[International Journal of Computer and Communication Technology](https://www.interscience.in/ijcct)

[Volume 9](https://www.interscience.in/ijcct/vol9)
[Issue 1](https://www.interscience.in/ijcct/vol9/iss1) *Research on Computing and Communication Sciences.*

Article 11

July 2023

Forecasting Stock Market Indices Using Gated Recurrent Unit (GRU) Based Ensemble Models: LSTM-GRU

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

Surabi Parida
Gandhi Institute for Education and Technology, parida.surabi220@gmail.com

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

Recommended Citation

Tripathy, Nrusingha; Parida, Surabi; and Nayak, Subrat Kumar (2023) "Forecasting Stock Market Indices Using Gated Recurrent Unit (GRU) Based Ensemble Models: LSTM-GRU," *International Journal of Computer and Communication Technology*: Vol. 9: Iss. 1, Article 11.
DOI: 10.47893/IJCCT.2023.1443
Available at: https://www.interscience.in/ijcct/vol9/iss1/11

Forecasting Stock Market Indices Using Gated Recurrent Unit (GRU) Based Ensemble Models: LSTM-GRU

Nrusingha Tripathy[1], SurabiParida[2], Subrat Kumar Nayak[3]

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

[2]Department of Computer Science & Engineering, Gandhi Institute for Education and Technology, Baniatangi, Bhubaneswar, India-752060

[1]nrusinghatripathy654@gmail.com, [2]parida.surabi220@gmail.com, [3]subratsilicon28@gmail.com

***Abstract*—**A "time sequence analysis" is a particular method for looking at a group of data points gathered over a long period of time. Instead of merely randomly or infrequently, time series analyzers gather information from data points over a predetermined length of time at scheduled times. But this kind of research requires more than just accumulating data over time. Data in time series may be analyzed to illustrate how variables change over time, which makes them different from other types of data. To put it another way, time is a crucial element since it demonstrates how the data changes over the period of the information and the outcomes. It offers a predetermined architecture of data dependencies as well as an extra data source. Time Series forecasting is a crucial field in deep learning because many forecasting issues have a temporal component. A time series is a collection of observations that are made sequentially across time. In this study, we examine distinct machine learning, deep learning and ensemble model algorithms to predict Nike stock price. We are going to use the Nike stock price data from January 2006 to January 2018 and make predictions accordingly. The outcome demonstrates that the hybrid LSTM-GRU model outperformed the other models in terms of performance.

***Keywords*—**Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Recurrent Neural Network (RNN), Performance, prediction

I. INTRODUCTION

Time series analysis to maintain consistency and reliability, a lot of data is typically required. A huge data collection ensures that your analysis can filter through contradictory data and that your sample size is representative. Furthermore, it ensures that any documented trends or patterns are valid and that they can take seasonal fluctuation into account. The act of creating statements about what's to come according to the past is known as forecasting, and time series information may be used in this process. Financial markets are seen to be one of the best options for investors looking to make a significant profit by tracking market movements [1, 2]. The gathering and analysis of information play a key role in making the best decision possible with the least amount of risk. Trading professionals may greatly benefit from the idea of forecasting trends in financial market data using modelling approaches. Modeling approaches evaluate and process time series generated by markets using complicated mathematical and statistical frameworks that are linear, nonlinear, and linear in nonlinear ways. They are able to accurately predict market values, prices, and trends in the future. Human actions, monetary developments, governmental choices, and a variety of other variables all have a significant influence on the market. According to historical market statistics, all of these influence variables are reflected in price fluctuations. Financial sector time series are challenging to model or record due to their complicated nature [3].

A time series is a collection of sequentially ordered observations that are randomly dispersed over time. Time series data, as opposed to cross-sectional data, consist of data points that are arranged chronologically and may be displayed on line graphs. The majority of data built from measurements that have been seen and are time-aligned exhibit nonlinear and complicated structure, necessitating high nonlinear modelling approaches. Figure 1 shows the overall data which is divided in to traing and testing [4-9].

