International Journal of Smart Sensor and Adhoc Network

Volume 4 | Issue 1 Article 2

April 2023

Deep Learning-based Gated Recurrent Unit Approach to Stock Market Forecasting: An Analysis of Intel's Stock Data

Nrusingha Tripathy
Siksha 'O' Anusandhan, nrusinghatripathy654@gmail.com

Ibanga Kpereobong Friday
kpereib01@gmail.com

Dharashree Rath
GITA Autonomous College, dharashree96@gmail.com

Debasish Swapnesh Kumar Nayak
swapnesh.nayak@gmail.com

Subrat Kumar Nayak
Siksha 'O'Anusandhan, subratsilicon28@gmail.com

Recommended Citation

Tripathy, Nrusingha; Kpereobong Friday, Ibanga; Rath, Dharashree; Nayak, Debasish Swapnesh Kumar; and Nayak, Subrat Kumar (2023) "Deep Learning-based Gated Recurrent Unit Approach to Stock Market Forecasting: An Analysis of Intel's Stock Data," *International Journal of Smart Sensor and Adhoc Network*: Vol. 4: Iss. 1, Article 2.
DOI: 10.47893/IJSSAN.2023.1234
Available at: https://www.interscience.in/ijssan/vol4/iss1/2

Deep Learning-based Gated Recurrent Unit Approach to Stock Market Forecasting: An Analysis of Intel's Stock Data

Nrusingha Tripathy[1*], Ibanga Kpereobong Friday[1], Dharashree Rath[2], Debasish Swapnesh Kumar Nayak[3], Subrat Kumar Nayak[3*]

Department of Computer Science & Engineering, Institute of Technical Education and Research, Siksha 'O'Anusandhan (Deemed to be) University, Bhubaneswar, India-751030 [1*,1,3,3*]

Department of Computer Science & Engineering, GITA Autonomous College, Bhubaneswar, India-752054 [2]

nrusinghatripathy654@gmail.com[1*], kpereib01@gmail.com[1], dharashree96@gmail.com[2], swapnesh.nayak@gmail.com[3], subratsilicon28@gmail.com[3*]

Abstract— The stock price index prediction is a very challenging task that's because the market has a very complicated nonlinear movement system. This fluctuation is influenced by many different factors. Multiple examples demonstrate the suitability of Machine Learning (ML) models like Neural Network algorithms (NN) and Long Short-Term Memory (LSTM) for such time series predictions, as well as how frequently they produce satisfactory outcomes. However, relatively few studies have employed robust feature engineering sequence models to forecast future prices. In this paper, we propose a cutting-edge stock price prediction model based on a Deep Learning (DL) technique. We chose the stock data for Intel, the firm with one of the quickest growths in the past ten years. The experimental results demonstrate that, for predicting this particular stock time series, our suggested model outperforms the current Gated Recurrent Unit (GRU) model. Our prediction approach reduces inaccuracy by taking into account the random nature of data on a big scale.

Keywords— Stock market prediction, Deep learning, Machine learning, Gated recurrent unit

I. INTRODUCTION

The field of computational finance has seen a significant amount of attention in recent years, with a focus on the stock price prediction. The stock market has the potential for high returns and has grown in popularity among investors. However, one of the most challenging tasks in finance is predicting stock movement, and this is a topic of great interest to both institutional and private investors. The ability to predict stock prices can provide investors with the necessary information to make informed decisions and potentially increase their returns. Over the years, various techniques have been employed to capture these fluctuations, ranging from traditional models to more advanced techniques like Neural Network models [1, 2]. This is an area of ongoing research, with various techniques and models being studied across multiple fields, including physics, economics, computer science, and statistics. However, previous studies have focused mainly on historical stock prices and technical indicators, and there is a lack of research on incorporating other factors [3-8]. Among these techniques, Gated Recurrent Unit (GRU) models have gained significant attention in recent years due to their efficiency and potential for predicting stock prices. The fundamental difference between GRU and Long Short-term Memory (LSTM) models is that GRU lacks an output gate, whereas LSTM has one. This aspect of GRU architecture makes it more effective in capturing the variations in stock prices, as well as a reduced model training time and their potential for predicting stock prices [9]. Overall, this study aims to provide a comprehensive understanding of the application of GRU models in stock price prediction and how it differs from traditional models. We believe that the insights gained from this study will be useful for investors, traders, and researchers in the field of computational finance [10]. Therefore, in this study, we aim to forecast stock prices taking into account multiple factors, including both opening and closing prices. We will delve into the intricacies of GRU models and their application in stock price prediction. Lastly, we present our proposed GRU model for stock price prediction and evaluate its performance using different evaluation metrics [11-13].

