A Systematic Analysis And Review Of Stock Market Prediction Techniques Using Hybrid Approach

Bhagyashree Pathak, Dr. Manoj Kumar Rawat
Phd Research Scholar, Professor
Computer Science Department, Sage University, Indore, India

Abstract: The stock market is very volatile in nature. This topic itself has lots of challenges and interests too. So many researchers try to predict the stock market values. This paper has different research approaches to predict the trend in the stock market using various methods and algorithms. In technical and fundamental analysis techniques and sentimental analysis, they use data mining techniques to identify the pattern from a massive amount of market data. Neural networks, artificial intelligence, machine learning, deep learning, and many more approaches were used for the prediction of trends in the stock market.

Keywords: Stock market, technical analysis, fundamental analysis, ANN, machine learning, data mining, trend, prediction.

I. Introduction:

Prediction of trends in the stock market is an area of huge interest, and we usually find common people discussing the ups and downs in the stock market. The stock market is many a time used like gambling, especially by retailers. Huge profits in a short duration of time attract people putting money in stocks. But the case is just the opposite. 90% of people lose money in stock markets just to benefit the rest bigger players which are organized in their approach, unlike retail investors. The stock market is a very volatile and high-risk domain in nature, but everyone directly or indirectly is connected to it. It has its advantages and disadvantages. Everyone wants to earn profit from the stock market because it can be a good source of secondary income, but it has risk involved which needs to be calculated first. Due to the above reasons, researchers have shown interest in forecasting the market trend by using different algorithms. Researchers proposed different models and techniques to identify the trend or to predict the market flow involving one or both fundamental and technical analysis techniques. They use data mining techniques to identify a pattern from a huge amount of market data. Neural networks, artificial intelligence machine learning, deep learning, and many more approaches were used for the prediction of trends in the stock market. Prediction of trends in the stock market is an area of huge interest, and we usually find common people discussing the ups and downs in the stock market. The stock market is many a time used like gambling, especially by retailers. Huge profits in a short duration of time attract people putting money in stocks. But the case is just the opposite. Generally speaking, 90% of people lose money in stock markets just to benefit the rest bigger players which are organized in their approach, unlike retail investors. The stock market is a very volatile and high-risk domain in nature, but everyone directly or indirectly is connected to it. It has its advantages and disadvantages. Everyone wants to earn profit from the stock market because it can be a good source of secondary income, but it has risk involved which needs to be calculated first. Due to the above reasons, researchers have shown interest in forecasting the market trend by using different algorithms. Researchers proposed different models and techniques to identify the trend or to predict the market flow involving one or both fundamental and technical analysis techniques. They use data mining techniques to
identify a pattern from a huge amount of market data. Neural networks, artificial intelligence machine learning, deep learning, and many more approaches were used for the prediction of trends in the stock market.

II. Related work:
In this section we discuss about the work done by many researchers in the field of stock trend prediction.

Different analysis techniques:

![Types of Analysis Diagram](Figure 1 - Types of Analysis)
1. Fundamental analysis
2. Technical analysis
3. Sentiment or news based analysis.

1. **Fundamental analysis:** Fundamental analysis is a method of measuring intrinsic value of share, share or any security by examining company sales, earnings, profits and other economic factors for forecasting. Financial analysis depends on the firm, employees, the board of directors, financial status of company, and company balance sheet etc. [16] Systematic review of fundamental and technical analysis.

On the basis of fundamental analysis, we can’t perform system automation; it’s difficult to forecast the stock values. It is useful for long term price movements but not for short term price change [31] [khan 2011 review paper]

**Financial analysis also considers different ratios**

- **Earnings per share or EPS:** EPS is the amount of profit that is assigned to each stock of the company. It is calculated by dividing the total revenues or gain of the company by the total number of outstanding shares.
  
  To put it in a formula:
  
  \[
  \text{EPS} = \frac{\text{Net income of the company after tax}}{\text{total outstanding shares}}
  \]

  As EPS is a symbol of the health of the company, a higher EPS means higher returns for the investor.

- **Price-to-earnings (P/E) ratio:** P/E is one of the essential tools of fundamental stock analysis. It reflects the company’s payouts as compared to its stock price. With this, you can know if the share of stock pays well for the price you pay. P/E ratio can be calculated by dividing the share price by the EPS.

  If a company’s share price is Rs 50 and the EPS is 5, then the P/E ratio is 10. A lower P/E ratio signifies the possibility of higher earnings compared to the stock price. A meager P/E ratio may mean a lower price per share compared to earnings. This signifies the stock is undervalued and shows potential to rise in future. The opposite is the case for a higher P/E ratio.

