



Ai-Driven Predictive Models For Stock Market Forecasting: A Comparative Study

Kunwar Aditya Singh

Student

GITAM (Deemed to be University)

ABSTRACT:

This study conducts a comparative analysis of stock market trend forecasting using AI, focusing on Machine Learning (ML) and Deep Learning (DL) techniques. It examines various models, including traditional approaches like neural networks and advanced deep learning methods such as LSTM, CNNs, and RNNs, to enhance prediction accuracy using historical and real-time data. The research also assesses short-term and long-term market trends by integrating technical indicators and macroeconomic factors, highlighting AI's role in revolutionizing investment strategies and optimizing financial market efficiency.

Keywords: Artificial Intelligence, Machine Learning, stock market forecasting, predictive models.

1. INTRODUCTION:

The stock market is the collective trading network involving companies and investors to sell and buy the shares of the company for profit. Investors buy and sell the shares of the company based on the company's value, financial conditions, price of the share, and market value of the company. Buying the shares of ownership is investing in the company for mutual profit, as more people buy the shares the demand for investing in the company increases so as the profit of the company. The Initial public offering enables the shares in the public market for the investors to buy and the potential investors invest in the company if it seems profitable in investment. The main shareholders offer the Initial public offering when a company is established and enables the shares in the public market for the investors to buy and the potential investors invest in the company if it seems profitable in investment, the invested money is utilized for the growth of the company. As the more people tend to buy the shares, the more the investment, the more the profit for the investors which in return increases the value of the company. The invested amount is beneficial in providing

capital for the company to expand business without taking loans from the bank, other way benefitting the investors by providing regular dividends. The price of the stock often relies on demand and supply, which depends on various factors like fluctuations in the cost of materials, changes in production techniques, labour cost, shipping cost, increases in marketing strategies etc., which might affect the value and profit of the company. These factors often directly impact the profit, demand and value of shares of the company in the market. These fluctuations might be problematic for potential investors as analysing the trends is complicated due to the unstable, non-linear, and noise concerning the market. The stock prices are unstable to predict the changes as they depend on numerous factors like market conditions, past earning reports and their changing behaviour.

1.1 Factors in predicting stock markets are:

Fundamental analysis evaluates the true value asset by analysing the various factors like the financial and income statements of the company, cash flow and balance sheet that influence company value and shares. Fundamental analysis briefs the financial health of the company and tells whether buying the shares would benefit the investor in the long term or not. The financial health of the company depends on various metrics like earnings, revenue, valuation and market conditions. This analysis method determines the performance and true value of the stock using various predictive analysis methods. This analysis is mainly for long-term investments, predicting the fair price of the stock, understanding business models, evaluating the health and performance of the company to predict future price movement and analysing whether the stock is overvalued or undervalued. The economic, industry and company analysis is calculated as the true value of the company or the intrinsic values of the shares. Investors are recommended to buy the shares if the value of the shares is higher if the true value is high and vice versa.

Technical analysis is the understanding of the patterns in the share price of the stock and predicting the share price by examining the charts and past data. It determines the right time to invest and exit the market for an investor. It focuses on the short-term benefits of investing to get benefitted. The tools used to predict past data are chart patterns and technical indicators. The charts and patterns of past data help to identify the present price spike in the market, line charts, bar charts, point figures and candle stick charts are some types of charts used to analyse the past data. Technical indicators determine the entry and exit signals relational to price movement, which accurately inform the investors about future price spikes and determine whether to buy, sell and hold their assets.

Time series analysis refers to a set of tools and techniques that help researchers in examining the time series data. Time series forecasting is a kind of predictive analysis that forecasts future values using the data that has been collected in the past throughout a period. In a time series, sequential data taken at regular intervals is observed. Time-series analysis evaluates and handles the series of data points over a period to know a model to predict the future value and knowledge of various strategies.