II. THEROTICAL FRAMEWORK

Basically, seven number of attributes we taken for this model except open and closing prices. To avoid overfitting, we use deep learning model. The train and test samples (20% and 80%) randomly generated.

A. Least Absolute Shrinkage and Selection Operator (LASSO)

It is crucial to recognize and make use of the most relevant and significant aspects that influence stock prices to make correct predictions. However, traditional methods may not always be able to capture all of these factors, leading to variations in the model. One solution to this problem is to use multidimensional variables, which capture a broader range of information and reduce the impact of missing data. Additionally, variable selection is a crucial aspect of statistical modelling in finance. Identifying the most informative and independent variables is essential to making accurate predictions. One estimation technique that can effectively condense a large number of variables is LASSO (Least Absolute Shrinkage and Selection Operator) [14, 15].The core concept of LASSO is the use of a penalty function that compresses regression coefficients and sets others to zero, resulting in a more precise model. This approach also addresses the issue of collinear data, which can occur when multiple variables are highly correlated. Minimizing the average of squares of the remainder while guaranteeing that each one of the actual values of the coefficients of regression is less than a constant is the fundamental idea behind LASSO. to acquire a comprehensible model as shown in Fig 1 Incorporating LASSO into the standard linear regression model allows for a more robust and accurate prediction of stock prices [9].

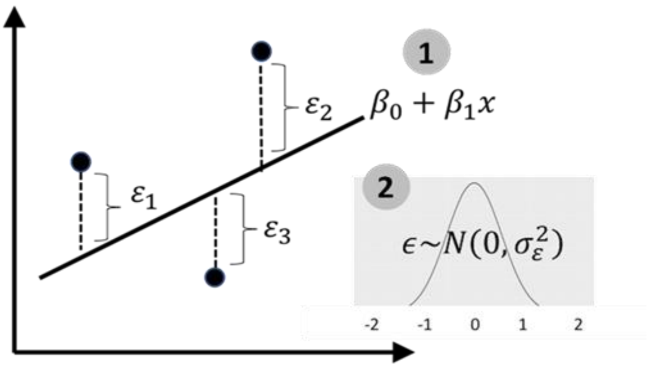

Fig 1. LASSO Linear Model

$$\hat{\beta} = \arg\min_{\beta} (y - X\beta)^{T} (y - X\beta) \quad (1)$$

The formula for LASSO is:

$$\hat{\beta} = \arg min_{\beta} \{(y - x_{\hat{\beta}})^{T} (y - x_{\beta}) + \lambda\beta_1\} \quad (2)$$

Where λ is called the penalty parameter that is specified by the user of LASSO. The larger the parameter λ, the more zeros in $\hat{\beta}$.

B. Principal Component Analysis (PCA)

Large amounts of data are more common than ever and are often difficult to understand. By minimizing information loss, the PCA approach reduces the dimensionality of huge datasets while improving user interpretability [16]. This is accomplished by creating new, independent variables that each individually optimize variance. Since it is fundamentally a statistical approach, statisticians have contributed significantly to its evolution even though it is used and occasionally reinvented in many other fields as seen in Fig 2, Accordingly, "keeping as much variability as feasible" entails identifying new uncorrelated variables, linear functions of the original dataset's variables, and sequentially maximizing variance. Solving an eigenvalue/eigenvector problem is all it takes to discover these additional variables, the main components (PCs) [17].

