Factors affecting in Stock Market Prediction by Machine Learning

Umakant Singh[1], Ankur Khare[2]

1Research Scholar,Computer Science, Rabinadranath Tagore University,Raisen, Madhya Pradesh, 464993

2Assistant Professor,Computer Science & Information Technology, Rabinadranath Tagore University,Raisen, Madhya Pradesh, 464993

ABSTRACT

Forecasting is a challenge to predict how a future result will occur. The main aim of forecasting the occurrence of this result is for decision makers to make better decision. Stock markets areas typically generate large volumes of data through trading, transactions or operations. Stock market helps traders to gain an insight into the economy, stock analysis, or securities. It involves studying the historical and current market data and developing methodologies for selecting the right stocks to trade. In today's financial world, the stock market has become one of the most important events. People make decisions to invest in the stock market based on previous research knowledge and predictions. Stock forecasting plays an important role in the increasingly difficult stock market business as people often look for tools and methods to minimize risk and maximize profits when it comes to forecasting. Employing traditional techniques such as fundamental and technical analysis does not seem to guarantee the consistency and accuracy of forecasts. As a result, machine learning techniques have become the latest trend in stock market forecasting, which is ultimately based on existing stock market values results of training on previous values. In this paper we study the factor who affects the stock market.

Keywords: Stock prediction, Machine learning, EPS, SMA, RSI

INTRODUCTION

About Stock Market-

Shares or stocks, also called shares of a company. Stock ownership means that the shareholder is the owner of a part of the company which is equal to the number of shares that is held as a fraction of the company's total outstanding shares. For example, an individual who is the owner of a hundred thousand shares of a company with a million outstanding shares would be having a ten percent stake ownership in it.

Stock exchanges are nothing but the secondary markets wherein the present owners of the shares could transact with the potential buyers.

The stock market is known for its volatility, randomness, and unpredictability. The share market is based on the concept of demand and supply. If the demand for a particular company's stock is higher and the supply is low then that company's share price would tend to increase and if the demand for company's share is low then the company share value tends to decrease. The successful prediction of a stock's price by its analysis could lead to a significant profit. We use very large historical data sets to represent a variety of conditions,

and we believe that time series patterns have great predictive power and produce high probabilities of profitable trades and high returns on corporate investment [14].

About Machine Learning-

Machine learning is a sub-part of artificial intelligence. Machine learning allows applications to learn from experience in the same way that humans do. As these applications receive data, they learn to grow and change as they gain experience. This is done by using algorithms that learn from data in an iterative process. Applications use pattern recognition to process a variety of data. Machine learning is the ability of applications to react to new data using iterations A machine learning algorithm learns how to predict an output based on previous examples of the relationship between input data and output, called training data. A model of the input-output relationship is improved over time by testing its predictions and correcting them if they are wrong. Machine learning is a set of computerized techniques for recognizing patterns in data. It is a way of generating something like the line of best fit. Automating this process can be useful if your data is rich and complex [15].

Types Of Machine Learning-

Supervised Learning:

Supervised learning requires a training data set containing either labeled data or data with known output values (such as rural/non-rural or house prices). Classification and regression problems are solved by supervised learning.

Unsupervised Learning:

Unsupervised learning techniques find patterns and structures in the data itself without using a training set. The clustering problem can be solved with an unsupervised approach.

Semi-Supervised Learning:

Semi-supervised learning uses mostly unlabeled and small amounts of labeled input data. Using small amounts of labeled data greatly improves the efficiency of unsupervised learning tasks. A model needs to learn structure in order to organize data and make predictions.

Reinforcement Learning:

Reinforcement learning uses input data from the environment as stimuli for how the model responds. Feedback is generated as a reward or punishment in the environment rather than through a training process like supervised learning. This type of method is used in robotic control.

How does machine learning work?

The machine learning process begins by feeding data into an algorithm. This is also called training the algorithm. To test whether the algorithm works, data is fed into the algorithm and the results are checked. If the desired results are not obtained, the algorithm is retrained and tested again. This process is repeated until the desired result is achieved. This allows the machine learning algorithms to learn and provide the desired output and improve the accuracy of the results [15].

The objective of this paper to determined the impact of fundamental factors (EPS, NAVPS, and P/E) and technical factors (GDP, CPI & IRS) on market prices. This paper also aim to creating awareness among the present and potential investors to consider these factors prior and avoid

unexpected loss.

