STOCK PREDICTION WITH TWITTER SOCIAL NETWORK ANALYSIS

SACHIN NEGI(MT19DCS005)

TABLE OF CONTENT

1. Project Title
2. Introduction
3. My work
4. Result

WHAT IS STOCK PREDICTION?

It can defined as predicting whether the stock price will go up or down in the future(can be the immediate next day).

CONVENTION METHOD FOR STOCK PREDICTION

- Technical Analysis
- Fundamental Analysis.

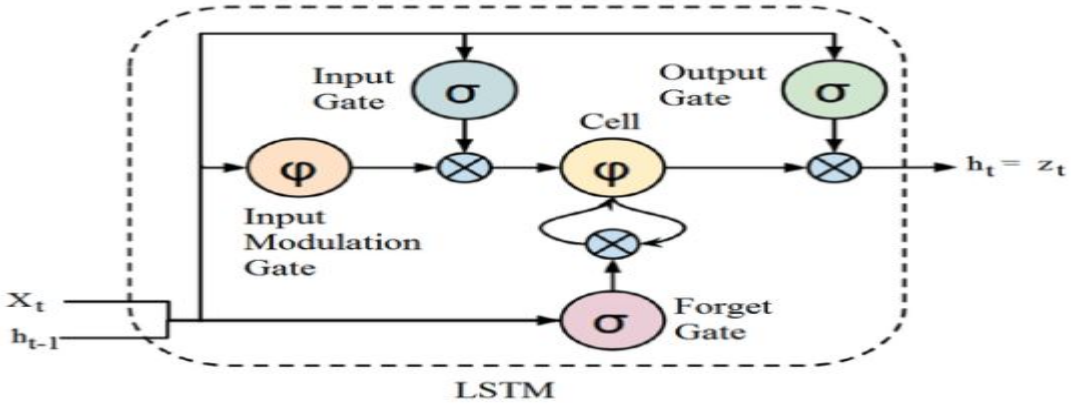
The technical analysis involves use of various Technical Indicators like simple Moving Average, Relative Strength Index, Exponential Moving Average etc . It basically involves finding the trend or pattern in stock price behaviour.

Fundamental Analysis involves analyzing balance sheets, statement of cash flow, income statement etc.

MACHINE LEARNING FOR STOCK PREDICTION

LSTM(Long -Short-Term-Memory) Model

The main advantage of LSTM Model over the conventional model is that it able to discard non-important information from history and retain important one.



SENTIMENT ANALYSIS

It involves taking the news in the social media. And based on opinion to predict the mood of the people. The sentiment analysis combined with the above method give good prediction.

What I have done?

I have worked on the aspect of sentiment analysis.

- Collected the tweets data from twitter through twitter API.
- A Tweet analysis is done. A tweets contains a lot of information:

  [user_id, user_name, user_screen_name, user_location, user_description, user_follower_count, user_friend_count, retweeted_data]

- Sentiment Analysis is done on the tweets. It is calculated if the tweets is positive or negative.
- Weightage to a positive or a negative tweet is given based on impact on the tweet.
- The mood of the people regarding a stock can be ascertain.

Preprocessing Step-Collecting the data

- I particularly have worked on COAL INDIA LIMITED stock.
- From twitter latest tweet related to it can be observed.
- I have used Twitter API and Postman.
- Each tweet has a unique id to it. Using Postman all information related to a tweet can be obtained.

Data Cleaning

- The data I obtained was in JSON format and converted to CSV for working.
- The CSV file obtained had 151 columns containing different information about a tweet. Some important parameters were:
  [user_id, user_name, user_screen_name, user_location, user_description, user_follower_count, user_friend_count, retweeted_data]
- The parameters I have chosen for analysis are:
  [total_followers, retweets, favourite_count]

Method Used

- Firstly each tweet sentiment analysis is done to know whether the tweet is positive, negative or neutral.
- Scoring the tweets: To score the tweet:
  The number of followings, retweets, likes are considered:

  $$score=(tweet\_sentiment*(log(\Sigma\ followers)+retweets+likes))$$

  tweet_sentiment: 1=positive,-1=negative, 0=neutral

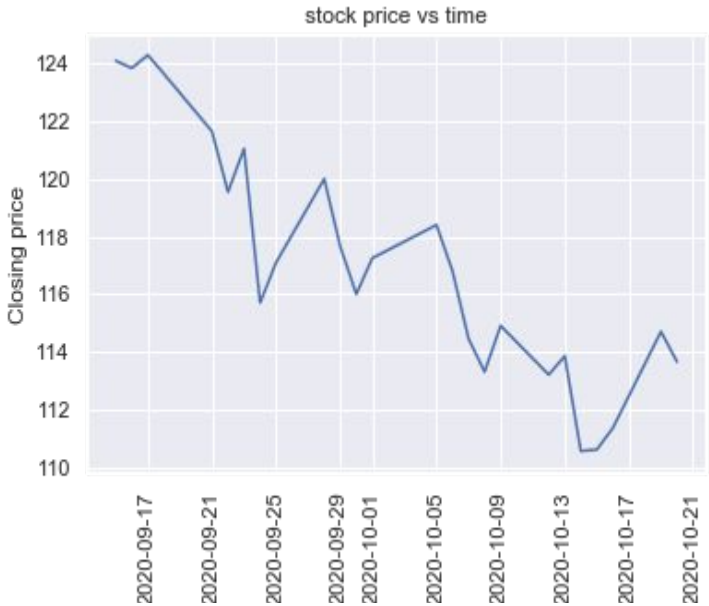
- It can be ascertain the overall perception about a stock at that moment of time among the people

Result

- The tweets are obtained between 10 Oct ‘20 to 20 Oct ‘20.
- On 15 Oct a positive tweet was widely shared and liked and for the subsequent days tweets like that were most dominating tweet in term of likes, retweet, comments.The exact tweet was:

  Coal India Limited declares Rs 68,000 performance-linked reward for its employees

- The stock price hit at its bottom upto 14 Oct but the following day the closing price increases and the trend continued upto 20 Oct.



Future scope

- Tweets streaming for the latest tweet related to a stock.
- Sentiment analysis of comments as they are available in Hindi language as well.