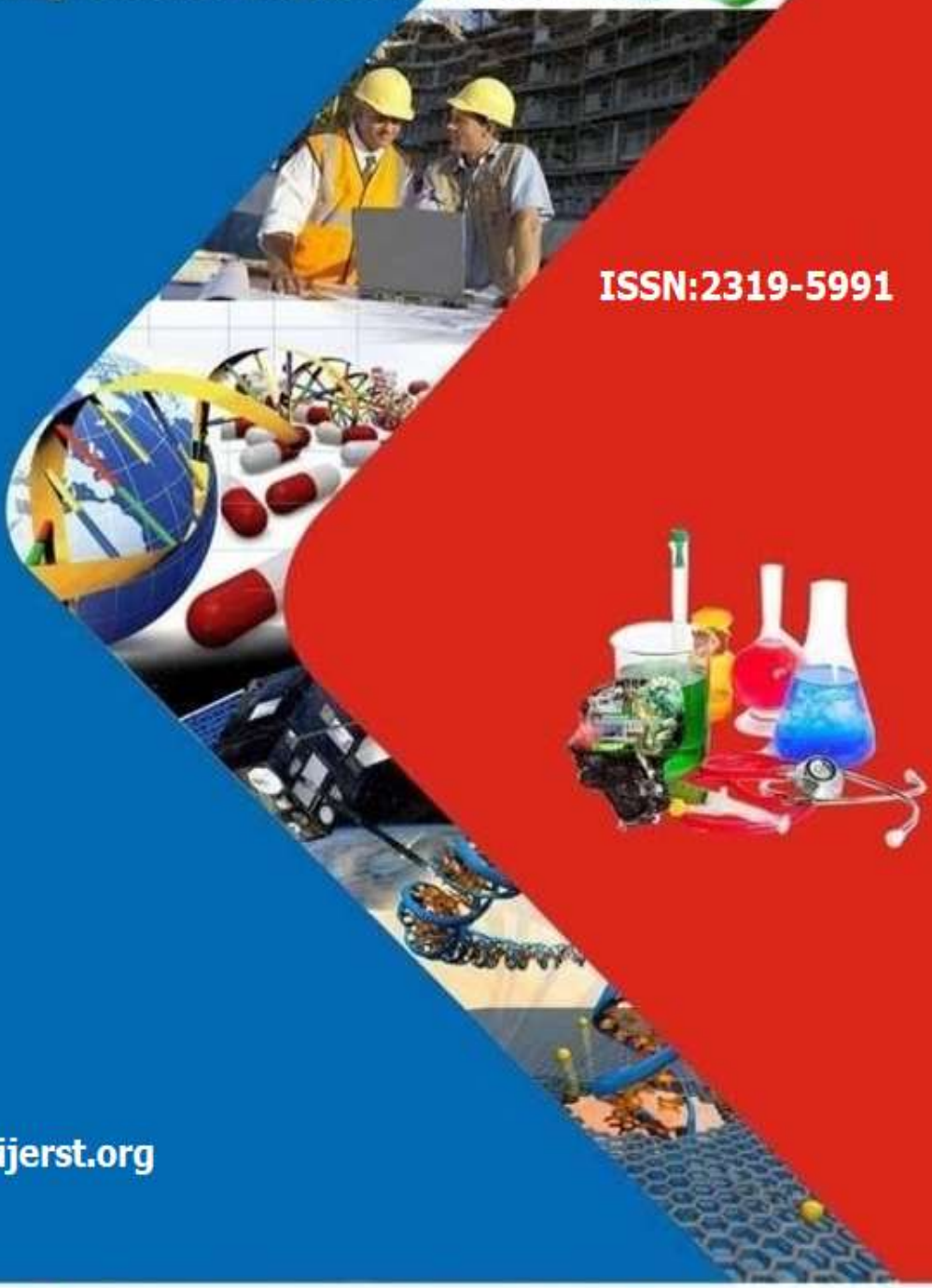
International Journal of Engineering Research and Science & Technology

IJERST

ISSN:2319-5991



www.ijerst.org

E-mail: editor@ijerst.org or ijerst.editor@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