1.2 Prediction of Market Earnings:

The prediction depends on the previous year's stock exchange and other economic criteria. The prediction of this unstable and vast variety of data is hardly possible to be accurate in forecasting stock market trends. The technology development in the current world helps in the analysis of this huge amount of data for the similar patterns from past data and present trends can be possible with the latest models of Artificial Intelligence.

1.2.1 Artificial Intelligence:

Artificial Intelligence (AI) is a field within computer science focused on creating systems capable of emulating human cognitive functions. These intelligent systems are designed to perform tasks such as logical reasoning, critical analysis, problem-solving, understanding and processing natural language, and interpreting sensory information. The learning of different patterns and analysing the data from huge amounts of data can be done by Machine Learning methods within AI. Machine learning analyse and understand the patterns in data over a period and predict the patterns based on the data insights. The various models of ML are used to train the machine on labelled and not labelled outcomes. Machine Learning involves in development of statical models and algorithms where computers learn from predictions and make predictions. Deep learning is a specific area of ML where large datasets and complex patterns are analysed using neural networks with many layers of interconnected neurons just like the human brain. The deep learning interconnections are inspired by the structure and functions of the human brain to do powerful tasks. This interconnected neural network contains a large amount of data with many layers.

1.2.2 AI in finance:

With the rapid evolution of AI in every field, implementation of AI in financial markets for risk evaluation and prevention, forecast stock prices and prediction of the stock market are crucial for investors, traders and researchers. AI algorithms have been increasingly introduced to stock market predictions to leverage the machines' ability to process and analyse vast datasets based on past and real-time data. Perks of AI enhancements in the stock market are predictive analysis, risk management, fraud detection, and improved trading algorithms.

1.2.3 AI in Stock Market Forecasting:

Forecasting stock market trends can be accomplished using a range of techniques, which includes deep learning, machine learning, and natural language processing. These cutting-edge methods utilize complex algorithms and financial data to predict market behavior with greater precision than traditional techniques like fundamental analysis, technical analysis, and time series analysis. The effectiveness and sophistication of AI-based approaches position them as a powerful alternative to conventional methods in stock market prediction.

2. LITERATURE REVIEW:

2.1.HISTORY OF STOCK MARKET:

Since the first days of structured stock exchanges in the 18th and 19th centuries, stock market raises have a long history. The first firm that issued stock that could be exchanged on an exchange of stocks was the Dutch East India Firm, which was established in 1602. Public stock offerings become a vital source of fundraising for businesses looking to generate funds for expansion. To engage in trade with South America, the South Sea Company was founded in England in 1711. In 1720, it released shares that sparked a financial boom. When the New York Stock Exchange (NYSE) was established in 1792, it started trading bank stocks and government bonds. As early as 1792, the Federal Reserve Bank of New York became the first corporation to be listed on the NYSE.

2.2 The Rise in the 19th Century:

In Europe and the US, public offers and stock market activity surged in the 19th century as a result of the industrial revolution and railroad expansion. To accommodate the development of railroads and industry, the New York Stock Exchange (NYSE) grew westward in the United States, establishing exchanges in places like San Francisco, Chicago, and Kansas City. To assess the performance of industry stocks traded on the New York Stock Exchange (NYSE), Charles Dow created the Dow Jones Industrial Average, which was first introduced in 1896.

2.3 Reforms of the Early 20th Century:

Following the 1929 stock market crash and the subsequent Great Depression, the United States took significant measures to reform its financial markets. One key step was the establishment of the Securities and Exchange Commission (SEC) in 1934, aimed at curbing fraudulent activities and speculative practices. Earlier, the Securities Act of 1933 had mandated that companies register new securities and provide transparent financial disclosures. Furthermore, the Glass-Steagall Act played a crucial role in reinforcing the banking sector's stability by founding the Federal Deposit Insurance Corporation (FDIC)

2.4 Stock Offerings in the Modern Era:

Stock markets have expanded globally, switched to electronic trading, listed thousands more firms, and offered a greater range of assets since the middle of the 20th century. The capitalization of the stock market has increased dramatically, and as trading expenses have decreased, so has public involvement. The years 1990 to 2000 saw a rise in initial public offerings (IPOs) for high-tech firms; but, since then, the market has changed, demanding sales, profits, or other encouraging results before going public.

