ISSN No- 2584-2706

Forecasting Stock Market Trends: A Comparative Study of Stocknet and Traditional Predictive Algorithms

Vaduguru Venkata Ramya
Gautam Dasaka

Abstract

This research proposal examines the performance output of machine learning (ML) and deep learning (DL) models in predicting stock prices, with a focus on model architecture and training outcomes. Nine models, including traditional ML algorithms and novel DL architectures, are evaluated on a dataset comprising Indian stock market data. Notably, Linear Regression, Random Forest, and BiLSTM exhibit superior performance on both training and testing datasets, with low Mean Squared Error (MSE) scores and high predictive accuracy (R^2). Linear Regression achieves a train MSE of 20.59 and a test MSE of 16.21. Decision Tree achieves a train MSE close to zero but significantly higher test MSE at 50.90. The Support Vector Machine shows high MSEs at 283,842 for training and 274,734 for testing, indicating poor generalization. Random Forest demonstrates competitive performance with a train MSE of 5.68 and a test MSE of 30.26. XGBoost, although performing well on training data (MSE: 5.50), struggles on unseen data (MSE: 141.47), indicating overfitting. Among the DL models, BiLSTM exhibits promising results with a train MSE of 8.22 and a test MSE of 0.64. Our proposed architecture, StockNet, shows encouraging performance with a train MSE of 8.06 and a test MSE of 0.07. However, traditional LSTM performs slightly worse, with a train MSE of 7.54 and a test MSE of 0.00. Lastly, the ResNLS Architecture performs well with a train MSE of 0.00 and a test MSE of 0.07. These findings underscore the importance of model selection and architecture design in stock price prediction tasks and provide insights into the strengths and limitations of various ML and DL approaches in financial forecasting.

Keywords - Stock price prediction, Machine learning, deep learning, Model evaluation, Indian stock market.

Introduction - The financial market is an exciting field for students and researchers, offering opportunities to develop skills in trading, analysis, and professional expertise. Investors study the market in various ways, such as analysing market behaviour, identifying key factors, and predicting market trends [1]. The equity market, with its many stock exchanges worldwide, constantly sees share prices fluctuate due to the forces of supply and demand [2, 3]. In this ever-changing environment, being able to predict future market events is very valuable. One way to achieve this is by using machine learning (ML) and artificial intelligence (AI) for stock market predictions [4 - 6]. Artificial neural networks (ANNs), a part of ML research, have been used in many areas, including finance and signal processing. However, traditional ANN algorithms sometimes struggle with stock market predictions due to issues like local optima [5].Deep learning (DL), a branch of ML and AI, excels at handling large datasets and recognizing complex patterns [7]. DL's strength lies in its ability to learn features automatically and model abstract concepts at multiple levels, which shows promise for stock market forecasting. Recent research highlights

DOI: https://doi.org/10.5281/zenodo.15110587

the growing interest in applying DL to stock prediction, inspired by its success in other time series data fields like speech recognition [8].This study focuses on using DL techniques to predict stock market movements using data from India's National Stock Exchange (NSE). By examining the detailed dynamics of stock prices and market behaviour, we aim to enhance the understanding of ML and DL applications in financial forecasting. Our goal is to provide insights that can guide investment strategies and deepen the comprehension of the financial market's complexities.

I. **Literature Review**

In recent years, significant attention has been given for the development and application of predictive models for stock market prices. Consequently, researchers have explored various methodologies, ranging from traditional statistical approaches to cutting-edge machine learning and deep learning techniques. This literature review synthesizes key findings from a variety of studies investigating the effectiveness of these methods, putting light on their implications for enhancing stock market forecasting accuracy.

Authors in their research investigated the use of nonlinear models and machine learning techniques such as neural networks and support vector machines (SVM) for stock price prediction [9–13]. Their findings include successful application of these techniques in predicting stock price time series, with insights into methodologies, ML algorithms, and significant features. In conclusion, they suggest the potential of these approaches in improving stock market forecasting accuracy.