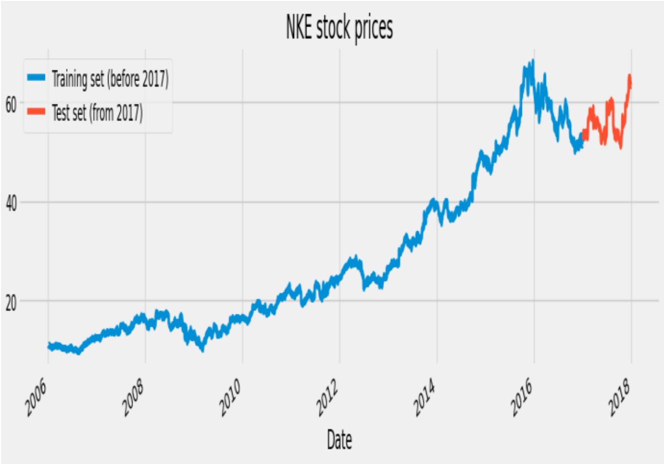

Figure 1. NIKE Training and Testing Data

II. PRELIMINARIES

A. RECURRENT NEURAL NETWORK (RNN)

The performance of several deep learning designs, including recurrent neural networks, deep neural networks, and deep belief networks, has outperformed the majority of data modelling approaches in numerous application fields [10]. Deep recurrent neural networks are used to build deep learning architectures by modifying the network weights based on both the most recent and previous input data. By cascading over the input data sequences, the network carries out the training process. Network hidden states have the ability to store information and use it for training purposes. Recurrent neural networks employ the training technique backpropagation through time (BPTT). By connecting data items at each time step to data elements from earlier time steps, it analyses data sequences in accordance with its own order [11-16]. Figure 2 shows the basic RNN model, where X_t is the input state, h_t is the current state.

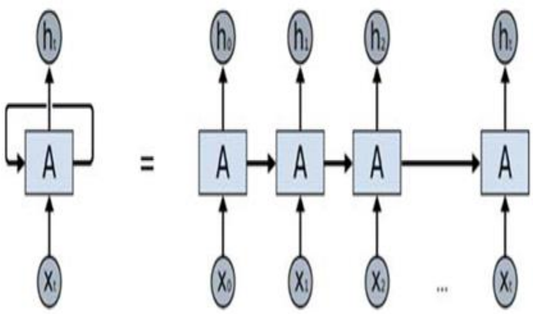

Figure 2. Basic RNN Model

B. LONG-SHORT TERM MEMORY(LSTM)

To solve the issue of absent and blowing modifications in a deep Recurring Neural Network (RNN), several variants have been developed. One of the most well-known of these is the Long Short Term Memory Network (LSTM). Theoretically, an LSTM recurrence unit seeks to "forget" irrelevant information and "remember" all of the past data that the network has seen up to this point. To do this, a number of "gates" that activate different function levels are introduced [17]. The Internally Cell State vector, which each LSTM recurring unit also maintains track of, theoretically describes the information that the previous LSTM recurrence unit chose to keep. One of the drawbacks of deep RNN is the vanishing gradient issue. To solve this problem, RNN variants like LSTM and GRU were created. Hochreiter and Schmidhuber created the LSTM. Multiple gated cells make up its design, andthey can permit data to flow through in accordance with the import of the incoming data piece. Each gated unit has a weight value that is calculated during training via a backpropagation procedure. Outputs from the LSTM cell are Ht (current hidden state) and Ct (current memory state) [18-19]. The predicted weights serve as a cutoff point for whether to save or remove data from cells. Figure 3 shows the basic LSTM model where C_{t-1} is the cell state, H_{t-1} is the hidden state and X_t is the current input. The following are the LSTM transition equations:

