Forecasting Stock Price by LSTM-CNN Hybrid Model and Compares Deep Learning Models

Aditi Arora
Dept. of Computer Science and Engineering
ABES Engineering College
Ghaziabad, India
aditi.arora@abes.ac.in

Yashaswi Upadhyay
Dept. of Computer Science and Engineering
ABES Engineering College
Ghaziabad, India
Yashaswi.20b0101129@abes.ac.i n

Satvik Shukla
Dept. of Computer Science and Engineering
ABES Engineering College
Ghaziabad, India
satvik.20b0101094@abes.ac.in

Saket
Dept. of Computer Science and Engineering
ABES Engineering College
Ghaziabad, India
Saket.20B0101076@abes.ac.in

2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE) | 979-8-3503-7700-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICSPCRE62303.2024.10675145

***Abstract— T*he The accurate forecasting of stock prices holds significant importance in devising effective trading strategies and making informed decisions about stock transactions. The inherent volatility in stock prices, driven by factors like geopolitical tensions, corporate earnings, and commodity costs, underscores the need for precise predictions. Domestic factors such as central bank policies, governmental decisions, inflation, and global market uncertainties also trigger stock price fluctuations. Estimating this volatility is a complex challenge for traders. A novel approach is presented in this research, proposing a hybrid model that integrates two distinct algorithms – Long Short-Term Memory-Convolutional Neural Network (LSTM-CNN). This model harnesses diverse data representations from stock time series to enhance price prediction accuracy. Comparative analyses are performed with alternative models like LSTM (Long Short-Term Memory), CNN (Convolutional Neural Networks) and Naive Bayes to ascertain the most effective prediction paradigm with minimal error. The study demonstrates that combining temporal and image features extracted from the same dataset diminishes prediction errors compared to using these features independently. The hybrid model outperforms singular prediction models, highlighting a high enhancement. Among the individual models, the LSTM model exhibits superior performance in terms of prediction accuracy.**

Keywords: **Stock Market Prediction, Machine Learning, CNN, LSTM, Learning Regression, Naïve Bayes, MAE, RMSE**

I. INTRODUCTION

The development of a stock market prediction system is a complex challenge that involves creating algorithms and models to forecast movements in stock prices or market indices. In aims to help traders, investors, and financial professionals make wise decisions when purchasing, disposing of, or holding financial assets, this system seeks to offer insightful information. Despite advances in machine learning and data analysis techniques, accurately predictingstock market movements remains a formidable activity due to the volatility and complexity of financial stock markets. [6]

The motivation behind this project is to obtain a competitive edge and make wiser financial judgments in the fast-paced world of investing. To maximize earnings, reduce risk, and improve portfolio performance, investors, traders, and financial institutions are motivated to develop prediction algorithms. These systems seek to offer insights into potential market trends by utilizing historical data, complex algorithms, and data analysis techniques, assisting users in spotting buying and selling opportunities. We intend to develop a model to minimize errors in the stock market prediction system [7].

These days, complex analytical methods derived from either technical or essential analysis are employed to forecast stock market prices. In particular, the data is vast and non-linear. when it comes to stock market analysis. To handle this diversity of data, an effective model that can recognize intricate relationships and hidden patterns is required within this sizable dataset. In this domain, machine learning approaches have demonstrated a 60– 86% increase in efficiency in contrast to earlier techniques (Y. Wu et al., 2017) [10].

CNN uses hierarchical feature extraction, they are especially good at capturing complex correlations seen in sequential data, such as changes in stock values over time. They differentiate between short- and long-term fluctuations and trends, as well as between local and global patterns, to make predictions. [14] Stock forecasting uses long-term relationships and temporal patterns that can be extracted from sequential data, like historical stock prices. LSTMs overcome the limitations of traditional models by retaining data for longer periods, which is essential for understanding market dynamics and trends. [13] The LSTM-CNN hybrid model helps investors make informed decisions in the context of market dynamics by utilizing the retention of memory of LSTM and the hierarchical feature extraction of CNN to create an effective forecasting framework that can recognize both short- and long-term trends.