In another research, authors implemented deep learning algorithms including LSTM and CNN with heuristic optimization in stock price prediction [14]. Their findings revealed LSTM's higher total profit compared to CNN, with certain combinations of DL models yielding specialized profits in investing simulations. They concluded by emphasizing the effectiveness of DL algorithms in capturing complex patterns in stock market data, leading to improved prediction outcomes.

Authors introduced hybrid models incorporating ANFIS, SVM, and ABC optimization for stock price forecasting [15] in their research. There findings demonstrated the effectiveness of these techniques in integrating technical indicators and reducing prediction errors. In conclusion, they suggested the potential of hybrid approaches in enhancing stock market prediction accuracy by leveraging diverse factors.

Innovative strategies like DLSMP and SPS, integrating DL architectures with technical indicators and news sentiments for asset value prediction [16, 18], are proposed by the authors in their research. Their findings indicated the superior performance of DL models over baselines, with the suggested strategies outperforming traditional methods. In conclusion, the authors highlight the significance of incorporating diverse data sources and employing advanced DL techniques for improved stock market forecasting.

Authors introduced advanced techniques like IPSO for hyperparameter tuning in LSTM models and fuzzy time series-based forecasting [21, 23]. Their findings demonstrated the superiority of IPSO in setting LSTM hyperparameters and the effectiveness of fuzzy time series in predicting nonlinear financial data. The research concluded with the importance of innovative approaches in addressing challenges in stock price prediction, leading to enhanced forecasting accuracy.

The performance of various ML and DL models, such as ANN, CNN, and BiLSTM, for stock market prediction [19, 26] is evaluated by the authors in their study. Their findings indicated CNN's higher accuracy compared to ANN and BiLSTM's superior performance over other RNN variants. In conclusion, the authors suggested the suitability of specific models for different prediction tasks,

DOI: https://doi.org/10.5281/zenodo.15110587

emphasizing the importance of model selection in achieving accurate forecasts.

II Aim and Objective

Aim

To develop and evaluate robust and adaptive predictive models for accurately forecasting stock market prices, leveraging advanced machine learning and deep learning techniques, including hybrid approaches, and incorporating diverse data sources to enhance prediction accuracy and reliability across different market conditions.

Objectives

1. Model Evaluation: To systematically evaluate the effectiveness of various machine learning and deep learning algorithms in forecasting stock market prices, comparing their performance to traditional statistical methods.
2. Hybrid Model Development: To investigate and develop hybrid models that combine multiple machine learning and deep learning approaches, assessing their performance relative to individual models in predicting stock market trends.
3. Data Integration: To explore the impact of integrating diverse data sources, such as technical indicators and global market indices, on the accuracy and effectiveness of stock market predictions.
4. Performance Optimization: To conduct extensive hyperparameter tuning and performance analysis of the predictive models to optimize their accuracy and reduce overfitting, ensuring consistent performance across different market conditions.

I. **Research Methodology**

This research methodology outlines the detailed steps and processes involved in analyzing and predicting stock prices using various machine learning (ML) and deep learning (DL) models. The dataset consists of stock price data for multiple companies, with a specific focus on the stock prices of Reliance. The methodology includes data preprocessing, exploratory data analysis (EDA), feature engineering, model training, evaluation, and comparison of different

ML and DL models.

3.1: **Data Loading and Initial Exploration**

1. Loading the Data:

- The dataset is loaded from a CSV file into a Data Frame for analysis.
- Initial exploration involves checking the shape of the dataset, the names of the columns, the presence of duplicates, and missing values.

