# COMPARATIVE ANALYSIS OF STOCK MARKET PRICE BEHAVIOR THROUGH MACHINE LEARNING APPROACHES

KHAN, S. N.[1*] – KHAN, S. U.[2] – SHAFIQUE, O.[3] – ANSAR, Z.[3] – IMRAN, P.[3] – ALTAMISH, M. H.[3] – HAMZA, A.[3]

[1] *Department of Graphics Computer Vision and Digital Systems, Silesian University of Technology, Gliwice, Poland.*

[2] *Faculty of Electrical and Electronic Engineering, Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia.*

[3] *Department of Information and Technology, University of the Punjab Jhelum Campus, Punjab, Pakistan.*

**Corresponding author*
*e-mail: sndskhan87[at]gmail.com*

(Received 05th January 2025; revised 01st April 2025; accepted 10th April 2025)

**Abstract.** Stock prices reflect the inherent fluctuations of the stock market, making stock trading a challenging endeavor that often results in monetary loss. This research examines the ability of machine learning algorithms to predict stock prices, focusing specifically on K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Naive Bayes (NB). The study aims to predict stock prices and compare the outcomes of these algorithms based on their accuracy levels to determine the best-performing algorithm for future use. It addresses common limitations in similar studies, such as small sample sizes, limited time frames, and the neglect of external factors that can influence stock prices. By acknowledging these limitations, the study provides a more comprehensive and reliable analysis, highlighting the potential of machine learning algorithms for predicting stock market prices. However, it underscores the importance of considering economic indicators, market sentiment, and geopolitical events when applying machine learning to stock price prediction. This research demonstrates that while these algorithms have significant potential, their application must be guided by a careful evaluation of limitations and external influences to effectively reduce trading risks and enhance investment strategies.
**Keywords**: *machine learning, weighted rate, macro rate, KNN, SVM, NB*

## Introduction

Stock market forecasting has drawn interest from both economists and computer scientists as a classic yet difficult problem. Over the past two decades, both linear and machine learning methods have been investigated to create an efficient prediction model. Deep learning models have just been proposed as new horizons for this subject, and the research is moving too quickly for anybody to keep up. To provide a current evaluation of recent efforts on deep learning models for stock market prediction along with classifying the various data sources, neural network architectures, and widely used evaluation metrics, they also consider implementation and reproducibility (Chong et al., 2017). Machine learning is used to perform technical analysis on historical stock price data, and sentiment analysis is used to perform fundamental research on social media data. Social media data is more important than ever today and can help anticipate the direction of the stock market. The process entails gathering information from the news and social media and extracting sentiments that people have expressed. The relationship between emotion and stock values is then examined. Future stock value forecasts can


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

then be made using the taught model. The ability of the strategy to forecast stock performance, sentiment, and the correlation between recent news and social data can all be demonstrated (Attigeri et al., 2015). The survey provides the results about the future of the stock market prediction are a very difficult process, and several aspects should be considered (Gandhmal and Kumar, 2019). A promising field is the use of machine learning methods and other algorithms for stock price analysis and forecasting. Shah provides a brief overview of stock markets and a taxonomy of stock market prediction techniques. Then, the researchers concentrate on a few scientific developments in stock analysis and forecasting. They go over the short-, long-, and technical-term methodologies utilized for stock research. Finally, they outline some issues and future directions for this field of study (Shah et al., 2019).

The stock market forecast patterns are increasingly successful and are regarded as a crucial activity. Stock prices will thus result in significant rewards from wise investment choices. Stock market projections are a significant difficulty for investors due to the stale and noisy data. As a result, predicting the stock market presents a significant difficulty for investors looking to maximize their return on investment. Predictions of the stock market are made using mathematical techniques and study aids. A comprehensive assessment of 30 research publications that propose various approaches, such as computation techniques, machine learning algorithms, performance metrics, and top journals are presented in Kumars article. As a result, this chosen research is assisting in the discovery of machine learning methods and their dataset for stock market prediction. The most popular artificial neural networks (ANN) are employed to provide accurate stock market predictions. Despite significant effort, the most recent stock market-related prediction approach has several drawbacks. In this study, it may be assumed that stock market forecasting is a comprehensive process and that specific factor for forecasting the stock market need to be thought of as more accurate (Kumar et al., 2018).