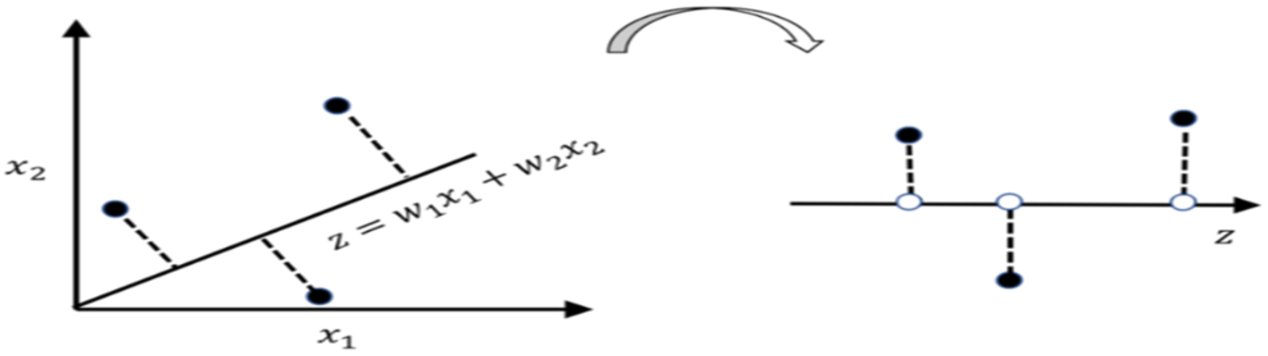

Fig 2. PCA Linear Model

The formula for LASSO is:
$w_{(1)} = \arg\max \{w^T_{(1)} X^T X w_{(1)}\}$ (3)

Where w_1 and w_2 are the free parameters, and x_1 and x_2 are the variables.

C. Gated Recurrent Units (GRU)

The Gated Recurrent Unit (GRU), as shown in Fig 3, is the most modern addition to the sequence modelling techniques after RNN and LSTM. Consequently, it offers a benefit over LSTM and RNN. The GRU is a simplified version of the LSTM and offers benefits over the traditional RNN. The GRU makes use of several gates to control information flow, much like an LSTM. Because of this, they improved on LSTM and used a straightforward design [12]. The GRU merely possesses a hidden state (Ht), compared to the LSTM which also has a discrete cell state (Ct). They are simpler to teach because of their clear structure. For each timestamp t, it receives an input Xt and the concealed state Ht-1 from period t-1. The subsequent timestamp is then assigned a new concealing state, Ht, which is printed later [18, 19].

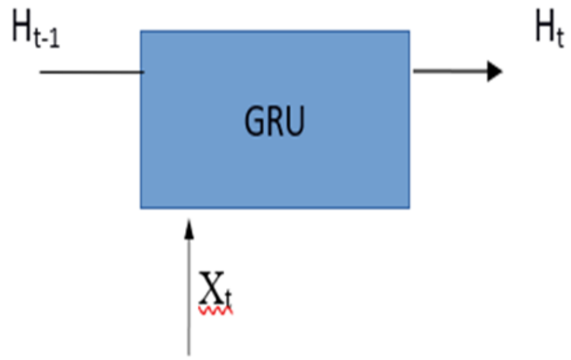

Fig 3. Architecture of GRU

Where: H_t = hidden state, X_t = input, H_{t-1} = previous state
Compared to an LSTM, a GRU has two gates. Reset gate and Update gate are the first two.

- Reset gate (Short term memory) – It is in charge of the network's short-term memory (hidden state). The reset gate's equation is shown below.
  $rt = \sigma\,(X_t * U_r + Ht{-}1 * W_r)$ (4)
  Here U_r A W_r are weight matrices.
- Update gate (Long Term Memory) – The equation is presented below. It also features an Long-term memory update gate.
  $u_t = \sigma\,(X_t * U_u + H_{t-1} * W_u)$ (5)
  Here Uu A Wu are weight matrices.

It is expected that recursive networks will continue to take on new forms. In addition, GRU seeks to resolve the vanishing gradient issue. The cell state and output gate seen in LSTM are not present in GRU. GRU is trained using Backpropagation Through Time (BPTT), a variation of the backpropagation technique. To lessen the discrepancy between the production that was expected and what was actually produced, the model's weights are gradually adjusted during the training phase [20]. The model is presented with unseen input data during testing, and the output is used to evaluate the model's performance. The input sequence's hidden state is transmitted from one one-time step to the next and stored in the hidden layer of GRU. The hidden state is updated at each time step using the reset and update gates.

III. METHODOLOGY

A. Data Selection

The dataset was obtained from INTEL and it was divided into two sections, with 80% covering the period from July 2012 to July 2020 being meant for training the model and the remaining 20% being meant for testing the model. Fig 4 shows the overall stock open and close price as year basis. The month-wise description of the data showing the open and close prices is described in Table I.

Fig 4. Overall Stock Open and Close Price

Table I. Month-wise Open and Close

Date	Open	Close
January	41.321429	41.338029
February	41.267592	41.318900
March	41.488853	41.556697
April	42.379856	42.375337
May	40.880569	40.915782
June	41.186056	41.110892
July	39.746368	39.761981
August	37.891810	37.872353
September	38.589606	38.617192
October	39.027738	39.041493
November	39.508039	39.559314
December	40.357667	40.350524

The prediction method was employed utilizing a loss and prediction function. A bar diagram was created to compare the monthly open and close prices of the stock, as well as the monthly high and low stock values as shown in Fig. 4Graphs were also created to compare comparing anticipated and actual closing prices, and the study concludes with a representation of the total closing price and predicted prices.