2. Exploring the Dataset:

- Shape of the Dataset: The number of rows and columns of the dataset is: (4094387, 8)
- Column Names: The features available for analysis are: 'Date', 'Stock', 'Open', 'High', 'Low', 'Close', 'Volume', 'Change Pct.
- Checking for Duplicates: To ensure data integrity by checking for duplicate records. There are no duplicate records present in dataset.
- Checking for Missing Values: Identification and handling any missing values to prevent issues during analysis. There are no null values present in the dataset.
- Dataset Information: Gathering information about data types and non-null counts for each column.
- Statistical Summary: A summary of the dataset's statistics, such as mean, median, standard deviation, etc.
- Unique Stock Count: The number of unique stocks in the dataset are 1749.
- Top 5 Stocks by Frequency: The five most frequently occurring stocks in the dataset are:

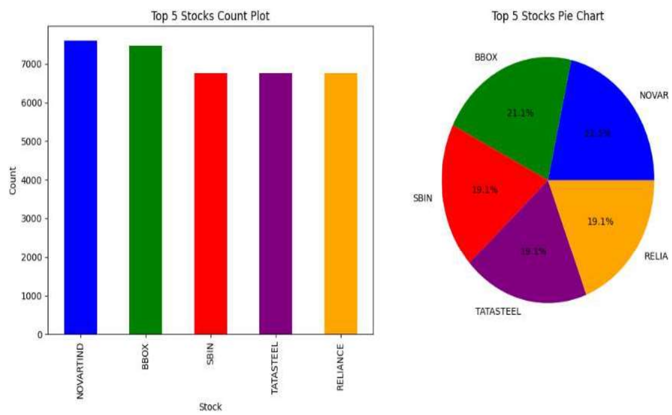
2.1 Exploratory Data Analysis(EDA)

Step 1. Visualizing Top 5 Stocks:

- Bar Plot: A bar plot to visualize the frequency of the top 5 stocks.
- Pie Chart: A pie chart to visualize the proportion of the top 5 stocks in the dataset.

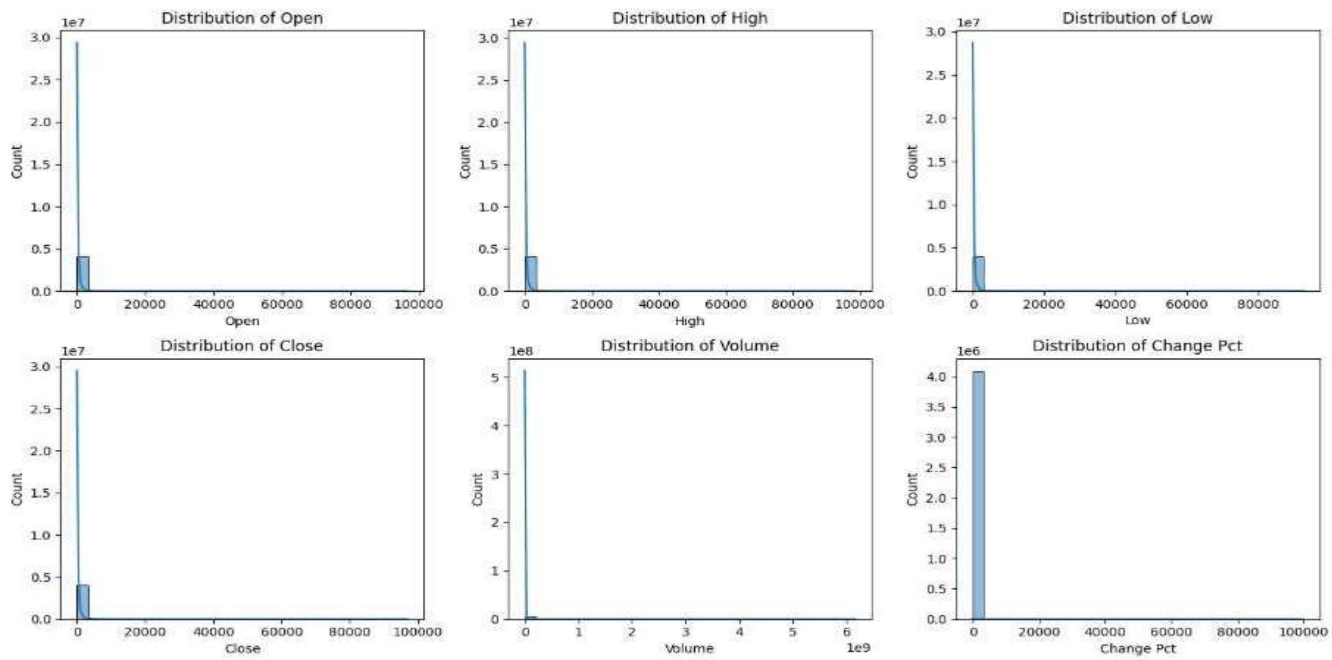
DOI: https://doi.org/10.5281/zenodo.15110587

