Project coordinator: Sita Jothishankar (slj2@illinois.edu), Asha Agrawal (ashaa2@illinois.edu)

What is the function of the tool?

This tool will scrape Twitter for a stock ticker name and then perform sentiment analysis to see how much positive and negative news there is regarding the term.

Who will benefit from such a tool?

Our main audience will be banks, investors, and shareholders. After finding the ratio of positive/negative news, they can incorporate this data into their views to make more robust and preemptive decisions for different commodities, stocks, companies, etc.

Does this kind of tool already exist? If similar tools exist, how is your tool different from them? Would people care about the difference?

We were not able to find an actual tool that found the sentiments, but found a study that showed that there is a correlation between between tweets and stock performance: https://towardsdatascience.com/stock-prediction-using-twitter-e432b35e14bd and http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf. With this supporting study, we believe that our project has a potential to make an impact.

What existing resources can you use?

https://github.com/cjhutto/vaderSentiment/blob/master/vaderSentiment/vaderSentiment.py and https://github.com/satishrath185/Movie-Review-Sentiment-Analysis/blob/master/Sentiment%20Analysis.ipynb can give us a launching pad for a sentiment analyzer. We will be adding to it to create a more robust analyzer.

What techniques/algorithms will you use to develop the tool? (It's fine if you just mention some vague idea.)

We will use cross validation to create a sentiment analyzer on a dataset from kaggle: https://www.kaggle.com/yash612/stockmarket-sentiment-dataset

Once we have created the analyzer we will use the twitter API to perform sentiment analysis on current tweets to provide users with the most up to date information. As mentioned above, we will be using existing resources and studies to help us figure out which algorithms would be best to use.

How will you demonstrate the usefulness of your tool?

We will want to show correlation between stock price and the sentiment of news stories. With this correlation, we will be able to demonstrate the usefulness as the analyzer will provide our audience a preemptive measure for their stocks.

A very rough timeline to show when you expect to finish what. (The timeline doesn't have to be accurate.)

Part 1: Oct 26 - Nov 2	Research algorithms and learn more about sentiment analysis
Part 2: Nov 2 - Nov 16	Create and train sentiment analyzer
Part 3: Nov 16 - Nov 25	Perform sentiment analysis on current tweets using our sentiment analyzer and twitter API
Part 4: Nov 25 - Dec 9	Create user interface for user to interact with analyzer + write out documentation of our project