### *Research statement and objectives*

Investors are very interested in accurate stock market forecasts, however, because the stock market is influenced by unstable elements like microblogs and news, it is difficult to forecast the stock market index using only historical data. The huge volatility of the stock price highlights how important it is to evaluate the influence coming from external factors on stock prediction. Machine learning algorithms can forecast stock markets using data from different datasets and financial news since these sources can influence the actions of investors. In this research, the researchers will employ algorithms to analyze datasets to determine how well stock market predictions for 10 days ahead will be accurate. They will use classifiers on the datasets to increase prediction accuracy and performance. Additionally, they will do tests to identify stock markets that are hard to forecast and those that are more impacted by financial news. To locate a reliable classifier, they will compare the outcomes of several algorithms. The experimental findings demonstrate that classifiers and financial news predict the New York Stock Exchange (NYSE), the greater effect of social media on New York and IBM stocks, and the greater influence of financial news on London and Microsoft stocks.

Due to outdated and noisy data, stock market predictions are a substantial challenge for investors. As a result, investors who want to get the best return on their investment find it extremely difficult to predict the stock market. To solve this problem, the researchers conduct a comparative analysis of stock market price behavior using


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

machine learning techniques. On multiple datasets (Google, IBM, Microsoft) of the stock market the researchers apply different classifiers (KNN, SVM, NB) to analyze the performance of algorithms to increase the accuracy of stock price. The stock market activity is unpredictable and subject to several influences. The major contribution of this research is based on the analysis of the classifiers which are described below as: (1) To analyze the machine learning algorithms to see how well stock market predictions will be accurate; and (2) To provide a comparative analysis of price behavior of the stock market by applying classifiers on the datasets to increase prediction accuracy and performance.

## *Related work*

Multiple studies have attempted to predict future price changes by studying the complicated and dynamic nature of the stock market. Machine learning techniques have been growing in popularity as a tool for predicting stock market prices because of their capacity for processing and analyzing massive volumes of data and identifying complex patterns. The efficiency and accuracy of these techniques in comparison to other traditional methods like technical and fundamental analysis are still up for debate. The purpose of this literature study is to compare and contrast various machine learning approaches that are used in stock market price prediction. Accurately predicting stock market price changes is a major economic benefit. Another strategy that has lately received a lot of attention is the use of machine learning to develop a predictive algorithmic model. The latter strategy must be used to educate robots to make trading judgments in such a short amount of time. Deep neural networks, the most outstanding invention in machine learning, were used to create a short-term prediction model. The reflective study intends to forecast stock prices in the short future. This assessment takes into account ten distinct stocks listed on the NYSE. The review primarily focuses on the forecast of these short-term prices using technical analysis which helps the framework to grasp the patterns in the previous prices given into it and attempts to probabilistically anticipate the stock's volatile future values. The study divides ANN into two types: feed-forward ANN and recurrent neural network (RNN). According to the assessment, feed forwards multilayer perceptron outperforms long near-term memory in forecasting stock values in the short term (Khare et al., 2017).

To forecast daily directional movements of a stock price utilizing financial news headlines and technical indicators as input are used in deep learning methods. The following technical indicators are compared: set 1: stochastic %K, stochastic %D, momentum, rate of change, William's %R, accumulation/distribution (A/D) oscillator, and disparity 5; set 2: exponential moving average, moving average convergence-divergence, relative strength index, on balance volume, and Bollinger bands. The study may find and analyze complex patterns and relationships in data, enabling more precise trading. Experiments have demonstrated that convolutional neural networks (CNN) are better than RNNs in extracting semantic information from texts, although RNNs are better at extracting context information and modelling complicated temporal aspects for stock market forecasting. Two models are also compared: SI-RCNN, a hybrid model made of a CNN for financial news and a long short-term memory (LSTM) network for technical indicators, and I-RNN, an LSTM network just for technical indicators. When the model predicts that the price will rise, the trading agent buys stocks on the current day and sells the next day; otherwise, the agent sells stocks on the current day and buys the following day. When comparing different sets of technical indicators, the suggested


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

technique reveals a significant role for financial news in stabilizing the findings and nearly no improvement (Vargas et al., 2018).