Our research work aims to predict stock prices and their behavior accurately. We will be using various machines and deep learning algorithms such as CNN, Naïve Bayesian and LSTM and hybrid model of the LSTM-CNN model. Authors will be evaluating the mean absolute error in each of the models and will give utmost priority in reducing the error in our models. Authors used the data set of Yahoo Finance and will implement various machine learning algorithms like LSTM, Naive Bayesian, and LSTM-CNN hybrid model [8].

II. RELATED PREVIOUS WORK

The research employs artificial intelligence and machine learning approaches. By offering stock price predictions in advance, the main objective is to increase profits and reduce losses for investors. It works on stock price forecasting utilizing **SVM** (Support Vector Machine), **RNN** (Recurrent Neural Networks), and RF approaches are cited, demonstrating the expanding trend of using these techniques for financial analysis. Yahoo Finance was used to compile historical data for five companies: Goldman Sachs, Nike, Johnson & Johnson, Pfizer, and JP Morgan Chase, covering the period from April 5, 2009, to April 5, 2019. Anovel method for predicting stock prices is demonstrated by the usage of **ANN** and **RF** models with new variables. Limited evaluation using only **RMSE** and **MAE** metrics, lack of a detailed dataset description, absence of comparisons with baseline models, and insufficient discussion of model limitations and future directions. Further advancements could result from investigating more complex deep learning models like **CNN** or **LSTM** (M. Vijh et al., 2020) [1].

It has long been understood that changes in stock price trends are a major problem. A stock price prediction approach based on **LSTM-CNN** was suggested in the paper. The advantages of adopting hybrid models for stock market forecasting. From theShanghai Composite Index, they got the dataset. For this study, stock prices from 1/July/1991 to 31/August/2020, which includes 7127 trading days, were employed as the data source. The **MAE** and **RMSE** are utilized as performance measurements. The primary focus of future research will be to enhance stock sentiment analysis to guarantee the precision of stock price forecast (W. Lu et al, 2020) [2].

In this paper, ANN for stock prediction is explored. The stock price of a firm is predicted in this study utilizing historical data using four deep learning models: **MLP**, **RNN**, **LSTM** and **CNN**. Both the National Stock Exchange (**NSE**) of India and the New York Stock Exchange (**NYSE**) have been considered. The tata motors dataset was extracted from NSE. The training dataset includes 4861 days worth of closing prices and spans the time from 1/1/1996 to 30/06/2015.The dataset covered the time from 3/Jan/2011 to 30/Dec/2016. In the suggested study, **CNN** performed better than the other three models and demonstrated the ability to identify abrupt price fluctuations. The use of hybrid models for prediction has not, however, been investigated (E. A. Gopalakrishnan et al, 2018) [3].

Due to their volatility and nonlinearity, stock market returns are difficult to anticipate with any degree of accuracy.The usage of **ANN** and **SVR** to predict financial time series with impressive accuracy has been covered by the author. The author has explored the **LSTM**, **CNN,** and models with varying degrees of accuracy. For the study,a financial dataset taken from the Jane Street Market Prediction competition on Kaggle was taken into consideration, recall, Precision and F1 scores for every model were utilized to compare the models. In their work, the **LSTM** 256x128 model is superior to all others and poses the fewest risks. Based only on the identification of significant opportunities, the reward function might also be used to examine the combination of Reinforcement Learning and the **LSTM** model (H. Hamoudi et al, 2021) [4].

The challenging task of foreseeing stock values in real time is addressed in the article, with a particular emphasis on businesses listed on the Nepal Stock Exchange Limited. The authors acknowledge that it is challenging for the public to understand the market. The objective is to create a technique that can forecast stock values every two minutes. They use a feed-forward neural network architecture that contains a hidden layer that contains a configurable number of neurons. The proposed system has shown promise accuracy for some businesses, but it might not perform as well. The causes of variances in performance are not discussed in the paper. The study offers a direction for future improvement by arguing that deepening the hidden layer of the neural network may enhance performance (A Shakva et al, 2018) [5].

