



# Stock Market Prediction Using Machine Learning

<sup>1</sup>Sanket Gajare, <sup>2</sup>Shraddha Garje, <sup>3</sup>Arati Lokhande, <sup>4</sup>Tejas Jamdar,

<sup>5</sup>Prof. Priyanka Kumbhar

<sup>1,2,3,4</sup>Student, <sup>5</sup>Faculty

<sup>1</sup>Information Technology,

<sup>1</sup>PG Moze College of Engineering Wagholi, Pune, India

## ABSTRACT

*The familiar problem of interest is predicting stock market analysis. Stock market is the integral part of country's economy. The aim of the stock market prediction is to predict the future financial stocks of the company. Sentimental analysis and technical analysis are the two approach used in predicting the stock price. Social media plays the role of representing the public sentiments about a current event. The approach through sentimental analysis helps in observing how well the changes occur in stock prices. The rise and fall of the stock prices is correlated with the opinion of people expressed on Twitter. With the help of technical analysis the stock price movement acts as a function of input feature. It helps on solving the regression problem in multiple kernel learning regression framework. The machine leaning is coupled with fundamengtal analysis. It evaluates the models for taking the buy-sell decision. The paper focuses on the use of Regression ,LSTM and naive bay's algorithm to predict stock values.*

## I. INTRODUCTION

Background :

Sentiment analysis is a use of natural language processing, text analysis and identify, extract, quantify, and study affective states and subjective information. This helps the business in understanding the social sentiments of their brand, product or service. There is Major correlation between individual reactions and stock returns. The valuable data in domain of stock market include several features like time, targeted audience and the brand. One of the most important features for the decision makers that are looking for investing in the stock market are time and brand.

Relevance :

Stock market is free market economy. It permits the companies to raise money by offering stock shares and allows investors to participates in the financial achievements. The company makes profit through capital gains and earn income through dividends. Stock price prediction gain insight about market behavior over time, spotting trends. The public stock dataset is too limited for any machine learning algorithm alone to work with. In a Project we perform sentimental analysis of twitter data to gain insight. To predict the stock price, the stock dataset is fetched from yahoo finance API. For this study secondary data has been collected. From

the website of KSE the monthly stock prices for the sample firms are obtained from Jan 2010 to Dec 2014. And from the website of SBP the data for the macroeconomic variables are collected for the period of five years. The time series monthly data is collected on stock prices for sample firms and relative macroeconomic variables for the period of 5 years. The data collection period is ranging from January 2010 to Dec 2014. Monthly prices of KSE -100 Index is taken from yahoo finance.

## II. LITERATURE REVIEW :

Several researchers have studied on various types of techniques for Concept of sentimental analysis is described and a model is proposed along with some ML methods. Sayavong Lounnapha[2], proposes a stock price prediction model based on convolution neural network, which has obvious self-adaptability and self-learning ability. Combining the characteristics of CNN (Convolution Neural Network) and Thai stock market, the data set is trained and tested after pre-treatment. On this basis, three stocks (BBL, CAPLL&PTT) listed on the Thai Stock Exchange are tested and compared with the actual stock price. The results show that the model based on CNN can effectively identify the changing trend of stock price and predict it which can provide valuable reference for stock price forecast. The prediction accuracy is high, and it is worth further promotion in the financial field. In 2016, researchers used genetic algorithm and simulated annealing to find initial weights of an ANN, and then back-propagation algorithm is used to train the network [7]. This hybrid approach outperformed the standard ANN based methods in prediction of Nikkei 225 index return. With slight modifications [8], genetic algorithm was successfully utilized to find optimized weights of an ANN in which technical indicators were utilized to predict the direction of Nikkei 225 index movement. Deep ANNs, that are basically neural networks with more than one hidden layer, among the first deep methods used in the domain. In 2016, some researchers [9] predicted NASDAQ prices based on the historical price of four and nine days ago. ANNs with different structures, including both deep and shallow ones, were examined in order to find appropriate number of hidden layers and neurons inside them. The experiments proved the superiority of deep ANNs over shallow ones. Ze Zhang[3] propose method to adopts self-adapting variant PSO algorithm to optimize the weights and thresholds of network. Afterwards, the optimized data, regarded as initial weight and threshold value, is given to Elman network for training, accordingly the prediction model for opening price of stock market based on self-adapting variant PSO- Elman network is formed. Finally, they verifies that model by some stock prices, and compares with BP network and Elman network, so as to draw the result that shows the precision and stability of this predication model both are superior to the traditional neural network. Heng Liu[4]introduce a deep residual network (ResNet) model for prediction, using the stock price graph as input. The results show that the ResNet model has the average accuracy of 0.40, which is higher than the stochastic indicator of 0.33. Recently more and more researchers have applied deep learning models to stock prediction problems, and many studies have shown that deep learning models have strong learning ability and prediction accuracy [11-13]. In 2015, He et al. [6] proposed a special convolutional neural network called deep residual neural network, which can increase the number of network layers to hundreds or even thousands easily, while training time is within the acceptable range, greatly improving the accuracy of image recognition.

