International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Stock Market Analysis using LSTM, Twitter and Technical Indicators

Prof. Omprakash Yadav*, Rahul Patel[1],Yash Shah[2], Saneesha Talim[3]

*Professor, Department of Computer Engineering, Xavier Institute of Engineering, Mumbai, Maharashtra, India
[1,2,3]B. E student, Computer Engineering, Xavier Institute of Engineering, Mumbai, Maharashtra, India

---

Abstract - *Stock value unpredictability is a profoundly mind twisting. The stock's volatile parameters influence the stock's current and the futures performance. LSTM (Term Memory Long-Short) is a sort of time intermittent neural system, which is appropriate in predicting the futures of a particular stock. In term of the qualities of stock and LSTM neural system calculation, this paper utilizes its repetitive neural systems to channel, examine the stock information, and set up the model for stock market. A mathematical approach knows as technical indicators is used to gain insights about the future trend of stock. On the other hand, a non-technical-approach used to analyze the future market is through analyzing sentiment through tweets using Twitter.*

***Key Words*: Close Prediction, LSTM, Neural Network, Twitter Sentiment Analysis, Technical Indicators.**

1. INTRODUCTION

Stock market forecasting has the most enthusiasm among researchers and analysts from different scholarly fields irrelative from their business advance. From the time the stock market has come into existence, it has been a great challenge for the investors to predict the Stock Market.

Based on this, the stock market is generally regarded by scholars from all walks of life as an intuitive reflection of the economic development of a country or region in a certain period. One of the main reasons lies in the stock market trading prices that can objectively reflect the stock market supply and demand relations. Moreover, the stock market is oftenregarded as an indicator of stock prices and quantities. However, due to the complexity, variability, and uncertainty of the stock market, the stock price formation mechanism presents the characteristics of complexity and unpredictability. Stock prices not only from the political, economic, market, technology, and investor behavior aspects such as individual factors, influenced by various factors in the interaction between the role at the same time, these will lead to changes in stock prices, the existence of the various uncertain factors lead to the complexity of the stock price changes.

Every penny that a stock investor invests in the stock exchange, he hopes to profit in a way or the other. Every company's rise and fall affect the behavior of the stock of the company Nevertheless, the mining stock market trend is generally considered to be both an interesting and challenging task due to its uncertainty, nonlinear, dynamic, characteristics. To normalize market for non-technical background people, our aim is to provide a secure, personalize application solution using the concepts of machine learning to provide an informative solution to venture into the depths of the market with a stepwise solution from buying trending or most likely to trend stocks to predict the precise time to sell those stocks.

2. LSTM

This study proposes an attention-based long short-term memory model to predict stock price trend. The model consists of four parts: input layer, hidden layer, attention layer, output layer. The input layer cleans the input data to meet the input requirements of the model. The hidden layer is connected to the line model network through LSTM unit. The attention layer weighted the feature vector. The output layer gets the calculated results. The model training is solved by gradient descent algorithm.

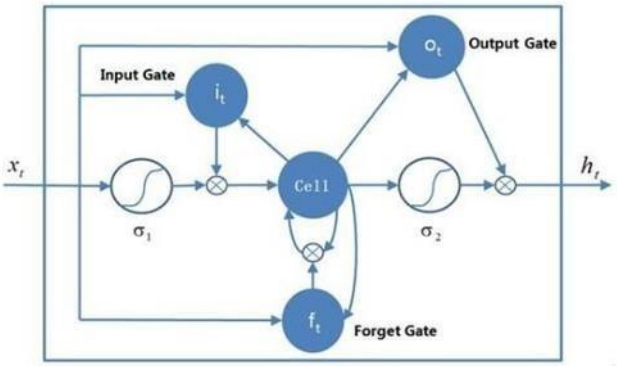


Fig-1: Conceptual breakdown of LSTM logic

The core concept of LSTM's is the cell state, and its various gates. The cell state act as a transport highway that transfers relative information all the way down the sequence chain. You can think of it as the "memory" of the network. The cell state, in theory, can carry relevant information throughout the processing of the sequence. LSTM has a new structure called a memory cell. A memory cell contains four main elements: an input gate, a forget gate, an output gate and a neuron unit which is a central linear unit with a fixed self-recurrent connection. The memory cell makes the decisions about what information to store, and when to allow reading, writing and forgetting, via the three gates that open and close. The three gates are implemented using the activation function to compute a value between 0 and 1.