  The P/E ratio can be categorized as:

  1. **Trailing P/E ratio** which means the P/E ratio of the past 12 months
  2. **Forward P/E ratio** which is the P/E ratio of the next 12 months

  If the forward P/E ratio is higher than the trailing one, then there may be a decrease in earnings. If the forward P/E ratio is lower than the trailing P/E ratio, then there could be an increase in the profits of the company.

  The significance of the P/E ratio differs from one investor to another. The P/E ratio shows how much you want to pay for the company’s earnings. Your willingness can be different from another investor.

- **Return on Equity**

  Return on Equity or ROE shows the efficiency of a company to generate profits on its shareholder’s investment. It is calculated by dividing net earnings after tax by shareholders’ equity. If the company has made Rs 50 lakh this year with shareholders’ equity at 5 lakh, then the ROE is $5000000/500000 = 10\%$.

  ROE is expressed in percentage terms. A higher ROE signifies a more efficient company. It means the company can increase its profitability without any additional capital. However, a company without many assets can also have a higher ROE. Therefore, not all companies with higher ROE are suitable for investment. It is best to compare ROE of companies within the same industry. An ROE within the range of 13 to 15 is considered good.
d. price-to-book (P/B) ratio
Also known as “stockholders equity”, the price to book ratio is the comparison of a stock’s book value to its market value. Book value is the cost of each asset minus its cumulative depreciation. The P/B ratio can be calculated by dividing the last closing price by the previous quarter’s book value per share. It tells us what the company will be left with if it repays all its liabilities and liquidates its assets. If the P/B ratio is less than one, then the stock is undervalued. If the rate is more than one, then the stock is overvalued. The P/B ratio is essential as it tells you if the company’s assets are comparable to the stock’s market value. The ratio is more significant for companies with higher liquid assets such as insurance, banking, investment and finance companies. Companies with more fixed assets and expenditure of R&D do not get any help from the P/B ratio.

e. beta:
The Beta is the correlation of the stock price with its industry. You can calculate the Beta by comparing the stock to the benchmark index. The Beta mostly oscillates between -1 and 1. However, it can have a value above or below this mark. Any beta value above 0 signifies the stock correlates with the benchmark index. Beta values below 0 mean shares are inversely correlated. A higher beta means higher volatility signifying greater risk of assets. The lower the Beta, lesser is the volatility.

f. Price-to-sales (P/S) ratio:
Price-to-sales ratio compares a company’s stock price with its revenue. You can calculate the P/S ratio by dividing market capitalisation by income or using the formula:

\[
P/S \text{ ratio} = \frac{\text{Per share stock price}}{\text{Per share revenue}}
\]

A lower P/S ratio indicates undervaluation, while anything above average suggests overvaluation. A lower P/S ratio is preferable as it means investors are ready to pay less for each unit of sale. The limitation of this indicator is that it does not take into account the expenses and debt of the company. As such, a company with a higher P/S ratio can also be unprofitable.

g. Dividend payout ratio
A dividend payout ratio tells us how much the company has earned and what portion of it is being given out as a dividend. It can be calculated by dividing the total amount of dividend by the net income of the company. A company can choose to distribute its profit as a dividend because there may be little room for growth. Dividend payout ratio accounts for the amount of income that a company retains for future growth, debt payoff and cash reserve.

h. Dividend yield ratio
Dividend yield ratio is what the company pays to its shareholders as a dividend relative to its share price. Expressed in percentage terms, the dividend yield ratio can be calculated by dividing the annual dividend of stock by the current share price. The dividend yield ratio is important for investors who are looking for earning dividends from a company. This measure for fundamental analysis is not available for all companies as not every company uses its profit to pay dividends. Some companies retain the profit to use for future growth.

i. Projected earnings growth (PEG) ratio
Projected earnings growth indicates how much you have to pay for each unit of future growth of earnings of the company. It is calculated by dividing the P/E ratio by projected growth in revenues. A lower projected earnings growth indicates a lesser amount to be paid for each unit of future earnings growth. A stock with a smaller PEG ratio is fundamentally stronger as it has higher projected growth in earnings. A stock with a higher PEG ratio is generally avoided by investors.

