European Journal of Electrical Engineering and Computer Science
Vol 9 | Issue 5 | September 2025
ISSN 2736-5751

RESEARCH ARTICLE

Prediction of Profitable Stock using Candlestick Patterns with ML

Md. Siam Ansary*

ABSTRACT

Individuals have access to a trading platform through the stock market. The economy expands thanks to these interactions. An intriguing study issue is predicting the profitability of a company. If such forecasts can be produced effectively, people will be able to invest more methodically. Using a variety of candlestick chart patterns that may be broadly categorized as bearish, bullish, or neutral, we addressed this issue in our study. K Nearest Neighbor, Decision Tree, Random Forest, Support Vector Machine, AdaBoost, and Multilayer Perceptron are just a few of the categorization models that have been evaluated and show promising results.

Keywords: Candlestick pattern, machine learning, stock.

Submitted: June 08, 2025

Published: September 01, 2025

10.24018/ejece.2025.9.5.738

Department of Computer Science and Engineering, Ahsanullah University of Science and Technology, Bangladesh.

Corresponding Author:
e-mail: siamansary.cse@gmail.com

1. Introduction

Research on artificial intelligence (AI) has grown in significance across many disciplines. Machine learning (ML) applications are employed in a wide range of industries, including engineering, science, health, finance, education, business, marketing, stock market, and medicine. There are self-driving cars on the roads, autonomous robots making our daily lives easier, as well as global financial institutions and investment banks. Due to the significant financial returns that can be obtained by investing in various businesses, the stock market is a particularly popular choice among investors. Although investing in the stock market can be rewarding, it can also exhibit unpredictable behavior, which makes predictions more difficult. Many academics and business professionals have tried to identify patterns in the historical data that could yield the most financial gain in an effort to solve this problem. Machine learning and other contemporary technology can be very helpful in predicting this kind of circumstance, perhaps saving many people from potential financial loss. An essential goal in the world of finance is predicting lucrative stocks since a reasonable prediction has the potential to produce significant financial profits, function as a hedge against market risks, and help design efficient stock exchange trading tactics. Due to the better long-term returns offered by the stock market, people frequently spend money there. Trading on the stock market has become incredibly popular as a way to generate significant profits in the most significant financial markets in the world. Therefore, any knowledge of potential information regarding the profitability of a specific share will absolutely guarantee enormous gains in this market. As a result, for investors, buyers, sellers, fund managers, policymakers, researchers, applied employees, and many other market participants, accurate market prediction is crucial. Researchers are currently looking for trustworthy predictive models due to the availability of data, the development of AI, and machine learning.

In order to make it simple to create an investment decision assistance system, the goal of this research is to detect profitable stocks. In technical analysis, a candlestick pattern is a representation of a stock that indicates the high, low, open, and closing values over a given time period. Since the nature of stocks may be predicted using candlestick pattern data, it is possible to choose the best stocks to invest in. It is possible to create a decision support system that can ensure profitable investment chances by combining ML approaches with Candlestick patterns.

1.1. Literature Review

Numerous studies on the stock market have been conducted by scholars, and various projects are still in progress.

In order to determine whether to purchase or sell particular stocks in the steel industry, Venkatesh et al. [1] experimented with technical analysis. In 2018, India was the second-largest steel manufacturer in the world, and that year, the country's steel consumption was just 8%; however, in 2019, that number was predicted to increase by 7%. In order to do this, they examined the applicability of technical analysis, the movement of market share prices in

large cap firms, the success of particular steel sectors, and the ability of technical analysis to predict future changes in share prices. The statistics and information were gathered from a variety of websites, including the corporate website, NSE websites, newspapers, and magazines. Techniques like Candlestick Charts, Simple Moving Average, ROC, and RSI were employed for the analysis.

Iqbal and Roy [2] explored the impact of the day-of-the-week effect on the Dhaka Stock Exchange (DSE) by analyzing the DSE market index from June 2004 to March 2015. Their study revealed several behavioral patterns among investors: a tendency to take profits on Thursdays and reinvest on Sundays; higher average trade volume, value, and market capitalization on Sundays; average returns peaking on Thursdays and dipping on Tuesdays; greater investor activity on Sundays and lower activity on Thursdays; and increased market volatility on Mondays with reduced volatility on Wednesdays.

A domain-specific programming language was introduced by Anand et al. [3] to help construct patterns for financial analysis. Such language can aid in data investigation because chart patterns have a geometrical shape and are influenced by stock price movement. Since trading on the stock market carries a significant level of risk, it is crucial to have an effective analysis technique for forecasting. A stock's price can fluctuate at any time depending on a number of factors and events, and over time, the value of a stock can changes. Yan and Yeng [4] found through experiments that the Long Short-Term Memory (LSTM) model can be particularly effective in stock price-related analysis works for such scenarios.