The introduction of machine learning techniques is a result of the difficulty of traditional approaches to accurately capture the subtleties of market behavior. The necessity of utilizing ensemble approaches, which combine many models to improve forecast accuracy and reduce model biases, is stressed throughout the introduction. In the study's conclusion, stacking is shown as themost successful ensemble technique for stock market prediction. These findings are especially important for professionals looking to improve their forecasting algorithms when applied to the financial markets. It promotes additional research into supporting vector machine feature selection and parameter optimization, potentially boosting the field of financial forecasting. (A. F. Adekoya at al, 2020) [9].

III. METHODOLOGY

A. *Description of Data*

Yahoo Finance has provided historical data for the five companies. The dataset includes a 13-year dataset from 1 Jan 2010 to 31 Oct 2023 of Google, Amazon and Apple. The stock's high, low, open, close, adjacent close, and volume are all included in the dataset. All that has been extracted is the stock's closing price daily.

Dataset: Jan/2010 – Oct/2023

Training Dataset: Jan/2010 – Dec/2020

Testing Dataset: Jan/2021 – Oct/2023

B. *Convolutional Neural Network*

CNN is an excellent choice among deep learning neural networks. **CNN** can capture time series information froma stock market dataset and extract features from photos and videos using a sequence of convolution and pooling layers. The convolutional layers of a CNN make it extremely efficient in processing and identifying trends in grid-like data structures, such as photographs. However, for efficient training, a substantial amount of labelled data and processing power are needed [11]. Key components of **CNN** are

i) ***Convolutional Layers:*** *These layers are used to identify patterns and relationships in input data.*

ii) ***Pooling layer****: These layers are used to store the most essential information about the data and utilizeit.*

iii) ***Fully connected layers****: These layers include the features retrieved from the preceding levels to providea classification or prediction. They resemble the layers found in classic neural networks.*

C. *Long Short-Term Memory*

Recurrent neural networks can have their memory increased by using the **LSTM** model. **RNN** are superior to short-term memory networks because they enable the current neural networks' utilization of previously discovered data. For urgent tasks, the older data is used. It is possible that we do not have a complete history of the neural node. **LSTM** is extensively utilized in **RNN** as well as other neural networks. The memory cell structure of LSTM networks makes them excellent at collecting dependencies over time in sequential data. But compared to simpler models, they require more computing power and take longer to train [15]. The components of **LSTM** are:

i) **Input Gate(I):** The input gate's purpose is to measure the significance of the fresh data that the input

$$I = \sigma(x * u + h * w) \qquad (2)$$

- ***x****: Data entered at timestamp t.*
- ***u****: the input weight matrix*
- ***h****: A timestamp's hidden state from before*
- ***w****: The input weight matrix connected to the hidden state.*

ii) **Forget Gate (F):** The amount of the prior state that must be kept, or how much of the state ought to be forgotten, is determined by combining the input with the previous output to create a fraction between 0 and 1. This is the forget gate equation.

$$F = \sigma(x * c + h * r) \qquad (3)$$

Where,

- ***x****: the current timestamp's input.*
- ***c****: The input's associated weight*
- ***h****: The prior timestamp's hidden state*
- ***r****: The weight matrix connected to the concealed state is what it is.*

iii) **Output gate**: The current state is obtained by combining output of the tanh block with the recently generated scaling fraction, which is gated as before at the output gate using the input and previous state. After that, this output is given out.

D. *Naive Bayes Alogrithm*

It is a classification model that utilizes the Bayes theorem to create Bayesian networks for a given dataset. It is presumed that a specific characteristic in a class that is independent of every other feature is present in the provided dataset. For instance, some characteristics make an object deemed to be A. Its name, "Naïve," comes from the factthat all these traits, regardless of how they relate to one another or other features, independently increase the likelihood that this item is A. Text classification problems benefit greatly from the simplicity, speed, and efficacy of Naive Bayes, especially when dealing with small amounts and high-dimensional data. But it makes the assumption that features are independent, which is rarely the case in practical situations, which could result in errors. The following is the Bayesian theorem:

$$P(A|B) = P(B|A) \times P(A) / P(B) \qquad (4)$$

In which,

- Posterior probability, or the chance of hypothesis A considering observed event B, is represented by the symbol P(A|B).