Analysts use fundamental analysis tools to determine an estimated future value of a company’s stock price. If analysts expect a higher future value than the current market price, then the chances of buying a
stock are higher. If analysts find the stock’s intrinsic value lesser than the current market price, there may be sell recommendations for the stock since it’s overvalued.[2]

2. **Technical analysis:** Analysis is a method used in trading discipline employed to evaluate investments and identified trading opportunities in price trends and patterns seen on charts. Technical analysis is done on past historical data through which one can identify trading opportunities in price trends and patterns seen on the chart. There are different types of indicators used in stock market for prediction.

   a. **SMA (Simple Moving Average):** A simple moving average (SMA) calculates the average of a selected range of prices, usually closing prices, by the number of periods in that range.

   A simple moving average (SMA) is an arithmetic moving average calculated by adding recent prices and then dividing that figure by the number of time periods in the calculation average

   **Formula for SMA:**

   Where:

   \[ A_n = \text{the price of an asset at period } n \]

   \[ n = \text{the } \text{SMA} = \frac{A_1 + A_2 + \ldots + A_n}{n} \]

   number of total periods.

   b. **EMA (Exponential Moving Average):**

   An exponential moving average (EMA) is a type of moving average (MA) that places a greater weight and significance on the most recent data points. The exponential moving average is also referred to as the exponentially weighted moving average. An exponentially weighted moving average reacts more significantly to recent price changes than a simple moving average simple moving average (SMA), which applies an equal weight to all observations in the period.

   **Formula for EMA:**

   \[ \text{EMA} = \text{Closing price} \times \text{multiplier} + \text{EMA (previous day)} \times (1 - \text{multiplier}) \]

   Where multiplier \(=\left[\frac{2}{(\text{number of observations} + 1)}\right]\). For a 20-day moving average, the multiplier would be \(\frac{2}{(20+1)}=0.0952\).

   c. **MACD:** Moving average convergence divergence (MACD) is a trend-following momentum indicator that shows the relationship between two moving averages of a security’s price. The MACD is calculated by subtracting the 26-period exponential moving average (EMA) from the 12-period EMA.

   **Formula for MACD:**

   \[ \text{MACD} = 12-\text{Period EMA} - 26-\text{Period EMA} \]

   **Relative Strength Index (RSI):** The relative strength index (RSI) is a momentum indicator used in technical analysis that measures the magnitude of recent price changes to evaluate overbought or oversold conditions in the price of a stock or other asset. The RSI is displayed as an oscillator (a line graph that moves between two extremes) and can have a reading from 0 to 100.

   \[ \text{RSI} = 100 - \left(\frac{100}{1 + \text{RS}}\right) \]

   Where RS=average gain /average loss.
d. **Bollinger Band (BB):** Bollinger Bands are a type of statistical chart characterizing the prices and volatility over time of a financial instrument or commodity, using a formulaic method propounded by John Bollinger in the 1980s.

Bollinger Bands are envelopes plotted at a standard deviation level above and below a simple moving average of the price. Because the distance of the bands is based on standard deviation, they adjust to volatility swings in the underlying price.

![Figure 2 Bollinger Bands](image)

e. **ADX:** The average directional index (ADX) is a technical analysis indicator used by some traders to determine the strength of a trend.

The trend can be either up or down, and this is shown by two accompanying indicators, the negative directional indicator (-DI) and the positive directional indicator (+DI). Therefore, the ADX commonly includes three separate lines. These are used to help assess whether a trade should be taken long or short, or if a trade should be taken at all.

f. **Trending line:** A trend line (shortened to "trend line" elsewhere on this website) is simply a momentum indicator. It measures the rate of increase in the share price over time and alerts you to any acceleration or deceleration of the trend.

All the above technical indicators are used in stock market for technical analysis. There are many indicators which work on different stock but these are commonly used by technical analysts.
2.1 Analysis based on prediction techniques:

A brief introduction about 3 types of Machine Learning Techniques

I.1 Supervised Learning Algorithms

This algorithm consists of a target/outcome variable (or dependent variable) which is to be predicted from a given set of predictors (independent variables). Using this set of variables, we generate a function that maps inputs to desired outputs. The training process continues until the model achieves a desired level of accuracy on the training data. Examples of Supervised Learning: Regression, Decision Tree, Random Forest, KNN, Logistic Regression etc.

I.2 Unsupervised Learning Algorithms

In this algorithm, we do not have any target or outcome variable to predict / estimate. It is used for clustering populations in different groups, which is widely used for segmenting customers into different groups for specific interventions. Examples of Unsupervised Learning: Apriori algorithm, K-means.