Guo et al. [5] tested neural network-based stock pattern identification. They noted that algorithms based on neural networks can produce better outcomes than rule matching and template matching approaches. They consider feature extraction to be important. A three-layer feedforward neural network was employed for the categorization work. Data from the Shanghai Stock Exchange was used in the study.

Shen and Shafiq [6] used LSTM to anticipate short-term stock market price trends. The Chinese Stock Market provided them with data spanning two years. In their experimental work, recursive feature elimination and principal component analysis were used to find the appropriate features.

Xu et al. [7] conducted a study focused on feature selection for predicting stock price trends. Given that stock price fluctuations are influenced by numerous factors, identifying the most significant features is essential for effective machine learning analysis. The researchers employed two Recursive Feature Elimination (RFE) methods—SVM-RFE and RF-RFE—based on Support Vector Machine (SVM) and Random Forest (RF), respectively. They used data from the Shanghai Stock Exchange for their experiments. The findings indicated that both RF and SVM are capable of predicting trends accurately, but SVM may offer superior performance. Moreover, while feature elimination appears necessary for RF, it may not be required when using SVM.

According to Chen and Chen [8], forecasting bullish turning points in stock research is more important because regular investors stand to gain more from them. For the research, datasets from the TAIEX and NASDAQ were used. For locating bull flag patterns, the study used chart patterns and trade indicators.

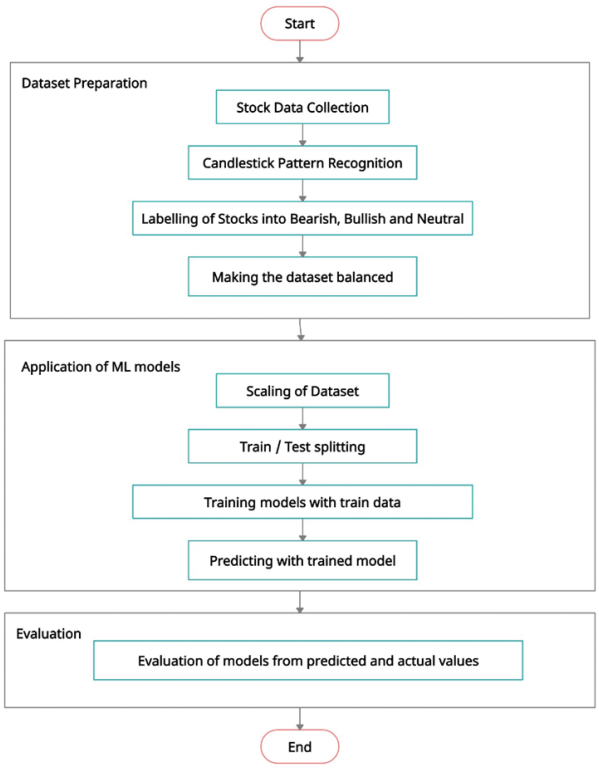

Fig. 1. Working steps of research approach.

In their research on candlestick pattern recognition, Lin et al. [9] They utilized a model that combined four machine learning techniques in their research. Logistic regression, k closest neighbor, limited Boltzmann machine, and random forest are the techniques employed. They discovered that using machine learning techniques, two- or three-day candlestick patterns can produce the greatest prediction analyses.

Go and Hong [10] focused on pattern matching-based stock value prediction. Data from the Korean Stock Market was used for the experiment. Different stock data patterns were first clustered, and Deep Neural Network (DNN) prediction was then carried out. The DNN model of Google's TensorFlow was employed by the researchers.

2. Method

The candlestick pattern of a stock can be used to determine whether it is bullish, bearish, or neutral. Bullish equities are successful since they reflect the market's rising trend. Conversely, negative equities signal a downward trend, suggesting they are not lucrative. Thus, using information on candlestick patterns and machine learning classification algorithms, successful stocks can be found.

2.1. Dataset Preparation

For a graphical summarization of the whole methodology, please refer to Fig. 1.

1) Collection of Stock Data: Initially, stock data was gathered through the Yahoo! Finance API [11], which

provided candlestick values for each stock. The dataset initially contained 7,502 records.

2) Recognition of Different Candlestick Patterns: Using the Technical Analysis Library [12], each stock's candlestick pattern has been identified. We have considered the following patterns.