•



Distribution of Numerical Features:

Histogram with KDE: Generate histograms with kernel density estimates (KDE) to visualize the distribution of numerical features such as Open, High, Low, Close, Volume, and Change Pct



3: Data Preprocessing

1. Filtering Data for a Specific Stock The dataset is filtered to focus on the stock prices of Reliance for a detailed analysis.
2. Time Series Analysis:

- Converting 'Date' to Date time: Convert the 'Date' column to date time format for time series analysis.
- Visualizing Stock Prices Over Time:
- Close Price: Plotting the close prices over time to observe trends and patterns.
- Open Price: Plotting the open prices over time to observe trends and patterns.
- High Price: Plot the high prices over time to observe trends and patterns.
- Low Price: Plot the low prices over time to observe trends and patterns.
- Volume: Plotting the trading volume over time to observe trends and patterns.

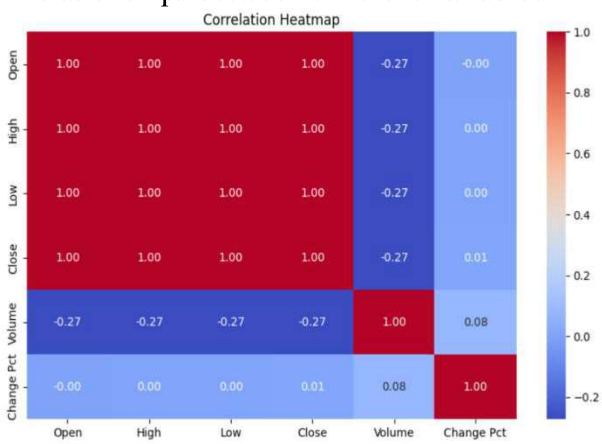
3. Year-wise Aggregation:

DOI: https://doi.org/10.5281/zenodo.15110587

- Maximum Values: Aggregation of the data by year to find the maximum values for each year.
- Minimum Values: Aggregation of the data by year to find the minimum values for each year.

4. Correlation Analysis:

- Heatmap: A heatmap is created to visualize the correlation between numerical features, which helps in understanding the relationships between different variables.



Step 4: Feature Selection and Data Splitting

1. Selecting Features and Target Variable:

- Identifying relevant features for prediction and the target variable, which is the closing price of the stock.
- Preparing the feature set (X) and target variable (y).

2. Splitting Data into Training and Testing Sets:

- Splitting the dataset into training and testing sets to evaluate the model's performance on unseen data. A typical split is 80% for training and 20% for testing.

Step 5: Model Training and Evaluation

In this paper, we develop the multi-model where we used algorithms like Linear Regression, Decision Tree, Support Vector Machine (SVM), Random Forest, XGboost, Long Short-Term Memory (LSTM), Bidirectional LSTM(BiLSTM) , and ResNets