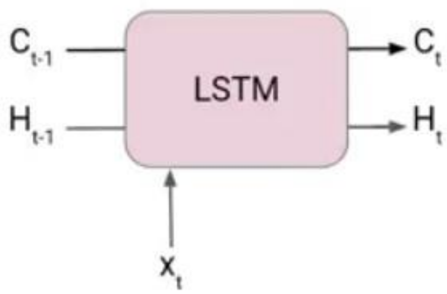

Figure 3. Basic LSTM Model

C. GATED RECURRENT UNIT(GRU)

The core idea of GRU is to use gating methods to continuously and continuously modify the network's hidden state. The gating mechanisms control how information enters and leaves the computer system. Two of the two gating methods used by the GRU are the refresh mechanism and the process of updating gate[20, 21]. The reset gate regulates the amount of the prior hidden state has to be deleted, while the update gate determines which percentage of the new input should be used to alter the hidden state. The output of the GRU is computed using the rationalized secret state. GRU is an improved version of LSTM. Blocks of gated recurrent units are used in the network design to regulate memory reset and refresh. GRU employs fewer parameters, which speeds up training, and achieves performance that is equivalent to LSTM's. The update and reset gates are the sole gates utilized in GRU. The update-gate is in charge of updating the network's existing memory, allowing the network to remember specific data input dependent on its import. The network can forget specific values at any time step since the reset gate is in charge of clearing the network's current memory. The following is the transition equation in GRU's hidden units. Figure 4 shows the basic GRU model, where $H_t$is the hidden state and $X_t$is the current input function.

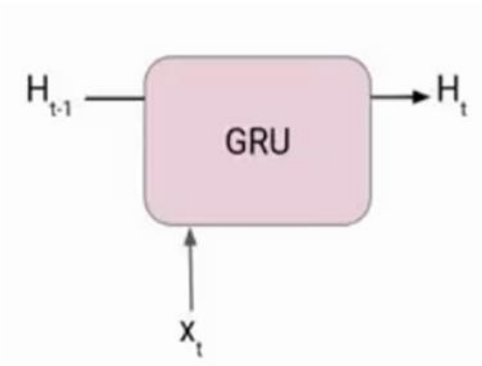

Figure 4. Basic GRU Model

D. HYBRID LSTM-GRU

Deep learning with uncertainties may now be incorporated into a wider range of business situations thanks to hybrid learning models. Better performance and model explain ability can be achieved in this way, which might lead to more broad adoption. The method for reducing model uncertainty that blends many varieties of neural networks that are deep with probabilistic methods [22, 23]. Different deep learning network types, such GANs or DRL, have demonstrated excellent agreement in regards to their efficacy and broad use with different sorts of data. However, unlike Bayesian or probabilistic techniques, algorithms for deep learning are unable to simulate uncertainty in the same manner. Hybrid learning models mix the two to take use of each type's advantages. The information that was processed is then split into training and test sets, and the GRU-LSTM Hybrid Model is subsequently used to train the model using Deep Learning methods. Model Performance will be measured as the Root Mean Squared Error (RMSE).

III. RELATED WORK

In this investigation, we came to two conclusions regarding the LSTM models. The univariate model outperforms the standard LSTM strategy when compared to the baseline and hybrid models, according to the first finding. External factors, however, have a considerable effect in the forecast of petroleum prices when the baseline model is improved. In this case, we discovered that the layers of convolution are able to capture and integrate convolutional characteristics of the petroleum price as well as its related parameters more effectively compared to the vanilla LSTM model, and this lacks the capability to combine numerous characteristics into one during processing, by adopting the hybrid model tackle and adding an CNN framework to each baseline model. During the 10-day prediction period, the multidimensional hybrid model outperformed the simple vanilla LSTM, producing substantially lower error scores with higher R2 ratings. Our second finding is that, despite the fact that the suggested univariate and multivariate models perform better than the competition, they are comparable in terms of prediction time. The results specifically reveal that the univariate suggested model is more precise than the multivariate suggested one for time frames spanning up to five days over the 10-day timeframe for the forecasts. However, the efficacy of the multivariate suggested model is significantly superior over durations of six days or more [2].

This section discusses the connection between open innovation and deep learning-based forecasting of stock prices techniques. Projecting daily returns on the stock market has not received much attention, particularly when using important machine learning techniques like deep neural networks (DNNs). Two-factor and three-factor financial evaluations are needed in these generated economic models' operations to look at the dynamics of the firm's profitability. The recommended model is based on the convergence of simulation techniques that enable analysis with random components and deterministic financial analysis techniques that are part of the DuPont model. An accurate forecast of a stock's future price might result in a sizable profit. Numerous approaches were employed in previous years to forecast stock patterns [3].