I.3 Reinforcement Learning:

Using this algorithm, the machine is trained to make specific decisions. It works this way: the machine is exposed to an environment where it trains itself continually using trial and error. This machine learns from past experience and tries to capture the best possible knowledge to make accurate business decisions. Example of Reinforcement Learning: Markov Decision Process.
3. Sentimental Analysis:

Sentiment analysis (or opinion mining) is a natural language processing (NLP) technique used to determine whether data is positive, negative or neutral. Sentiment analysis is often performed on textual data to help businesses monitor brand and product sentiment in customer feedback, and understand customer needs.

In stock market it can be identify, when some news occur on particular company. Now as a result its stock price value moves either in positive direction or in negative direction or even neutral.

III. Literature Survey (Survey papers are sorted according to the method they used)

There are various techniques were used for prediction the trend in stock market, for this machine learning algorithms are used on available data or past data for prediction of any stock value or trend. To obtain the accurate result researcher uses various data source and for prediction uses machine learning model such as SVM (support vector machine), CNN (convolution neural network), RNN (Recurrent neural network), ANN (Artificial neural network), Random forest, Naïve Bayes, LSTM (Long-short term memory), linear regression, Back propagations, sentiment analysis and many more. In this section we are discussing various researchers work for prediction of trend in the filled of machine learning, data mining techniques, neural network and technical indicators.

[1] In this paper author used soft computing analysis for predicting the stock trend. He designed a block diagram for a hybrid intelligent system for stock market analysis, with the help of stock index value, data preprocessor, neural network trend using the scaled, conjugate algorithm, and neuron Fuzzy system to obtain a stock trend and analysis.

[3] In this paper, the author tries to predict the trend in the stock market with the help of both news and Technical indicators. He uses SVM that is (a support vector machine) to solve classification and prediction problems. He also designs system architecture for prediction: price forecast and gives results by applying all the three aspects that are by news prediction, SVM, and technical indicators to find out the accuracy of prediction.

[4] The paper proposed the hybrid stocks trading framework with technical analysis in the field of machine learning techniques. The author also proposed a decision support system with the ANN technique. In this technique, he used computational FLANN and 6 technical indicators, by which he predicted the trend in the stock market. Here they used three basic values of the stock market i.e. buy, sell and hold. The Paper also defines the concept of an extreme machine learning algorithm for a single hidden layer. The researcher preferred a multi-layer model and took a combination of technical indicators with computational intelligence tools.

[5] In this paper, the author tries to design a predictive model by joining machine learning applications with fundamental technical analysis. In this paper, they use optimal long short term memory that is (o-LSTM), deep learning, and stock technical indicator. The author tries to develop an automated system with the help of machine learning and deep learning approaches and also uses a correlation tensor with the stock indicator to speed up the deep learning model. All three approaches are used jointly to make an effective forecast in the stock market.

[11] This paper focuses on various application of AI, machine learning and data mining techniques to financial markets; they use a genetic algorithm for Optimization and offers support to evidence for the sustainability of an algorithm in many economic applications.

[20] In this research, the author did a comparative study of four supervised machine learning algorithms. He uses SVM linear regression algorithm, KNN (k-nearest neighbor), and random forest algorithm to predict stock market movement for historical data. The author found that out of these algorithms random forest gives the best result approximately 95%, after that KNN performs well than linear regression and SVM performs very poorly.

[7] In this paper, the author works on stock prediction and how to manage the portfolio in the stock market by using curve fitting techniques. Different types of curve equations and use of data mining techniques to mine some good and different types of mutual funds in the Indian stock market and identify the trend or different sectors and try to find out a systematic portfolio management framework to increase the return over a long period with less risk. For this author try to find out the different statistical methods such as standard deviation, linear regression, and nonlinear regression correlation, time series analysis for prediction. He used all of these, one by one, and found that regression analysis is very much useful for prediction. The author tries to write an
algorithm with the regression curve fitting method and try to find out values in different sectors. From this algorithm, they found out that in the starting three months, the stocks do not perform well but after six months all stocks show a positive result.


[10] The paper suggests a system that uses multi-Kernel learning to integrate both market news and stock price to improve the accuracy of stock market predictions.

[11] This paper focuses on various application of AI, machine learning and data mining techniques to financial markets; they use a genetic algorithm for Optimization and offers support to evidence for the sustainability of an algorithm in many economic applications.

[13] Paper presents that AI and data mining can be used together to identify the knowledgeable patterns to predict the trend in the stock market to generate results more accurately.

[2] In this paper, authors suggest a model with 3 NN model that is multiple layer perception (MLP) generalized feedback networks, radial basis function with some technical indicator like RSI, MA, stochastic oscillator, moving average convergence divergence (MACD), and sensitivity analysis for finding the cause and effect relationship between input and output of the network. He takes 8 inputs and observes sensitivity and compares the actual data with the forecast of the next day closing stock price and finds out neural network help to reduce the error.