- P(B|A) stands for potential probability, which is the likelihood that the evidence presented could support a hypothesis.
- The probability of the theory prior to examining the data is known as Prior Probability, or P(A).
- P(B) is the marginal probability or the probability of evidence.

E. LSTM-CNN Hybrid Algorithm

In this model, we combine **CNN** and **LSTM** together. This approach uses **LSTM** for data forecasting and **CNN** for extracting the temporal characteristic from the data. It can fully utilize the chronological order of stock price data to provide more accurate forecasts.[12] We selected to employ the **CNN-LSTM** neural network technique because it allows us to train on patterns while **CNN** tracks the characteristics of the dataset. For problems containing both temporal and spatial data, such as analysing videos or sequential text categorization, LSTM-CNN combines the best features of both LSTM and CNN, resulting in an extremely successful solution. But compared to either LSTM or CNN alone, it requires more computing power and is more difficult to create and train.

IV. RESULT

In our research work, we have obtained historical data of three highly capitalized companies that are listed on the New York Stock Exchange for our research project. Google (**GOOG**), Amazon (**AMZN**) and Apple (**AAPL**) are the names of the stocks. To assess the overall accuracy of themodel, we employed **RMSE** and **MAE** as performance metrics. Next, we implemented a variety of machine learning algorithms, including naïve bayes, **CNN**, **LSTM**, and hybrid **LSTM-CNN** models. The results are as follows:

F. ***Google(GOOG)***

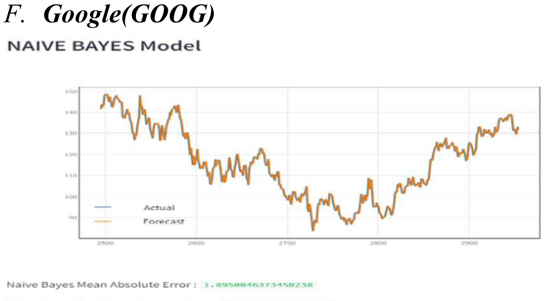

Fig.1. Graph of Naive Bayes Models of Google Stock

The **MAE** and **RMSE** on the naïve bayes model on Google stock is 1.89 and 2.50 as shown in the Fig.5. Underscoring the significance of customizing ML models to the distinct qualities of every financial instrument to attain the best outcomes.

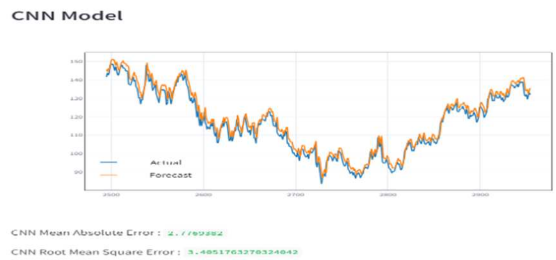

Fig.2. Graph of CNN Models of Google Stock

In Fig.6 the **MAE** and **RMSE** are 2.77 and 3.405, respectively. This observation highlights a common shortcoming of **CNN** models in managing abrupt and severe changes in the market, which may result in differences between the actual and expected values at such instances.

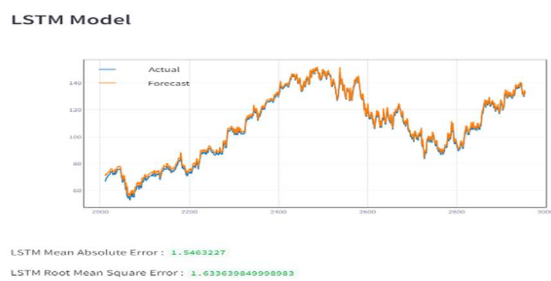

Fig.3. Graph of LSTM Models of Google Stock

The **MAE** and **RMSE** errors in the **LSTM** model are 1.54 and 1.63, respectively as shown in Fig.7.Additionally, it is noted that there is a slight difference between the actual and predicted values. Becauseof their strong feature extraction and context preservation capabilities, **LSTM** models are well-suited for identifying and deriving insights from temporal patterns in stock price fluctuations.