3. METHODS:

With the rapid evolution of AI in every field, implementation of AI in financial markets for risk evaluation and prevention, forecast stock prices and prediction of the stock market are crucial for investors, traders and researchers. AI algorithms have been increasingly introduced to stock market predictions to leverage the

machines' ability to process and analyse vast datasets based on past and real-time data. Perks of AI enhancements in the stock market are predictive analysis, risk management, fraud detection, and improved trading algorithms. Predicting stock market trends has increasingly relied on various advanced techniques, including deep learning, machine learning, and natural language processing. Given the vast amount of financial data available and the advancements in artificial intelligence, these AI-driven methods are often favoured over traditional approaches such as fundamental analysis, technical analysis, and time series forecasting. The superior performance of AI techniques in handling complex datasets and uncovering patterns contributes to their growing popularity in market predictions.

With the rapid fluctuation of complex patterns in the stock market, it is often challenging for individuals to predict accurate changes. The advancement in Artificial Intelligence and Machine Learning paved a new path for active prediction of stocks with the availability of vast historical data and algorithms. The data collected from various sources are processed to remove unnecessary inputs and missing values. This feature refines the information quality for accurate prediction. The general steps follow normalization, scaling and feature selection.

3.1 The methods used in stock market prediction:

A *Recurrent Neural Network (RNN)* is designed to handle sequential data by retaining memory of previous inputs, forming a directed cycle of information. It is particularly effective for tasks where the sequence of data is crucial, such as natural language processing, time series prediction, and speech recognition. RNNs predict future values in a time series by leveraging past data. This capability makes RNNs highly valuable for time series analysis, as they are able to comprehensively recall and utilize information from earlier inputs over a period of time.

Convolutional Neural Network (CNN) is the model of Deep Learning which recognizes patterns and features in images, mimicking human brain operations. The 4 main components in the CNN are convolutionary layers, fully connected layers, pooling layers and rectified linear units.

- a. Convolution layers: this layer helps in recognizing a specific pattern from images like curving of the digits and edges, the shape of the digit, where small grids called filters are present in each layer. These small grids which act like magnifying glasses move over the images to look for specific patterns like curves, lines and shapes. Different filters combine to identify the patterns, so each layer's pass on the image creates a new grid.
- b. Activation function or Rectified Linear Unit (ReLU) enables the network to capture non-linear relationships among the features in the image. Teaching the difference between the patterns makes the network robust enough to identify patterns.
- c. Pooling layer identifies the features from the convoluted matrix, which is done by reducing the dimensions of the convoluted matrix also called the featured map, to reduce the memory while teaching the network. So, features like Max pooling, sum pooling and average pooling help in the featured map.

- The max value of the featured map is called Max pooling.
- The sum values of all in the feature map are called Sum pooling.
- The average of total values in the feature map is Average pooling.

The dimension of the small grids of the featured map become smaller as the pooling functions are applied. The last pooling layer is flattened to process the next layer called the fully connected layer.

- d. The fully connected layer is the last layer of CNN. The neurons are connected to every neuron in the previous and subsequent layers creating a highly connected network, that applies a linear transformation to the input vector. These layers process the input data directly into output data. It is mainly used in complex pattern recognition.

Long short-term memory (LSTM) is the enhanced variant of a Recurrent neural network. While traditional RNN has a hidden state, which passes through time and makes it complex to learn long-term dependencies for the network. The problem can be overcome with the implementation of a memory cell in the network which acts as a container to hold the information for an extended period. LSTM learn long dependent sequences of data, which makes it possible to do tasks such as translation of language, recognition of speech, anomaly detection, video analysis and forecasting time series.

Linear Regression is a widely utilised machine learning technique used to determine the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the data. When the model involves just a single independent variable, it is referred to as simple linear regression. On the other hand, when multiple independent variables are included, it is called multiple linear regression. This approach is known as multivariable regression when multiple independent variables are used, and as univariable when only one dependent variable is being analyzed. Due to its efficiency in processing large datasets, linear regression is particularly advantageous for real-time data analysis.