Performance is improved when many deep learning models are combined. In particular, CNN-LSTM, GRU-CNN, and combination models were proposed as RNNs to be integrated [24-28]. The suggested models were assessed using different stock market indices, look-back times, optimizers, features, and learning rates to anticipate both one-time-step and multi-time-step closing costs of stock market indices. According on the experimental findings, the suggested models that integrate RNN variants typically outperform more conventional machine learning models like RNN, LSTM, GRU, and WaveNet. For one-time-step forecasting in specific, the ensemble model generated noteworthy results. Furthermore, our models' performance increased by using the suggested unique characteristic, which is the mean of the high and low prices, when compared to that of earlier research that employed the open, high, and low prices, as well as the volume of trading of stock market indices, as features. In addition, our simulations with MV characteristics frequently produced positive outcomes. Notably, it is possible to interpret cutting back on features as avoiding overfitting [4].

In this paper we taking Nike stock price data from January 2006 to January 2018 and taking this data as input to different machine learning, deep learning and and their hybrid formulations. We compare the models with their Root Mean Square Error (RMSE) values, the RMSE value of ensemble model that is LSTM-GRU outperforms all model with a minimized value 1.744.

IV. METHODOLOGY

We must include the serial connection characteristic when modelling sequential data. Serial linkage between data points suggests that an earlier data point's effect may manifest itself at a much later time [29-31]. In modelling sequential data, neural networks have greatly improved modelling capabilities. They have the capacity to forecast throughout several time periods. Many professional language translation applications operate by neural network models, which can instantly translate across languages and comprehend the significance of a word in a given context. These neural networks are capable of handling univariate time series with ease. They have the ability to foresee not just one time in the future, but several. Basically, in this work we taking RNN, LSTM, GRU and hybrid LSTM-GRU models. The real and predicted NIKE stock price forecasting figures given below according to their predictions,

