**Elon Musk’s Tweets vs. TSLA Stock Price (Sentiment Analysis)**

1. Abstract:

In this project we would like to find the relationship between tweets of one important Twitter user CEO Elon Musk and the corresponding stock price behavior. The tweets of Elon Musk, who is the CEO of Tesla, and the change of Tesla stock price are used as data in our project. We have tried to derive Sentiment Analysis of Elon’s Tweets on Tesla’s stock movement using VADER Sentiment Analysis.

1. Objective:

The main objective/problem definition is to derive the sentiment of Elon Musk’s Tweets, try to evaluate the twitter sentiments and correlate it with the stock fluctuation/volatility caused due to his tweeting pattern.

1. Methods:
   1. Data Mining Process:

3.1.1 Tweet Data:

One challenge of this project is to get the data from Twitter, since the large data set available to use earlier are no longer available according to Twitter’s privacy policy. Therefore, we use the Twitter developer tool tweepy to obtain the content of tweets of Elon Musk’s Twitter account. At the most 200 tweets information could be collected for a certain user using Tweeter API. In this project, information of 200 tweets of Elon Musk was collected via Tweepy.

3.1.2 Finance Data:

With reference to a pre-requisite, API being used for this project (link: <https://api.iextrading.com/1.0/stock/TSLA/book>) to derive stock data set for Tesla. We used Python module wrapper to retrieve stock data from the Investors Exchange (IEX) Developer API platform. IEX provides real-time financial data from the various IEX endpoints, as well as historical time-series data. This data includes stock quotes, fundamentals, actions, and information. In addition, support for IEX market data and statistics is provided:

Stocks

Reference Data

IEX Market Data

IEX Stats

3.1.3 Sentiment Mining

We have used VADER sentiment analysis. Sentiment analysis is simply the process of working out (statistically) whether a piece of text is positive, negative or neutral. VADER produces four sentiment metrics from these word ratings, which you can see below. The first three, **positive, neutral and negative**, represent the proportion of the text that falls into those categories. The final metric, the **compound score**, is the sum of all of the lexicon ratings which have been standardized to range between -1 and 1.

3.1.4 Match of Twitter and Finance Data

* Creating data frames for each of the above data sets: tweet data, stock data, sentiment data.
* Filtering out the tweet data based on the indexes to match the stock data retrieved on the corelated trading dates.
* Resetting the index and sorting the data frame to prove a correlation between the dates of Elon Musk’s tweet with the change in tesla stock price.
* Creating a dictionary in a final data frame of tweets and corelate change in tesla’s stock price

1. Validation and Visualization

Using matplotlib plotting various graphs to demonstrate the analysis we tried to evaluate the twitter sentiments and correlate it with the change in Stock (Volatility) We will primarily focus on Compound Sentiments based on Elon’s tweets of the a particular time period in comparison to fluctuation in stock price and volumes generated.

* 1. Sentiment Analysis on Tweet: Compound V/s Positive V/s Negative score on tweet
  2. Compound Score Vs Stock Volume (Scaled in Millions)
  3. Change in Closing price of tesla stock Vs Compound Score
  4. Percentage Change Over Time vs compound score
  5. Percentage Change Over Time vs Tweet Date
  6. Retweets VS Compound Score

1. Conclusion:

In this project we have tried three pairs of features and labels to investigate the relationship of an important twitter account holder Elon Musk’s tweets and the stock price behavior:

1) The most frequent tweets and the “jump” / fluctuation of tesla stock price,

2) The sentiment scores and the direction of movement, and

3) Most recent aggregation of sentiment based on retweets.

We conclude that with the help of this project, the Social media sentiment score along with some other tools and advanced machine learning implementation can predict the movement of stocks based on the behavior/reactions of sentiments analyzed over Social Media Platforms