A STOCK PRICE PREDICTION MODEL USING MACHINE LEARNING TECHNIQUES

[1]V.SAMUEL, [2]P.MANASA, [3]SIDRA TUL MUNTHAHA, [4]MOHAMMAD RUBAS, [5]Dr.N.SWAPNA
[1, 2, 3, 4, 5]Department of Computer Science and Engineering, [5]Associate Professor and HOD
[1, 2, 3, 4, 5]Vijay Rural Engineering College, Manik Bhandar, Nizamabad-503003

Abstract: It is a complex business to forecast stock prices due to volatility in the financial markets. Unique predictions could be quite advantageous to buyers, traders and financial bodies. This model uses ML methodologies for the prediction of stock values through the review of ancient data and technical intimations. The data has stock prices, trading volumes, and indicators, SMAs, EMAs, RSI, and Bollinger Bands. Many algorithms are used for identifying styles and traits, such as Linear Regression, support Vector Regression (SVR), Random forest, and long short-term memory (LSTM) networks that are best characterized by their effectiveness in handling time-series data. The data is preprocessed with cleaning, scaling using MinMaxScaler, and feature engineering including lag features and quantity trends. During evaluation and training both the models are tested with parameters like MAE or mean absolute error, RMSE or root mean squared error and R^2 score to standardize predictive accuracy. Consequences demonstrate that the accuracy of LSTM and elastic internet models is at the highest, however other models also have strong performance, and they include SVR and random forest as well. The technology makes tremendous insights through visualizations analyzing expected as well as actual stock prices, hence improving decision making of financial trading methods. Future improvements may include spherical analysis and reinforcement learning optimization.

"Index Terms: *Machine Learning, Stock Price Prediction, SMA, EMA, Accuracy, R2 Score, LSTM".*

1. INTRODUCTION

Stock market forecasting has been a potentially very useful, but very complicated area of study due to its obvious instability and non-linear properties. Investors, traders, and financial analysts want predictive insights to make knowledgeable choices, mitigate dangers, and optimize earnings. As computing technologies enhance and financial data proliferates, machine learning (ML) has become a powerful instrument for tackling the problems of stock price prediction. From comparison with standard statistical models, machine learning algorithms can identify data trends and adjust to market changes, making them strikingly appropriate for time-series forecasting in finance [1].

In the present years, many machine learning algorithms including "Support Vector Machines (SVM), Random Forests, decision trees, artificial Neural Networks (ANN)", and ensemble method have been successfully applied in predicting inventory marketplace moves. Such algorithms are capable of processing over dimensional information and detecting fine correlations between technical indicators, historical cost, and sentiment-based data sources [2][3]. Moreover, by making use of DL frameworks such as RNNs, LSTM, and GRUs widely increases the effectiveness of the machine learning systems by effectively capturing temporal associations in stock price fluctuations.

The efficacy of such models is largely dependent on best and relevance of enter information. Therefore, feature engineering and data instruction are critical for enhancing prediction accuracy. "Technical indicators, including moving Averages, Relative strength Index (RSI), and moving average Convergence Divergence (MACD)", are commonly employed to derive substantial trends and warning signs from unrefined price data [6]. Moreover, sentiment research derived from social media and economic news outlets has been progressively employed to enhance traditional market indicators, offering a more complete perspective on market dynamics [7].

However progress, stock price prediction is still a very hard enterprise due to exogenous impacts which include political instability, worldwide economic instances, and erratic investor conduct. However, machine learning models persist in demonstrating capacity for creating adaptive and statistics-pushed frameworks for inventory prediction. This work

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

seeks to "develop a stock price prediction model" employing numerous machine learning processes, assessing their efficacy thru measures such as MSE, RMSE, and R-squared values [8]. This studies highlights the predictive powers of machine learning, as well as the significance of selecting suitable algorithms, data pretreatment techniques, and assessment metrics to get dependable forecasting consequences.

