An Analysis of Deep Learning and Machine Learning Models for Predicting Stock Prices

1st Shubbh Mewada
Department of Computer Engineering
SAL Institute of Technology and Engineering Research
Ahmedabad, India
https://orcid.org/0009-0008-8170-294X

2nd Devshri Pandya
Department of Computer Engineering
SAL Institute of Technology and Engineering Research
Ahmedabad, India
https://orcid.org/0000-0002-5041-4311

3rd Hetvi Vora
Department of Information Technology
Thadomal Shahani Engineering College
Mumbai, India
https://orcid.org/0009-0009-2149-3944

4th Jignesh Thaker
Department of Mechanical Engineering
Adani University
Ahmedabad, India
https://orcid.org/0000-0002-2599-3622

***Abstract*—The stock market, also known as the market for investments, holds significant influence over the present-day economy. It impacts investor's gains and losses as share prices fluctuate. The stock market presents several challenges, including the complexity of predicting future market dynamics accurately. Inaccurate prediction in financial data analysis hampers decision-making and maximizing investor benefits. Capturing the intricate nature of the stock market poses difficulties in understanding its behavior and making reliable predictions. Enhancing accuracy in short-term price prediction is crucial for informed investment decisions. This research tackles the challenge of accurately predicting future market dynamics in the complex financial market to maximize investor benefits. To achieve this, the paper employs Machine Learning (ML) and Deep Learning (DL) techniques. Various factors are analyzed, and valuable information is extracted to enhance stock price prediction. The research utilizes ML models along with a range of DL models. Empirical analysis conducted on randomly selected NASDAQ stocks demonstrates notable accuracy in short-term price prediction. The findings highlight the effectiveness of ML and DL models in capturing the intricate nature of the stock marketplace. This research contributes to the advancement of financial data analysis and emphasizes the potential for more precise prediction by leveraging ML and DL techniques.**

***Index Terms*—Stock Price Prediction, Time Series, Artificial Intelligence, Machine Learning, Deep Learning, Neural Networks, Moving Averages, Support Vector Machines, Random Forest, Multi-Layer Perceptron, Convolutional Neural Network, Long Short-Term Memory, Bi-Directional Long Short-Term Memory, Gated Recurrent Unit, Bi-Directional Gated Recurrent Unit**

I. Introduction

The ability to accurately predict fluctuations in the stock market is of great interest to shareholders and financial organizations due to its potential for generating substantial financial gains. However, the unpredictable nature of time series data poses a significant challenge in achieving precise stock price prediction. In response, fundamental and technical analysis using Moving Averages (MA) [1], along with the utilization of Machine Learning (ML) and Deep Learning (DL) techniques, to develop effective predictive models for stock market analysis [2]. DL employed models have gained popularity in this domain mainly because of their adaptability and capacity to handle convoluted trends in data [3].

While DL based models seems promising, they also come with limitations. Their large size and lack of prior knowledge about input data make them susceptible to over-fitting, wherein the prediction becomes too specific to the training data and is unable to generalize well to new data. The aim of this research is to explore and utilize identified ML and DL models to enhance the overall performance for stock market prediction [4].

A. Motivation and Contribution

The motivation behind this research is to enhance stock market prediction using various ML and DL models such as Support Vector Machine (SVM), Random Forest (RF), Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Bi-Directional Long Short-Term Memory (Bi-LSTM), Gated Recurrent Unit (GRU) and Bi-Directional Gated Recurrent Unit (Bi-GRU). The goal is to explore these models to help investors, traders, and financial institutions make well-informed decisions in the dynamic and volatile financial landscape, ultimately contributing to more sustainable and profitable trading strategies.

The primary contribution of this paper is mentioned below:

- Exploration of different Moving Averages (MA) to calculate the past characteristic states, evaluating their impact on predicting the stock opening price of the next trading day [1].
- Analysis of time sequences and association of stock price data, leading to the proposal of ML model namely SVM, RF and DL models such as MLP, CNN, LSTM, Bi-

979-8-3503-3439-5/23/$31.00 © 2023IEEE

LSTM, GRU, Bi-GRU for predicting the stock opening price for the following trading day.