Forecasting algorithms may be classified into linear and non-linear models. This study uses four types of deep learning architectures to forecast a company's stock price based on past prices. The researchers used the closing prices of two separate stock exchanges, the National Stock Exchange (NSE) of India and the NYSE. The network was trained using the stock price of a single NSE business and forecasted the stock prices of five distinct NSE and NYSE companies. CNN has been shown to outperform the other models. Despite being trained on NSE data; the network was able to forecast for NYSE. This was achievable since both stock markets have some internal characteristics in common. When the acquired findings were compared to the linear model, it was discovered that the non-linear model outperformed the current linear model (Hiransha et al., 2018). The goal of stock market prediction is to forecast the future value of a company's financial stocks. The use of machine learning, which produces forecasts based on the values of current stock market indices by training on their prior values, is a new trend in stock market prediction technology. The research focuses on the application of regression and LSTM-based machine learning to forecast stock prices along with its factors (Althelaya et al., 2018). As a result, if it becomes feasible to expect the future direction of the stock market using proper procedures, investors might optimize their return on investment. This research aims to assess the predictive potential of machine-learning algorithms in a stock market. This analysis relied on daily close price data from the iShares MSCI UK exchange-traded fund from January 2015 to June 2018. Four machine-learning algorithm models are used in the prediction procedure. The findings show that the deep learning approach predicts better than the other methods, and the SVM method ranks second to the neural network and random forest methods in terms of error (Nikou et al., 2019).

Lee gives financial network metrics that may be used to develop global stock market investing strategies. Utilization of a vector auto-regressive model is, to create both undirected and directed volatility networks of the global stock market based on simple pair-wise correlation and system-wide connectivity of national stock indexes, an achievement in the prediction. Impact investigation and use of network indicators are the most valuable tasks as inputs for defining strategies in logistic regression, SVM, and random forest. Two techniques based on stock price indices are developed: (1) a worldwide stock market prediction approach and (2) a regional allocation strategy for developed/emerging markets. According to the findings of the performance analysis, network indicators are useful supplemental in predicting the global stock market and regional relative directions (up/down). These indicators were especially beneficial during market downturns. The first attempt is to develop methods for global portfolio management based on financial network indicators and to indicate how network indicators may be applied in practice (Lee et al., 2019).

Due to various key factors, the nature of stock market movement has always been confusing for investors. With machine learning algorithms, the studies intend to drastically minimize the risks in prediction. For experimental assessments, four stock market groups from the Tehran stock exchange are chosen: diversified financials, petroleum, non-metallic minerals, and basic metals. This research investigates and compares nine machine learning models and two powerful deep learning methods. The researchers input values are 10 technical indicators drawn from ten years of historical data, and they are designed to be used in two ways. First, the indicators are calculated using stock trading values as continuous data, and then secondly, they are converted to binary data before use. Based on the input methods, each prediction model is assessed using three metrics. The assessment findings show that for continuous data, deep learning methods beat other prediction models significantly. Furthermore, the findings reveal that those deep learning approaches are the best in binary data evaluation also; nevertheless, the difference becomes less significant due to the notable gain in model performance in the second way (Nabipour et al., 2020). Another researcher used the LSTM model to take the highest accuracy in prediction and CNN for the fastest speed execution (Mehtab and Sen, 2020).

Zhang's model uses a quantitative investment approach to predict stock prices. Both the backpropagation neural networks (BPNN) prediction model and the comparative analysis of the errors of various techniques are examined (Zhang and Lou, 2021). Based on the capabilities of three ANNs, this study proposed nine novel integrated models for forecasting intraday stock prices. The particle swarm optimization-BPNN model has the highest prediction accuracy in projecting intraday stock prices, according to the results. Other models delivered poor results in terms of prediction accuracy (Chandar, 2021; Kumar Chandar, 2021). Predicting stock market prices is a difficult and complex task, and researchers have investigated several methods for providing accurate predictions. The key methods, conclusions, and limitations from recent studies on stock market price prediction are summarized here in *Table 1*. The chosen studies on stock market price prediction using machine learning methods, as shown in *Table 1*. It provides a summary of each research method, conclusions, and limitations. These researches mostly focus on providing various successful predictions about the stock market prices of certain stocks or indices. Overall, the studies show how machine learning algorithms may be used to predict stock market prices, but they also emphasize the need to carefully take limitations and outside influences into consideration.