B. Feature Selection

The study utilized multiple feature selection techniques to analyze stock trading input data, such as LASSO and PCA. The results of the study indicated that PCA was a more effective method for identifying linear relationships between variables in the input data. Unlike LASSO, which can only capture linear relationships between variables and the output variable, PCA is able to identify a wide range of linear relationships within the data. Additionally, PCA does not require any assumptions about the distribution of the data, making it a more versatile and robust method for feature selection. The results of the study demonstrate that using PCA can lead to better results and improved performance in stock trading models as compared to LASSO alone.

IV. DISCUSSION

In this study, a type of RNN model known as GRU was employed to predict the closing price of INTEL stock. The GRU model utilizes a three-layer network planning, which entails of an input layer, hidden layers, and an output layer. To evaluate the performance of the GRU model, two common evaluation metrics, Mean Absolute Error (MAE) and Mean Square Error (MSE) was used. The MAE measures the average difference between the predicted and actual values, while the MSE trials the average formed alteration between the predicted and actual values. The best value of MAE and MSE achieved by the model were 0.82821 and 1.12731 respectively, indicating that the model performed reasonably well in terms of accuracy. Another evaluation metric that was used in this study is the R2 score. The R2 score obtained in this study was 0.94703, which indicates that the model has a good fit for the data. It is worth noting that these values are relative to the specific dataset and problem used in this study and the results may not generalize to other datasets or problems. Additionally, it's important to mention that these values alone don't give us the complete picture of the model's performance, other evaluation techniques, and real-world performance should be considered as well. In addition to this, Fig 5 presents a comparison of the market movement prediction, comparing the last 15 days and the prediction of the next 30 days' close price. This allows for an understanding of how well the model can predict future market movements. Lastly, Fig 6, shows the entire plot of entire trained set-in combination with the predicted price values. This allows for a comprehensive analysis of the model's performance and its ability to predict prices over a longer period. The results from the figures demonstrate that the model's performance is very good, and it can accurately predict the closing prices of INTEL stock. The results also show that the price movement from our model shows the market moving in an upward trend, which indicates that the model can capture the market trends and patterns. In this study, in addition to the performance metrics used to evaluate the model's performance, the results were also validated using Mean Gamma Deviance (MGD) and Mean Possession Deviance (MPD). These metrics were chosen because they provide additional information about the potential risks and deviation of the model's predictions. MGD is used to analyse the volatility of the change in delta, which can indicate the potential risks of a particular trade. Mean Possession Deviance (MPD) is a measure of the deviation of a portfolio's asset weights from their target weights. According to the results, the MGD and MPD values obtained were 0.00054 and 0.02456 respectively. The model can forecast an alteration in delta with an elevated level of consistency and with little volatility, according to the low value of MGD. The comparatively low value of MPD also shows that the portfolio's asset weights deviate little from their target weights, demonstrating the model's high degree of accuracy in price prediction and close alignment with the target portfolio.