- Comparative analysis for stock price prediction models used in this research, demonstrating the superior performance and effectiveness of MLP, BiGRU, and SVM methods. These findings underscore the suitability of these models for stock market forecasting.

This research utilizes ML and DL techniques to advance stock market's price prediction, benefiting investors and financial institutions.

The paper is organized as follows: Section II covers related work on ML and DL in stock market prediction. Section III presents an overview of key principles and models along with their architectures. Section IV presents experimental result analysis. Lastly, Section V represents concluding remarks of this paper.

II. Related Work

The literature on stock market prediction is abundant, with numerous studies exploring diverse methodologies to tackle this complex challenge. This section presents a comprehensive review of the research work conducted in the field of stock market prediction, encompassing various ML and DL models.

Numerous studies have shown the potential of technical analysis in aiding trading decisions [1]. Among these methods, MA stands out as a key indicator, reflecting stock fluctuations [1]. Pioneers an unrestricted approach, leveraging various forms of MA, including Weighted Moving Average (WMA) and Exponential Moving Average (EMA), to devise advanced trading strategies. The experimental results highlight the system's ability to pinpoint optimal trading opportunities, surpassing conventional methods and even outperforming the buy-and-hold strategy, thereby yielding significant profits.

Stock market prediction faces challenges due to complex parameters [5]. New variables are created to enhance accuracy since the historical data lacks sufficient features. The study compares Artificial Neural Network (ANN) and RF, with ANN surpassing RF in predicting the subsequent day's closing stock price. Future work may involve incorporating financial news articles and other parameters using DL models for improved predictions [5].

ML algorithms for anticipating the value of stock market prediction using Decision Tree (DT), and RF is presented in [6]. Future work may involve exploring algorithmic approaches to deep learning consisting of Long Term Memory (LTM), Short Term Memory (STM), and CNN to enhance capabilities.

Utilization of two methods for stock price prediction, Linear Regression (LR) yields the most precise findings, followed by SVM among K-Nearest Neighbors (KNN), SVM, DT, and RF models [7]. The mixed approach with LR as the second algorithm achieves highly accurate predictions, outperforming both LR and SVM independently.

A comparative study between LSTM and MLP models, with feature engineering is presented in [8]. The study challenges the Efficient Market Hypothesis (EMH) to some extent and supports research on Inefficient Market Hypotheses. Experimental results shows that MLP model outperforms in short-term stock price prediction highlighting the potential of neural networks [8].

The research conducted in the study [9] demonstrated that both the GRU and LSTM models outperformed models like ANN and SVM in terms of prediction. The GRU model was effective in predicting trends and maintaining a robust strategy. However, it exhibited slightly lower confidence in its predictions compared to the LSTM model. While the LSTM model achieved optimal results, it faced difficulties in accurately capturing dynamic nature of price fluctuations [9].

III. Proposed Approach

In this study, we aim to enhance stock market prediction by leveraging the capabilities of both ML and DL models. A step by step process carried out for stock market prediction is explained in this section. Moreover, Fig. 1 displays the graphical representation for proposed approach.