***Table 1**. Summary of the selected research articles along with their methods, outcomes, and limitations.*

| Ref. | Year | Methods | Objectives | Limitations |
|---|---|---|---|---|
| Khare | 2017 | Multi-Layer Perceptron and LSTM | Multi-Layer Perceptron outperforms the LSTM model in short-term stock price prediction.<br>Neural networks are effective tools for forecasting in a chaotic framework like the stock market. | Extend the model to work with tick-by-tick data for stock markets.<br>Build a platform to execute trades in real time based on these forecasts. |
| Vargas | 2018 | RCNN | SI-RCNN model outperforms SI-RCNN-2, which employs the first and second sets of indicators, respectively, while the IRNN model outperforms I-RNN-2. | Use of test techniques to create better embedding vectors for news headlines.<br>Train the suggested model on market simulation using a reinforcement learning technique. |
| Hiransha | 2018 | Multi-layer Perceptron, RNN, LSTM and CNN | CNN outperformed the other three networks because it can catch sudden changes in the system. | This study did not investigate the benefits of employing a hybrid network, which mixes two networks to create a prediction model. |
| Althelaya | 2018 | Regression and LSTM based Machine learning | The main contribution is the implementation of the LSTM model for stock price estimation, with highly encouraging findings. | Using a bigger dataset can further improve the accuracy of the stock market prediction system. |
| Nikou | 2019 | ANN, SVM, random forest, and LSTM | The SVM algorithm is more accurate than random forest and neural network methods.<br>Recurrent network with LSTM block outperforms other methods in predicting close price. | Future studies ought to take other influencing factors into account.<br>Compare the findings to the current research.<br>The use of time series data in this study prevented the examination of other influential factors on stock price prediction. |


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

| | | | | |
|---|---|---|---|---|
| Lee | 2019 | Logistic regression, SVM, and random forest | SVM outperformed logistic regression and random forest in the prediction model. With the use of non-linear connections between variables, SVM is more suited for market forecasting. | Model improvement or deep learning will be explored in the future as an alternative approach. |
| Nabipour | 2020 | Machine learning methods and two deep learning algorithms | RNN and LSTM, both deep learning algorithms, were found to be superior models in both techniques. | Despite efforts, no valuable studies on the same stock market were found. This deficiency is a novelty of the research. The study can serve as a baseline for future studies. |
| Mehtab | 2020 | CNN and LSTM | LSTM model: highest forecasting accuracy. CNN model: fastest execution speed. | Investigate the feasibility of developing GANs-based predictive models to further increase the forecasting accuracy. |
| Zhang | 2021 | Backpropagation and neural network | Quantitative investment, classification, and prediction of stock price patterns, analysis of errors. | More stock technical indicators can be improved, the prediction accuracy of the law, and external factors (economic development momentum, government policies, other emergencies). |
| Chandar, Kumar Chandar | 2021 | Nine new integrated models based on the potential of the ANNs | Forecasting of the intraday stock price through models and high prediction accuracy. | Enhance prediction accuracy, and reduce the computation time. |

## Materials and Methods

By analyzing the complex and dynamic nature of the stock market, several researchers have attempted to forecast future price changes. A rising number of people are using machine learning techniques to forecast stock market prices due to their ability to process and analyze large amounts of data and identify complex patterns. Consequently, the researchers provide a brief review of several machine learning forecasting algorithms with their comparative analysis in this section.

### *Support Vector Machine (SVM)*

The SVM is used in this study to anticipate the stock market. It is regarded as one of the best algorithms available for time series prediction. The supervised approach may be used for regression as well as classification. The SVM includes visualizing data as a point in an n-dimensional space. The separating hyperplane officially defines an SVM, which is a discriminative classifier. In other words, given labelled training data (supervised learning), the algorithm generates the best hyperplane for categorizing fresh samples. In 2D space, a hyperplane is a line that divides a plane into two sections, with each class lying on each side. These are the properties that are plotted on certain coordinates. As illustrated in *Figure 1*, the SVM algorithm builds a border across the data set called the hyper-plane, which divides the data into two groups. The hyper-plane is a decision boundary that may be stretched or maximized on any side of the data points (Reddy and Sai, 2018). In the same image, if µ is an unknown data point and w is a vector perpendicular to the hyper-plane, the SVM decision rule will be Eq. (1):


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

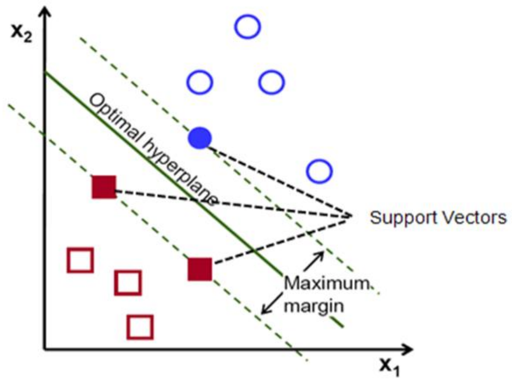


***Figure 1.*** *The SVM decision-making boundary.*

$$\bar{\omega}\bar{\mu} + b \geq 0 \qquad \text{Eq. (1)}$$

The width w of the hyper-plane must be maximized the spread in Eq. (2):

$$w = \left[\frac{2}{||w||}\right] = \left(max\left[\frac{2}{||w||}\right]\right) \qquad \text{Eq. (2)}$$