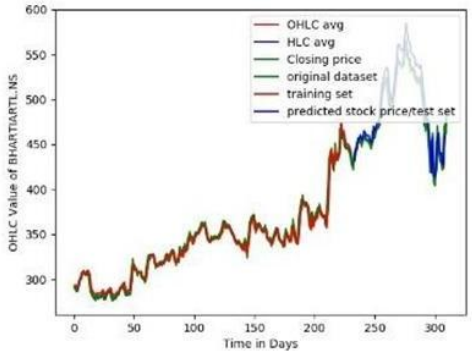
A. Input Gate:

(1) Take the date, closing price, opening price, maximum price and minimum price of the stock as input data to form a time series; (2) split the input data into training set and test set according to the ratio of 8:2(3) convert each component of input data into the interval [0,1] after standardization.

$$i_t = \sigma\ (W_i\ .\ [\ h_{t-1}, x_t] + b_i)$$

$$\mathcal{C} = \tanh(\ W_C\ .\ [h_{t-1}, x_t] + b_C)$$

B. Forget Gate:



TRAINING RMSE IS : 6.82529578381203 45

TEST RMSE IS : 16.170935869748 1

LAST DAY VALUE IS : 460.0045471191406

NEXT DAY VALUE IS : 493.37335205078125

This gate decides what information should be thrown away or kept. Information from the previous hidden state and information from the current input is passed through the sigmoid function. Values come out between 0 and 1. The closer to 0 means to forget, and the closer to 1 means to keep.

$$ft = \sigma\ (W_f.\ [\ h_{t-1}, x_t] + b_f)$$

C. Output Gate:

After the model training is completed, stock time series is input for prediction, that is, stock data is input for N days to predict the stock trend on the N+1 day. The trained model uses the trading data of the first four trading days to predict the closing price of the fifth trading day.

$$o_t = \sigma\ (\ W_o[\ h_{t-1}, x_t] + b_o)$$
$$h_t = o_t * \tanh(C_t)$$

Chart-1: LSTM model implementation using Airtel Limited stock for 2019-01-01 to 2020-01-04

3. Twitter Sentiment Analysis

An opinion or view that is expressed based on emotion is often referred to as a sentiment. These passionate feelings from various gatherings of users, when examined and determined, their mentality towards the total relevant extremity or enthusiastic reaction to a report, correspondence, or occasion is done in sentiment analysis. Sentiment analysis, permits user to know the circumstance of what the review are made regarding specific product or ongoing matters rather than understanding deep thoughts of the reviewer.

Twitter, being one among several popular social media platform, where people communicate their feelings and opinions about a brand, an item or an assistance. Examining sentiments for tweets is exceptionally useful in deciding individuals' opinions as positive, negative or neutral. If individuals discover points important or fascinating, at that point they would want toshare their insight about the subject. The topic could be an item or an assistance or a social message or some other article.

Taking everything into account, with everyone's web-based beneficial experience being so isolated and individualized, it might be difficult to check twitter assessment and whether it is sure or negative from a single channel. Dismembering Twitter data and doing estimation assessment with tweets can be significantly more straightforward than you may speculate - read on to find with our twitter supposition examination model. Investigation of online networking realities like tweets encourages individuals to get a thought of a specific item or theme which makes it simple to make choices on it. Estimation investigation additionally makes a difference in individuals to change their demeanor about the wrongful conviction on an item, administration or a point. It causes individuals to pick which is best by dissecting the remarks or on the other hand tweets of a specific subject or an item.

The client's tweets from twitter are gathered dependent on input indicated by the client as hash tags. The way toward characterizing tweets is started by first gathering the tweets. Utilizing Twitter API, it is reachable to gather twitter information. A library named RAuth is utilized in performing validation by giving in the keys. Purchaser Key, Consumer Secret, Access Token and Access Token Secret for twitter application and perform Handshake convention. After this, a declaration is downloaded and PIN is produced for the application to get to tweets.

Right now, live twitter information is ordered dependent on the estimations. This is possible in 4 stages. In stage I tweets are gathered by taking the contribution to the type of hash tags and the number of tweets to be considered is limited somewhere in the range of 5 and 1000. The tweets gathered are the ones that are gushing live on the web. In stage II the gathered tweets are pre-prepared. Each word in the tweet is tokenized. Tokenization alludes to the demonstration of separating a string arrangement into pieces, for example, phrases, words, watchwords, images and different components called tokens which assumes a

significant job in expelling the undesirable words in the content, such as evacuating unique image related with username and hash tag in a tweet. Stop words being words that don't adjust the significance of a sentence is evacuated, which additionally limits the exertion of ordering every single expression of a tweet by lessening the number of words to be looked at.

