International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 6.4 Website: https://ijarmt.com ISSN No.: 3048-9458

Review Paper on Stock Price Prediction using Supervised Machine Learning Technique

Vicky Kumar
M. Tech. Scholar
Department of Computer Science and Engineering
MITS, Bhopal

Prof. Atul Kumar Mishra
Head of Dept.
Department of Computer Science and Engineering
MITS, Bhopal

Abstract
Stock market prediction has become a significant research domain due to its potential to support informed investment decisions, risk management, and financial planning. However, predicting stock prices remains challenging because of the dynamic, nonlinear, and highly volatile nature of financial markets. Supervised Machine Learning (ML) techniques have emerged as powerful tools for modeling complex price patterns by learning from historical, labeled stock data. This review paper provides a comprehensive analysis of various supervised ML algorithms used in stock price prediction, including Linear Regression, Support Vector Regression (SVR), Decision Trees, Random Forest (RF), Gradient Boosting methods such as XGBoost, and Neural Network-based models. The study examines the capabilities of these algorithms in capturing nonlinear market behaviors, handling noise, and improving prediction accuracy.
Key factors such as feature engineering, technical indicators, preprocessing strategies, and hyperparameter optimization are discussed to highlight their role in enhancing model performance. A comparative assessment from existing literature indicates that ensemble learning methods and deep neural models outperform traditional statistical approaches in forecasting accuracy and robustness. Despite significant progress, challenges such as market volatility, data sparsity, and susceptibility to overfitting persist. This review concludes that while supervised ML techniques cannot completely eliminate market unpredictability, they substantially enhance forecasting reliability and serve as effective tools for financial analysis and data-driven decision-making.
Keywords: - Machine Learning, Stock Market, Financial Planning

1. INTRODUCTION

The stock market plays a crucial role in the global economy, influencing financial growth, investment strategies, and corporate operations. Predicting stock prices, however, remains one of the most challenging tasks due to the market's nonlinear, highly volatile, and stochastic nature. Stock prices are influenced by numerous factors, including economic indicators, political events, company performance, investor sentiment, and global market trends. The complex interdependence of these variables makes traditional statistical forecasting models such as ARIMA, GARCH, and

Moving Average increasingly insufficient, as they rely on assumptions of linearity and stationarity that rarely hold true in real-world financial environments [1, 2].

In recent years, the rapid growth of data availability, computing power, and artificial intelligence has opened new possibilities for stock market forecasting. Among these advancements, Supervised Machine Learning (ML) techniques have gained significant attention due to their ability to learn intricate patterns from vast historical datasets. Supervised learning algorithms use labeled data—typically historical stock prices and technical indicators—to map inputs to outputs and generate predictive models capable of forecasting future trends. Techniques such as Linear Regression, Support Vector Regression (SVR), Decision Trees, Random Forest (RF), Gradient Boosting Machines (GBM), and Artificial Neural Networks (ANN) have been widely explored for this purpose [3].

Supervised ML models are particularly effective in capturing nonlinear relationships, reducing prediction errors, and demonstrating adaptability in highly dynamic market conditions. Ensemble learning methods like Random Forest and XGBoost offer improved robustness and generalization by combining multiple weak learners, while deep neural networks such as LSTM and GRU excel in modeling time-dependent patterns and long-term dependencies. These models outperform traditional statistical methods in many comparative studies, especially when enhanced through feature engineering, hyperparameter tuning, and appropriate data preprocessing techniques [4, 5].

Despite these strengths, stock price prediction remains inherently difficult due to the unpredictability of financial markets, sudden price fluctuations, and noise within the data. External shocks, such as economic crises or unexpected news events, can dramatically alter market behavior. Therefore, machine learning models must be designed with careful consideration of data quality, model complexity, and evaluation metrics to avoid overfitting and ensure reliability in real-time applications [6].

This review paper provides an in-depth analysis of supervised machine learning techniques used for stock price prediction. It highlights their operational mechanisms, advantages, limitations, and suitability for various financial forecasting tasks. The aim is to guide researchers and practitioners toward selecting appropriate models, understanding their performance characteristics, and identifying future research opportunities in the rapidly evolving domain of stock market prediction [7, 8].

2. LITERATURE REVIEW

Recent research on stock price prediction using machine learning and deep learning techniques shows a rapid evolution toward more accurate, hybrid, and data-driven forecasting systems. Barua et al. (2024) conducted a comparative analysis of deep learning models for the Indian stock market and demonstrated that architectures such as LSTM and GRU outperform traditional ML models due to their strong ability to capture temporal dependencies in financial time-series data. Similarly, Fathali et al. (2022) analyzed the NIFTY 50 index using machine learning models and highlighted that algorithms like Random Forest and Support Vector Regression can effectively reduce prediction error when appropriate technical indicators and feature selection strategies are applied. Piyush et al. (2023) further emphasized the potential of supervised learning algorithms in stock prediction and concluded that preprocessing techniques and hyperparameter tuning play a crucial role in improving model performance.