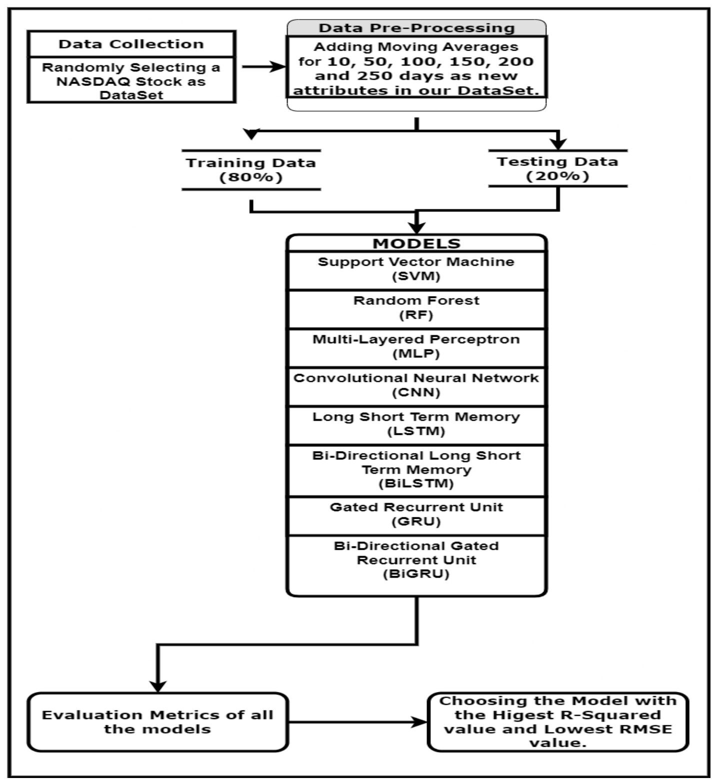

Fig. 1. Flowchart for our proposed methodology.

A. Data Collection

This research focuses on enhancing the stock price prediction and considers the granularity of the data used. The data-set used in this research comprises historical daily prices for all tickers currently trading on NASDAQ[1]. It is structured with key fields comprises features attributes like date, open price,

[1]https://www.kaggle.com/code/jacksoncrow/download-nasdaq-historical-data/input?select=stocks

high price, low price, close price, adjusted close price, and volume.

B. Data Pre-processing

Data pre-processing is crucial for stock price prediction as it addresses missing data, handles noise and outliers, scales and normalizes features, enables feature engineering to extract relevant patterns, deals with time series dependencies, removes redundancy. By applying these pre-processing techniques, the data becomes suitable for prediction models, enhancing their accuracy and effectiveness in forecasting future stock prices [5]. MA is a statistical procedures used in data analysis, especially when predicting stock prices. In this research study MA is used because to their capacity to lessen the effects of short-term market volatility. MA provide a smoothed trend-line by computing average values over specific duration, helping researchers to find underlying patterns and trends beneath the noise on the surface. This makes it easier to spot the direction of the market at large and potential price changes for stocks. MA also help in identifying chart patterns which provide insight into potential future price moves [1].

- **Principle of Moving Averages (MA):** MA are a statistical technique used to smooth out data by calculating the average of a specified number of consecutive data points [1].

$$MA_n = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{1}$$

In equation (1), MA_n represents the moving average of the data points. $x_1+x_2+\cdots+x_n$ are the data points to be averaged. n is the total amount of points of information considered for the average.

MA over 10, 50, 100, 150, 200, and 250 days were calculated and added as attributes to each data point. The choice of duration was standardized for consistency across NASDAQ datasets, providing short, medium, and long-term trends applicable for major US exchange-traded stocks. Short windows (10 and 50 days) capture transient noise, while long-term windows account for momentum and support/resistance levels. Although not perfect for every dataset, these smoothed trend features enhance predictive performance compared to raw values. Fig. 2 displays stock prices with their corresponding MA.

C. Machine Learning (ML) Models

- **Support Vector Machine Model (SVM):** SVM, a supervised ML algorithm, handles high-dimensional data and complex decision boundaries by integrating information into a higher-dimensional space of features [7]. SVM for regression (SVR) predicts continuous variables, such as stock prices, by finding the best-fitting hyperplane with a specified margin of error and captures non-linear patterns using support vectors and kernel functions, enabling accurate prediction of complex and non-linear stock price movements [9].

