# Stock-Prediction-Using-Twitter-Sentiment-Analysis

In this project, we have applied sentiment analysis and two statistical machine learning models, Random Forest and Support Vector Regression. These models are used to depict the correlation between the tweets extracted from Twitter and the stock market movements of a company. We have performed sentiment analysis of the Twitter data based on a whole day as well as on an hourly basis to analyze the effect it has on stock market prediction. RMSE (root mean squared error) values evaluate and compare the models.

An overview of the process flow of our project is as follows:

**1. Data Collection & Parsing:** The stock market data is collected using Yahoo Finance API and tweets are fetched from Twitter using GetOldTweets API. In this step, the preprocessing of the tweets such as removing stop words, hyperlinks, and other steps are carried out.

**2. Sentiment Analysis:** The sentiment analysis of the tweet is carried out using VADER. Here each tweet is given a sentiment score which determines if the tweet is positive, negative, or neutral.

**3. Processing:** The rows that have missing values such as price values are further processed. The data along with the sentiment scores is divided into train and test data and is fed to the model.

**4. Applying Regression Models:** To predict stock Market prices, we have used Random Forest and Support vector regression models in this project. We will then use RMSE scores to validate the efficiency of our models and to analyze which model works better for the used dataset.