Applying Lagrange's multiplier as $\alpha_i$in Eq. (3):

$$L = 0.5\ ||w||^2 \rightarrow -\sum \alpha_i[y_i\ (\omega_i x_i + b) - 1]{=}\sum \alpha_i - 0.5 \sum i \sum j\ \alpha_i\alpha_j y_i y_j x_i x_j \qquad \text{Eq. (3)}$$

The updated decision rule will be Eq. (4):

$$(\sum \alpha_i\ y_i x_i)\mu + b \geq 0 \qquad \text{Eq. (4)}$$

High-dimensional and non-linearly separable data benefit most from the SVM. *Figure 2* illustrates the major steps of the SVM workflow.


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

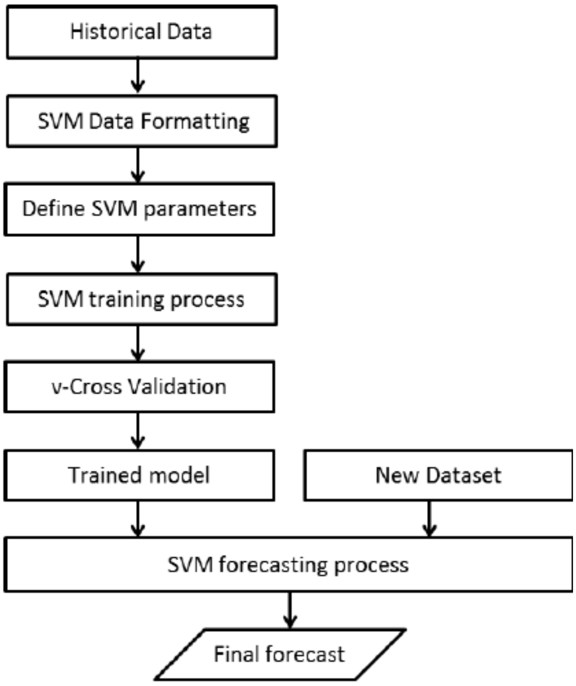


***Figure 2.*** *The working flow of the SVM.*

### *K Nearest Neighbors (KNN)*

Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units. An exception would be the use of English units as identifiers in trade, such as 3.5-inch disk drive. The KNN is a simple algorithm for predicting numerical targets based on a similarity metric. A classification system based on similarity can be used to map the stock prediction problem. A pair of vectors is created by mapping the test data and historical stock data together. Each vector refers to an N-dimensional stock feature. After that, a judgment is made by computing a similarity measure like the Euclidean distance (Bhattacharjee and Bhattacharja, 2019). KNN computes the average numerical goal of the K-closest neighbors. Eq. (5) is used to calculate the distance between two points:

$$D(x_i, y_i) = \left(\sum_{i=1}^{n} |x_i - y_i|^q\right)^{\frac{1}{q}} \quad \text{Eq. (5)}$$

The class label is then decided by a majority vote among the chosen k records, and it is subsequently applied to the query record. In *Figure 3,* a simple flow chart of KNN is shown. KNN is used to calculate the following closing price prediction for the stock market: (1) The number of nearby neighbors, *k*, should be calculated, (2) Calculate the separation between the query record and the training samples, (3) Assign the class labels of the KNN that received the most votes as the prediction value for the query record, (4) Organize all training records by distance values.


