

A Hybrid Analysis of Stock Price Prediction using Vector Auto-Regressive (VAR), Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU)

Krishna Veni.M*(Reg No: 20214542152020), Arumugam. P** and Pradeep Nijanthan. P***
Research Scholar*, Professor**and Research Scholar***, Department of Statistics,
Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu
krish20393@gmail.com;sixfacemsu@msuniv.ac.in; pradeepnijanthanp@gmail.com

ABSTRACT

Railway sector in India vast developing industry for which government of India creates the stocks IRFC and Railvikas to contributes the railway sector by make the investors invest in these stocks. This study, forecast the future trends of these two stocks at the same time using the statistical method and the machine learning methods. The performance evaluation metrics MAE and MSE are used in the model to conform the model accuracy. In addition to providing useful solutions for the issues faced by stock capitalists, accurate stock forecasting would surely be extremely advantageous for stock exchange governments.

1. INTRODUCTION

Indian Railways Finance Corporation (IRFC), under the Ministry of Indian Railways, which provides the financial arm to the railway projects, plays a crucial role in meeting the 'Extra Budgetary Resources' (EBR) needs of Indian Railways through market borrowing. Over the years, IRFC has financed over Rs. 4 lakh crores, facilitating the procurement of essential rolling stock for Indian Railways. Additionally, it has contributed to funding railway schemes, infrastructure projects, and capacities, ensuring the sustainable expansion of the railway sector. As a prominent infrastructure financier, IRFC has been publicly listed on the National Stock Exchange (NSE) and Bombay Stock Exchange (BSE) since January 29, 2021. With the President of India as its promoter, IRFC is a key financial arm of Indian Railways. Its transformation from a public limited company to a recognized Non-Banking Financial Company (NBFC) underscores its essential contribution to national infrastructure. Adhering to prudential norms, IRFC continues to finance railway projects, supporting the sustainable development of India's transportation infrastructure. IRFC also been lending to avarious entity in the railway sector like Rail Vikas Nigam Limited (RVNL), RailTel, Konkan Railway Corporation Limited (KRCL), Pipavav Railway Corporation Limited (PRCL), etc. Nowadays traders and investors invest more in this government stocks to develop more infrastructure in country. More Technical analysis is done by a researcher and the investors to buy a stock and invest in the governmental stocks. A fundamental statistical time series approach, utilizing the Vector Autoregression (VAR) model, forecasts stock prices by predicting both opening and closing prices concurrently. And the deep learning models used to evolve, stock market forecasting has transitioned from conventional approaches to cutting-edge deep learning methodologies. Moreover, stock market forecasting has evolved from traditional methods to advanced deep learning techniques, contemporary methods now encompass Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs). Recently, there has been a growing interest in utilizing Transformer-based models and reinforcement Learning (RL) techniques for

predicting stock market trends. However, numerous surveys on stock market forecasting have limitations, such as focusing on outdated technologies and using vague classifications for models. Previous research has tackled numerous challenges and unresolved issues in stock market prediction, leading to notable advancements. By the LSTM and GRU models, which make the prediction of the IRFC stocks open and close price at the same time, these multiple variable cases. The error metrics in the statistics are used to find the perfect fit of the model they are Mean Squared Error (MSE) and Mean Absolute Error (MAE) both gives the model error in terms of differencing the absolute and predicted values.

2. REVIEW OF LITERATURE

Shiri, F. M., Perumal, T., Mustapha, N., & Mohamed, R. (2023), conducted extensive research toward deep learning models, which include CNN, RNN, LSTM, and GRU, analyzing their sequence modeling capabilities. which displayed how well LSTM and GRU capture long-term dependencies compared to traditional models, making them suited to financial time series forecasting. The study shows GRU's effectiveness in lowering computational complexity without sacrificing accuracy while outlining each model's strengths and weaknesses. In the context of learning symbolic sequences. *Cahuantzi, Roberto & Chen, Xinye & Güttel, Stefan (2023)* compared LSTM and GRU. In response to their findings, GRU is a strong candidate for financial forecasting tasks since it can perform comparably to LSTM while using less computing power. A hybrid GRU-LSTM network for foreign exchange currency rate prediction found by *Islam, M. S., & Hossain, E. (2021)*. The study concluded that applying the advantages of GRU's computational efficiency and LSTM's long-term memory retention, combining the two architectures yields a predictive performance. For stock market forecasting, where it's critical to capture both short-term swings and long-term patterns, this hybrid approach might prove advantageous. Multivariate LSTM and GRU networks were used by *Primananda, S. B., & Isa, S. M. (2021)*, to forecast gold prices in Rupiah, demonstrating that the inclusion of multiple economic indicators improves prediction accuracy. The study emphasizes the value of multivariate analysis, which adds more financial variables to forecasting models to increase their robustness. The findings also demonstrate that deep learning methods capture nonlinear market patterns more accurately than traditional models.