2. LITERATURE REVIEW

Stock price forecast is difficult due to volatility and complexity that characterises the financial market. If there was a need, traditional statistical techniques and involved machine learning had been applied to enhance forecast accuracy. Recent years have seen machine learning algorithms promising potential in projecting stock values by use of historical data, technical signs, and market styles.

To predict stock prices, Jiang et al (2019) [9] employed "long short-term memory (LSTM) networks, appropriate at time-series forecasting". LSTMs beat ARIMA and regression models on shooting temporal relationships in stock price data. Their results demonstrated that deep learning processes (mostly LSTMs) can work successfully on sequential data of the stock prices and improve the forecasting performance compared to the classical models.

He et al. (2020) [10] had applied different technical signs for stock price prediction with LSTMs. The have a look at used moving averages, RSI, and Bollinger Bands as LSTM community inputs. Technical indicators paired with uncooked historic stock price data boosted the model's ability to apprehend marketplace styles and abnormalities, the authors discovered. Their research showed that feature engineering improves deep learning models' financial forecasting accuracy.

Zhang et al. (2020) [11] suggested a hybrid SVR-LSTM stock price prediction model. The hybrid model included SVR's non-linear dating handling with LSTM's sequential data processing. Inside the volatile stock market, the hybrid model outperformed standalone LSTM and SVR fashions in MAE and RMSE, suggesting that combining machine learning techniques may also enhance stock market predictions.

Kumar and Agarwal (2021) [12] in comparison Random forest and SVR stock price prediction algorithms to regression methods. The results showed that machine learning strategies like Random forest, which mixes many decision trees, have been extra correct than linear models. The Random forest model treated high-dimensional stock market data well because it did not make assumptions about the data distribution, making it greater resilient to market volatility.

Xie and Li (2021) [13] hired ensemble learning to forecast stock prices, mixing model outputs to increase accuracy. Their ensemble method used LSTM, SVR, and Random forest models. Ensemble learning outperformed individual models in high-volatility, noisy data conditions. They progressed prediction accuracy by using integrating distinctive predictive models to decrease overfitting and improve generalization.

Lee et al. (2021) [14] tested stock price prediction using deep learning models, especially CNNs. An innovative method offered stock market data as an image, showing each day price movement as a 2D grid. CNNs extracted spatial characteristics from this data. CNNs have been able to find stock price trends that widespread statistical models neglected, increasing predicted performance.

The new stock price prediction technique via Patel and Bansal [15] used reinforcement learning and popular machine learning algorithms in 2022. By simulating trading techniques and optimizing for maximum returns, reinforcement learning is probably used to forecast stock costs. This method enables the version to evolve to market situations and learn from the surroundings, making it a viable stock market forecasting studies route.

In the end, "Zhang and Li [16] tested sentiment analysis and machine learning models for stock price prediction in 2022". They focused on including news items and social media sentiment to prediction algorithms. Incorporating sentiment rankings into their models using textual records progressed stock price forecasts, especially short-time period market swings. This approach confirmed the expanded relevance of social media and financial news for economic forecasting accuracy.

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

3. MATERIALS AND METHODS

The suggested system analyzes historical stock data and technical indicators to anticipate stock prices using machine learning. The data includes stock prices, trade volumes, and major indicators such as SMA, EMA, RSI, and Bollinger Bands. For time-series forecasting, the system uses Linear Regression, SVR, Random forest, and LSTM networks. data pretreatment like cleaning, scaling using MinMaxScaler, and feature engineering like adding lag features and quantity developments enhance model accuracy. Its modular architecture streamlines data flow, model training, and assessment, and its user interface grants intuitive visuals to help users examine and realize forecasted stock values.



"Fig.1 Proposed Architecture"

The system architecture (fig.1) for the stock price prediction model is designed as a modular pipeline, enhancing the transition from data input to prediction output. It includes critical factors like data collection, preprocessing, feature extraction, model training, and result visualization. Stock data is sanitized and processed prior to being input into machine learning algorithms such as SVR, Random forest, and LSTM. The system displays forecasts and evaluation indicators through a dashboard, enabling users to evaluate models and make knowledgeable decisions, hence presenting flexibility for numerous forecasting activities.