- Hammer: A single-candle bullish reversal pattern called the hammer can be seen at the bottom of a downtrend. Such a pattern indicates that the bulls are stepping up their involvement in the game and are suggesting that there might be a change in the market direction after a downtrend in which the price action produced a sequence of lower lows and lower highs. the high close, which indicates that the bulls have recently taken control of the price movement.
- Inverted Hammer: A particular sort of candlestick pattern that appears after a downtrend and is typically interpreted as a trend-reversal indication is the inverted hammer.
- Dradonfly Doji: The Dragonfly Doji is seen as a bullish reversal candlestick chart pattern that portends a recovery.
- Bullish Spinning Top: A spinning top is a candlestick formation that denotes uncertainty about the direction of the next trend. A spinning top at the bottom of a downtrend indicates that the bearish is gaining ground and that the bullish may eventually take control.
- Bullish Engulfing: At the bottom of a downtrend, a bullish engulfing candle forms, signaling an increase in purchasing pressure. When the share opens lower than the previous trading session and finishes higher than the prior closing, this pattern is created.
- Bullish Harami: A bullish harami is a fundamental candlestick chart pattern that denotes the possibility of a market or asset's bearish trend reversing.
- Piercing: A bullish candlestick reversal pattern occurs with the appearance of the Piercing Pattern. When the middle of the bearish candle from the day before is closed above, a pattern known as this one is formed.
- Morning Star: A bullish candlestick pattern in a price chart is known as a morning star. It denotes the beginning of an ascent. It develops at the bottom of a downtrend and serves as a warning sign that the downtrend is about to change direction.
- Three White Soldiers: A bullish candlestick pattern called the three white soldiers appears near the end of a bear market. It denotes the end of the current downward trend. Strong purchasing pressure shown by this candlestick results in a trend reversal.
- Doji Star: A reversal candlestick pattern happens with the Doji Star pattern. Long candle at the beginning, gaps to doji, then turns around and moves the other way.
- Morning Doji Star: A downtrend occurs when the Morning Doji Star candlestick pattern first appears. A large- bodied candlestick signals the continuation of the decline.
- Three Inside Up: Three Inside Up Candlestick Chart Pattern is a highly reliable bullish trend reversal pattern. It forms during a downward trend. A massive down candle, a smaller up candle contained within the previous candle, and then another up candle that closes above the close of the second candle make up this bullish reversal pattern.
- Three Outside Up: On the candlestick chart, a pattern known as the three outside up trading pattern develops over the course of three trading sessions. It is a reversal pattern that starts with a candle in the trend's direction and manifests itself during a decline.
- Doji: Dojis, which are frequently parts of patterns, are sessions in which the candlestick for a security has an open and close that are almost equal.
- Marubozu: A Marubozu is a specific kind of candlestick charting pattern that shows that the price of an asset did not move outside of the range between its opening and closing prices.
- Spinning Top: The pattern of a spinning top is seen as neutral. A short true body that is vertically centered between extended upper and lower shadows characterizes this candlestick pattern.
- Hanging Man: Investors utilize the hanging man candle- stick pattern, a bearish reversal candlestick pattern, to help them decide whether to enter or exit a trade.
- Shooting Star: A shooting star is a bearish candlestick that is close to the day's bottom and has a small actual body with little to no lower shadow. At the peak of uptrends, it is regarded as a bearish reversal candlestick pattern.
- Gravestone Doji: The Gravestone Doji is a bearish candlestick pattern that displays the candle opening and closing at the day's low.
- Bearish Spinning Top: When the stock opens sharply lower and then buyers start to enter again, a bearish spinning top pattern forms.
- Bearish Engulfing: The bearish engulfing pattern is a technical chart formation that serves as a warning signal for potential price declines. Classified as a bearish reversal pattern, it commonly emerges at the apex of an upward trend, indicating a possible shift from bullish to bearish market sentiment.
- Bearish Harami: When a day has a large bullish candle and the next day has a smaller bearish candle, the candlestick is known as a bearish harami.
- Dark Cloud Cover: Technical analysis's term for a candlestick pattern that indicates a bearish reversal is the Dark Cloud Cover. It occurs when a down candle in a candlestick chart opens above the close of the preceding up candle and then moves on to close below the up halfway of the candle.
- Evening Star: Technical analysts employ the Evening Star candlestick formation on stock price charts as a reversal indicator signaling a potential shift in market trend di- rection. This pattern is characterized as a bearish reversal signal, suggesting an impending downturn following an uptrend.