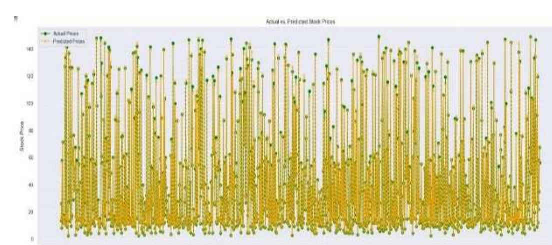

Fig.4. Graph of LSTM-CNN Model on Google Stock

The **LSTM CNN** hybrid model's **MAE** and **RMSE** values are 0.57 and 1.07, respectively. The performance matric is remarkably close to 1. When it comes to tasks requiring spatial processing, **CNNs (Convolutional Neural Networks)** are superior. The advantages of both models can be combined in hybrid **LSTM-CNN** models to improve performance on tasks requiring both spatial processing and long-term memory.

G. *Apple(AAPL)*

Fig.5. Graph of Naive Bayes Models of Apple Stock

The **MAE** and **RMSE** value of Apple stock on naïve bayes model is 2.26 and 2.96, respectively as shown in Fig.10. The values of performance metrics are close to three in this model. The actual and predicted values are close to each other even when there is a sudden rise or spike in the prices of stock.

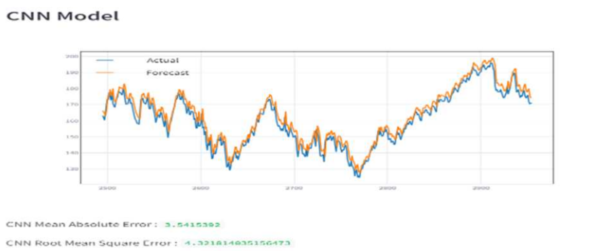

Fig.6. Graph of CNN Models of Apple Stock

MAE at RMSE values of **CNN** model on Apple stock is 3.54 and 4.32, respectively as shown in the Fig.11. That naive Bayes model has performed better than **CNN** model in case of Apple stock. We can also notice the delay in the actual and predicted values.

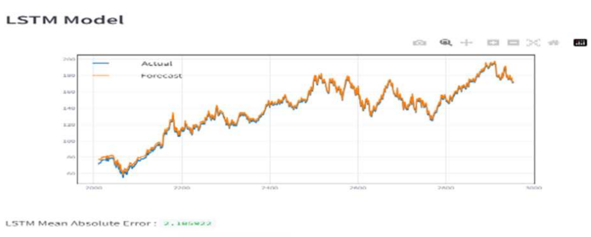

Fig.7. Graph of LSTM Models of Apple Stock

From Fig.12 the **MAE** and **RMSE** value of Apple stock on **LSTM** model is 2.10 and 2.22. The values are remarkably close to 2. However. The predicted price is always less than the actual price for apple stock.

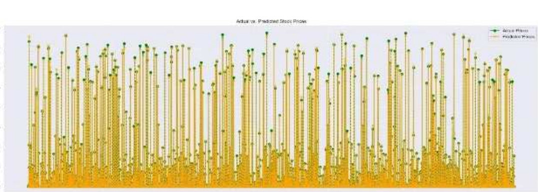

Fig.8. Graph of LSTM-CNN Model on Apple Stock

An **LSTM-CNN** hybrid model's **MAE** and **RMSE** errors on Apple stock are 0.57 and 1.07, respectively. There is a correspondence between the observed and anticipated values. The highest degree of broken line fitting that is consistent with one another is found in the **LSTM-CNN** model. As it utilizes **CNN**'s power in spatial feature extraction to improve short-term trend recognition and **LSTM**'s time-series skills to capture long-term dependencies, this hybrid model combines the best aspects of both models.