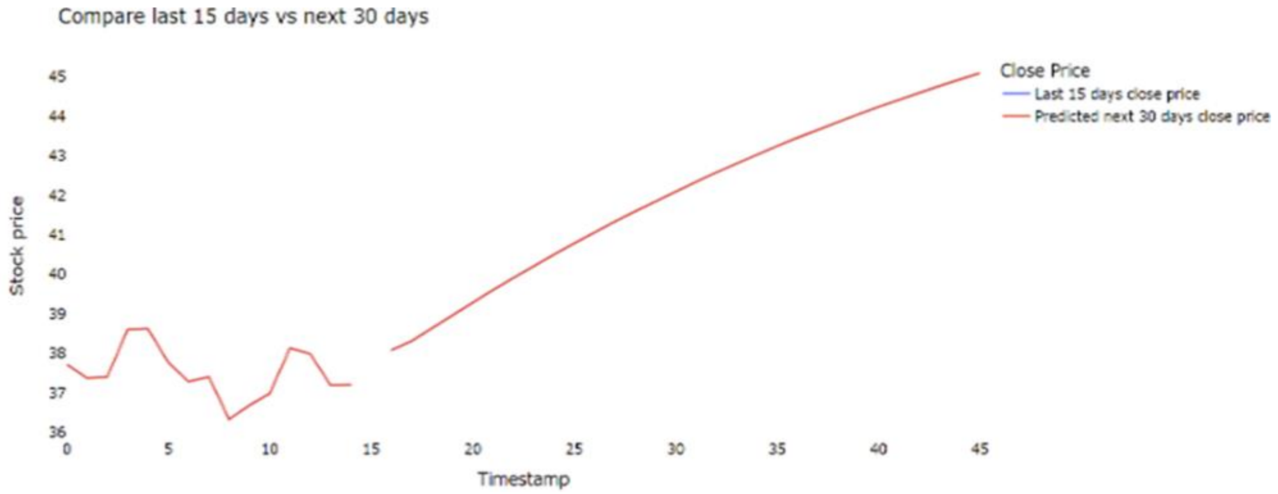

Fig 5. Comparing the Prediction with last 15 days vs 30 days

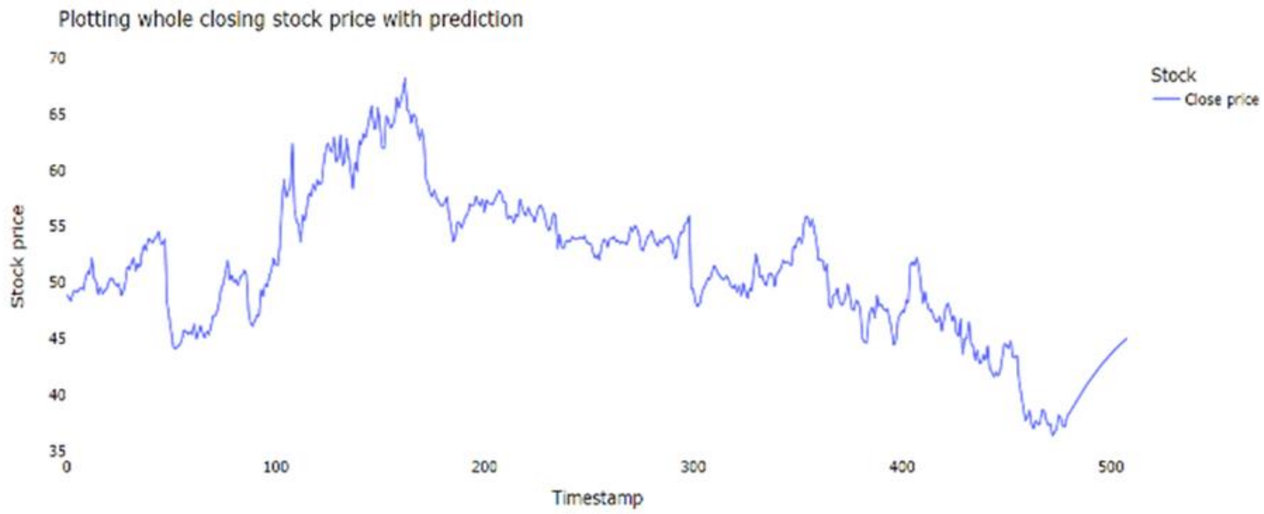

Fig 6. Plotting Whole Closing Price with Prediction

V. CONCLUSION

The conclusions drawn from this study are based on the specific dataset and research period examined. It is important to note that the results may vary for different companies and industries. The study utilized a Gated Recurrent Unit (GRU) model to visually display the future forecast using a 10-year historical dataset of the Intel Corporation. However, it should be noted that the models could potentially be adapted for short-term predictions and smaller datasets in future extensions of this research. Additionally, this study could be expanded to include the prediction of stock prices for a wider range of companies and industries across different nations. It is important to recognize that the stock market is a highly dynamic and constantly changing system, and predicting stock prices is a highly complex and challenging task. Therefore, it is crucial to also take into account the specific characteristics and traits of the stocks being analyzed. Additionally, to adapt to the ever-changing conditions of the stock market, it is essential to continually update and train the network system using the most recent data generated by the stock market.

Funding
The authors did not receive support from any organization for the submitted work.
Data availability Datasets available online https://www.kaggle.com/code/zifnab/starter-bitcoin-historical-data-d32c6d6d-5/data (Accessed on 05 July 2022).

Conflicts of interest: The authors declare no conflicts of interest. The views expressed are personal.

