Machine Learning Technique for Stock Market Prediction Using Financial News Feeds, Tweets and Technical Indicators

[1] Sreelekshmi S Nair; [2] Dr.Radhakrishnan B

[1] CSE Department, BMCE
APJ Abdul Kalam Technological University, Sasthamcotta, Kerala, India

[2] CSE Department, BMCE
APJ Abdul Kalam Technological University Sasthamcotta, Kerala, India

Abstract - Stock Market Prediction is used to predict and identify the future value of a company share or stock on exchange. Prediction of stock market performance accurately is a challenging task and the successful prediction can yield profit by predicting the future price. The proposed prediction model includes different attributes as input. One is numerical input and another one is textual input. Numerical inputs include technical indicators and historic prices. Textual inputs include financial news feeds and Tweets. Here we aim to improve the stock market prediction using machine learning technique .Prediction model consist of Recurrent Neural Network (RNN) to combine the three aspects for enhancing the predictability of the daily stock market trends.

Keywords - *Recurrent Neural Network, ANN, Technical Indicators.*

1. Introduction

To grow or start a business, capital is the most important part. To attain this, the companies issue shares and those shares are sold to public. Stock Market is a place where shares of public listed companies are traded. When computer technologies were least used, the buying and selling of shares were done physically like auctions etc. But now everything can be done online. One can buy or sell shares sitting from home.

A stock exchange is a place where stock brokers and traders can buy, sell shares of stock, bonds and other securities. Many large companies have their stocks listed on a stock exchange. This makes the stock more liquid and thus more attractive to many investors. The two famous stock markets in India are NSE-National Stock Exchange and BSE-Bombay Stock Exchange. The companies which issue the shares, firstly have to register into these markets. From these registered companies, the public can buy or sell the corresponding shares.NSE consist of 1700 registered companies and BSE of 5400 registered companies. The up and down of stocks in BSE is called SENSEX and the up and down of stocks in NSE is called NIFTY .SEBI or Security Exchange Board of India have issued some guidelines and only based on these guidelines ,the investments and trading of stocks and shares takes place in these markets.

Stock Market Prediction is trying to understand or determine the future value of a company share or stock on exchange. The successful prediction will lead to yield profit by predicting the future price. Stock Market Prediction can be divided into three, Fundamental analysis and technical analysis and Technological Method.

Fundamental analysis involves analyzing the profit of the company on the basis of its current performance. What fundamental analysis in stock market is trying to achieve, is finding out the true value of a stock, which then can be compared with the value it is being traded with on stock markets and therefore finding out whether the stock on the market is undervalued or not.

In technical analysis, a technical indicator is a mathematical calculation based on historic price, volume, or (in the case of futures contracts) open interest information that aims to forecast financial market direction. Technical indicators are a fundamental part of technical analysis and are typically plotted as a chart pattern to try to predict the market trend. Indicators generally overlay on price chart data to indicate where the

IJCSN

price is going, or whether the price is in an "overbought" condition or an "oversold" condition. Many technical indicators have been developed and new variants continue to be developed by traders with the aim of getting better results. New Indicators are often back tested on historic price and volume data to see how effective they would have been to predict future events.

The Technological Method is machine learning, the most prominent technique involves the use of Artificial Neural Network (ANN) and Genetic Algorithms (GA). Machine learning is expected to bring heavy change in the world of technology. It helps to handle multi dimensional data which have various varieties of data types in a dynamic environment. It helps the user to utilize the resources more efficiently by using previous processed data. Predicting stock market will be useful to investors to invest to stock based on various factors. It will result to gain profit by predicting the future price.

2.   Literature Survey

Anthony R. Calingo et al [1] uses the general mood of the public and tweets from specific news sources is causative to the movement of the PSEi. It can be assumed that twitter sentiments have a predictive relation to the closing Index of the PSE. Using the Naïve Bayes algorithm in assessing the sentiment of the tweets have proven, based on the p-values, to have a predictive power over the stock market closing Index and it shows 73% accuracy.

Mustain Billah et al [2] proposed the improved Levenberg Marquardt (LM) algorithm of Artificial Neural Network (ANN). This algorithm has been applied for stock market closing price prediction and it shows 53% more accuracy in stock prediction than ANFIS. It also requires less memory allocation and computing time. This improved LM training algorithm proves neural network to be better computing tool for predicting closing stock price in Bangladesh Stock Exchange perspective.

L. Minh Dang et al [3] proposes a novel framework to predict the directions of stock prices by using both financial news and sentiment dictionary. This study includes the proposal of a novel two-stream Gated Recurrent Unit Network and Stock2Vec - a sentiment word embedding trained on financial news dataset and Harvard IV-4. The proposed Two-stream Gated Recurrent Unit (TGRU) overall accuracy was 66.32% which outperforms the performance of the previous model including GRU, LSTM, our model has two states of learning including backward and forward so it is able to

learn more useful information, especially for text processing problem.