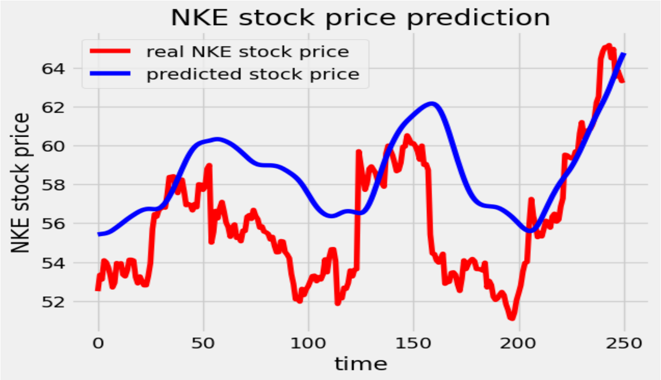

Fig 5. Recurrent Neural Network (RNN) Predicted NIKE Stock Price

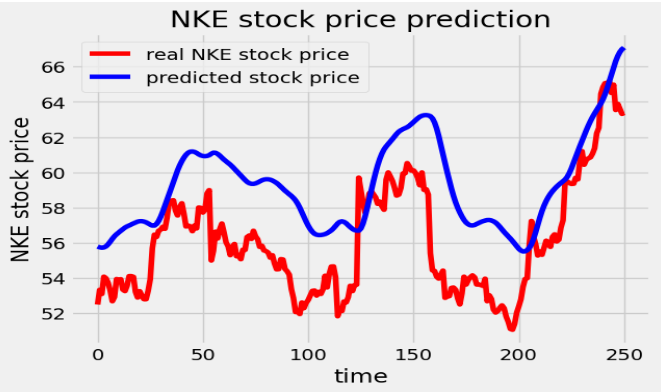

Fig6. Long Short-Term Memory (LSTM) Predicted NIKE Stock Price

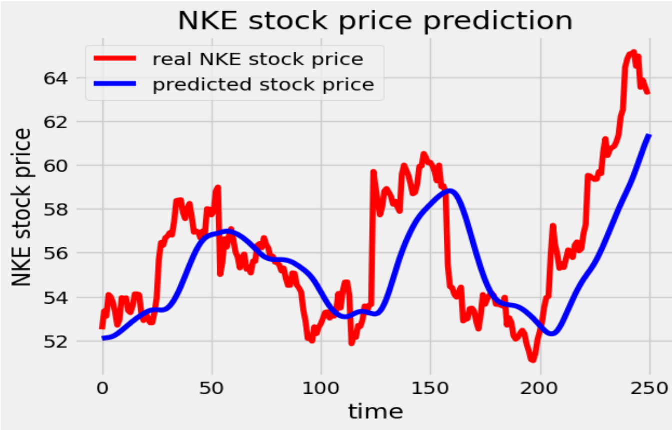

Fig 7. Gated Recurrent Unit (GRU) Predicted NIKE Stock Price

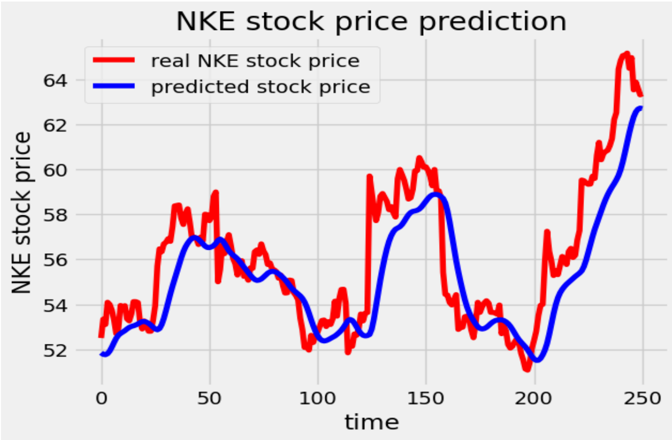

Fig 8. Hybrid LSTM-GRU Predicted NIKE Stock Price

We taking time constraint in X-axis and NIKE stock price in Y-axis. Figure 5 displays the Recurrent Neural Network (RNN) predicted graph with RMSE 3.95, similarly Figure 6 and Figure 7 shows LSTM, GRU predicted graph with RMSE 3.029, 2.014. Figure 8 shows the hybrid LSTM-GRU model predicted graph with RMSE 1.744 which is minimal as compare to other three model.Table 1.contains the model names with the Mean Square Error(MSE), Root Mean Square Error (RMSE) values.

Table 1. Error Calculation Table

Models	MSE Score	RMSE Score
RNN	15.602	3.95
LSTM	9.174	3.029
GRU	4.056	2.014
Hybrid LSTM-GRU	3.041	1.744

V. CONCLUSION

The focus of earlier research on companies' stock price prediction varied and included techniques for identifying external factors that connect with stock price, traditional machine learning techniques, and hybrid models of collectively learned algorithms. Few research, have tried to enhance both stock price prediction techniques. The hybrid LSTM-GRU model gives better result in this work. The LSTM- and GRU-based model, we suggested outperformed the current models namely, LSTM- GRU ensemble model. Our suggested model also provides improvements by taking into account other elements that are crucial for forecasting stock price. In order to clearly show that the suggested approach outperforms competing approaches and to ensure that the outside information we gathered does, in fact, the finer features required for stock price prediction, we tested the efficiency of the model using four assessment measures. We intend to integrate other variables that include text-related data such as headlines and social media-sourced phrases that may give detailed and contextual information on stock price in a more comprehensive way since incorporating external data can enhance prediction performance. We reserve further study in this area for the future.

REFERENCES

[1] Yan, K.; Wang, X.; Du, Y.; Jin, N.; Huang, H.; Zhou, H. Multi-step short-term power consumption forecasting with a hybrid deep learning strategy. Energies 2018, 11, 3089.

[2] Kim, Gun Il, and Beakcheol Jang. "Petroleum Price Prediction with CNN-LSTM and CNN-GRU Using Skip-Connection." *Mathematics* 11, no. 3 (2023): 547.

[3] Alkhatib, Khalid, HuthaifaKhazaleh, Hamzah Ali Alkhazaleh, Anas RatibAlsoud, and Laith Abualigah. "A new stock price forecasting method using active deep learning approach." Journal of Open Innovation: Technology, Market, and Complexity 8, no. 2 (2022): 96.

[4] Song, Hyunsun, and Hyunjun Choi. "Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models." *Applied Sciences* 13, no. 7 (2023): 4644.

[5] Jin, Z.; Yang, Y.; Liu, Y. Stock closing price prediction based on sentiment analysis and LSTM. Neural Comput. Appl. 2019, 32, 9713–9729.

[6] Shakya, A.; Michael, S.; Saunders, C.; Armstrong, D.; Pandey, P.; Chalise, S.; Tonkoski, R. Solar irradiance forecasting in remote microgrids using markov switching model. IEEE Trans. Sustain. Energy 2016, 8, 895–905.