I. Reseult Analysis

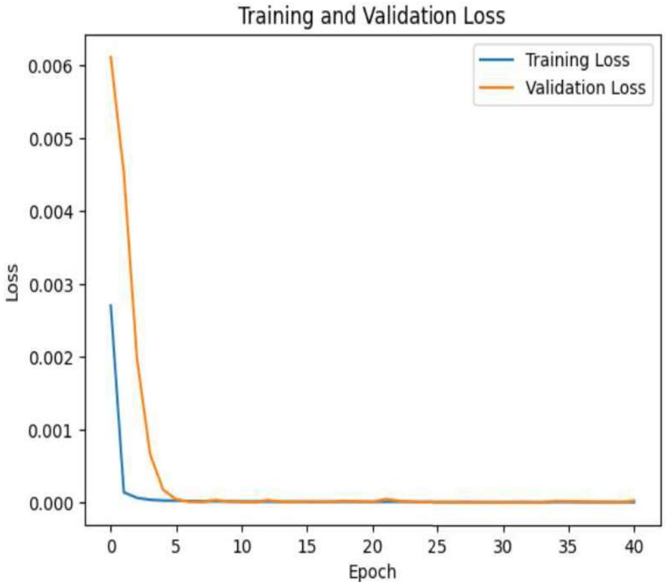
Scores of ML Model.

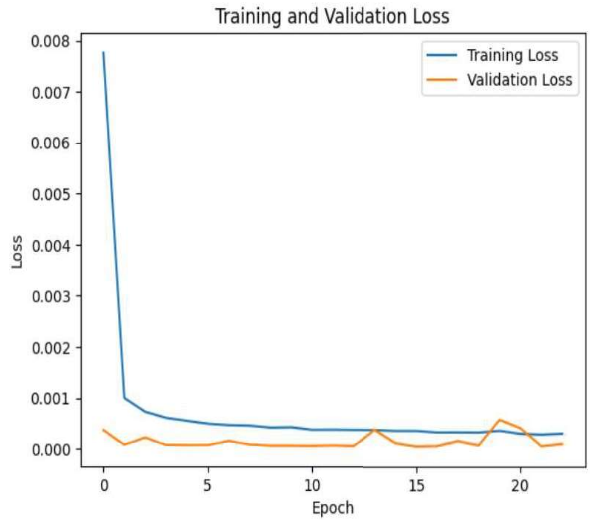
Model	Train MSE	Test MSE
Linear Regression	20.590445688311544	16.20969020627725
Decision Tree	7.001724010837383e.33	50.89520717455622
Support Vector	282842.0344986097	274734.0638769868
Random Forest	5.675907350789453	30.262481510021896
XGBoost	5.502510702478244	141.47446985417758

Score of DL Model

Model	Train MSE	Test MSE
LSTM	7.538344544605806e-06	0.0006653494401058559
BiLSTM	8.22429443473728e-06	0.0006376361814938557
ResNLS Architecture	8.055745582125002e-06	0.06800398842944297
Our Architecture	2.006282315323118e-05	0.0013173271723083625

DOI: https://doi.org/10.5281/zenodo.15110587





Conclusion

In this study, we evaluated several machine learning and deep learning models to predict stock prices. The models included Linear Regression, Decision Tree, Support Vector Machine (SVM), Random Forest, XGBoost, LSTM, BiLSTM, ResNLS Architecture, and our custom StockNet architecture. The performance of each model was assessed based on Mean Squared Error (MSE) and R^2 scores on both training and testing datasets.Among the models tested, Linear Regression, Random Forest, and BiLSTM demonstrated excellent predictive capabilities with low MSE and high R^2 scores, indicating strong model performance and good generalization to unseen data. LSTM and the custom StockNet architecture also showed promising results, especially in capturing temporal dependencies and specific stock price characteristics.On the other hand, the Decision Tree and XGBoost models exhibited signs of overfitting, performing exceptionally well on training data but less effectively on testing data. The SVM model, however, struggled with this dataset, displaying high MSE and low R^2 scores, indicating poor performance in capturing the underlying patterns.

Future work

In future work, extensive hyper parameter tuning should be conducted for all models to optimize their performance and potentially reduce overfitting. Additionally, exploring additional features, such as sentiment analysis from news articles, macroeconomic indicators, and technical indicators, could improve model accuracy. Investigating more advanced deep learning architectures, such as Transformers and hybrid models, will further enhance predictive performance. Developing and evaluating ensemble methods that combine predictions from multiple models could improve overall accuracy and robustness. Regularization techniques should be implemented to prevent overfitting, particularly for models prone to this issue, such as Decision Trees and XGBoost. Using cross-validation techniques will ensure that models are evaluated more robustly, leading to more reliable performance estimates. Moreover, methods for real-time stock price predictions should be developed, integrating models with streaming data to provide up-to-date forecasts. Finally, focusing on improving the interpretability and explainability of complex models will help better understand the decision-making process and gain insights into the factors influencing stock prices.