3. METHODOLOGY

Vector Autoregression (VAR) is a forecasting method suitable for situations where two or more time series have mutual influences on each other, indicating that the relationship between the involved timeseries is reciprocal. It is classified as an Autoregressive model because each variable (time series) is represented as a function of its past observations, meaning that the predictors are essentially the lags (time-delayed values) of the series. In contrast to other Autoregressive models like AR, ARMA, or ARIMA, which are unidirectional where predictors only affect the Y variable and not the other way around, VAR models are characterized by their bi-directional nature, where variables affect one another.

$$Y_t = \sum_{j=1}^p \Phi_j Y_{t-j} + a_t \dots (1)$$

$$E(a_t a_t') = \begin{bmatrix} E(a_{1t}^2) & E(a_{1t}a_{2t}) & \cdots & E(a_{1t}a_{nt}) \\ E(a_{2t}a_{1t}) & E(a_{2t}^2) & \cdots & E(a_{2t}a_{nt}) \\ \vdots & \vdots & \ddots & \vdots \\ E(a_{nt}a_{1t}) & E(a_{nt}a_{2t}) & \cdots & E(a_{nt}^2) \end{bmatrix} \text{ with } E(a_t a_s') = 0 \text{ for } t \neq s$$

It written in the form,

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + a_t \dots (2)$$

This VAR model can be expressed in matrix notation as VAR(p). Also, Y_t , a_t , Φ_j and Y_{t-j} can express in the matrix form

$$Y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix}, a_t = \begin{bmatrix} a_{1t} \\ a_{2t} \\ \vdots \\ a_{nt} \end{bmatrix}, \Phi_j = \begin{bmatrix} \Phi_{11}^j & \Phi_{12}^j & \cdots & \Phi_{1n}^j \\ \Phi_{21}^j & \Phi_{22}^j & \cdots & \Phi_{2n}^j \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{n1}^j & \Phi_{n2}^j & \cdots & \Phi_{nn}^j \end{bmatrix} \text{ and}$$

$$Y_{t-j} = \begin{bmatrix} y_{1t-j} & y_{2t-j} & \cdots & y_{nt-j} \\ y_{1t-j-1} & y_{2t-j-1} & \cdots & y_{nt-j-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1t-j-p} & y_{2t-j-p} & \cdots & y_{nt-j-p} \end{bmatrix} \dots (3)$$

3.1 LSTM

Long - Short Term Memory (LSTM) Network is a special Recurrent Neural Network (RNN) specifically designed to process and predict time series data. Compared with standard RNNs, LSTM captures and remembers long term dependencies effectively through a gating mechanism suitable for modelling and predicting long sequences.

A typical LSTM cell consists of three gates: the forget gate f_t (1), which determines the extent of the previous information to be discarded; the input gate i_t (2), which assesses the information that will be stored in the cell memory; and the output gate o_t (3), which decides how to compute the output based on the current information.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \dots (4)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \dots (5)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \dots (6)$$

Here, the W , R , and b variables denote matrices and vectors of trainable parameters. The LSTM unit is defined by

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \dots (7)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \dots (8)$$

$$h_t = o_t \odot \tanh(C_t) \dots (9)$$

3.2 GRU

The Gated Recurrent Unit (GRU) is a different variant of the RNN architecture that tackles the problem of short-term memory while providing a more straightforward structure compared to LSTM. Instead of having separate input and forget gates like LSTM, GRU merges them into a single update gate, creating a more efficient design. In contrast to LSTM, GRU lacks an independent cell state. A GRU unit is comprised of three primary components: an update gate, a reset gate, and the current memory content. These gates allow the GRU to selectively refresh and harness information from earlier time steps, enabling it to understand long-term dependencies within sequences.

The update gate controls the extent to which previous information is preserved and integrated with the current input at a given time step. It is computed based on the concatenation of the previous hidden state h_{t-1} and the current input x_t , followed by a linear transformation and a sigmoid activation function.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \dots \quad (10)$$

The reset gate decides how much of the past information should be forgotten. It is computed in a similar manner to the update gate using the concatenation of the previous hidden state and the current input.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \dots \quad (11)$$

The present memory state is determined by the reset gate along with the combination of the altered previous hidden state and the current input. This outcome is then processed through a hyperbolic tangent activation function to generate the candidate activation.