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

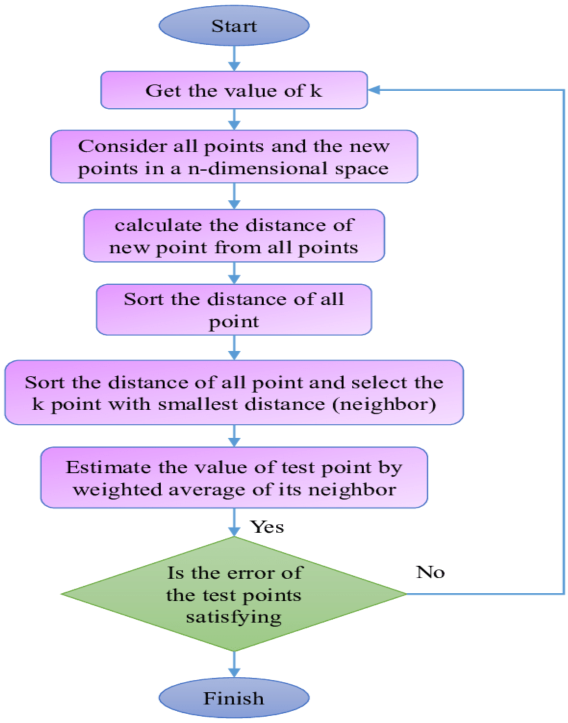


***Figure 3.*** *The working flow of KNN.*

## *Naïve Bayes (NB)*

This algorithm is a classification method that creates Bayesian networks for a given dataset based on the Bayes theorem. It is based on the assumption that the provided dataset has a specific characteristic in a class that is unconnected to any other feature. For instance, a certain property of an object causes it to be classified as A. The existence of these attributes may depend on one another or other features, but regardless of how they interact, they all individually improve the probability that this item is A, which is the reason it is said to as "Naïve." The NB method has the advantages of being simple to construct, effective for very big datasets, and even being able to outperform more advanced classification approaches (Mahajan Shubhrata et al., 2016). The Bayes algorithm estimates the probability of an occurrence based on past knowledge of conditions that could be relevant to the event in Eq. (6).

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)} \qquad \text{Eq. (6)}$$

Where; P(A/B)=The probability of A being true given that B is true; P(B/A)=The probability of B being true given that A is true; P(A)=The probability of A being true; P(B)=The probability of A being true. The important steps that need to be taken in this algorithm are listed: (1) A frequency table must be created from the provided dataset; (2) Calculate the probability of the occurrences and produce a table of possibilities using the probabilities; (3) Determine each class's posterior probability using Eq. (6); (4) The result of the prediction is the class with the highest posterior probability.


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

## Results and Discussion

Many studies have tried to predict future price movements by examining the complicated and dynamic character of the stock market. Machine learning techniques are increasingly being used to predict stock market prices because of their capacity to process and analyze vast amounts of data and recognize complicated patterns. In this section of the research, the researchers use machine learning algorithms namely KNN, SVM, and NB, and perform a comparative analysis to identify the best performance algorithm with the highest accuracy to predict the price of the stock market (*Table 2*). Researchers gathered monthly data sets of Google, Microsoft, and IBM stocks from January 2020 to December 2022 for our trials from Yahoo Finance, regarding Date, Open, Close, High, Low, and Adjacent Close are the dataset properties used in studies. An open, high, low, and close price of a financial instrument used to demonstrate change over time. (1) Date: This attribute represents the event's related date. (2) Open: This characteristic represents the opening price of the stock for the trading month. (3) Close: This feature represents the month's closing price of the stock deal. It is the reference point used by investors to compare the performance of a company over time. (4) Low: This feature represents the stock trade's lowest price during the month. (5) High: This property represents the highest price of the month's stock trading. (6) Adjacent close: This is used to keep track of or evaluate previous results.

***Table 2****. The result of KNN, SVM and NB in a comparative analysis.*

| Category | Google Dataset | Microsoft Dataset | IBM Dataset |
|---|---|---|---|
| Weighted average for SVM decision-making algorithm | | | |
| Precision | 0.33 | 0.61 | 0.24 |
| Recall | 0.58 | 0.52 | 0.49 |
| F1-score | 0.42 | 0.53 | 0.32 |
| Support | 45 | 89 | 147 |
| Macro average for SVM decision-making algorithm | | | |
| Precision | 0.29 | 0.55 | 0.24 |
| Recall | 0.50 | 0.55 | 0.50 |
| F1-score | 0.37 | 0.51 | 0.33 |
| Support | 45 | 89 | 147 |
| Weighted average for the KNN approach | | | |
| Precision | 0.59 | 0.62 | 0.54 |
| Recall | 0.60 | 0.52 | 0.54 |
| F1-score | 0.59 | 0.52 | 0.54 |
| Support | 68 | 134 | 221 |
| Macro average for the KNN approach | | | |
| Precision | 0.58 | 0.57 | 0.54 |
| Recall | 0.57 | 0.57 | 0.54 |
| F1-score | 0.57 | 0.52 | 0.54 |
| Support | 68 | 134 | 221 |
| Weighted average for the NB approach | | | |
| Precision | 0.33 | 0.44 | 0.21 |
| Recall | 0.58 | 0.60 | 0.46 |
| F1-score | 0.42 | 0.53 | 0.29 |
| Support | 45 | 89 | 147 |
| Macro average for the NB approach | | | |
| Precision | 0.29 | 0.33 | 0.23 |
| Recall | 0.50 | 0.50 | 0.50 |
| F1-score | 0.37 | 0.40 | 0.32 |
| Support | 45 | 89 | 147 |