H. *Amazon(AMZN)*

Fig.9. Graph of Naive Bayes Models of Amazon Stock

MAE and **RMSE** values of naïve bayes model on amazon stock is 2.44 and 3.37, respectively as shownin the Fig.15. Because they are straightforward and simple to train, naive Bayes classifiers are a goodoption for tasks involving a small amount of training data. They can classify new data quickly, which makes them appropriate for real- time applications.

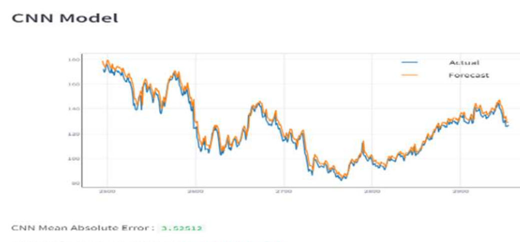

Fig.10. Graph of CNN Models of Amazon Stock

According to Fig.16, the **CNN** model's **MAE** and RMS values for Amazon stock are 3.52 and 4.504, respectively. In terms of Apple stock, the **CNN** model was outperformed by the naïve Bayes model.

Fig.11. Graph of LSTM Models of Amazon Stock

On applying **LSTM** model on amazon, the **MAE** and **RMSE** error is 1.88 and 2.00 as shown in Fig.17. When used on Amazon stock, it performs better than **CNN** and the naïve bayes model. **LSTM** models can be used to forecast the direction of changes in stock prices, which can help investors make more informed decisions.

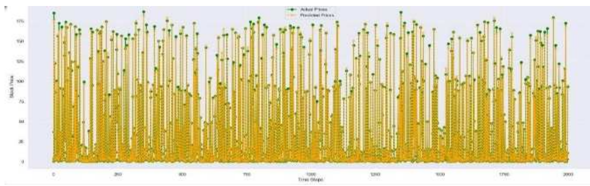

Fig.12. Graph of LSTM-CNN Model on Amazon Stock

The hybrid model has an **RMSE** error of 1.29 and an **MAE** error of 0.55. Additionally, compared to other models, it is more precise in predicting stock values for Amazon stock. It is therefore capable of providing extremely accurate stock price predictions. For tasks requiring both long-term memory and spatial processing, hybrid **LSTM-CNN** models combine the best features of both models.

Our team's analysis of the results using RMSE and MAE as performance metrics is shown in the overall table below:

Model\ Errors	AAPL		AMZN		GOOG	
	MAE	*RMSE*	*MAE*	*RMSE*	*MAE*	*RMSE*
LSTM Model	2.1031	2.2188	1.8881	2.0043	1.5463	1.6336
CNN Model	3.5368	4.3161	3.5251	4.5042	2.7769	3.4051
NAÏVE BAYES Model	2.2578	2.9650	2.4455	3.3746	1.8950	2.5072
LSTM-CNN Hybrid Model	0.2758	0.8101	0.5520	1.2927	0.5735	1.0778

Table 1. Four highly capitalised companies' RMSE and MAE

III. CONCLUSION

In our research work, we utilized a variety of machine learning and deep learning algorithms on popular blue-chip stocks, such as AAPL, AMZN and GOOG, and analyzed the outcomes using **RMSE** and **MAE** as performance metrics. After applying various algorithms in the given dataset required for our research paper, we realized that the **LSTM-CNN** hybrid model outperformed all the other algorithms we have applied. When **LSTM-CNN** was applied, the stock of Apple (AAPL) outperformed all other stocks and produced **MAE** and **RMSE** values of 0.2758 and 0.8101, respectively. The Amazon (AMZN) stock performed worst when the **LSTM-CNN** hybrid model was applied with **MAE** and **RMSE** values of 0.552 and 1.2927, respectively.

