**CPTS 570 Project Proposal Document**

**Team members:** Kunal Sanghvi, Dheeraj Vurukuti

**Problem Statement:** Twitter Sentiment Analysis, Sentiment analysis, is a machine-learning task where we want to determine which is the general sentiment of a given document. We can extract subjective information from a text and attempt to categorize it according to its polarities, such as positive, neutral, or negative, using machine learning techniques. It is a very useful analysis since we might ascertain the general perception of news or forecast stock picks for a specific company, for example, if most people have a favorable opinion of it, perhaps its stock prices would rise, and so on. Due to the complexity of the language (vocabulary, grammar, etc.), sentiment analysis is still far from being solved, but this is exactly why it is so fascinating to work on.

**Methodology:**

Extract data from Twitter and use data wrangling techniques, we will further process it and then do sentimental analysis using a logistic regression model and further derive information on the tweets.

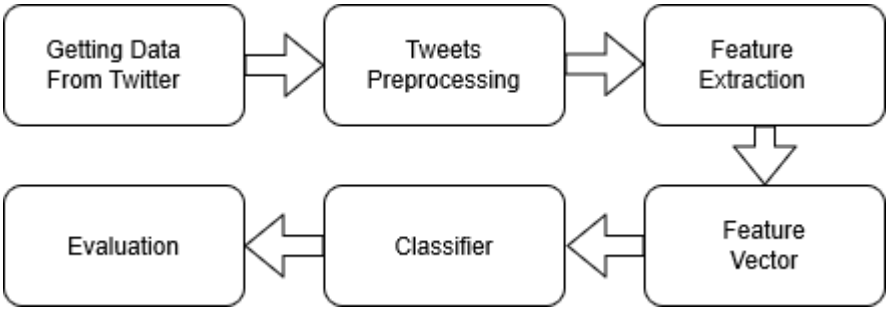


Fig 1. General Flow chart of Methodology

We are planning to use it on various models like SVM and Naïve Bayes as well. The in-depth flow chart is as follows:

**Final Product:** Based on the different tweet data of different sources like financial analysts, news channels, editors, writers, general people, etc., we will predict the sentiment of tweets, and based on that we may predict top stock picks for that day. We will also show a comparison of various models based on which we will derive that the model which we have used has a good accuracy over others.