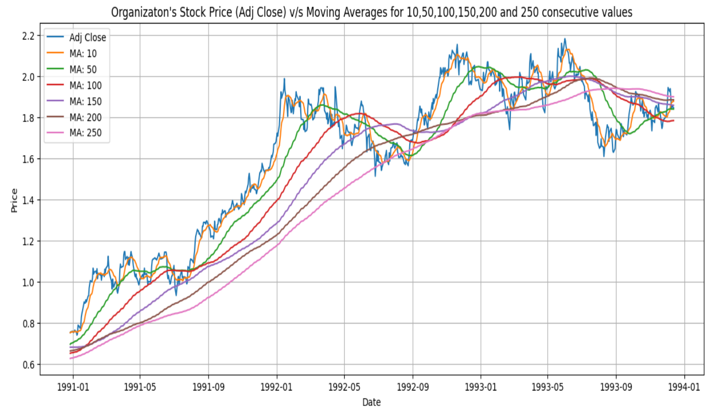

Fig. 2. Moving Averages for better Pattern Recognition.

- **Random Forest (RF):** RF is an ensemble ML technique for regression and classification tasks as it reduces model variance and improves accuracy by combining multiple decision trees [7]. Novel variables are introduced for guiding decision-making in each tree. Because of the noise in stock market data, the algorithm treats stock market analysis as a regression issue and predicts the following day's opening stock price based on training variables [5]. The parameters selected for ML models are presented in Table I.

TABLE I
PARAMETERS FOR MACHINE LEARNING (ML) MODELS.

Model	Parameter	Value
RF	random_state	42
	max_depth	10
	min_samples_leaf	2
	min_samples_split	5
	n_estimators	50
SVM	C-value	10
	epsilon	0.01
	kernel	'linear'

D. Deep Learning (DL) Models

- **Multi-Layered Perceptron (MLP):** MLPs are frequently employed for estimation of stock prices mainly because of their propensity to comprehend extensive relationships and pattern recognition [8]. Their capability to capture non-linear mappings, consider multiple input features, handle temporal dependencies, offer flexibility in architecture, and use training algorithms like back-propagation for optimization [10].
- **Convolutional Neural Network (CNN):** In stock price prediction, CNN offer advantages such as local pattern detection, hierarchical feature learning, parameter sharing, translation in-variance, and handling multi-dimensional data with multiple channels [10]. However, it's important to complement CNNs with additional information and techniques to consider all relevant factors influencing stock prices.

TABLE II
PARAMETERS FOR DEEP LEARNING (DL) MODELS

PARAMETERS	MODELS					
	MLP	CNN	LSTM	BiLSTM	GRU	BiGRU
Dense Layers	5	2	1	1	3	3
Number of LSTM / BiLSTM Layers	-	-	4	4	-	-
Number of GRU / BiGRU Layers	-	-	-	-	2	2
Convolutional One-Dimensional	-	2	-	-	-	-
MaxPooling size	-	2	-	-	-	-
Flatten	-	1	-	-	1	1
Number of Filters used	-	256, 128	-	-	-	-
Number of Units used	512, 256, 128, 96, 1	96, 48	96, 96, 96, 96	96, 96, 96, 96	64, 128	64, 128
Kernel Size	-	3	-	-	-	-
Activation Function	Rectified linear unit	Rectified linear unit	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent
Dropout Value	-	-	0.2	0.2	0.3	0.3
Loss Function	Mean Squared Error	Mean Squared Error	Mean Squared Error	Mean Squared Error	Mean Squared Error	Mean Squared Error
Optimizer	Adam	Adam	Adam	Adam	Adam	Adam
Batch Size	32	32	64	32	64	32
Epochs	50	50	100	50	125	50
Total Parameters Trained	185,409	122,033	260,065	741,313	260,161	429,441

