Term Project Proposal: Analytical Tool for Social Media

Rose Lauture, Brian Collins & Kaija Petrone
Prof. Zhi Li | MIS3640: Problem Solving & Software Design | 16 November 2016

Big Idea:

The big idea for our project is to create an analytical tool that social media managers can use to compare their company's activity with competitors' activities on various platforms. Specifically, we aim to create a dashboard that returns key metrics, such as number of posts posted by a given company, average number of likes per post, average number of shares per post, the sentiment of the posts, the average number of post views, and so on. To format the results, we can strive to create a webpage that neatly and effectively displays the metrics for the social media managers to easily reference.

Learning Goals:

Our learning goals will include learning how to compile different sources of companies' social media activity into a file to analyze. In addition, we plan to learn what code we can use that will measure the data and give us an output where we can make assumptions. Furthermore, an important learning goal will be for us to learn how to incorporate the information we will receive from the output and print it out on a webpage both neatly and efficiently.

Implementation Plan:

To complete this project, we plan to follow the following steps:

1. Select social media platform to utilize (e.g. Twitter, Facebook, Instagram, Snapchat, LinkedIn, etc.).
2. Determine which metrics we plan to measure.
3. For the selected platform, obtain a list of posts with corresponding information, including post location, time, username, age of user, gender of user and other demographic information about user.
4. Write loops to analyze the sentiment of the posts.
5. Extract the locations of the users.
6. Aggregate data to provide insights about groups of users and allow for comparisons.
7. Utilize Flask to create a webpage on which to return the metrics.
8. Consider requiring users to login by creating a username and password.

Project Schedule:

(Bold: Actual Deliverables; Non-Bold: Team Goal Deliverables)

Task	**Target Deadline**	**Actual Deadline**
Project Proposal	**16 Nov**	**16 Nov**
Select Social Media Platform	21 Nov	

Determine Preferred Metrics	21 Nov	
Research How to Extract Data from Selected Platform	28 Nov	
Design Review	**TBD**	**TBD**
Write Pseudocode	28 Nov	
Write Actual Code	30 Nov	
Create Detailed List of Questions	30 Nov	
Meet Prof. Li with Questions	01 Dec	
Create Project Website	30 Nov	
Code Review	**TBD**	**TBD**
Outline Presentation	02 Dec	
Complete PowerPoint	~30 Nov	
Update Project Website	01 Dec	
Mid-Project Presentation	**TBD**	**TBD**
Finalize Project Website	04 Dec	
Project Website	**TBD**	**TBD**
Finalize Code (Clean Up and Add Docstrings)	02 Dec	
Code Submission	**TBD**	**TBD**
Outline Presentation	02 Dec	
Complete PowerPoint	04 Dec	
Final Demo/Presentation	**06 Dec**	**08 Dec**

Collaboration Plan:

To collaborate effectively on this assignment, we plan to utilize various online productivity tools, including Slack for communication, GoogleDrive for simultaneously working on document and GitHub to work on the same repository and allow for files to be automatically updated. A notable benefit of Slack includes the fact that it allows various other platforms to be integrated into it; for example, if we deem it necessary, we can can integrate Trello into Slack.

Trello is an online platform used to organize tasks and ideas into various lists. There are many features within the software, such as the ability to assign users to tasks, create checklists, upload links, and specify deadlines.

Perhaps one of the most important measures to ensure effective use of time is to schedule our meetings weeks in advance so that we can be appropriately prepared for them. It would also be effective to create lists of what needs to be accomplished before each meeting and what needs to be accomplished by the end of each meeting.

An agile development methodology can certainly be effective for this project. It can encourage us to work quickly and iteratively and not spend too much time working on a feature that we may later realize is not desired by our end users. In the execution of any project, feedback is incredibly helpful because it highlights areas of improvement that developers can use to work even more efficiently and effectively.

Risks:

A significant risk that we foresee for this assignment is not smoothly extracting the code from the data source. This can be problematic because if we do not have a good foundation upon which to analyze the data, our project will not be as strong. In the effort to avoid this risk, we should perform research on how to properly extract data from data sources. Perhaps we can begin by utilizing the resources provided by Professor Li for Assignment 3 on Text Mining. Additionally, we can utilize online and in-person resources for help.

Additional Course Content:

Particularly relevant topics we have discussed in class include:

- Harvesting Data (from Assignment 3)
- Pickling Data
- Counting Numbers of Words
- Natural Language Processing