Support Vector Machine (SVM) is used for linear classification or non-linear classification. The SVM algorithm divides data points into multiple classes in the feature space by locating the best hyperplane in an N-dimensional space. With regard to the closest points of different classes, the hyperplane aims to preserve the biggest buffer feasible. The size of the hyperplane is based on the number of features. If two input features are present, the hyperplane is effectively a line.

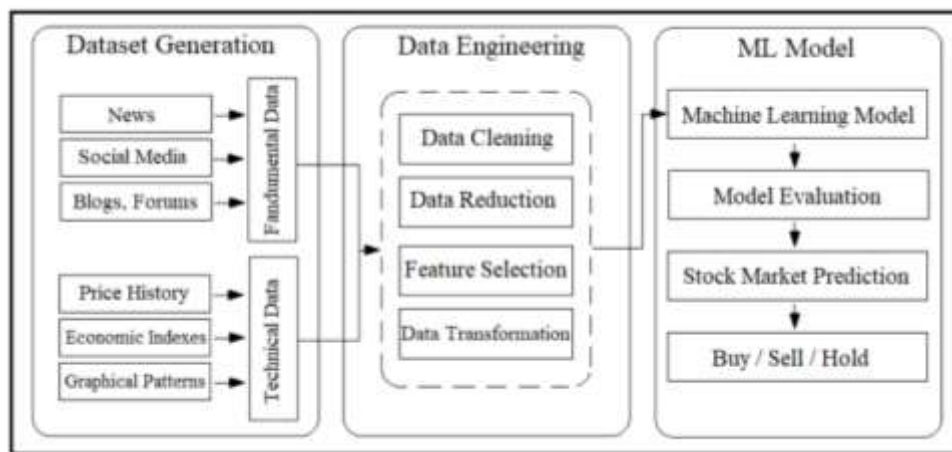


Fig. 1: The model training framework to predict the stock market.

3.2 Process of Predicting Stock Market Using AI/ML Techniques:

Data set generation: A huge amount of Data sets is collected from different sources like news articles, and social media to identify emerging trends and include historical data like past stock prices, financial data and trading volumes. It also has data on macroeconomic factors like interest rates, GDP growth, and inflation rate to analyse the stock movements.

Data Processing: The collected data from the different sources is processed to filter the noise, missing values and ubiquitous present in the raw data. Filtering the raw data to create new features to move averages and sentiment scores to understand the trends better.

Model Selection: Analysing the refined data and choosing the ML model to improve the prediction accuracy. Selecting the appropriate model involves choosing from various deep learning and machine learning techniques, such as Recurrent Neural Networks, linear regression, Support Vector Machines, Long Short-Term Memory models, and Convolutional Neural Networks.

Prediction: The models are trained to analyse using various techniques in neural networks based on historical models to minimize error while predicting. After the analysis of the model, AI models use sentiments for future stock prices.

Evaluation: The parameters like no.of layers, and learning rate are adjusted to improve the accuracy in prediction and performance by continuously analysing and updating the data with new datasets.

This process of analysis helps investors in making wise decisions in forecasting the stock market. The insights on future stock movement using this AI / ML technique help the improving stock market forecast.

4. RESULTS:

The graphical representation of various techniques in the prediction of stock to analyse the accuracy. Promising advances and mixed degrees of success have been observed in the application of artificial intelligence (AI) in stock market prediction. When it comes to stock price prediction, AI systems have proven to be able to achieve accuracy rates of more than 80%. The stock market is inherently volatile and complex, impacted by a wide range of factors such as investor emotions and global economic conditions. For this