- **Long Short Term Memory (LSTM):** LSTM encapsulates temporal dependencies and patterns in sequential data. They take historical price and volume records as input and learn to predict the future price of stocks [9]. LSTM cells with gates control information flow and memory within the model. During training, the model grows accustomed to its internal weights using optimization algorithms to minimize prediction errors which is beneficial during validation phase [10].
- **Bi-Directional Long Short Term Memory (BiLSTM):** A BiLSTM combines two LSTM layers to capture both forward and backward dependency in data sequences [10]. The forward LSTM handles the sequence from the beginning to the end, while the backward LSTM handles it in reverse. The outputs generated by the two layers are combined to form the final result. By considering dependencies in both directions, the BiLSTM captures long-term dependencies and provides a comprehensive understanding of the input sequence [12].
- **Gated Recurrent Unit (GRU):** GRU, a Recurrent Neural Network (RNN) architecture, captures long-term dependencies in stock price patterns by retaining relevant information over time as it addresses the vanishing gradient problem through gating mechanisms, enhancing learning of long-term dependencies [9]. GRU offers computational efficiency and faster training compared to LSTM, making it suitable for large stock price data-sets. Its dynamic memory adaptation enables adjustment to varying trends and market dynamics [10].
- **Bi-Directional Gated Recurrent Unit (BiGRU):** The BiGRU architecture combines bidirectional processing and the GRU to enhance stock price prediction. By considering past and future time steps simultaneously, BiGRU captures dependence over time while providing a holistic view of the data [11]. Its ability to process the input sequence in forward as well as reverse directions enhances context awareness and enables the identification of intricate relationships and meaningful features. BiGRU is especially beneficial for handling noisy or incomplete data, as it leverages information from both directions to mitigate the impact of missing or noisy data points [13].

The parameters selected for DL models are presented in Table II. For the MLP model, the first dense layer contains 512 units, the second dense layer contains 256 units, the third dense layer contains 128 units, the fourth dense layer contains 96 units, and the final dense layer contains 1 unit. For CNN model, the first convolutional layer has 96 units, and the second convolutional layer has 48 units with 256 filters in each layer. For the LSTM, all four LSTM layers have 96 units each. In the BiLSTM model, all four Bidirectional LSTM layers have 96 units each. In the GRU model, the first GRU layer has 64 units, and the second GRU layer has 128 units. In the BiGRU model, both Bidirectional GRU layers have 64 units in the forward and backward directions, totaling 128 units.

IV. EXPERIMENTAL RESULT ANALYSIS

To demonstrate the efficacy of the ML and DL models employed, the evaluation of the effectiveness of the models is dependent on the Root Mean Square Error (RMSE) and R-square (R^2) metrics, which serve as evaluation criteria to assess the prediction accuracy and efficacy of the various methods. The following section present a comparative analysis of different ML and DL models used.

TABLE III
EXPERIMENTAL RESULT ANALYSIS

Models	RMSE Value	R^2 Score
SVM [7]	0.23713669	0.994762
RF [7]	0.1525962	0.832489
MLP [9]	0.001052	0.874004
MLP	0.0081381	0.9991108
CNN	0.0108570	0.9984175
LSTM	0.0781960	0.9938853
BiLSTM	0.0215556	0.9937620
GRU	0.0181805	0.9955625
BiGRU	0.0108726	0.9984129

The table III provides evaluation of the different approaches used for forecasting stock prices, based on their RMSE values and R^2 scores. Among the models mentioned, SVM [7] demonstrates an RMSE of 0.23713669 and an R^2 score of 0.994762, indicating a relatively strong performance in terms of accuracy and fitting the data. The RF [7] obtains an RMSE of 0.1525962 and an R^2 score of 0.832489, suggesting a reasonable prediction capability, although it might not capture the variability as effectively.

MLP [9] stands out with a comparatively low RMSE of 0.001052, showcasing high accuracy in predictions. The associated R^2 score is 0.874004, indicating a good fit to the data. Furthermore, the MLP used in this study achieves an RMSE of 0.0081381 and a relatively high R^2 score of 0.9991108, highlighting its noteworohty predictive capabilities and strong data fitting.