Advancements in deep learning have led researchers to explore hybrid architectures. Kanwal et al. (2022) proposed the BiCuDNNLSTM-1dCNN model, combining LSTM and CNN layers, and reported significant improvements in prediction accuracy due to the joint extraction of spatial and temporal features. Raviraj et al. (2021) also used deep learning, particularly LSTM networks, to model time-series data of Indian stock prices and verified their superiority over conventional regression models. Another dimension of research integrates textual sentiment and market news. Duarte et al. (2021) developed a predictive model based on news analytics, demonstrating that sentiment extracted from financial news can effectively detect stock price declines in the Brazilian market.

Apart from deep learning, probabilistic and generative approaches have also been explored. Su and Yi (2022) introduced an HMM-based model for stock prediction and showed that hidden Markov structures can capture latent market states effectively. Staffini (2022) utilized a Deep Convolutional Generative Adversarial Network (DCGAN) to forecast stock prices, proving that generative models can learn complex distributions of financial data and generate realistic price sequences. Moreover, Liu et al. (2021) investigated stock volatility factors using a FA-ANN-MLP model and concluded that hybrid feature-optimized neural networks significantly enhance volatility prediction. Finally, Lv et al. (2021) proposed a hybrid transfer learning framework for stock index forecasting, showing that transferring knowledge from related markets improves prediction accuracy, especially when training data is limited.

Overall, the literature indicates a strong shift from traditional ML methods toward deep learning and hybrid frameworks. Studies consistently highlight that integrating temporal modeling, feature engineering, sentiment analysis, and transfer learning significantly enhances prediction performance. This body of research confirms that supervised ML and deep learning are highly effective in capturing nonlinear, dynamic patterns in stock market data.

3. MACHINE LEARNING

Machine Learning is a subset of Artificial Intelligence concerned with “teaching” computers how to act without being explicitly programmed for every possible scenario. The central concept in Machine Learning is developing algorithms that can self-learn by training on a massive number of inputs. Machine learning algorithms are used in various applications, such as email filtering and computer vision, where it is difficult or infeasible to develop conventional algorithms to perform the needed tasks [4]. Machine learning enables the analysis of vast amounts of information. While it usually delivers faster, more precise results to identify profitable prospects or dangerous risks, it may also require additional time and assets to train it appropriately. Merging machine learning with AI and perceptive technologies can make it even more effective in processing vast volumes of information. Machine learning is closely associated with computational statistics, which focuses on making predictions using computers. Machine learning approaches are conventionally divided into three broad categories, namely Supervised Learning, Unsupervised Learning & Semi-supervised Learning, depending on the nature of the "signal" or "feedback" available to the learning system.

Machine Learning (ML) has emerged as a transformative technology in the stock market, offering advanced tools for analyzing complex financial data and predicting future price movements. Stock markets are highly dynamic and influenced by a wide range of factors, including company fundamentals, macroeconomic indicators, investor sentiment, and global events. Traditional statistical methods often struggle to capture these nonlinear and rapidly changing patterns. ML addresses this challenge by learning from historical data, identifying hidden trends, and generating predictive models that adapt to market behavior.

In stock market applications, ML algorithms are used for a variety of tasks such as price prediction, trend classification, volatility estimation, portfolio optimization, algorithmic trading, and risk assessment. Supervised learning techniques—such as Linear Regression, Support Vector Machines (SVM), Random Forest, Gradient Boosting Models, and Neural Networks—learn from labeled historical data (e.g., past prices, technical indicators) to forecast future stock values. These models can handle complex, non-linear relationships and often outperform traditional time-series models when trained on high-quality data. Deep learning approaches, especially LSTM and GRU networks, are particularly effective for time-series forecasting because they capture long-term dependencies and temporal patterns in financial data.

Supervised Learning

A model is trained through a process of learning in which predictions must be made and corrected if those predictions are wrong. The training process continues until a desired degree of accuracy is reached on the training data. Input data is called training data and has a known spam / not-spam label or result at one time.

Unsupervised Learning

By deducting the structures present in the input data, a model is prepared. This may be for general rules to be extracted. It may be through a mathematical process that redundancy can be systematically reduced, or similar data can be organized. There is no labeling of input data, and there is no known result.

Semi-Supervised Learning

Semi-supervised learning fell between unsupervised learning (without any labeled training data) and supervised learning (with completely labeled training data). There is a desired problem of prediction, but the model needs to learn the structures and make predictions to organize the data. Input data is a combination of instances that are marked and unlabeled.

4. STOCK MARKET

A SM commonly known as Equivalent market is a public market which operates virtually to ensure smooth buying and selling of stocks at a reasonable price and are often governed by an Apex body that regulates all the transaction and sets up procedures for effective trade. Stocks are the units of ownership in a company and are also known as equities. In order to impact development these SM act as a facilitators for companies to access funds from the common public investment. The stock market can serve two very important objectives:

- Providing Capital to companies which they can utilize to expand their business flow. Lets assume a company issues 10 million stocks priced at Rs 10 each, then this has the potential to make 100 million for the company in form of public investment and this can add huge amount of working capital to the business, it can easily be used to fund operations and affect growth.
- Giving Investors an opportunity to trade public stocks and share company profitability, making this trade a win -win for both Investors and companies.