[19] This paper focuses on various neural network models for forecasting the stock trend by observing a neural model. They supposed that long short term memory algorithm for predicting the financial information in the current era; they also suggest that this algorithm must be used for forecasting.

[27] In this paper author design a system based on the hybrid system. He uses a decision tree rough set to predict the next day's trend in BSE. They use various technical indicators for feature extraction like volume-based indicator, price trend indicator, and overlays indicator. They compare the proposed system with an artificial neural network and naive bays trend predictor. Their proposed system outperformed both ANN and nave-based trained prediction systems. Few author used technical indicators for trend prediction like [2] suggest a model with 3 NN model that is multiple layer perception (MLP) generalized feedback networks, radial basis function with some technical indicator like RSI, MA, stochastic oscillator, moving average convergence divergence (MACD), and sensitivity analysis for finding the cause and effect relationship between input and output of the network. He takes 8 inputs and observes sensitivity and compares the actual data with the forecast of the next day closing stock price and finds out neural network help to reduce the error.

[12] The paper presents a technical analysis of the Indian stock market on Nifty. It used moving averages in a variety of ways to generate profit in the market.

[14] This paper is about the descriptive analysis of MACD indicators on five randomly chosen stocks of NSE. MACD is both a Trend and Momentum indicator hence wide use for investors. Centerline cross-over generates a strong signal both on buy and sell. The limitation of MACD is it generates signals in overbought and oversold conditions so the use of fundamentals is suggested by the author to overcome it.

[25] In this paper, the author proposed a comparative study of different technical indicators to find out the efficiency, reliability, and sensitivity correlation of each technical indicator to predict which sector is efficient in the future market for buying or selling. After the studies, he finds that the Bollinger band stands a high success ratio.

[29] The paper presents the use of twelve different technical indicators as input in MLP and identifies or calculates the impact of individual technical indicators on a stock's daily close price for forecasting. For this, they use the Brazil stock index IBOVESPA index and high volume of daily trading. They calculate daily closing price value with moving average. For this they use swing trade operation they identify the impact of the individual technical indicator. After that, they also find the combination of Technical indicators.

[15] The paper presents the combination of 4 NN (AR, MA, ARIMA, and ARMA) and four deep learning architectures that are multilayer perception (MLP), recurrent neural network (RNN), long short term memory (LSTM), and convolution neural network (CNN) for predicting the stock price, the study shows that ARIMA model and obtained results when compared it shows that neural network outperforms the existing linear model (ARIMA).

[17] In this research paper author had used Firefly methods for forecasting the stock prices as an input data set for this author collect data from 4 different websites and this data is inserted into some well defined
mathematical formulas by using recurrent back propagation neural network for feature reduction online sequential extreme learning machine algorithm many other are used for they use a time horizon of 1 day, 3 days, 5 days, 7 days, 15 days, and 30 days strategies and apply these strategies data set. They found that OSELM outperforms the rest experimental models.

IV. Conclusion:

After reading and analyzing the above papers, I found that most researchers have given a single strategy to apply to particular stocks or indices. Either they have used technical analysis or a combination of technical as well as fundamental analysis. Few also added one more leg by using news and sentiment analysis to technical and fundamental, making it a multi-layered tool. Also, they compared their strategy to already available strategies or tools and outperformed them. But after studying and attending live trading sessions in the stock market, I conclude that there are a variety of stocks like humans with different natures. They may be less or more volatile, traded with less/more volume, its correlation with the index, company size (small cap, mid-cap, or large-cap), in different sectors (financial, agro, realty, service, etc), inclusion/exclusion in the index, etc. So trying to fit all of them in a single strategy may not be a good idea, and it will not generate the same results with all types of stocks. Here, my gap finding in the above literature is, there should be a bunch of different strategies/tools which will work on different types of stocks according to their flow type. Backtesting will reveal, which tool works best on a particular stock. Trading that stock with the selected tools may generate better results in future trades. Later, we can also categorize tools according to a particular nature of stock say less volatile or highly traded, etc.

REFERENCES


[8]e. w. saad, v. danil, v. prakhor, c. donald and wunsch, "comparativev study of stock trend prediction using


market predictor," 2021.


[27] B. B. Nair, V. P. Mohandas and N. R. sakthivel, "a decision tree rough set hybrid system for stock market trend prediction," *international jornal of computer application*, vol. 6, no. 9, 2010.


[32] https://www.investopedia.com

[33] https://www.analyticsvidhya.com