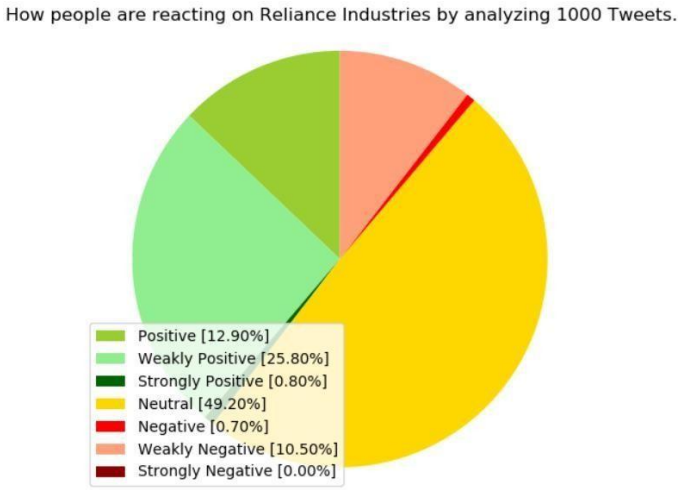


Chart-2: Sentiment Analysis Result of Reliance Industries analyzing 1000 tweets

4. Technical Analysis using Indicators

A Technical analysis using technical indicators is an art of mathematics & statistics of using historical stock data like price movements, volumes and market trends. Thus, anticipating future stocks trends in the form of graphs and charts helping the investors to predict what is more likely to happen to the prices over a shorter period of time so that he/she can take an informed decision. The study of technical analysis, mostly comprises of a person deciding an optimal time or moment to enter/invest in a particular security and exit to generate maximum returns or incur minimal losses. Such decisions depend upon the price functioning, the indications given by them and the critical turning points of the market. Technical analysis majorly consists varieties of technical indicators, which helps the investors to identify such movements and trends.

Technical Indicators are mathematical formulas applied on the parameters like open, high, low, close of a specific stock which taken from various financial institutions or firms and then it is plotted over in form of a graph. There are various types of technical indicators which are categorized by what they are used for.

The four major types of technical indicator are: 1) Trend 2) Volatility 3) Momentum 4) Volume

1) Trend: These types of technical indicators help to determine the direction and the strength of the stock. Usually the plotting value of the price is levelled out and represented by a single line. Because of such process the indicator lacks the abrupt price change thus called as trends. One of the major disadvantages of this is such indicators loses the money when the market is unsteady. One of the most important Trend indicators is: 1) MACD

a) MACD Stands for moving average convergence divergence. It is a trend-momentum indicator which helps the user to identify the oversold and overbought condition as well as the bullish-bearish moments of the stock.

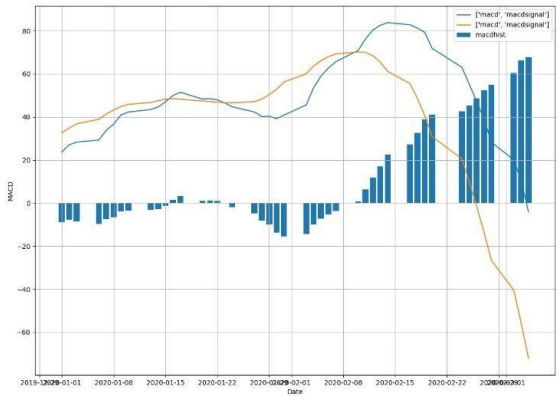


Chart-3: Macd for Bharti Airtel

Formula for MACD
MACD = 12 - Period EMA – 26 - Period EMA

To read from the MACD graph is very simple. MACD sets off technical indications when it crosses above it means to buy and to sell if they crosses below them. The frequency of the crossovers also determines the oversold-overbought instances.

2) Volatility Indicators - Volatility indicators available to the analysts helps them in determining the specific entry-exit points for each trade. They are used to find out rate of movement disregarding its direction. One of the prime volatility indicators is Bollinger Bands.

a) Bollinger Bands - Bollinger Bands are analytical tools developed by John Bollinger. It comprises of 3 lines namely a simple moving average and an upper and lower band. The upper and lower bands are simple positive and the negative difference of 20 days from the base line, can be modified based on the user preferences. It is believed that when prices move towards the upper band indicates the overbought condition and when the prices tries to go towards lower bands shows oversold condition. A squeeze is term which is used in Bollinger. When the lower and the upper band comes closer tightening then moving average, it is known as squeeze giving the signal for possible profitable trading opportunities.
Contradicting, when the band widens further apart,

Followed by decreasing volatility means the major possibilities to leave the trade.

