provide a guide for how the application should be used.

3.

  Currently, I am not facing any challenges. Using PRAW to pull data from Reddit and utilizing NLTK to perform sentiment analysis have proven to be doable tasks, so thankfully I haven't had any trouble so far. I do anticipate having some potential challenges when configuring the UI, as I haven't used Flask (the tool I plan to use) in quite some time, so I may be a bit rusty. If it does pose a significant challenge, I could move to a command-line based tool, but I should likely figure out Flask quick enough that it shouldn't be a blocker for the development of my project.