DOI: https://doi.org/10.5281/zenodo.15110587

References

[1] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock closing price prediction using machine learning techniques," *Procedia Computer Science*, vol. 167, pp. 599–606, 2020.

[2] W. Khan, M. Ghazanfar, M. Azam, et al., "Stock market prediction using machine learning classifiers and social media, news," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 3433–3456, 2022, doi: 10.1007/s12652-020-01839-w.

[3] . M. Nabipour, P. Nayyeri, H. Jabani, S. Shahab, and A. Mosavi, "Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis," *IEEE Access*, vol. 8, pp. 150199–150212, 2020, doi: 10.1109/ACCESS.2020.3015966.

[4] M. Nikou, G. Mansourfar, and J. Bagherzadeh, "Stock price prediction using deep learning algorithm and its comparison with machine learning algorithms," *Intelligent Systems in Accounting, Finance and Management*, vol. 26, no. 4, pp. 164–174, 2019.

[5] D. K. Sharma, H. S. Hota, K. Brown, and R. Handa, "Integration of genetic algorithm with artificial neural network for stock market forecasting," *International Journal of System Assurance Engineering and Management*, vol. 13, suppl. 2, pp. 828–841, 2022, doi: 10.1007/s13198-021-01209-5.

[6] H. H. Htun, M. Biehl, and N. Petkov, "Survey of feature selection and extraction techniques for stock market prediction," *Financial Innovation*, vol. 9, no. 1, p. 26, 2023.

[7] W. Jiang, "Applications of deep learning in stock market prediction: recent progress," *Expert Systems with Applications*, vol. 184, p. 115537, 2021.

[8] R. Singh and S. Srivastava, "Stock prediction using deep learning," *Multimedia Tools and Applications*, vol. 76, pp. 18569–18584, 2017, doi: 10.1007/s11042-016-41Payal.

[9] . R. Jamous, H. ALRahhal, and M. El-Darieby, "A new ann-particle swarm optimization with center of gravity (ann-psocog) prediction model for the stock market under the effect of covid 19," *Scientific Programming*, 2021, pp. 1–17, doi: 10.1155/2021/6656150.

[10] A. Thakkar and K. Chaudhari, "Fusion in stock market prediction: a decade survey on the necessity, recent developments, and potential future directions," *Information Fusion*, vol. 65, pp. 95–107, 2021.

[11] D. Kumar, P. K. Sarangi, and R. Verma, "A systematic review of stock market prediction using machine learning and statistical techniques," *Materials Today: Proceedings*, vol. 49, pp. 3187–3191, 2022

[12] L. N. Mintarya, J. N. Halim, C. Angie, S. Achmad, and A. Kurniawan, "Machine learning approaches in stock market prediction: a systematic literature review," *Procedia Computer Science*, vol. 216, pp. 96–102, 2023.

[13] C. A. Krishnapriya and A. James, "A survey on stock market prediction techniques," in *2023 International Conference on Power, Instrumentation, Control and Computing (PICC)*, 2023, pp. 1-6.

[14] C. Stoean, W. Paja, R. Stoean, and A. Sandita, "Deep architectures for long-term stock price prediction with a heuristic-based strategy for trading simulations," *PLoS ONE*, vol. 14, no. 10, p. e0223593, 2019.

[15] M. Sedighi, H. Jahangirnia, M. Gharakhani, and S. Farahani Fard, "A novel hybrid model for stock price forecasting based on metaheuristics and support vector machine," *Data*, vol. 4, no. 2, p. 75, 2019.