reason, the accuracy level attained is noteworthy. With three-year cumulative returns of 34%—nearly three times the worldwide industry average—AI-driven hedge funds have trounced conventional investment techniques. According to this, trading in stocks may benefit from artificial intelligence (AI). Larger volumes of unstructured data, including news articles and social media, can be effectively analyzed by AI systems using natural language processing (NLP) to determine market sentiment. With real-time sentiment analysis at their disposal, investors may now make more informed predictions thanks to these capabilities. Trading within milliseconds is now possible because to AI systems that have improved high-frequency trading tactics. By taking advantage of small price differences, dealers can increase their total trading efficiency and profitability thanks to this quick response capabilities. Making better investing decisions is made possible by AI models, which offer data-driven insights that lessen the need for intuition. Through real-time analysis of diverse market conditions, this strategy facilitates improved risk management.

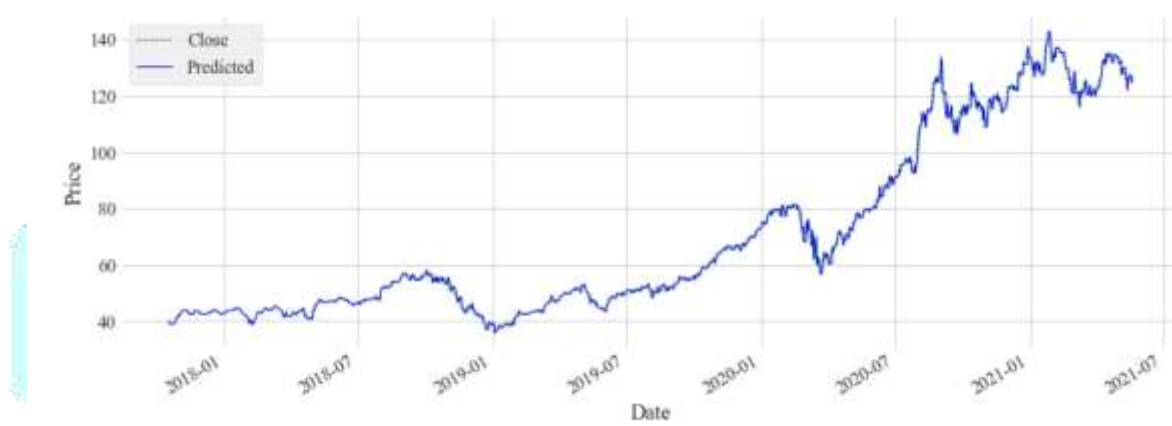


Fig. 2: Prediction by linear regression model

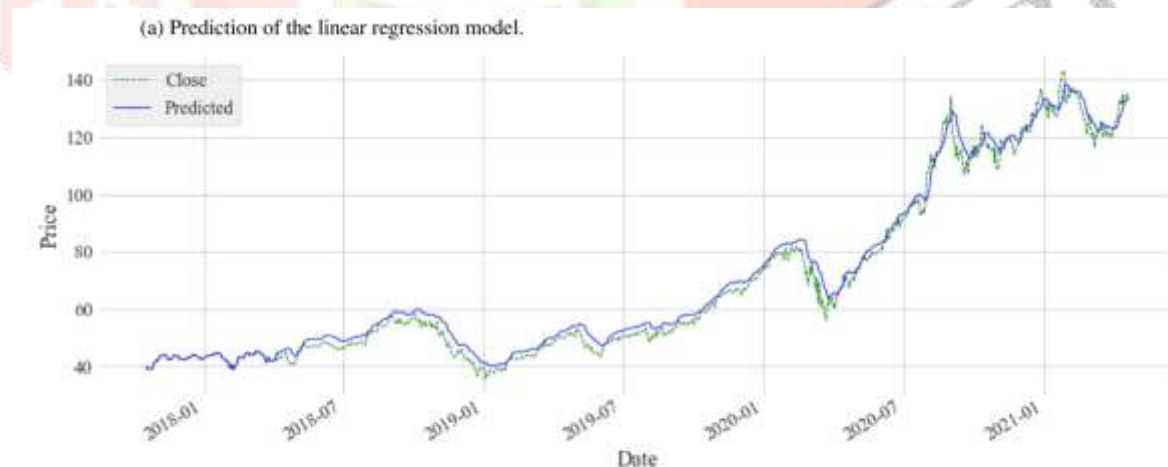


Fig. 3: Long Short Term Machines model Prediction

AI has, all things considered, demonstrated a great deal of promise for improving the efficiency and accuracy of stock market predictions. The stock trading industry has changed as a result of the incorporation of cutting-edge methods including natural language processing, machine learning, and sentiment analysis. For both investors and researchers, however, issues with algorithmic biases, market efficiency, and data quality continue to be major concerns. As artificial intelligence (AI) technology advances, it will probably play a bigger position in stock market forecasting, providing traders with more options and information.

5. CONCLUSIONS:

The difficulties in predicting the stock market using traditional methods can be leveraged using ML techniques. Stock market analysis depends on two factors i.e., fundamental analysis and technical analysis. The data sets from various sources are used to analyse the accuracy of the prediction. Labelled datasets are utilized to analyse the supervised learning algorithms and other evaluation metrics to examine multiple factors. Leveraging the ML algorithms, deep learning models process the vast amount of past and present data from different sources for efficient accuracy in prediction. Other factors like macroeconomic factors like stock market price, and real-time data are also considered in prediction. Stock market prediction is complex and influenced by unpredictable factors, so various AI models and techniques are implemented in active prediction. The growth of AI in the stock market continues to evolve to provide investors with adaptive forecasting tools.