## III. PROPOSED SYSTEM

### IV. METHODOLOGY

Natural Language Processing:

NLP, ability of computer program to understand human language. Just as humans have different sensors computers have programs to read. It is frequently trained on historical data using machine learning models to predict future stocks. Large amount of data is processed with the help of AI. Hence, Natural language processing (NLP) enables to analyze financial documents. It helps in the way to program computers to process and analyze large amount of data.

Naïve Bayes Algorithm:

Naive Bayes Algorithm is a supervised learning algorithm. It is used for solving classification problems. It is used for solving multipleclass prediction problems. The probability of an event is the key insight of bayes theorem. Ever knowledge attribute is independent of each other.

Time Series Prediction:

Predicting the future events by analyzing the past trends is time series prediction. The involvement of historical data in predicting the future values and make observation to drive future strategic decision making. This helps in making the trend chart and provides guidance for the investors. It gives a idea of price pattern movement.

Auto Regression:

A time series model that makes use of the observation from the previous steps as a input to regression equation to predict the value of the upcoming steps. It ends in accurate result and a useful way for predicting the future price of company stock.

Arima From Stats Model:

Arima stands for “Auto Regressive Integrated Moving Average”. It consists of measuring of events that happen over the period of time. It understands the past data and presicts the future state by checking the differences between values in the time series instead of the actual stock value.

Forward Moving Object:

It can be called as gthe gtechnical indicator. It is used in finding momentum of the trnding stock. Fixed time frame is calculated by adding next successive things to the perviously calculated average.

## VI. IMPLEMENTATION TECHNIQUES

1.Raw data preprocessing component:

It takes the raw data like content of news and comments, historical price and volume data as for further process. Here the cleaning helps in ccreating sentimental analysis.

2.Features Extraction Component:

Every single word is broken down during the fetching of tweets and comaprison is done with bag of words. The assigning of the labels is from those respective bag of words.

3.Sentiment Analysis Component:

The whole sentiment is the analysis with the threshold value for a sentiment to be classified. The process is of the actual percentage of positive and negative sentiments from the fetched data.

4.Technical Analysis Component:

The generation of signals of movements in market and combining the outputs of every signal using the simple combination of neural network combination.

## VII. PROBLEM STATEMENT

We proposed a method for stock prediction based on a set of technical trading rules and sentimental analysis, which are optimized by a optimization algorithm. The aim of the reserch is to check whether its possivble to obtain a set for trading rules using the sentiments of people in the market which can be used in decision making process such as Buy, Wait or Sell.

## VIII. CONCLUSION

ML has the potential to revolutionize stock market prediction by providing accurate and reliable forecasts. ML algorithms can analyze vast amounts of data quickly and accurately, identify patterns and trends that may not be apparent to human analysis and reduce the risk of human error and bias. However, there are also several challenges that need to be addressed, such as obtaining large amounts of data, identifying relevant variables for prediction, and adressing overfitting. Despite these challenges, uses of ML in stock marker prediction is a promising area of research that could provide significant benefits for investors, traders and financial institutions.

## IX. REFERENCES

- Pahwa, K., & Agarwal, N. (2019). Stock Market Analysis using Supervised Machine Learning. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing.
- Akba, F., Medeni, I. T., Guzel, M. S., & Askerzade, I. (2020). Assessment of Iterative Semi-Supervised Feature Selection Learning for Sentiment Analyses: Digital Currency Markets. 2020 IEEE 14th International Conference on Semantic Computing.
- Gupta, R., & Chen, M. (2020). Sentiment Analysis for Stock Price Prediction. 2020 IEEE Conference on Multimedia Information Processing and Retrieval.