### *SVM*

The researchers used SVM to predict the stock price of all three datasets. In terms of precision, SVM performed best for the Microsoft dataset for both weighted average and macro average. In terms of weighted average, *Table 2* shows SVM performed best for the Microsoft dataset with F1-score and precision of 0.52 and 0.61 respectively.


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

Whereas recall was higher for the Google dataset at 0.58. SVM gave the worst result for the IBM dataset with an F1-score and precision of 0.32 and 0.24 respectively. On the other hand, the same is true for the macro average. SVM again performed better on the Microsoft dataset with an F1-score of 0.51 (*Table 2*). Precision and recall were higher for the Microsoft dataset as well at 0.55 for both. SVM gave the lowest F1-score for the IBM dataset at 0.33.

### *KNN*

Of all the three algorithms, the KNN algorithm performed the best. It accurately predicted stock prices for both weighted average and macro average. In terms of weighted average, *Table 2* shows KNN algorithm gave a higher recall and F1-score for the Google dataset whereas precision was higher for the Microsoft dataset. Overall, it performed best on the Google dataset. KNN performed the same for IBM datasets for both macro and weighted averages. If the researchers look at the macro average, *Table 2* shows that KNN performed better on the Google dataset again with an F1-score of 0.57. Its precision and recall were at 0.58 and 0.57 respectively. For the IBM dataset, all three parameters were at 0.54 with support being 221.

### *NB*

The NB algorithm was the least-performing algorithm of all three algorithms mentioned. If the researchers glance at the weighted average, *Table 2* shows NB performed better on the Microsoft dataset with an F1-score of 0.53. It performed the worst on the IBM dataset with an F1-score of 0.29. Its precision and recall were at 0.21 and 0.46 respectively, which were the lowest of all three datasets. In terms of macro average, *Table 2* NB performed equally well on both the Google and the Microsoft data sets with a little better performance for the latter. Its F1-score was 0.4 and 0.37 for the Microsoft and Google datasets. Recall was the same for all datasets at 0.50.

### *Comparative analysis*

A comparison of all three algorithms is shown in *Table 3* and demonstrates that KNN fared best for accuracy across all three datasets, with the highest accuracy being 66.2% for the Microsoft dataset. However, SVM ultimately fell short of the level of prediction reached by KNN. While NB was more accurate in its predictions of the Microsoft dataset, it was less successful in its predictions of the other two datasets. In terms of prediction accuracy and performance, a comparison of machine learning methodologies has been conducted. Following an analysis of each method separately, it is discovered that the machine learning method KNN was the best-performing algorithm whereas SVM was slightly ahead of NB in terms of prediction accuracy.

***Table 3**. Comparison of classification accuracy of the machine learning approaches across the selected datasets.*

| Accuracy | Google | Microsoft | IBM |
|---|---|---|---|
| KNN | 60.9% | 66.2% | 61.9% |
| SVM | 59.2% | 61.5% | 52.7% |
| NB | 57.7% | 66.2% | 46.2% |


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

## Conclusion

The study's research contributions are noteworthy because they are relevant to the stock market and price behavior. Furthermore, this study aided in the improvement of classifier prediction accuracy and performance. This study addressed the significant issue of stock market prediction. First and foremost, both technical and fundamental analyses are considered. It proved difficult to predict stock prices using only historical data. Secondly, the researchers selected multiple stock market data sets and applied different classifiers to those datasets to analyze algorithm performance whereas SVM, NB, and KNN are examples of classifiers. The efficacy of these solutions has been demonstrated by outperforming experimental findings such as SVM, NB, and KNN. Using the three classifiers mentioned above and applying them to the data sets produced the following results: (1) The researchers compared the results of all algorithms and chose the best-performing algorithm among others with the highest prediction accuracy level. (2) Following that, the researchers improved the performance of the chosen algorithm to achieve greater accuracy in stock market price prediction.