Mehak Usmani et al [4] proposed Multi-Layer Perceptron algorithm (MLP)of machine learning   for stock market prediction. This method performed best as compared to other techniques. The prediction model uses different attributes as an input and predicts market as Positive & Negative. The proposed model gives accuracy about 77% predicts correct market performance. MLP seems to be more efficient in predicting the market performance. While verifying the model's variants on test data set, MLP outperforms the other algorithms.

Pisut Oncharoen and Peerapon Vateekul et al [5] proposed the stock market prediction using event vectors and numerical information as input to a deep neural network. This model consists of CNN and LSTM architectures and uses event embedding vectors extracted from news headlines, historical prices and technical indicators as input. Experiments on three datasets and employ accuracy and trading simulation as performance metrics. On average, the proposed model gives 69.86% better prediction accuracy than all baseline models.

3.   Design and Implementation

Objective of this model is to improve prediction performance by adding both numerical and textual information as input to a neural network. The data input used for this model is divided into three parts consists of historical prices, news headlines and Tweets. Data are split into training data and test data. We train the models using training data and then the performance of each model is evaluated using the test data. The historical Stock market price data's are extracted from Yahoo! Finance website. Financial news feeds data from different news sources namely Reuters, Reddit and Intrinio. The tweets were collected by using Twitter's API and Python. All scripts for extracting the tweets were written and executed using Python programming language. Python has readily available packages and libraries that can be easily accessed in order to perform several processes that need to make use of APIs over the Internet. All textual input data's are performing sentiment analysis for prediction. The NLTK tokenizer performed these tasks and broken down the text into tokens and each token is a word in order to have a more accurate classification process. Words are divided into positive and negative and each word is designated with a weight. Numerical Information uses historical prices and a set of technical indicators derived from price data. Using historic price data's, various technical

251

IJCSN
www.IJCSN.org

indicators are used like Stochastic oscillator (%K), William (%R), Relative Strength Index (RSI).

- Stochastic oscillator (%K): a momentum indicator which compares the closing price of a stock to its price range over a period. It can be used to foreshadow reversals when the indicator reveals bullish or bearish divergences.

$$\%K = 100 \times \left(\frac{C - L_p}{H_p - L_p}\right)$$

where C is the closing price of the day under consideration, Lp and Hp are lowest and highest prices in the last p days, respectively.

- William (%R): a momentum indicator which supports investors in detecting overbought and oversold conditions. It is based on a comparison between the current close and the highest high for a user-defined look back period p.

$$\%R = 100 \times \left(\frac{H_p - C}{H_p - L_p}\right)$$

where C is closing price of the day under consideration, and Lp and Hp are the lowest and highest prices in the last p days after the day under consideration, respectively.

- Relative Strength Index (RSI): a momentum oscillator which evaluates speed and trend of price movements. RSI fluctuates between 0 and 100. In practice, investors usually sell when RSI value ≥ 80 and buy when it is ≤ 20.

$$RSI = 100 - \left(\frac{100}{100 + RS}\right)$$

where RS is the average gain of positive periods during a specified time frame / average loss of negative periods during the specified time frame.

All inputs are given to recurrent neural network to combine three aspects and then use softmax as the activation function. The structure of the proposed model is shown in Figure.
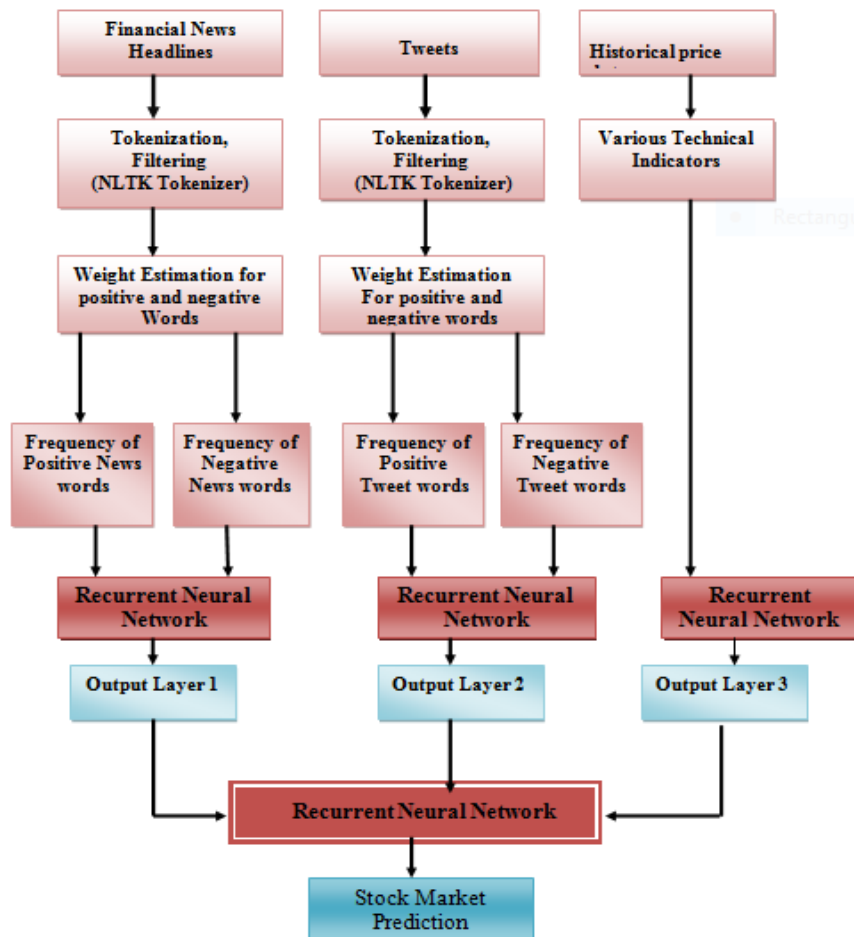