Although stock trading is an old age concept, the modern day stock trading began to spread its roots with the formation of East India company from London. A stock bazaar or a stock market or even equivalent market can be defined as the assemblage of those getting goods or service and those rendering this or making this possible. (Or a loose connection of transactions which are only economic and there is no physical institution or even digital entrance of goods and services); some of these could be safely written on goods exchange and still to those not sold in public. Supplies could be classified in different ways [17, 18]. One major method is by the residential country approach in which the case study is Nestle and another one called Novartis from Switzerland, and thus they could be regarded as a portion of Swiss supply market, though this supply could also be exchanged with neighboring or even countries from other regions.

5. CONCLUSION

Stock price prediction remains a challenging task due to the highly volatile, nonlinear, and dynamic nature of financial markets. This review highlights that supervised machine learning techniques have significantly advanced the accuracy and reliability of stock forecasting by learning complex patterns from historical data. Traditional methods such as Linear Regression and Support Vector Regression provide a baseline for prediction, but ensemble models like Random Forest and Gradient Boosting, along with deep learning architectures such as LSTM, GRU, and hybrid CNN-LSTM frameworks, demonstrate superior performance in capturing temporal dependencies and market fluctuations.

The literature clearly shows that model performance depends heavily on high-quality data preprocessing, appropriate feature engineering, and effective hyperparameter optimization. Incorporating alternative data sources such as news sentiment, technical indicators, and macroeconomic variables further enhances model robustness. While machine learning techniques cannot fully eliminate uncertainty caused by sudden market shocks or external events, they considerably strengthen data-driven investment decisions, risk management strategies, and financial forecasting.

Overall, supervised machine learning emerges as a powerful tool for stock price prediction, offering improved generalization, adaptability, and predictive accuracy compared to traditional statistical approaches. Future research should focus on integrating hybrid models, transfer learning, and explainable AI techniques to further enhance interpretability, stability, and real-world applicability in diverse market conditions.

REFERENCES

[1] Moumita Barua, Teerath Kumar, Kislay Raj and Arunabha M. Roy, "Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market", MDPI, FinTech, 3, 551–568, 2024.

[2] Fathali, Z.; Kodia, Z.; Said, L.B. Stock Market Prediction of NIFTY 50 Index Applying Machine Learning Techniques. Appl. Artif. Intell. 2022, 36, 2111134.

[3] Piyush; Amarjeet; Sharma, A.; Kumar, S.; Ansari, N.N. Stock Price Prediction Using Machine Learning. In Proceedings of the International Conference on Recent Trends in Computing, Jammu, India, 26–27 October 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 79–87.

[4] Kanwal, A.; Lau, M.F.; Ng, S.P.; Sim, K.Y.; Chandrasekaran, S. BiCuDNNLSTM-1dCNN—A hybrid deep learning-based predictive model for stock price prediction. Expert Syst. Appl. 2022, 202, 117123.

[5] Shravan Raviraj, Manohara Pai M M. and Krithika M Pai, "Share price prediction of Indian Stock Markets using time series data - A Deep Learning Approach", IEEE Mysore Sub Section International Conference (MysuruCon), IEEE 2021.

[6] J. J. Duarte S. M. Gonzalez and J. C. Cruz "Predicting stock price falls using news data: Evidence from the brazilian market", Computational Economics vol. 57 no. 1 pp. 311-340 2021.

[7] Zhi Su and Bo Yi, "Research on HMM-Based Efficient Stock Price Prediction", Mobile Information System, Hindawi, 2022.

[8] Alessio Staffini, "Stock Price Forecasting by a Deep Convolutional Generative Adversarial Network", Frontiers in Artificial Intelligence, 2022.

[9] Jingqi Liu;Xinzhen Pei;Junyan Zou, "Analysis and Research on the Stock Volatility Factors of Chinese Listed Companies Based on the FA-ANN-MLP Model", International Conference on Computer, Blockchain and Financial Development (CBFD), IEEE 2021.

[10] Chenkang Lv;Boyong Gao;Cui Yu, "A hybrid transfer learning framework for stock price index forecasting", IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, IEEE 2021.

[11] G. Ding and L. Qin "Study on the prediction of stock price based on the associated network model of lstm" International Journal of Machine Learning and Cybernetics vol. 11 no. 6 pp. 1307-1317 2020.

[12] S. T. Z. De Pauli M. Kleina and W. H. Bonat "Comparing artificial neural network architectures for brazilian stock market prediction" Annals of Data Science vol. 7 no. 4 pp. 613-628 2020.

[13] L. Qi M. Khushi and J. Poon "Event-driven lstm for forex price prediction" 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE) pp. 1-6 2020.

[14] D. Wu X. Wang J. Su B. Tang and S. Wu "A labeling method for financial time series prediction based on trends" Entropy vol. 22 no. 10 pp. 1162 2020.

[15] Zhihao PENG, "Stocks Analysis and Prediction Using Big Data Analytics", International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), IEEE 2019.A. Site D. Birant and Z. Isik "Stock market forecasting using machine learning models" 2019 Innovations in Intelligent Systems and Applications Conference (ASYU) pp. 1-6 2019.