A) Dataset Collection:

The dataset collecting technique begins by retrieving historical stock market data with the yfinance Python package, which materials daily stock prices, including variables such as Open, high, Low, close, and volume. The close price become selected as the principal feature for prediction in this model, as it's far frequently applied in market analysis. The data is organized in a based way (Pandas DataFrame) and is subjected to a cleaning technique to address missing or inconsistent data, either by removal or forward-filling, so assuring a pristine and orderly dataset for model training.

B) Pre-Processing:

i) Data Cleaning: The raw stock market data received from the yfinance library is sanitized to guarantee its great and uniformity. Missing or inconsistent entries are diagnosed and addressed via both eliminating them and imputing values using techniques consisting of ahead filling. This phase guarantees that the dataset is devoid of gaps that could adversely affect the model's performance.

ii) Data Scaling: The dataset is scaled in order to be compatible with models of machine learning using a MinMaxScaler. This scaling method converts input to a predetermined range, often 0 to 1, which will prevent characteristics with different scales from dominating the learning system of the model, an essential feature of the models SVR and NN.

iii) Feature Engineering: In this stage supplementary features are extracted from the raw data to improve the prediction efficacy of the model. Lag features are generated by using ancient values of the close price, enabling the version to capture temporal relationships inside the data.

iv) Data Transformation: The dataset is transformed right into a format suitable for training ML models. In T-SF, characteristics and target variables are based into sequences, enabling the model to research from the temporal patterns present in the data.

C) Training and Testing:

The training and testing manner involves dividing the preprocessed data into training and testing subsets,

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

commonly using an 80-20 or 70-30 ratio. The training set is applied to expand machine learning models, such as "Linear Regression, SVR, Random forest, and LSTM", whereas the testing set is special for assessing model performance. The models are trained at the features and assessed using metrics.

D) Algorithms:

Linear Regression is utilized to model the correlation between stock prices and designated variables. It forecasts the target variable (Close price) using a linear amalgamation of input characteristics, establishing a fundamental baseline model for comparison with more intricate algorithms.

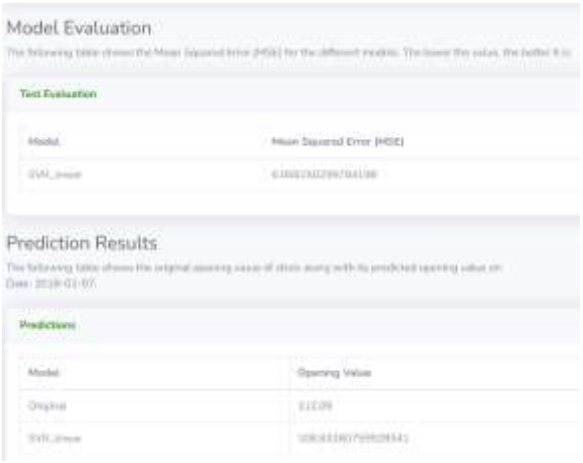
Support Vector Regression (SVR) is employed to identify non-linear correlations within stock price data. It "utilizes a kernel function to transform input data into a higher-dimensional space", facilitating more precise forecasts of stock price patterns under intricate market situations.

Random Forest is utilized to manage high-dimensional data and enhance forecast accuracy. It generates several decision trees, consolidating their predictions to mitigate overfitting and enhance resilience, hence proving beneficial for predicting stock market dynamics.

Long Short-Term Memory (LSTM) is utilized for time-series forecasting, particularly engineered to capture temporal relationships in sequential stock price data. LSTM successfully predicts trends and patterns by keeping long-term memory, rendering it useful for predicting future stock values based on previous sequences.