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h) \dots \quad (12)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t \dots \quad (13)$$

3.3 Mean squared error

Mean squared error (MSE) is an error metric used to measure the average squared difference between the predicted values and the actual values in the dataset. It is calculated by taking the average of the squared residuals, where the residual is the difference between the predicted value and the actual value for each data point.

$$MSE = \frac{1}{n} \sum_{i=1}^n [y_i - \hat{y}_i]^2 \dots \quad (14)$$

where, y_i is the i^{th} observed value, \hat{y}_i is the corresponding predicted value, n = the number of observations.

3.4 Mean Absolute Error

MAE is a statistical measure that quantifies the average magnitude of errors in a set of predictions, without considering their direction. It is the average of the absolute differences between predicted values and actual values.

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x| \dots \quad (15)$$

where, n = number of observations or errors, x_i = individual predicted or measured value, x = actual value, $|x_i - x|$ = absolute error for each observation

3.5 DATA SOURCE

The data collected from the date 03.01.2022 to 29.12.2023 of Railvikas and IRFC stocks, which consist of the variables open, high, low, previous day close, volume and close price. For this study, close price of the both stocks used. Data Collected from the national stock exchange (NSE) official website ([data](#))

4. RESULTS

The observed data for the trading period of 494 days. The Statistical prediction method Vector Autoregressive model (VAR) which handles two-time dependent series at the same time (Close price) of the both Indian stocks.

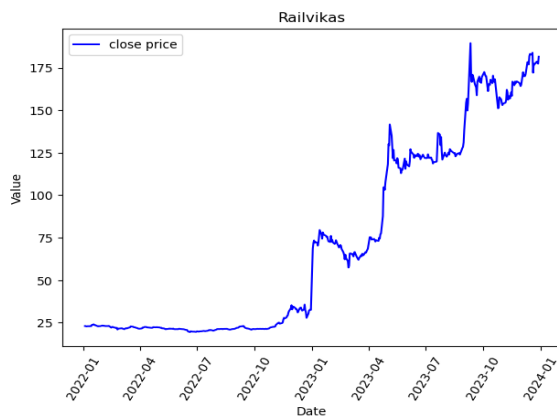


Fig 1. Time Plot of Railvikas close price

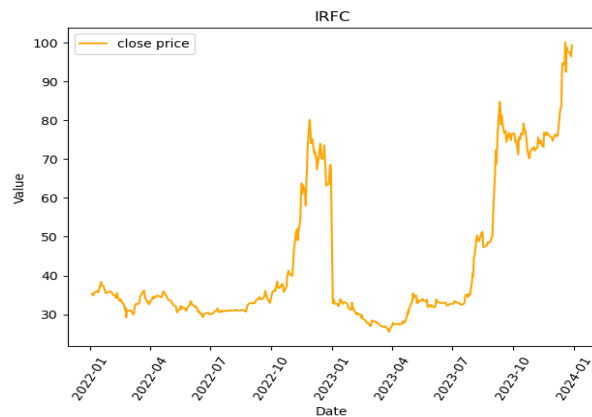


Fig 2. Time Plot of IRFC close price

The Fig 1 displays the close price of the Railvikas stock data from the historical of the trading days and similarly, the Fig.2 shows the IRFC stock close price. This data follows the Time series and the VAR model castoff into forecast the future prices, dates and values are plotted in the X-axis and Y-axis.

To frame the VAR model to the stock prices data should follows the stationarity. For that ADF tested for the stationarity of Stock open values and close values.

ADF Test	Railvikas	IRFC
P-Value	0.9780152357370996	0.9752345691390887
	Non-Stationary	Non-Stationary

Table 1. ADF Test (without differentiation)

ADF Test	Railvikas	IRFC
P-Value	0.0000000000000173	0.0000000000006906946
	Stationary	Stationary

Table 2. ADF Test (after 1st order differentiation)

Results from the ADF test resulted that values are not Stationary to make stationary make 1st order differentiation. After the 1st order differentiation all the values in data are stationary.

