

DPM project proposal

Names: Parth Vyas, Yashwanth Babu Dasari

For this project, we plan to create a plotting and data showcasing tool for tracking stocks using the twitter data over a period of time and also for the current day. We will use the tweepy library for initiating requests to the twitter data API. After inspection of the query parameters, we noticed that we can also get the location and associated hashtags from a particular tweet. This will enable us to also enable the user to get location based tracking of discussion on a particular day in general or a particular stock. In addition to the data processing and backend part, this will enable us to create an interactive end to end application.

The sentiment analysis will be divided into 2 parts:

1. The basic analysis using groupby in pandas and other methods
2. Using a library such as pytorch or tensorflow for NLP in tweets and grouping them into various types. A rough grouping will follow:

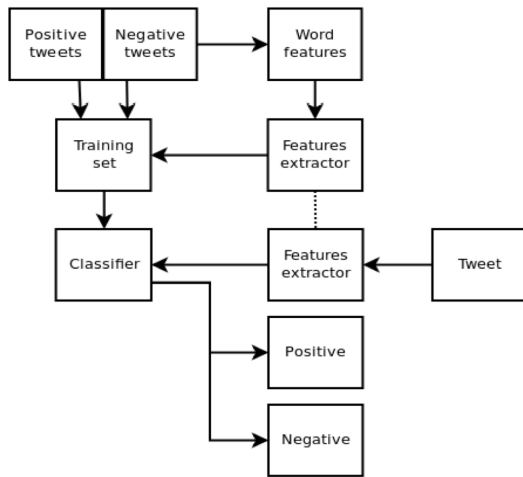
Term presence and their frequency

Part of speech information

Negations

Opinion words and phrases

Here is the rough roadmap that we want to work upon:



The input and output characteristics will be written more explicitly in the week 1 deadline.