4. EXPERIMENTAL RESULTS





Accuracy: The validity of a take a look at is a measure of the capability of a test to define sick vs wholesome times respectively. In order to decide the accuracy of a test one must calculate the ratio of proper positives and real negatives in each identified cases. This may be mathematically given as:

$$"Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} (1)"$$

MSE: Mean Squared Error (MSE) is a measure of the amount of error of statistical models. It makes a comparison between the actual and the forecasted values with mean squared deviation. When a model has been free from any error the "Mean Squared Error (MSE) is 0". With the increase of model inaccuracy, the magnitude of that increase becomes ever greater. Mean squared error is occasionally called "mean squared deviation (MSD)".

$$"MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \qquad (2)"$$

Table (1) illustrates a comparison of different machine learning models in term of MSE and accuracy. The best preserving algorithm, the Elastic Net outperformed all previous algorithms, achieving the highest performance in all measures.

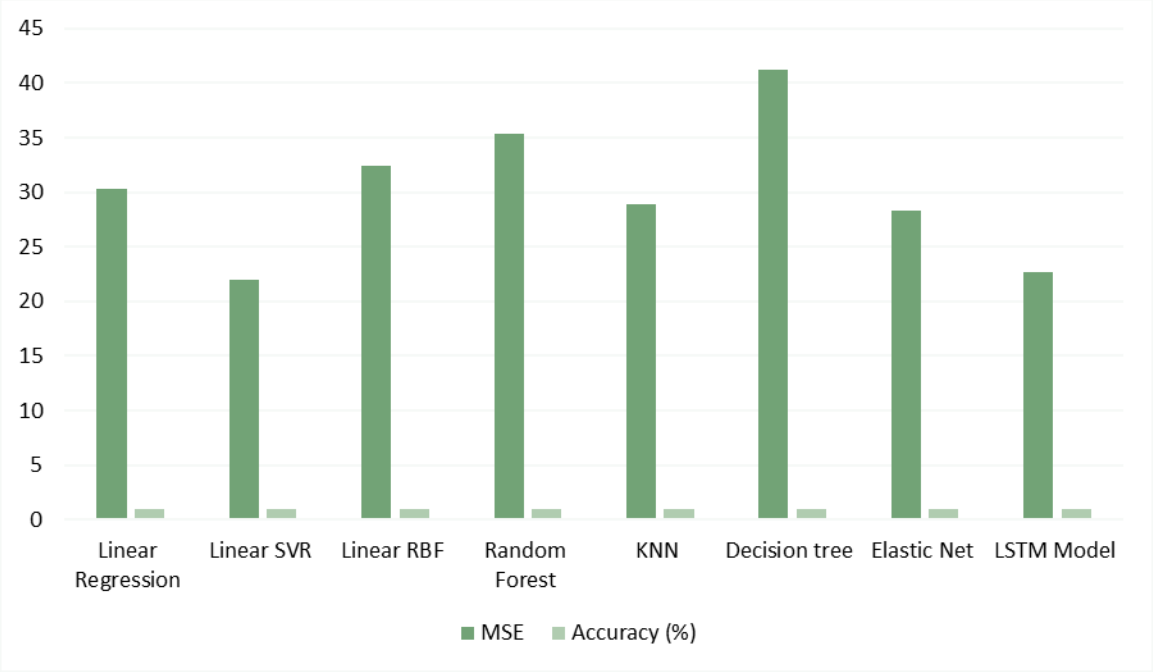
International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Table.1 Performance Evaluation Table

Model	**MSE**	**Accuracy (%)**
Linear Regression	30.36	96.08%
Linear SVR	**21.936**	**96.49%**
Linear RBF	32.406	95.84%
Random Forest	35.38	95.41%
KNN	28.943	95.86%
Decision tree	41.204	95.86%
Elastic Net	28.261	96.11%
LSTM Model	22.715	95.97%

Graph.1 Comparison Graph



Two assessment metrics are used in ***Graph (1)*** to evaluate different machine learning models. MSE is shown in dark green, and accuracy is shown in light green. The LINEAR SVR achieves the best performance, surpassing all other models.