	Railvikas Close price	IRFC close price
Railvikas Close price	1.000000	-0.018094
IRFC close price	-0.018094	1.000000

Table 3. Correlation matrix of residual

In the Table 3 shows the how the residuals are correlated to the both Railvikas and IRFC stocks in the matrix. The relationship between the residuals of Rail Vikas and IRFC is -0.018094 is very close to zero. This shows there is almost no linear relationship between the errors of these stock price predictions.

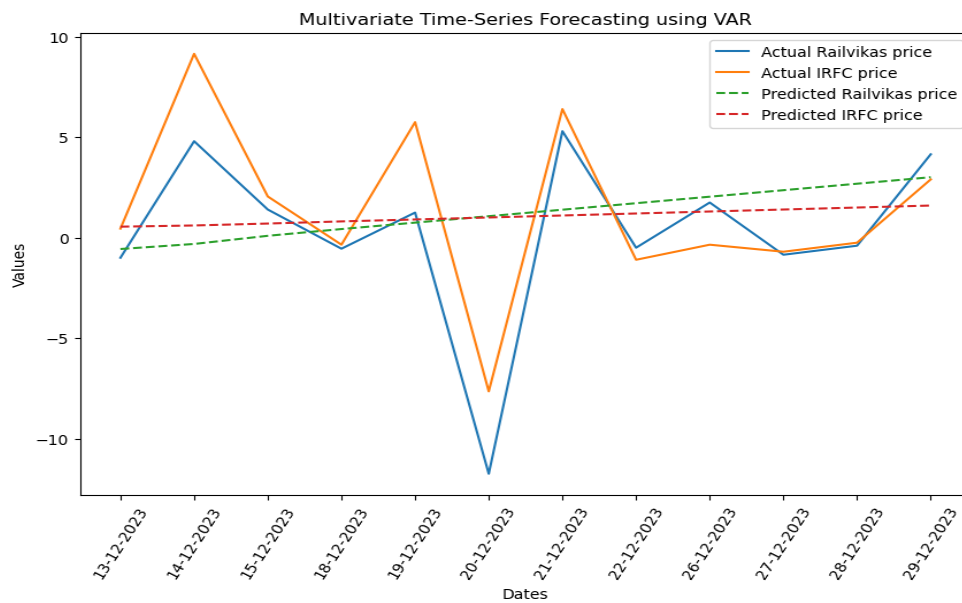


Fig 3. VAR model actual, predicted values for Rail vikas and IRFC close price

Table 4. Error metric of VAR model

Model	Error metric	Rail vikas	IRFC
VAR	MAE	2.9171	3.2541
	MSE	19.5803	18.3150

VAR model results in the line chart the actual and forecast values of the open and close prices of the IRFC stock. The Error metric for the VAR model resulted in the table 4. The MSE and MAE values of Rail vikas are 2.9171 and 19.5803 for IRFC are 3.2541 and 18.3150

The Deep learning techniques RNN (Recurrent Neural Network) methods LSTM and GRU techniques used. For this technique 80% of the data used as train model and 20 % as a test model.

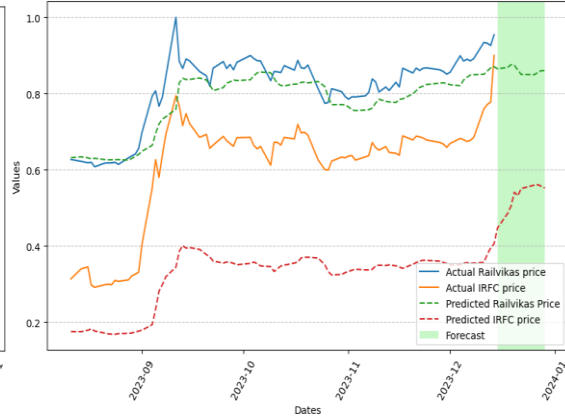
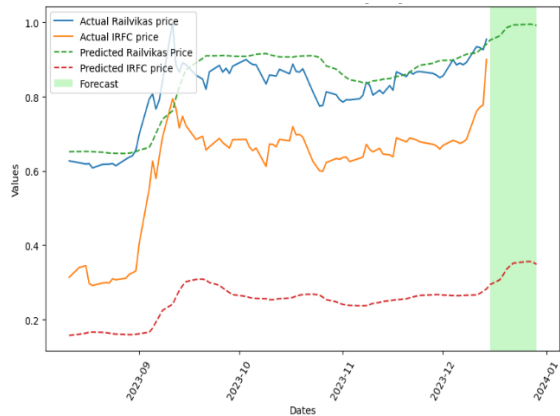


Fig 4. LSTM actual, predicted and forecast values of Railvikas and IRFC close prices