The CNN model used in this paper demonstrates an RMSE of 0.0108570 and an R^2 score of 0.9984175, indicating a high level of accuracy and data fitting. The LSTM model obtains an RMSE of 0.0781960 and an R^2 score of 0.9938853, suggesting robust predictive power while also fitting the data reasonably well. BiLSTM model achieves an RMSE of 0.0215556 and a R^2 score of 0.9937620, indicating its ability to make reasonably accurate predictions and fitting the data. The GRU model presents an RMSE of 0.0181805 and a R^2 score of 0.9955625, showcasing its strong predictive capacity and excellent data fitting. Finally, the BiGRU model demonstrates an RMSE of 0.0108726 and an R^2 score of 0.9984129, suggesting a high level of predictive accuracy and data fitting.

In summary, the findings from the presented table III indicates that DL models consistently outperform ML models in terms of predictive accuracy and data fitting. Specifically, the DL models exhibit lower RMSE values and higher R^2 scores, which signify better accuracy and goodness of fit to the data. This pattern suggests that DL models are more adept at capturing intricate patterns and relationships within the data, resulting in comparatively better predictive capabilities.

Fig. 3, 4, 5, 6, 7, 8, 9, 10 aid with comprehension of the effectiveness and precision of the different model employed.

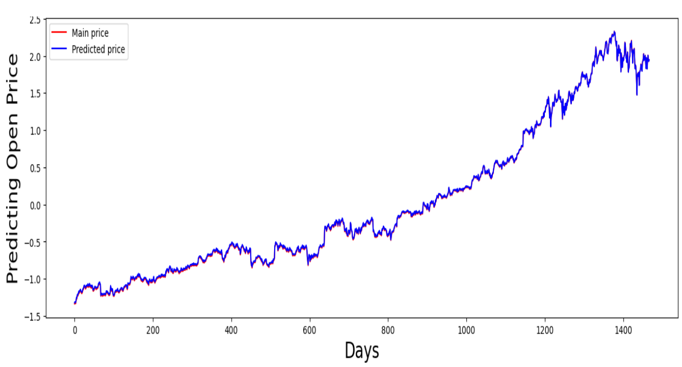

Fig. 3. Predicted Vs. Actual Price for SVM Model

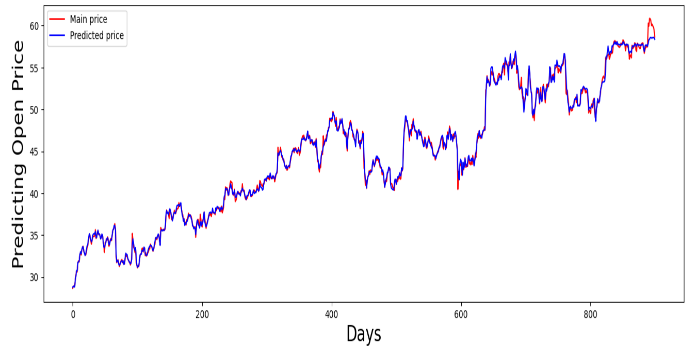

Fig. 4. Predicted Vs. Actual Price for RF Model

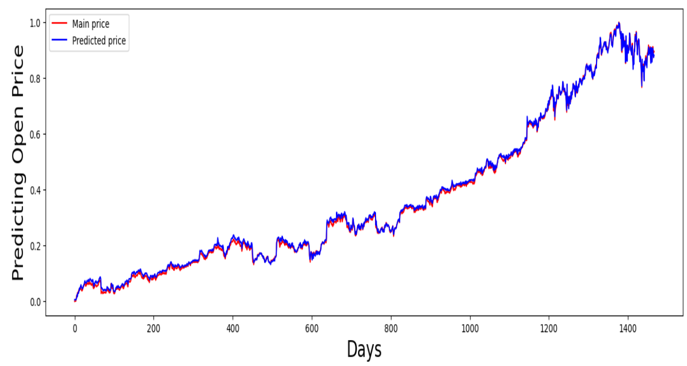

Fig. 5. Predicted Vs. Actual Price for MLP Model

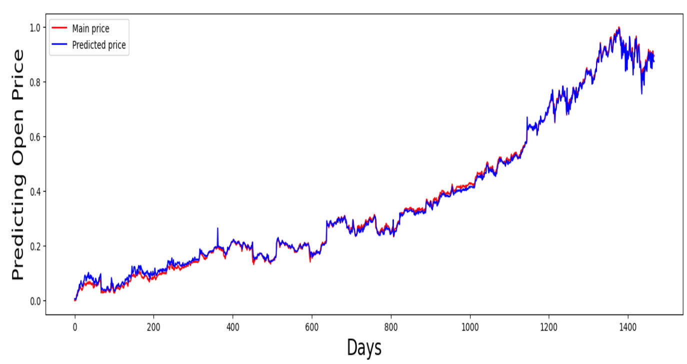

Fig. 6. Predicted Vs. Actual Price for CNN Model

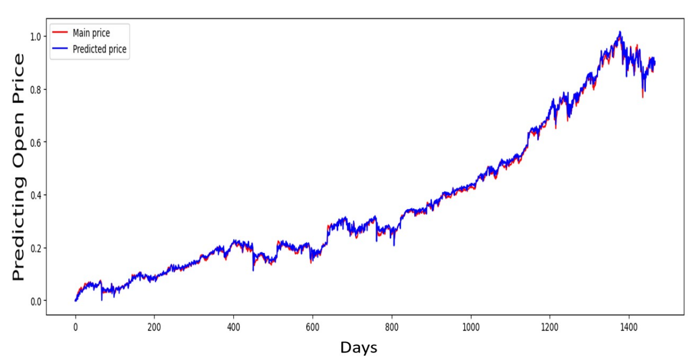

Fig. 7. Predicted Vs. Actual Price for LSTM Model

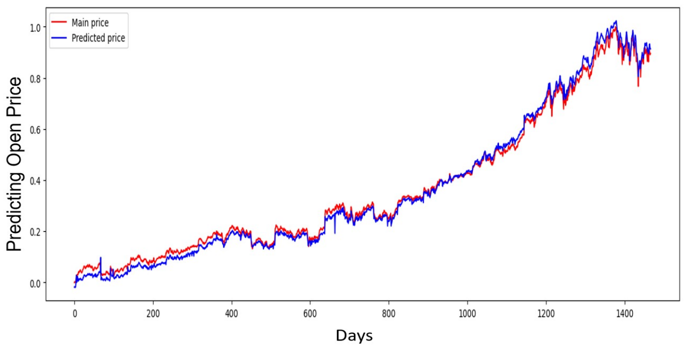

Fig. 8. Predicted Vs. Actual Price for BiLSTM Model