Out of the considered patterns, Neutral Candlestick Patterns are Doji, Marubozu, Spinning Top and Bearish Candlestick Patterns are Hanging Man, Shooting Star, Gravestone Doji, Bearish Spinning Top, Bearish Engulfing, Bearish Harami, Dark Cloud Cover, Evening Star while other Candlestick Patterns are Bullish Patterns.

3) Labelling of the Stocks: Stocks are divided into three categories when their candlestick patterns are identified. These three patterns are the bearish, bullish, and neutral candlestick patterns. Bullish patterns show an upward trend in the market, indicating that the stock is profitable. Market downward movement is indicated by a bearish pattern, while market stability is indicated by a neutral pattern. Stocks that are rising are profitable, those that are falling are not, and the remainder are steady.

4) Balancing of the Dataset: After the stocks are labeled as bearish, neutral and bullish, it has been observed that most of the stocks are of neutral candlestick pattern, after which are bullish pattern stocks with bearish pattern stocks being the least in number. Random under sampling is done on neutral candlestick pattern stocks so that the amount of stocks match the number with bullish pattern stocks and then random oversampling is done on bearish pattern stock so that all types of stocks have the same number of instances. This balanced dataset has 7050 samples.

2.2. *Application of Machine Learning Techniques*

1) Feature Scaling: Machine learning classification models can be used to find successful stocks after the dataset has been produced, as bullish stocks represent profit, bearish stocks imply loss, and neutral stocks denote market stability. Scaling has been done on the prepared dataset. The features can initially be in different ranges, but the scaling procedure makes sure that all the features in the same range which helps with mitigation of bias.

2) Train—Test Splitting: Train data and test data have been divided using K fold cross validation. K-1 folds are utilized for training and one-fold is used for testing in each iteration. The value has been set to five in this case for K.

3) The ML Classification Models: Six classification models are trained with training data of each iteration. The models are mentioned below.

- K Nearest Neighbor Classifier: The K-Nearest Neighbor (KNN) classification method retains all known instances and classifies new data points based on similarity measures. It assigns a class to a new instance by taking a majority vote from the K closest neighbors, which are identified using a distance metric such as Euclidean distance. The new instance is then classified into the most common class among these K nearest neighbors.
- Decision Tree Classifier: A decision tree constructs predictive models using a hierarchical, tree-like structure. It recursively partitions the input dataset into smaller, more homogeneous subsets based on feature-based splitting criteria, thereby forming a series of nested decision rules. The resulting structure comprises internal decision nodes, which represent feature-based splits, and terminal leaf nodes, which denote the final output classes or predictions.
- Random Forest Classifier: Random Forest is an ensemble learning method designed to improve upon the limitations of the traditional Decision Tree model. By combining the results of multiple decision trees, it offers greater reliability and accuracy. This collective approach helps reduce overfitting and increases the overall robustness of the predictions, making Random Forest significantly more dependable than individual decision trees.
- Support Vector Machine Classifier: Support Vector Machine (SVM) is a widely used algorithm for classification tasks. It works by analyzing the input data to find an optimal boundary that separates different classes. The main objective of the SVM technique is to identify the most effective decision boundary or hyperplane that maximizes the margin between the classes, ensuring accurate classification.
- AdaBoost Classifier: AdaBoost is an ensemble boosting algorithm that enhances the performance of weak classifiers by combining them into a stronger, more accurate model. Through an iterative process, multiple weak learners are trained and aggregated to form a robust classifier. The core idea involves assigning weights to training samples and updating them in each iteration to focus on instances that were previously misclassified. For AdaBoost to be effective, two key conditions must be met: the base classifier must be capable of being trained interactively on weighted data, and it should aim to accurately classify the training examples while minimizing the error during each iteration.
- Multilayer Perceptron Classifier: A Multilayer Perceptron (MLP) is a class of feed-forward artificial neural networks composed of an input layer, one or more hidden layers, and an output layer. Information propagates unidirectionally through the network—from input to output—without any feedback connections. During the training process, the MLP optimizes its internal parameters (weights and biases) using the backpropagation algorithm, which computes gradients of the loss function and updates the weights to minimize the discrepancy between predicted outputs and actual target values.

The trained models are applied on the testing data portion of each iteration. Then, the test results are evaluated with the classification evaluation metrics for analysis purposes.

3. Evaluated Results

Fig. 2 illustrates the performance metrics of the machine learning models, displaying values for accuracy, precision, recall, and F1 score.