Fig 5. GRU actual, predicted and forecast values of Railvikas and IRFC close prices

Fig.4& Fig. 5 Illustrates the actual and the predicted values in the line chart for the both LSTM and GRU models and it include the close price of the two stocks (Railvikas and IRFC) variables in the dataset at the same time it forecast the values of the close prices in the future. Blue and orange line in the graph shows actual values of the both stocks and the dotted lines in red and green as the predicted values of the both model in the tow figures and the green area shows the forecast values of the both models

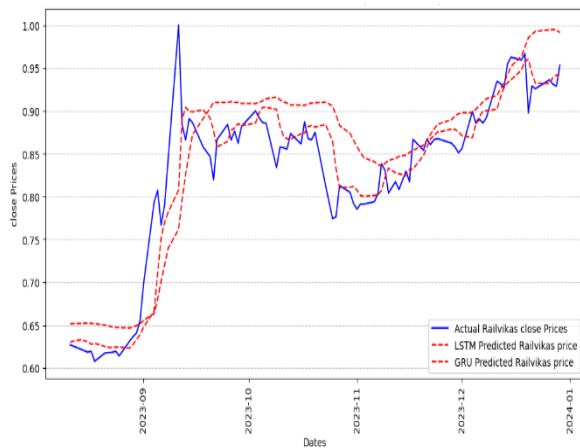


Fig 6. LSTM and GRU model actual and predicted values for Railvikas close price

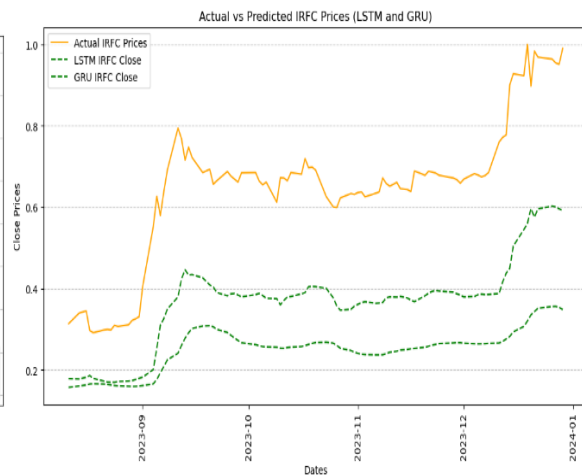


Fig 7. LSTM and GRU model actual and predicted values for IRFC close price

Furthermore, the Fig. 6 and Fig. 7 close price of the stocks divided and the model is compared with respect to actual close price with LSTM and GRU model forecasted values and the both red dotted lines in fig 6 shows the LSTM and GRU prospective values of the Railvikas stocks and green dotted lines in fig 7. displays the both models predicted values of the IRFC stocks

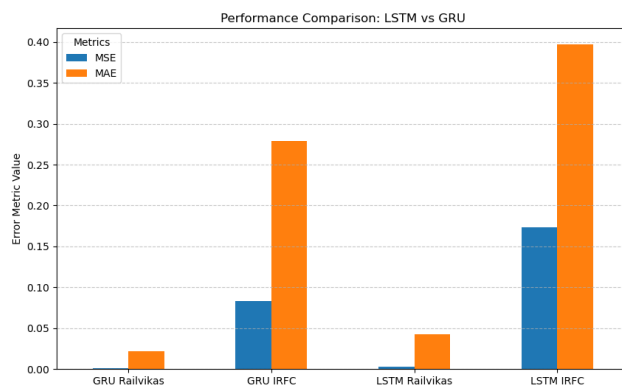


Table 5. Error metric of GRU and LSTM model

Model	Error metric	Railvikas	IRFC
GRU	MAE	0.0458	0.3072
	MSE	0.0032	0.1016

LSTM	MAE	0.0427	0.3967
	MSE	0.0031	0.1736

Fig. 8 MSE, MAE values of GRU and LSTM model

The error metric of the both GRU and LSTM model to the Rail vikas and IRFC stock resulted as in the *Table 5*. In this study, VAR, GRU and LSTM used for the comparison of multivariate time dependent dataset in single time, VAR model is a statistical technique is used to forecast and LSTM, GRU predicts the prices by the deep learning algorithm. Here open and close price of the stock are the two variables, this multivariate time dependent dataset applied these model and fitting of the best model is concluded