In this study, the three proposed classifiers significantly improve prediction accuracy and other performance assessment criteria. The classifiers developed because of this research were tested. All the generated models outperform well-known methodologies, according to the experimental data. These experimental results show that using binary data instead of continuous data improves model performance significantly. In the prediction model that combined the price index and network indicators, SVM, a non-parametric technique, outperformed KNN and NB. It could be argued that the SVM approach is better suited for market forecasting via a non-linear relationship between variables. The future directions are to reduce the operational complexity, hybridization of the prediction model, empirical testing for larger datasets with dimensions or range, real-world optimization problems, and validation of accuracy with other state-of-the-art models.

## Acknowledgement

This research is self-funded.

## Conflict of interest

The authors declare that there is no conflict of interest involved in this research study.

## REFERENCES

[1] Althelaya, K.A., El-Alfy, E.S.M., Mohammed, S. (2018): Evaluation of bidirectional LSTM for short-and long-term stock market prediction. – In 2018 9th International Conference on Information and Communication Systems (ICICS), IEEE 6p.

[2] Attigeri, G.V., MM, M.P., Pai, R.M., Nayak, A. (2015): Stock market prediction: A big data approach. – In TENCON 2015-2015 IEEE Region 10 Conference, IEEE 5p.

[3] Bhattacharjee, I., Bhattacharja, P. (2019): Stock price prediction: a comparative study between traditional statistical approach and machine learning approach. – In 2019 4th International Conference on Electrical information and Communication Technology (EICT), IEEE 6p.


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620

[4] Chandar, S.K. (2021): Forecasting intraday stock price using ANFIS and bio-inspired algorithms. – International Journal of Networking and Virtual Organisations 25(1): 29-47.
[5] Chong, E., Han, C., Park, F.C. (2017): Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. – Expert Systems with Applications 83: 187-205.
[6] Gandhmal, D.P., Kumar, K. (2019): Systematic analysis and review of stock market prediction techniques. – Computer Science Review 34: 15p.
[7] Hiransha, M.E.A.G., Gopalakrishnan, E.A., Menon, V.K., Soman, K.P. (2018): NSE stock market prediction using deep-learning models. – Procedia Computer Science 132: 1351-1362.
[8] Khare, K., Darekar, O., Gupta, P., Attar, V.Z. (2017): Short term stock price prediction using deep learning. – In 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE 5p.
[9] Kumar, I., Dogra, K., Utreja, C., Yadav, P. (2018l): A comparative study of supervised machine learning algorithms for stock market trend prediction. – In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), IEEE 5p.
[10] Kumar Chandar, S. (2021): Stock price prediction based on technical indicators with soft computing models. – In Image Processing and Capsule Networks: ICIPCN 2020, Springer International Publishing 15p.
[11] Lee, J., Kim, R., Koh, Y., Kang, J. (2019): Global stock market prediction based on stock chart images using deep Q-network. – IEEE Access 7: 167260-167277.
[12] Mahajan Shubhrata, D., Deshmukh Kaveri, V., Thite Pranit, R., Samel Bhavana, Y., Chate, P.J. (2016): Stock market prediction and analysis using Naïve Bayes. – International Journal on Recent and Innovation Trends in Computing and Communication 4(11): 121-124.
[13] Mehtab, S., Sen, J. (2020): Stock price prediction using CNN and LSTM-based deep learning models. – In 2020 International Conference on Decision Aid Sciences and Application (DASA), IEEE 7p.
[14] Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., Salwana, E., S, S. (2020): Deep learning for stock market prediction. – Entropy 22(8): 23p.
[15] Nikou, M., Mansourfar, G., Bagherzadeh, J. (2019): Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. – Intelligent Systems in Accounting, Finance and Management 26(4): 164-174.
[16] Reddy, V.K.S., Sai, K. (2018): Stock market prediction using machine learning. – International Research Journal of Engineering and Technology (IRJET) 5(10): 1033-1035.
[17] Shah, S., Ismail, I., Shahrin, A. (2019): Stock market manipulation: a comparative analysis of East Asian emerging and developed financial markets. – Management Science Letters 9(1): 183-192.
[18] Vargas, M.R., Dos Anjos, C.E., Bichara, G.L., Evsukoff, A.G. (2018): Deep leaming for stock market prediction using technical indicators and financial news articles. – In 2018 International Joint Conference on Neural Networks (IJCNN), IEEE 8p.
[19] Zhang, D., Lou, S. (2021): The application research of neural network and BP algorithm in stock price pattern classification and prediction. – Future Generation Computer Systems 115: 872-879.


eISSN: 2716-6481
https://doi.org/10.55197/qjssh.v6i2.620