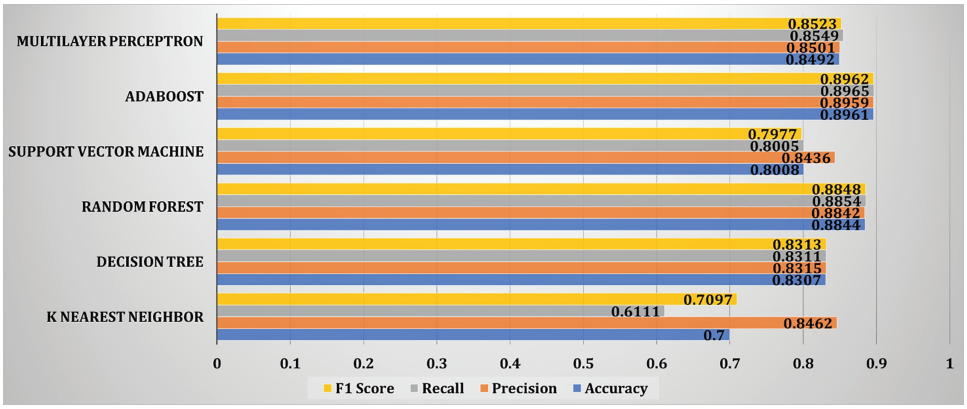

Fig. 2. Evaluated performance of ML models.

The AdaBoost technique has outperformed the other models. The AdaBoost classifier outperformed KNN, Decision Tree, Random Forest, SVM, and MLP in the experiment primarily due to its iterative boosting mechanism, which focuses sequentially on the misclassified instances by adjusting their weights, thereby improving the model's ability to handle complex decision boundaries and reduce bias. Unlike individual classifiers like Decision Trees or KNN, AdaBoost combines multiple weak learners to form a strong ensemble that is more robust to overfitting and noise. Additionally, compared to Random Forest and MLP, AdaBoost's adaptive weighting scheme allows it to concentrate learning on difficult samples, enhancing overall accuracy and generalization. Its ability to minimize training error iteratively while maintaining model simplicity often results in superior predictive performance.

4. Conclusion

In this research experiment, we have been able to detect profitable stocks with the help of Candlestick pattern using different classification approaches. Candlestick chart patterns have proved to be very accommodating in this task. The performances of the ML models have been quite satisfactory. We hope to conduct such researches using more datasets with many more classification techniques.

Conflict of Interest

The author declares that he does not have any conflict of interest.

References

[1] Palraj V, Sudheer K, Paramasivan S. A study on technical analysis using candlestick pattern of selected large cap stocks listed in National Stock Exchange (NSE), India with reference to steel sector. *GSI J Ser B Adv Bus Econ*. 2021;3(2):62–71.

[2] Iqbal MM, Roy JK. An analysis of day-of-the-week effects in Bangladesh stock market: evidence from dhaka stock exchange. *Daffodil Int Univ J Bus Econ*. 2015;9(1):13–25.

[3] Anand S, Chin WN, Khoo SC. Charting patterns on price history. *Proceedings of the Sixth ACM SIGPLAN International Conference on Functional Programming*, pp. 134–41, Florence, Italy, 2001 Sep.

[4] Yan Y, Yang D. A stock trend forecast algorithm based on deep neural networks. *Sci Program*. 2021;2021:1–9.

[5] Guo X, Liang X, Li X. A stock pattern recognition algorithm based on neural networks. *Proceedings of the Third International Conference on Natural Computation*, pp. 401–5, Haikou, China, 2007 Aug.

[6] Shen J, Shafiq MO. Short-term stock market price trend prediction using a comprehensive deep learning system. *J Big Data*. 2020;7:66.

[7] Xu Y, Li Z, Luo L. A study on feature selection for trend prediction of stock trading price. *2013 International Conference on Computational and Information Sciences*, pp. 1009–1012, Shiyang, China, 2013 Jun.

[8] Chen TL, Chen FY. An intelligent pattern recognition model for supporting investment decisions in stock market. *Inf Sci (Ny)*. 2016;346–347:261–74.

[9] Lin Y, Liu S, Yang H, Wu H, Jiang B. Improving stock trading decisions based on pattern recognition using machine learning technology. *PLoS One*. 2021;16(2):1–25. doi: 10.1371/journal.pone.0255558.

[10] Go YH, Hong JK. Prediction of stock value using pattern matching algorithm based on deep learning. *Int J Pure Appl Math*. 2018;120:10921–10938.

[11] Yahoo! Finance's API [Internet]. [Last Accessed on 15 May 2025]. Available from: https://pypi.org/project/yfinance/.

[12] Technical Analysis Library [Internet]. [Last Accessed on 15 June 2023]. Available from: https://mrjbq7.github.io/talib/.