LITERATURE SURVEY

Dou Wei et al.[1] has analysed and compared a variety of neural networks prediction method and finally chosen LSTM((Long Short-Term Memory).The data used for prediction is taken from Oriental fortune website by using crawler and stored in database after performing cleaning operations. The paper uses attention-based LSTM(long short-term memory) and consists of four parts : input Layer, hidden layer, attention layer and output layer. The inputs to the models were trading date, opening price, closing price, lowest price, highest price and daily volume of the stock .For standardization of input data following formula is used:
Standardized input data = (original data - minimum)/(maximum - minimum)

Dinesh Bhuriya et al.[2] used various regression techniques to predict the stock price of TCS .They have mainly implemented linear, polynomial and RBF regression model out of which linear regression model provides best result based on confidence value. They have used Open price,High price, Low price and Number of trend as input variable and Close price as output.

Omveer Singh Deora et al.[3] used two different machine learning models for stock market forecasting. LSTM is used for historical data and CNN for current data. For both these models their respective data sets are split in 80:20 ratios, where 80% is used for training and the remaining 20% is used for testing. The purpose of applying both these models is that LSTM models are better skilled to find any relations in the data set and use them to predict future prices whereas CNN primarily emphasize on current data for its prediction not requiring any information on historical data.

Dr. Devpriya Soni et al.[4] proposed a model that checks accuracy at each stage when analyzing stock trends. This model is simplistic and applies only to historical data. The accuracy provided by the authors is 96.60% which is a very high number. We can use this model and combine it with sentiment analysis for a more thorough and efficient prediction.
The proposed model:
Divide the dataset into two parts-

- Training Set
- Testing Set

1. In the training set:
a. Calculate the change in share price over a period of time for every day and take mean
b. Share price can increase or decrease
c. Prepare a matrix z for storing zeroes and ones having number of rows equal to equal to the size of training set and columns equal to number of factors included for the share
d. If share price increases assign 1 else 0
e. Count number of 0s and 1s-if number of 1s>0s then the share is a good share else a bad share
f. Create a prediction matrix y
g. If the stock shows positive growth then the value of the share price is increased by the change we calculated earlier and the first row of the newly created prediction matrix is marked as 1 otherwise if the stock shows negative growth then the value is decreased and correspondingly the first row of the prediction matrix is marked as 0.

2. **In the testing set:**
a. Prepare a similar matrix x by storing 1s and 0s on the basis of actual rise and fall of the share price
b. Now check first day's prediction by comparing values in x and y matrix
c. If the match is found, prediction is right and we further proceed for the next day accordingly by what we did for the previous one(i.e. if share price is increased in the previous prediction then for the next prediction as well ,we do the same and vice versa).
d. If the values of x and y are not matching then we correct the error by a factor of twice the change we calculated earlier of the share price value for the next prediction.
e. So for example if the share price falls and we predicted a rise so for the next prediction we now alter the predicted value of the previous day by a factor of double the change and reduce its(predicted price) value accordingly.
f. Also the next field of y matrix is marked as opposite to the previous one. In this way, the predicted price is calculated and subsequently the accuracy of prediction by comparing it to actual values.

Aparna Anant Bhat et al.[5] proposed a system consisting of modules such as data collection, technical analysis, neural network prediction and sentiment analysis. In data collection, they gathered data of historical prices and volumes of the stocks from NSE India and google finance. This data was used for technical analysis components and neural networks. News and blog articles were also collected with the date of publication for sentiment analysis. But the predictions for intra-day trading were not available everyday at times and the frequency of opinions, news articles and comments varied for different working days.
In technical analysis they implemented the technical indicators such as MACD,EMA and RSI using the historical data collected. They made predictions based on these technical indicators. For prediction using neural networks, the authors tested the neural networks model with different sets of input and hidden layer neurons and different activation functions. They used activation functions like sigmoid and hyperbolic tangent functions for triggering the neurons. They used back propagation to further optimize the prediction values. The training data consisted of closing prices and volumes of each day. In sentiment analysis, the authors tried to check the positive and negative sentiments from the reviews /news/ blog articles to further improve the accuracy.
In this research, the authors concluded that though combining Technical Analysis with technological methods like Neural Network gave better performance, addition of Sentiment Analysis further increased the results as stock prices very much dependent on day to day activities and emotions in the market and the fundamental analysis.