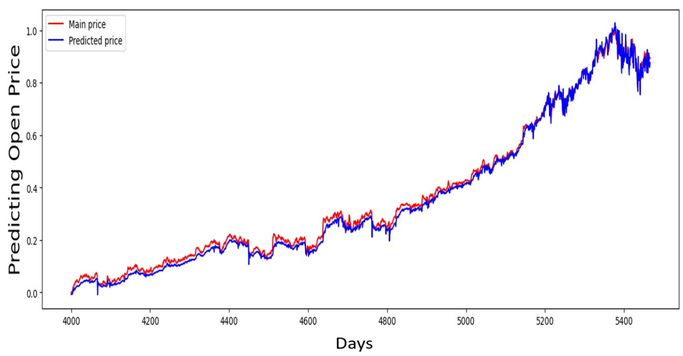

Fig. 9. Predicted Vs. Actual Price for GRU Model

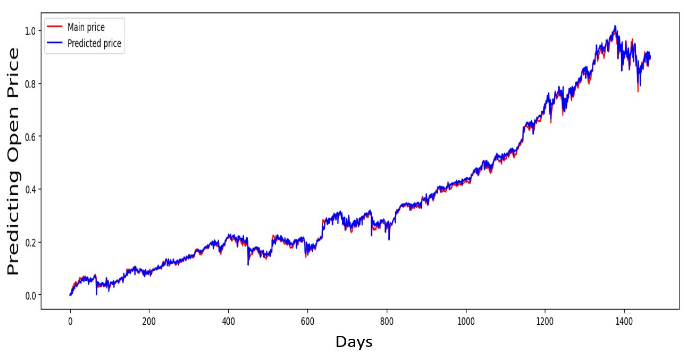

Fig. 10. Predicted Vs. Actual Price for BiGRU Model

V. Conclusion

In this study, we explored the potential of ML and DL models for stock market prediction. We collected and preprocessed historical financial data from NASDAQ, conducting relevant feature engineering to establish a robust foundation for our models. We conducted a thorough analysis of traditional ML algorithms, including SVM and RF, setting a strong baseline for comparison with more intricate DL architectures. Our investigation encompassed advanced DL models such as MLP, CNN, LSTM, BiLSTM, GRU, and BiGRU. These models showcased their adeptness in capturing sequential and temporal patterns within time-series data, rendering them highly suitable for stock market forecasting.

We evaluated our models using metrics like Root Mean Squared Error (RMSE) and the coefficient of determination (R^2 Score). Furthermore, by interpreting the models' predictions, we gleaned valuable insights into the underlying factors influencing the forecasts, thereby enhancing the interpretability and credibility of our predictions. Inherent risks and uncertainties persist in financial prediction, yet our comprehensive exploration of ML and DL models significantly contributes to refining forecasting capabilities and empowering decision-makers within the financial landscape.

References

[1] Shu-Yu Kuo, Yao-Hsin Chou, Building Intelligent Moving Average-Based Stock Trading System Using Metaheuristic Algorithms.

[2] Mendelsohn Louis B. (2000), Trend Forecasting with Technical Analysis: Unleashing the Hidden Power of Intermarket Analysis to Beat the Market, Marketplace Books.

[3] Ganesh, "Fundamental analysis as a method of share valuation in comparison with technical analysis."