REFERENCES

[1] Adams, R., Kewell, B., & Parry, G. (2018) Blockchain for Good? Digital Ledger Technology and Sustainable Development Goals. In Handbook of Sustainability and Social Science Research; Filho, W.L., Marans, R., Callewaert, J., Eds.; World Sustainability Series; Springer: Cham, Switzerland.

[2] Omitaomu, O. A., & Niu, H. (2021). Artificial Intelligence Techniques in Smart Grid: A Survey. Smart Cities, 4(2), 548-568.

[3] Hassani, H., Huang, X., & Silva, E. (2018) Big-Crypto: Big Data, Blockchain and Cryptocurrency. Big Data Cogn. Comput. 2, 34.

[4] Nizzoli, L., Tardelli, S., Avvenuti, M., Cresci, S., Tesconi, M., & Ferrara, E. (2020) Charting the Landscape of Online Cryptocurrency Manipulation. IEEE Access, 8, 113230–113245.

[5] Yap, K. Y., Sarimuthu, C. R., & Lim, J. M. Y. (2020). Grid integration of solar photovoltaic system using machine learning-based virtual inertia synthetization in synchronverter. IEEE Access, 8, 49961-49976.

[6] Sharma.Gautam. K., & P. N. Kumar. (June 2020) Empirical Analysis of Current Cryptocurrencies in Different Aspects. In Proceedings of the ICRITO 2020—IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), 4–5; pp. 344–348.

[7] Killer. C., Rodrigues. B., & Stiller, B. (May 2019) Security Management and Visualization in a Blockchain-based Collaborative Defense. In Proceedings of the ICBC 2019—IEEE International Conference on Blockchain and Cryptocurrency, Seoul, Korea, 14–17; pp. 108–111.

[8] Kethineni, S., & Cao, Y. The Rise in Popularity of Cryptocurrency and Associated Criminal Activity. Int. Crim. Justice Rev. 2019, 30, 325–344.

[9] Chen, X., Huang, Y., Zhang, S., Li, W., & Zhang, S. (2020). Topological Convolutional Neural Networks for Transient Stability Assessment on Massive Historical Online Power Grid Data. In 2020 12th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC) (pp. 1-5). IEEE

[10] Liu, Y., & Tsyvinski. (2020) A. Risks and Returns of Cryptocurrency. Rev. Financ. Stud., 34, 2689–2727.

[11] Yuneline, M.H. (2019) Analysis of cryptocurrency's characteristics in four perspectives. J. Asian Bus. Econ. Stud. 26, 206–219.

[12] Hitam, N.A., & Ismail, A.R. (2018) Comparative Performance of Machine Learning Algorithms for Cryptocurrency Forecasting. Indones. J. Electr. Eng. Comput. Sci. 2018, 11, 1121–1128.

[13] Andrianto, Y. (2017) The Effect of Cryptocurrency on Investment Portfolio Effectiveness. J. Financ. Account., 5, 229

[14] Derbentsev, V., Babenko. V., Khrustalev. K., Obruch. H., & Khrustalova. S. (2021) Comparative Performance of Machine Learning Ensemble Algorithms for Forecasting Cryptocurrency Prices. Int. J. Eng. Trans. A Basics 2021, 34, 140–148.

[15] Patel, M.M., Tanwar, S., Gupta, R., Kumar. N., (2020). A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions. J. Inf. Secur. Appl., 55, 102583.

[16] Yang, H., Niu, K., Xu, D., & Xu, S., (2021). Analysis of power system transient stability characteristics with the application of massive transient stability simulation data. Energy Reports, 7, 111-117.

[17] Kalogirou, S.A. (2000). Application of artificial neural network for energy systems. Applied Energy, vol.67, no. 1-2, pp. 17–35.

[18] Chen, Z., Li. C & Sun. W., (2019) Bitcoin price prediction using machine learning: An approach to sample dimension engineering. J. Comput. Appl. Math., 365, 112395.

[19] Althelaya, K.A., El-Alfy, E.-S.M., & Mohammed. S. (April 2018). Evaluation of bidirectional LSTM for short-and long-term stock market prediction. In Proceedings of the 9th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan., 3–5.

[20] S. Kumar Nayak, A.Kumar Nayak, S. .Mishra, and P. Mohanty, "Deep Learning Approaches for Speech Command Recognition in a Low Resource KUI Language", *Int J Intell Syst Appl Eng*, vol. 11, no. 2, pp. 377–386, Feb. 2023.