Meghna Misra et al.[6] discussed using various machine learning algorithms to predict the stock market.They had compared the Linear Regression, Support Vector Machine and Random Forest algorithms for getting higher accuracy and least error. The analysis shows that Support Vector Machine has high accuracy on non linear classification data whereas Linear Regression is the preferred algorithm if the available model is that of regression and the Random Forest shows high accuracy on binary classification model and multilayer perceptron offers the least error in prediction of the stock trend.

FUNDAMENTAL AND TECHNICAL ANALYSIS

Due to the fluctuations in stock prices, every investor wants to be aware of the future price trends of a company's stock in order to make educated investment decisions. Fundamental analysis and technical analysis are used to investigate the future price trend of the stock for this goal.

Fundamental Analysis

In this approach, a company's fundamental value is calculated based on its many assets and product value. Fundamental analysis looks at all factors that might influence a company's stock price in the future, including financial statements, management practices, industry, etc. The intrinsic value of the firm is looked at to ascertain if the stock is overvalued or underpriced. The worth of the corporation can be influenced by a variety of things.

(i)Stock Trading Ratios: This ratios is used to understand the earning of the company with respect to its competitors. The higher the earning, the value of the company is also the higher. This can be further subdivided into EPS and P/E ration.

Earnings Per Share (EPS)- This is the amount earned by a shareholder per share.

Price/Earnings (P/E ratio)- This is the ratio between the company's stock price to its earnings.

(ii)Market Size: The size of the market and the company's share are essential for valuation. If a company provides a vital product and it is the only one that offers that product, then it is likely that its value is very high.

(iii)Demand: The demand for the product manufactured by a company can have an impact on its stock price. The demand can be dependent on time and market.

(iv)Management: The Company's management team plays a vital role in the direction of the company's future. A team of proven individuals has more impact on stock prices than a team of unproven individuals.

(v)Economic Climate: This is a crucial factor in the stock prices of a company, if the economy is booming, then companies perform well, and hence their stock prices are high. A prime example is the financial crisis in 2008 during which a lot of the companies including banks lost their value. Long-term investors typically use fundamental analysis. Most of the factors in the fundamental analysis do not change day to day, hence they are more useful in a buy and hold approach.

Technical Analysis

Technical analysis uses past charts, patterns and trends to forecast the price movements of the entity in the coming time. Technical Analysis in stock trading uses charts and various mathematical formulas to identify trends to predict the stock price movements. In this approach, the trader looks at the price alone without paying attention to the underlying fundamentals.

The 4 Basic Principles of Technical Analysis

(i)Markets alternate between range expansion and range contraction-

This theory is based on the idea that prices trade in ranges and it is important to determine what the current range looks like and what the next range looks like. **(ii)Trend continuation is more likely than reversal-**

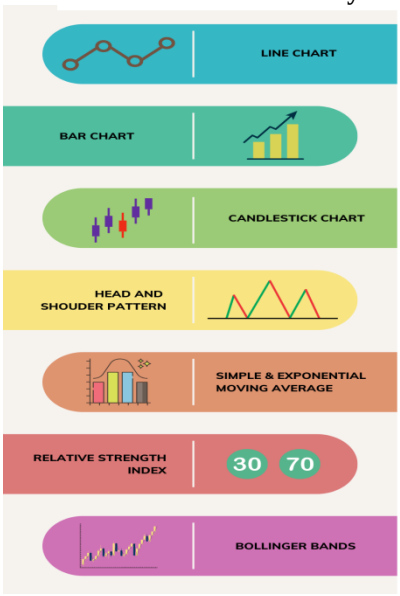
If the market is trending in a particular direction, then your best bet is to assume that direction will continue for the time being. Of course, it's important to also identify when the trend might be about to reverse.

(iii) Trends end in one of two ways: climax or reversal.
A trend which ends in a buying climax is one where the last buyers are willing to pay practically any price. Once that last buyer buys, a vacuum is created on the other side and the market collapses. A trend reversal, on the other hand, occurs when buying or selling simply stops losing momentum.

(iv)Momentum precedes price-

This rule means that when the market makes a sudden move ("momentum" or "momentum"), the price is likely to continue moving in that direction. Technical analysis mostly uses bar charts, candlestick charts, line charts, and point and figure charts. Technical analysts use these chart patterns to identify trading signals .