Fig 1 .Stock Market Prediction Model

3.1 Recurrent Neural Network (RNN)

A recurrent neural network is a class of artificial neural network where connections between nodes from a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior.RNN can use their internal state (memory) to process sequence of inputs.RNN were designed to work with sequence prediction problems. The RNN utilizes a supervised machine learning technique called back propagation for training.
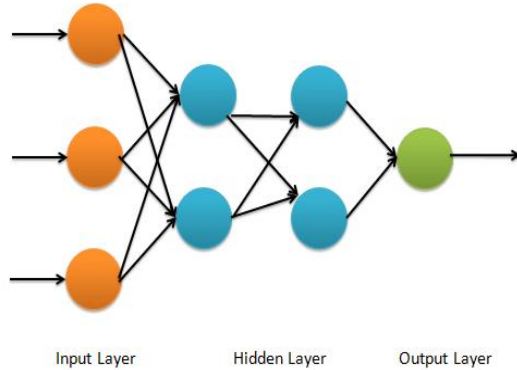

Fig 2.Recurrent Neural Network

4. Experimental Result

This experiment shows that the accuracy of stock market prediction can be improved by considering both numerical and textual information as input into a recurrent neural network. The experimental results of stock market prediction obtained are discussed .The table below shows the comparison of the outcomes obtained by using previous methods and the proposed method.

Table 1.Prediction Accuracy of Different Methods

Model	Accuracy
RNN	92%
TGRU	66.32%
LM	53%
CNN and LSTM	69.86%
NAIVE BAYES	73%
MLP	77%

5. Conclusion

In this paper, an improved method uses Recurrent Neural Network (RNN) for stock market prediction using Numerical and textual inputs. Historical prices and technical indicators are numerical inputs and Financial News headlines and public sentiments or mood from a large collection of Twitter data are textual inputs. The

successful prediction of the stock market will lead to yield profit by predicting the future trend of the market. The proposed system jointly combines news feeds, tweets and numerical values for prediction. On average, the proposed model gives better prediction accuracy than all baseline models and gives overall accuracy about 92% which outperforms the performance of the previous models and the result shows more accurate output.

Reference

[1] Anthony R. Caliñgo , Ariel M. Sison, Bartolome T. Tanguilig III.Prediction Model of the Stock Market Index Using Twitter Sentiment Analysis. I.J. Information Technology and Computer Science, 2016, 10, 11-21.

[2] Mustain Billah, Sajjad Waheed, Abu Hanifa.Stock Market Prediction Using an Improved Training Algorithm of Neural Network. 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)

[3] L. Minh Dang, Abolghasem Sadeghi-Niaraki, Huy D. Huynh, Kyungbok Min And Hyeonjoon Moon.Deep Learning Approach For Short-Term Stock Trends Prediction Based On Two-Stream Gated Recurrent Unit Network. Ieee Access.

[4] Mehak Usmani, Syed Hasan Adil, Kamran Raza,Syed Saad Azhar Ali.Stock Market Prediction Using Machine Learning Techniques. 2016 3rd International Conference On Computer And Information Sciences (ICCOINS).

[5] Pisut Oncharoen and Peerapon Vateekul.Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators.IEEE .

[6] Manuel R. Vargas, Carlos E. M. dos Anjos, Gustavo L. G. Bichara, Alexandre G. Evsukoff. Deep Learning for Stock Market Prediction Using Technical Indicators and Financial News Articles.IEEE Access.

[7] Mingyuan WU, Xiaotian DIAO. Technical Analysis of Three Stock Oscillators. 2015 4th International Conference on Computer Science and Network Technology (ICCSNT 2015).

Author Profile

Sreelekshmi S Nair received her B.Tech (CSE) degree from University of Kerala in 2017. She is currently pursuing her Masters in Computer Science & Engineering from KTU.

Dr.Radhakrishnan B is working as the Head of CSE Department. He has more than 14 years experience in teaching and has published papers on data mining and image processing. His research interest includes image processing, data mining and image mining.

IJCSN
www.IJCSN.org