[16] M. Nabipour, P. Nayyeri, H. Jabani, et al., "Deep learning for stock market prediction," *Entropy*, vol. 22, no. 8, pp. 1–23, 2020.

[17] X. Pang, Y. Zhou, P. Wang, W. Lin, and V. Chang, "An innovative neural network

approach for stock market prediction," *Journal of Supercomputing*, vol. 76, pp. 2098–2118, 2020.
[18] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: a case of Hong Kong," *Image Processing Management*, vol. 57, no. 5, pp. 1–19, 2021
[19] Y. Ji, A. W. C. Liew, and L. Yang, "A novel improved particle swarm optimization with long-short term memory hybrid model for stock indices forecast," *IEEE Access*, vol. 9, pp. 23660–23671, 2021
[20] S. Albahli, T. Nazir, A. Mehmood, A. Irtaza, A. Alkhalifah, and W. Albattah, "AEI-DNET: a novel densenet model with an autoencoder for the stock market predictions using stock technical indicators," *Electronics*, vol. 11, no. 4, p. 611, 2022.
[21] W. Hussain, J. M. Merigo, and M. R. Raza, "Predictive intelligence using ANFIS-induced OWAWA for complex stock market prediction," *International Journal of Intelligent Systems*, vol. 37, no. 8, pp. 4586–4611, 2022.
[22] P. Chhajer, M. Shah, and A. Kshirsagar, "The applications of artificial neural networks, support vector machines, and long short-term memory for stock market prediction," *Decision Analytics Journal*, vol. 2, p. 100015, 2022.
[23] A. Bhambu, "Stock Market prediction using deep learning techniques for short and long horizon," in *Soft Computing for Problem Solving: Proceedings of the SocProS 2022*, Singapore: Springer Nature Singapore, 2023, pp. 121-135.
[24] G. Sonkavde, D. S. Dharrao, A. M. Bongale, S. T. Deokate, D. Doreswamy, and S. K. Bhat, "Forecasting stock market prices using machine learning and deep learning models: a systematic review, performance analysis and discussion of implications," *International Journal of Financial Studies*, vol. 11, no. 3, p. 94, 2023.
[25] Chong, et al., "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," *Expert Systems with Applications*, 2017.
[26] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, 2018.
[27] M. Ballings, et al., "Evaluating multiple classifiers for stock price direction prediction," *Expert Systems with Applications*, 2015.
[28] J. Patel, et al., "Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques," *Expert Systems with Applications*, 2015.
[29] X. Zhong and D. Enke, "Forecasting daily stock market return using dimensionality reduction," *Expert Systems with Applications*, 2017a.
[30] X. Zhong and D. Enke, "A comprehensive cluster and classification mining procedure for daily stock market return forecasting," *Neurocomputing*, 2017b.
[31] A. Moghaddam, et al., "Stock market index prediction using artificial neural network," *Journal of Economics, Finance and Administrative Science*, 2016.
[32] E. Pehlivanli, et al., "Indicator selection with committee decision of filter methods for stock market price trend in ISE," *Applied Soft Computing*, 2016.
[33] D. F. Malagrino, et al., "Forecasting stock market index daily direction: A Bayesian Network approach," *Expert Systems with Applications*, 2018.
[34] O. Boyacioglu and D. Avci, "An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange," *Expert Systems with Applications*, 2010.
[35] .Yoon and G. Swales, "Predicting stock price performance: A neural network approach," in *Proceedings of the twenty-

DOI: https://doi.org/10.5281/zenodo.15110587

fourth annual Hawaii international conference on system sciences*, vol. 4, pp. 156-162, 1991.

[36] P. F. Pai and C. S. Lin, "A hybrid ARIMA and support vector machines model in stock price forecasting," *Omega*, vol. 33, no. 6, pp. 497-505, 2005.

[37] M. M. Akhtar, A. S. Zamani, S. Khan, et al., "Stock market prediction based on statistical data using machine learning algorithms," *Journal of King Saud University-Science*, vol. 34, no. 4, p. 101940,2022