Some Common Technical analysis Terms and Indicators-



Line chart – A single line which connect stock prices is called a line chart.

Bar chart – It is also referred to as an open-high-low-close chart.
A chart which has open, high, and low and closing data sets in a vertical line in the form of a bar.
Candlestick charts- It show four price points (open, close, high, and low) throughout the period of time the trader specifies. The advantage of candlestick charts is the ability to highlight trend weakness and reversal signals that may not be apparent on a normal bar chart. Candlestick charts serve as a basis of technical analysis.
Head and Shoulder Pattern –
The head and shoulders pattern is considered one of the most reliable trend reversal patterns. A head and shoulders pattern is used in technical analysis. It is a specific chart formation that predicts a bullish-to-bearish trend reversal. The pattern appears as a baseline with three peaks, where the outside two are close in height, and the middle is highest. The head and shoulders pattern forms when a stock's price rises to a peak and then declines back to the base of the prior up-move. Then, the price rises above the previous peak to form the "head" and then declines back to the original base. Finally, the stock price peaks again at about the level of the first peak of the formation before falling back down.

Simple moving average (SMA) and exponential moving average (EMA)-These are two types of moving averages that can help you figure future market patterns.
You can calculate simple moving average by adding the closing price of a company stock and then dividing it by the number of days. So, if you want to calculate a 50-day SMA, you will have to add the closing price of the stock for the last 50 days and then divide it by 50. You can similarly calculate a 100-day or a 200-day SMA.
EMA, on the other hand, gives more significance to recent trading data. This is why EMA is better at spotting price changes vis-a-vis SMA. Therefore, it is the moving average of choice for most traders.
The Relative Strength Index (RSI) –
It is a momentum indicator that measures the speed and change of price movements. The RSI oscillates between zero and 100. Traditionally the RSI is considered overbought when above 70 and oversold when below 30. Signals can be generated by looking for divergences and failure swings. RSI can also be used to identify the general trend.
Bollinger Bands –
This plot trading bands above and below a simple moving average. The standard deviation of closing prices for a period equal to the moving average employed is used to determine the band width. This causes the bands to tighten in quiet markets and loosen in volatile markets. The bands can be used to determine overbought and oversold levels, locate reversal areas, project targets for market moves, and determine appropriate stop levels.
Comparison Chart

Comparison Parameter	**Fundamental Analysis**	**Technical Analysis**
Machine learning	Fundamental Analysis is a practice of analyzing securities by determining the intrinsic value of the stock.	Technical analysis is a method of determining the future price of the stock using charts to identify the patterns and trends.
Relevant for	Long Term Investments	Short Term Investments
Function	Investing	Trading
Objective	To identify the intrinsic	To identify the right time to

	value of the stock.	enter or exit the market.
Decision making	Decisions are based on the information available and statistic evaluated.	Decisions are based on market trends and prices of stock.
Focuses on	Both Past and Present data.	Past data only.
Form of data	Economic reports, news events and industry statistics.	Chart Analysis
Future prices	Predicted on the basis of past and present performance and profitability of the company.	Predicted on the basis of charts and indicators.
Type of trader	Long term position trader.	Swing trader and short term day trader.

Other Factors

Political Factors

There have been multiple political factors affecting stock markets. For instance, the price of stocks goes down in case of risk of war, weak government, public outrage against the government, etc. Budget announcements or elections significantly impact the volatility of the market, affecting the stock prices. Moreover, the new government policies introduced regarding the Indian economy can affect the share market. In addition, the value of stocks is also reduced in case of riots or political turmoil in the country.

In addition, there is also a possibility that stock market(s) and traders respond in a biased manner for right-winged government, as compared to the left-winged [16].

Interest Rates

The governing body, the Reserve Bank of India (RBI), regulates interest rates, directly influencing the price of stocks. When the interest rate is low, the companies can borrow a considerable amount at a lower interest, resulting in their profits due to an increase in the stock price. On the other hand, higher interest rates lead to lesser profits and reduced stock prices [16].

Natural Calamities

Natural calamities and pandemics such as floods, earthquakes, and pandemics such as Yellow Fever, Ebola and the recent COVID-19 one too, can drastically affect the value of stocks. Due to the stock prices are bound to fall due to the destruction of property, finances, and other assets. It affects not only a company's performance but also people's capability to spend [16].