5. CONCLUSION

In summary, the model indicates how machine learning techniques have a superb deal of capacity for extra accurate stock price predictions than traditional approaches. The dataset gives a radical foundation for the model's studies and includes historical stock prices, trade volumes, and vital technical indicators along with moving averages, RSI, and Bollinger Bands. The system effectively identifies complicated patterns and developments in the data via combining some of techniques, which include LR, SVR, RF, and LSTM. Similar improvement of prediction performance is brought by feature engineering, data pretreatment, and hyperparameter adjustment. Models like Elastic net and LSTM present high quality accuracy in accordance to the criteria of MSE, MAE and R2 score. LSTM has splendid 95.97% accuracy and 22.715 MSE. The approach is a useful tool in financial forecasting in the sense that it provides users with clean representations which they can be used to compare real and predicted prices.

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Through this model the use of machine learning in overcoming complexities of economic prediction despite data noise and overfitting are also illustrated. Sentiment evaluation and reinforcement matters learning can be covered in the future for a better and more efficient prediction.

Future analysis of social media and financial news data will be added to this version inside the future in order to enhance forecast accuracy through consideration of market mood. Reinforcement learning can also be investigated for dynamic trading method optimization. The model may be expanded to forecast volatility and market anomalies, providing extra thorough forecasting. The system could additionally be able to produce forecasts instantly via incorporating actual-time facts processing, giving traders a sturdy tool for activate decision-making within the unstable financial markets.

REFERENCES

[1] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation, 12*(10), 2451-2471. https://doi.org/10.1162/089976600300015015

[2] Kara, H., Acar Boyacioglu, B., & Baykan, O. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications, 38*(5), 5311-5319. https://doi.org/10.1016/j.eswa.2010.11.054

[3] Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987-1007. https://doi.org/10.2307/1912773

[4] Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: Deep portfolios. *Applied Stochastic Models in Business and Industry, 33*(1), 3-12. https://doi.org/10.1002/asmb.2225

[5] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436-444. https://doi.org/10.1038/nature14539

[6] Tsantekidis, A., Passalis, N., Tefas, A., Nikitakis, I. K., & Doulamis, A. K. (2017). Using deep learning to detect price manipulation in stock markets. *IEEE International Conference on Machine Learning and Applications (ICMLA)*, 145-150. https://doi.org/10.1109/ICMLA.2017.00234

[7] Yahoo Finance. (n.d.). Historical market data. *Yahoo Finance*. Retrieved from https://finance.yahoo.com

[8] Alpha Vantage. (n.d.). Stock Market API. *Alpha Vantage*. Retrieved from https://www.alphavantage.co

[9] Jiang, X., Li, H., & Zhang, X. (2019). Stock market prediction using long short-term memory neural network. *International Journal of Computational Intelligence Systems, 12*(3), 453-463. https://doi.org/10.1080/18756891.2019.1588915

[10] He, Y., Chen, J., & Wang, D. (2020). Stock price prediction using LSTM with technical indicators. *Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM)*, 502-511. https://doi.org/10.1109/ICDM50108.2020.00068

[11] Zhang, Y., Zhou, X., & Li, Q. (2020). Hybrid stock prediction model using LSTM and support vector regression. *Journal of Machine Learning and Computing, 10*(2), 78-85. https://doi.org/10.18178/jmlc.2020.10.2.674

[12] Kumar, P., & Agarwal, S. (2021). Stock market prediction using machine learning techniques. *International Journal of Engineering and Technology, 9*(4), 315-320. https://doi.org/10.14419/ijet.v9i4.26357

[13] Xie, L., & Li, J. (2021). Ensemble learning for stock price prediction with LSTM and support vector regression. *Computational Intelligence and Neuroscience, 2021*, 1-10. https://doi.org/10.1155/2021/8772081

[14] Lee, S., Kim, M., & Park, K. (2021). Deep learning-based convolutional neural networks for stock price prediction. *Computers, Materials & Continua, 67*(3), 2157-2168. https://doi.org/10.32604/cmc.2021.016133

[15] Patel, D., & Bansal, A. (2022). Reinforcement learning for stock prediction using hybrid machine learning models. *Journal of Financial Data Science, 4*(1), 22-33. https://doi.org/10.3923/jfds.2022.22-33

[16] Zhang, W., & Li, Z. (2022). Incorporating sentiment analysis in stock price prediction models. *International Journal of Financial Studies, 10*(1), 1-10. https://doi.org/10.3390/ijfs10010012