5. CONCLUSION

The daily prices of the IRFC and Railvikas stocks' close prices predictions are used the past prices. The statistical model and the machine learning models used for the forecast by comparison of the models and the error metrics MSE and MAE used for the model fitness for the both stocks statistics and machine learning both models make different approach to predict the prices. The performance evaluation metrics MAE and MSE results in the (*Table 4*) for the VAR model for Railvikas, MAE and MSE are resulted the model is not good fit for both stock data. The machine learning technique LSTM and GRU model error metric shows in the (*Tables 5*) which is very lower in the error metric with compare to the VAR model. In comparison of the MSE and MAE values, of the three model's LSTM performs slightly better than GRU for Rail Vikas, as it has a lower MAE (0.0427 vs. 0.0458) and a lower MSE (0.0031 vs. 0.0032). GRU performs significantly better than LSTM for IRFC, as it has lower MAE (0.3072 vs. 0.3967) and lower MSE (0.1016 vs. 0.1736). If predicting Rail Vikas, LSTM is a better choice due to slightly lower errors. If predicting IRFC, GRU is significantly better and should be preferred.

6. DISCUSSION

Indian railways sector is crucial to grow the Indian economy which stocks drastically improving in the recent years the increase in the stock price makes the prediction and find the trend of the stocks using the statistical methods. In this research, machine learning techniques, and statistical techniques are used to finding the pattern and trends of the two stocks at a same time using the evaluation techniques MSE and RMSE to find the accuracy of the all techniques and resulted the best one. This study tells the accuracy of the model in further study researcher can do the future points prediction using the best model evaluated. A hybrid model (GRU-LSTM) could be explored to improve forecasting performance for both stocks. IRFC may have a more volatile or nonlinear price movement, making GRU's simpler gating mechanism more effective. LSTM is better at capturing long-term dependencies, while GRU is often preferred for shorter-term trends. The nature of each stock's price movements may favour one model over the other.

7. REFERENCES

1. Shiri, F. M., Perumal, T., Mustapha, N., & Mohamed, R. (2023). *A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU*. arXiv preprint arXiv:2305.17473.
2. Cahuantzi, Roberto & Chen, Xinye & Güttel, Stefan. (2023). *A Comparison of LSTM and GRU Networks for Learning Symbolic Sequences*. 10.1007/978-3-031-37963-5_53.
3. Pahlevi, M. R., Kusrini, K., & Hidayat, T. . (2023). *Comparison of LSTM and GRU Models for Forex Prediction*. Sinkron : Jurnal Dan Penelitian Teknik Informatika, 7(4), 2254-2263.
<https://doi.org/10.33395/sinkron.v8i4.12709>
4. Islam, M. S., & Hossain, E. (2021). *Foreign exchange currency rate prediction using a GRU-LSTM hybrid network*. Soft Computing Letters, 3, 100009.
<https://doi.org/10.1016/j.socl.2020.100009>
5. Ulina, M., Purba, R., & Halim, A. (2020, November 3). *Foreign Exchange Prediction using CEEMDAN and Improved FA-LSTM*. 2020 5th International Conference on Informatics and Computing, ICIC 2020.
<https://doi.org/10.1109/ICIC50835.2020.9288615>
6. Mochamad Ridwan, Kusman Sadik, Farit Mochamad Afendi (2023). *Comparison of ARIMA and GRU Models for High-Frequency Time Series Forecasting*. Scientific Journal of Informatics.
7. Zhang, S., & Zhou, N. (2023). *Extraction of Multiple Electrical Parameters From IP-Affected Transient Electromagnetic Data Based on LSTM-ResNet*. IEEE Transactions on Geoscience and Remote Sensing, 61, 1-14.
8. Zhu, F., Liu, L., & Lin, T. (2020). *An LSTM-based Traffic Prediction Algorithm with Attention Mechanism for Satellite Network*. Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition.
9. Primananda, S. B., & Isa, S. M. (2021). *Forecasting Gold Price in Rupiah using Multivariate Analysis with LSTM and GRU Neural Networks*. Advances in Science, Technology and Engineering Systems Journal, 6(2), 245–253. <https://doi.org/10.25046/aj060227>
10. Mishra, P.K. & J.R.Das, & Mishra, Santosh. (2010). *Gold Price Volatility and Stock Market Returns in India*. American Journal of Scientific Research. 9. 47-55.
11. Vinayakumar, R., Soman, K.P., Poornachandran, P., & Kumar, S.S. (2018). *Detecting Android malware using Long Short-term Memory (LSTM)*. J. Intell. Fuzzy Syst., 34, 1277-1288.