Inflation

Inflation directly affects the finances of people resulting in reduced capacity to invest. Moreover, increased inflation rates discourage people from investing, making companies suffer. Hence, inflation has a critical role in affecting one's investing power, purchasing power, and the country's overall economy[16].

Supply and Demand

One of the main factors affecting the share market is the imbalance between supply and demand which leads to the increase or decrease in the price of stocks. In addition, factors such as economic data and interest rates affect the demand for stocks leading to fluctuations in the value of stocks [16].

CONCLUSION

Forecasting of stock prices has been a challenging task for many researchers and analysts. The profits can be maximized if the market value is predicted keeping the risk low. The popularity of stock market trading is growing rapidly, which is encouraging researchers to find out new methods for the prediction using new techniques. The forecasting technique is not only helping the researchers but it also helps investors and any person dealing with the stock market. This paper introduces the concept of stock market and machine learning. We discuss about fundamental and technical factor that affect the stock price. The purpose of this is to paper to study the affecting factor of stock market which is used in stock market prediction.

REFERENCES

[1] Dou Wei, "Prediction of Stock Price Based on LSTM Neural Network" in International Conference on Artificial Intelligence and Advanced Manufacturing(AIAM), 2019.

[2] Dinesh Bhuriya, Girish Kaushal, Ashish Sharma, Upendra Singh, "Stock Market Prediction Using A Linear Regression" in International Conference on Electronics, Communication and Aerospace Technology ICECA, 2017.

[3] Omveer Singh Deora, Pawan Jha, S.T. Sawant Patil, T.B. Patil, S. D. Joshi**,** "Monitoring and Training Stock Prediction System For Historical & Live Dataset using Lstm & nn," (IJITEE)(2019)

[4] Dr. Devpriya Soni, Sparsh Agarwal, Tushar Agarwal , Pooshan Arora, Kopal Gupta, "Optimised Prediction Model For Stock Market Trend Analysis,"(I 3)(2018)

[5] Aparna Anant Bhat, Sowmya Kamath S, "Automated Stock Prediction and Trading Framework for Nifty Intraday Trading" in 4th ICCCNT - 2013 July 4-6, 2013, Tiruchengode, India.

[6]Meghna Misra, Ajay Prakash Yadav, Harkiran Kaur, "Stock Market Prediction using Machine Learning Algorithms: A Classification Study," (ICRIEECE) IEEE - (2018)

[7]Siyuan Liu, Guangzhong Liao, Yifan Ding, "Stock Transaction Prediction Modeling and Analysis Based on LSTM," - IEEE (2018).

[8] Juan Ricardo Rivera Peruyero , Pere Marti-Puig, "Webbased system for evaluating day trading strategies," 2011 7th International Conference on Next Generation Web Services Practices 253

[9] F. A. Gers and J. Schmidhuber (2000), "Recurrent nets that time and count", in Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, IEEE : 189-194.

[10] Hiransha Ma , Gopalakrishnan E.Ab , Vijay Krishna Menonab, Soman K.P(2018), "NSE Stock Market Prediction Using DeepLearning Models", Procedia Computer Science 132 : 1351–1362

[11] T. Kim and H. Y. Kim (2019), "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," PloS one, vol. 14(2) : e0212320.

[12] Alamir Labib Awad , Saleh Mesbah Elkafas , Mohammed Waleed Fakhr (2021)"ROLE OF MACHINE LEARNING IN PREDICTING STOCK PRICES: A LITERATURE SURVEY" Journal of Management Information and Decision Sciences Volume 24,Special Issue 1, 2021

[13] Mahinda Mailagaha Kumbure , Christoph Lohrmann , Pasi Luukka , Jari Porras (2022) "Machine learning techniques and data for stock market forecasting: A literature review"https://www.researchgate.net/publication/358731897

[14] Sathya S, Dr. N. Radha (2022)"STOCK MARKET PREDICTION USING MACHINE LEARNING" IJCRT, ISSN: 2320-2882.

[15] Vedhaant Jain, Aniket Kulkarni (2020)"Survey on Various Algorithms of Machine Learning and its Applications"International Research Journal of Engineering and Technology (IRJET)Volume: 07 Issue: 10 e-ISSN: 2395-0056

[16]https://www.kotaklife.com/insurance-guide/wealth-creation/factors-affecting-stock-markets