Google (GOOG) stock gave the best results when the linear model was applied with **MAE** and **RMSE** values as 8.4709 and 10.6690, respectively. The Naive Bayes model outperformed the **CNN** model in allstocks and gave promising results. But when the stock prices of the blue-chip companies suddenly increased or decreased, the **CNN** model was unable to provide reliable figures. Google stock also gave the best results when **LSTM** was applied with **MAE** and **RMSE** values as 1.5463 and 16336, respectively. In the proposed work, the **LSTM-CNN** hybrid model outperformed all the other algorithmswe have applied, and it was able to capture sudden variations in the movement of prices of the stocks.

REFERENCES

[1] M. Vijh et al, (2020). "Stock Closing Price Prediction using Machine https://doi.org/10.1016/j.procs.2020.03.326

[2] W. Lu et al, (2020). "A CNN-LSTM-based model to forecast stock prices." Volume 2020, Article ID 6622927.https://doi.org/10.1155/2020/6622927

[3] E. A. Gopalakrishnan et al, (2018). "NSE Stock Market Prediction Using Deep-Learning Models." Volume 132, Pages 1351-1362. https://doi.org/10.1016/j.procs.2018.05.050

[4] H.Hamoudi et al,"Stock Market Prediction using CNN and LSTM". https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Stock+Market+Prediction+using+CNN+and+LSTM++Hamdy+Hamoudi+%2C+Mohamed+A+Elseifi+++&btnG=

[5] A Shakva et al, (2018). "Real-Time Stock Prediction using Neural Network." 2018 8th InternationalConference on Cloud Computing, Data Science & Engineering (Confluence). https://doi.org/10.1109/CONFLUENCE.2018.8443057

[6] A. Moghar et al, (2020). "Stock Market Prediction Using LSTM Recurrent Neural Network ", Volume 170, Pages 1168-1173. https://doi.org/10.1016/j.procs.2020.03.049

[7] S. Banerjee et al, "Stock Market Prediction", International Journal of Innovative Technology and Exploring Engineering 9(9):2278-3075. http://dx.doi.org/10.35940/ijitee.I7642.079920

[8] Ampomah E.K et al, "Evaluation of Tree-Based Ensemble Machine Learning Models in Predicting Stock Price Direction of Movement," Information 2020, 11(6), 332. https://doi.org/10.3390/info11060332

[9] A. F. Adekoya et al, (2020). "A comprehensive evaluation of ensemble learning for stock-market prediction." Article number: 20. https://doi.org/10.1186/s40537-020-00299-5

[10] Ayitey Junior et al., 2022 **,** "Forex market forecasting with two-layer stacked Long Short-Term Memory neural network (LSTM) and correlation analysis," *J. Electr. Syst. Inf. Technol. 2022 91*, vol. 9, no. 1, pp. 1–24, doi: 10.1186/S43067-022-00054-1.

[11] Artha, A., Antoni, W., & Norhaslinda, Z. A. (2020). "Sectoral stock prediction using convolutional neural networks with candlestick patterns as input images", International Journal of Emerging Trends in Engineering Research. https://doi.org/10.30534/ijeter/2020/07862020

[12] Cavalli, S., & Amoretti, M. (2021). "CNN-based multivariate data analysis for bitcoin trend prediction",Applied Soft Computing, 101, Article 107065. https://doi.org/10.1016/j.asoc.2020.107065

[13] Abbasimehr H, Paki R (2022), "Improving time series forecasting using LSTM and attention model", J Ambient Intell Human Comput 13:673–691. https://doi.org/10.1007/s12652-020-02761-x

[14] Chen, W., Jiang, M., Zhang, W. G., & Chen, Z. (2021). "A novel graph convolutional feature based convolutional neural network for stock trend prediction", Information Sciences, 556, 67–94. https://doi.org/10.1016/j.ins.2020.12.068

[15] Ding, G., & Qin, L. (2020). "Study on the prediction of stock prices based on the associated network modelof LSTM", International Journal of Machine Learning and Cybernetics, 11 (6), 1307–1317. https://doi.org/10.1007/s13042-019-01041-1