[4] Cherkassky, "Practical Selection of SVM Parameters and Noise Estimation for SVM regression."

[5] Vijh, M., Chandola, D., Tikkiwal, V. A., Kumar, A. (2020). Stock Closing Price Prediction using Machine Learning Techniques. Procedia Computer Science, 167, 599–606. doi:10.1016/j.procs.2020.03.326

[6] Subba Rao Polamuri, K. Srinivas, A. Krishna Mohan. Stock Market Prices Prediction using Random Forest and Extra Tree Regression

[7] Vazirani, S., Sharma, A., Sharma, P. (2020). Analysis of various machine learning algorithm and hybrid model for stock market prediction using python. 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE). doi:10.1109/icstcee49637.2020.9276859

[8] Khare, K., Darekar, O., Gupta, P., Attar, V. Z. (2017). Short term stock price prediction using deep learning. 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT). doi:10.1109/rteict.2017.8256643

[9] Sethia, A., Raut, P. (2018). Application of LSTM, GRU and ICA for Stock Price Prediction. Smart Innovation, Systems and Technologies, 479–487. doi:10.1007/978-981-13-1747-7_46

[10] Lu, W., Li, J., Wang, J., Qin, L. (2020). A CNN-BiLSTM-AM method for stock price prediction. Neural Computing and Applications. doi:10.1007/s00521-020-05532-z

[11] Yuanshuai Duan,Yuanxin Liu, Yi Wang, ,Shangsheng Ren, Yibo Wang. Improved BIGRU Model and Its Application in Stock Price Forecasting. https://doi.org/10.3390/electronics12122718

[12] Lee, MC., Chang, JW., Yeh, SC. et al. Applying attention-based BiLSTM and technical indicators in the design and performance analysis of stock trading strategies. Neural Comput Applic 34, 13267–13279 (2022). https://doi.org/10.1007/s00521-021-06828-4

[13] M. E. Karim and S. Ahmed, "A Deep Learning-Based Approach for Stock Price Prediction Using Bidirectional Gated Recurrent Unit and Bidirectional Long Short Term Memory Model," 2021 2nd Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2021, pp. 1-8, doi: 10.1109/GCAT52182.2021.9587895.