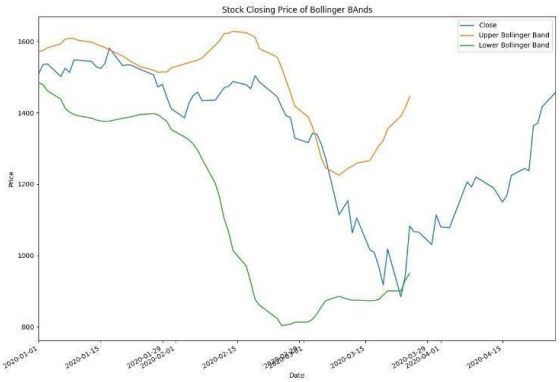


Chart-4: Bollinger for Reliance Industries

Formula for Bollinger Bands:
BOLU = MA (TP, n) + m ∗ σ [TP, n]
BOLD = MA (TP, n) – m ∗ σ [TP, n]

Where:
BOLU = Upper Bollinger Band BOLD = Lower Bollinger Band MA = Moving average
TP (typical price) = (High + Low + Close) ÷ 3
n = Number of days in smoothing period (typically 20)
m = Number of standard deviations (typically 2)
σ [TP, n] = Standard Deviation over last n periods of TP

3) Momentum: -Momentum indicators helps to determine the speed of fluctuations of the price by comparing prices over many different periods of time. It is calculated by comparing the current closing price with the past closing price. One of the trending momentum technical indicator is Ichimoku Kinko Hyo.

a) Ichimoku kinko Hyo - Ichimoku is used to calculate the momentum along with the areas for assistance and resistance levels. They consists of 5 lines namely, the tekan-sne, kijun- sen, senkou span A, span B and chikou span.

Tekan-sen: It is calculated by adding the high and the low of past nine periods and diving the result by 2.

Kijun-Sen: It is calculated by adding the high and low of past 26 period and dividing the result by 2.

Senkou span A: Also known as leading span Aiscalculated by averaging the Tenkan Sen and Kijun-Senline.

Senkou Span B: Also known as leading span B iscalculated by adding the highest and lowest of 52 periods and dividing by 2. Further by, plotting the outcome 26-periods ahead.

Chikou Span- Also known as lagging span. It is plotted by taking closing price of 26 days back on the chart. It is used to show probable sections of support and resistance.

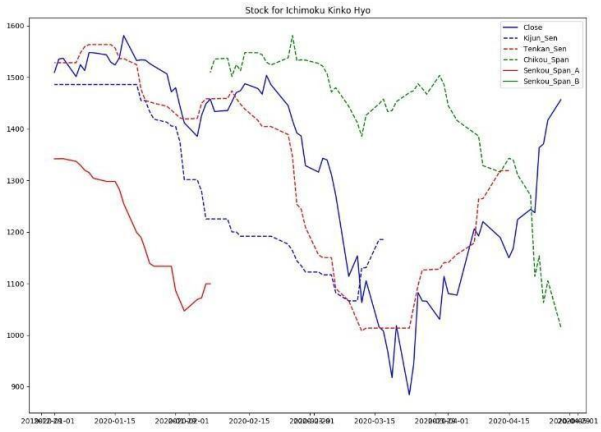


Chart-5: Ichimoku Kinko Hyo for Airtel

4)Volume: Volume is number of shares traded in a particular time frame. A time frame can range from 1 minute to 1 Year. So, volume plays an important role in deciding the direction or price for a particular stock. Volume based technical indicators majorly focuses on calculations performed on volume. One of the famous volume-based indicator is Chaikin Money Flow Index.

a) Chaikin Money flow index: -CMF is named after is developer and founder Marc Chaikin. It is based on the volume-weighted average over a specific period. The normal CMF period is 21 days. The main ideology behind the CMF is the closing price nearer to the high more accumulation has taken place. Conversely, the nearer the ending price is on the lower side means distribution has taken place.

Formula for CMF:
CMF = n-day Sum of [(((C - L) - (H - C)) / (H - L)) x Vol] / n- day Sum of Vol
Where:
n = number of periods, typically 21 H = high
L = low C = close
Vol = volume

5. CONCLUSION

The vogue of stock market trading is growing rapidly, which is encouraging researchers to find out new methods for the prediction using new techniques. The projection technique is not only helping the researchers but also helps investors and any person dealing with the stock market. In order to help predict the stock indices, a projection model with good accuracy is required. Right now, have utilized many gauging innovations which helps financial specialists, examiners or any individual keen on putting resources into the securities exchange by giving them a decent information on the future circumstance of the securities exchange.

REFERENCES

[1] T. Gao, Y. Chai and Y. Liu, "Applying long short-term memory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.

[2] Wei, "Prediction of Stock Price Based on LSTM Neural Network," 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), Dublin, Ireland, 2019, pp. 544-547.

[3] C, B. (2014). Technical Analysis in Select Stocks of Indian Companies. International Journal of Business and Administration Research Review, 2 (4).

[4] H. Irwin (2004). "The Profitability of Technical Analysis: Project Research Report No. 2004-04.

[5] M. S. Neethu and R. Rajasree, "Sentiment analysis in twitter using machine learning techniques," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, 2013, pp. 1-5.

[6] V. Prakruthi, D. Sindhu and D. S. Anupama Kumar, "Real Time Sentiment Analysis Of Twitter Posts," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, 2018, pp. 29-34.

[7] Investopedia site for Technical Analysis https://www.investopedia.com/technical-analysis-4689657