6. CHALLENGES:

- **Evaluation of Relative Features and Data Quality:** These two factors have a significant impact on how well AI predicts the stock market. Predictions that are off can be caused by imprecise data or poorly chosen features.
- **Market Efficiency:** Conventional financial theories believe that stock prices represent all available information, making forecasting difficult by nature. Examples of these theories include the Efficient Market Hypothesis and Random Walk Theory. The argument over whether AI is still capable of spotting patterns that defy these hypotheses is still being debated among experts.
- **Algorithmic Bias:** There are issues with the possible biases in artificial intelligence algorithms as well as the unreliability of their decision-making methodology. The dependability of forecasts and the requirement ability in AI models used for trading are called into question by this.

7. REFERENCES:

- [1] Mohammad Abedin, Sohrab Mokhtari, and Armin B Mehrabi. Bridged detection using machine learning algorithms. In Health Monitoring of Structural Systems XV, volume 11593, page 115932P. International Society for Optics and Photonics, 2021.
- [2] Robert D Edwards, WHC Bassetti, and John Magee. Technical analysis of stock market trends. CRC Press, 2007.
- [3] Alireza Abbaspour, Sohrab Mokhtari, Arman Sargolzaei, and Kang K Yen. A survey on the active fault-tolerant control systems. Electronics, 9(9), 2020.
- [4] Riswan Efendi, Nureize Arbaiy, and Mustafa Mat Deris. A new procedure in stock market forecasting based on fuzzy random auto-regression time series model. Information Sciences, 441:113–132, 2018.
- [5] Haiqin Yang, Laiwan Chan, and Irwin King. Support vector machine regression for volatile stock market prediction. In International Conference on Intelligent Data Engineering and Automated Learning, pages 391–396. Springer, 2002.
- [6] Eugene F Fama. Efficient capital markets: Ii. The journal of finance, 46(5):1575–1617, 1991.

- [7] Tom Howley, Michael G Madden, Marie-Louise O'Connell, and Alan G Ryder. The effect of principal component analysis on machine learning accuracy with high dimensional spectral data. In International Conference on Innovative Techniques and Applications of Artificial Intelligence, pages 209–222. Springer, 2005.
- [8] Girish Chandrashekar and Ferat Sahin. A survey on feature selection methods. Computers & Electrical Engineering, 40(1):16–28, 2014.
- [9] Ahmad Kazem, Ebrahim Sharifi, Farookh Khadeer Hussain, Morteza Saberi, and Omar Khadeer Hussain. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. Applied soft computing, 13(2):947–958, 2013.