[7] Rendon-Sanchez, J.F.; de Menezes, L.M. Structural combination of seasonal exponential smoothing forecasts applied to load forecasting. Eur. J. Oper. Res. 2019, 275, 916–924.

[8] Wang, K.; Qi, X.; Liu, H. Photovoltaic power forecasting-based LSTM-Convolutional Network. Energy 2019, 189, 116225.

[9] Wang, J.Q.; Du, Y.; Wang, J. LSTM based long-term energy consumption prediction with periodicity. Energy 2020, 197, 117197.

[10] Wang, L.; Wang, Z.; Qu, H.; Liu, S. Optimal forecast combination based on neural networks for time series forecasting. Appl. Soft Comput. 2018, 66, 1–17.

[11] Kim, K.; Kim, D.K.; Noh, J.; Kim, M. Stable forecasting of environmental time series via long short-term memory recurrent neural network. IEEE Access 2018, 6, 75216–75228.

[12] Heidari, A.; Khovalyg, D. Short-term energy use prediction of solar-assisted water heating system: Application case of combined attention-based LSTM and time-series decomposition. Sol. Energy 2020, 207, 626–639.

[13] Tovar, M.; Robles, M.; Rashid, F. PV Power Prediction, Using CNN-LSTM Hybrid Neural Network Model. Case of Study: Temixco-Morelos, México. Energies 2020, 13, 6512.

[14] Wu, L.; Kong, C.; Hao, X.; Chen, W. A short-term load forecasting method based on GRU-CNN hybrid neural network model. Math. Probl. Eng. 2020, 2020, 1428104.

[15] Tripathy, Nrusingha, Subrat Kumar Nayak, Julius Femi Godslove, IbangaKpereobong Friday, and Sasanka Sekhar Dalai. "Credit Card Fraud Detection Using Logistic Regression and Synthetic Minority Oversampling Technique (SMOTE) Approach." Technology 8, no. 4 (2022): 4.

[16] Nayak, Subrat Kumar, Ajit Kumar Nayak, Smitaprava Mishra, and Prithviraj Mohanty. "Deep Learning Approaches for Speech Command Recognition in a Low Resource KUI Language." International Journal of Intelligent Systems and Applications in Engineering 11, no. 2 (2023): 377-386.

[17] Trilok Nath Pandey, Nrusingha Tripathy, SarbeswarHota& Bichitrananda Patra (2023) Empirical analysis of machine learning techniques for prediction of indian exchange rate, Journal of Statistics and Management Systems, 26:1, 13-22, DOI: 10.47974/JSMS-943.

[18] Pati, Abhilash, ManoranjanParhi, Binod Kumar Pattanayak, Debabrata Singh, DebabrataSamanta, Amit Banerjee, SajalBiring, and Goutam Kumar Dalapati. "Diagnose Diabetic Mellitus Illness Based on IoT Smart Architecture." Wireless Communications and Mobile Computing 2022 (2022).

[19] Pati, Abhilash, ManoranjanParhi, and Binod Kumar Pattanayak. "IHDPM: An integrated heart disease prediction model for heart disease prediction." International Journal of Medical Engineering and Informatics 14, no. 6 (2022): 564-577.

[20] Roul, Abhinandan, Abhilash Pati, and ManoranjanParhi. "COVIHunt: An Intelligent CNN-Based COVID-19 Detection Using CXR Imaging." In Electronic Systems and Intelligent Computing: Proceedings of ESIC 2021, pp. 313-327. Singapore: Springer Nature Singapore, 2022.

[21] BjoernKrollner, Vanstone Bruce and Finnie Gavin, "Financial time series forecasting with machine learning techniques: A survey", 2010.

[22] J. G. Agrawal, V. S. Chourasia and A. K. Mittra, "State-of-the-art in stock prediction techniques", International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering, vol. 2, no. 4, pp. 1360-1366, 2013.

[23] Zachary C. Lipton, "A Critical Review of Recurrent Neural Networks for Sequence Learning", arXiv preprint arXiv:1506.00019, 2015.

[24] Rather Akhter Mohiuddin, Agarwal Arun and V. N. Sastry, "Recurrent neural network and a hybrid model for prediction of stock returns", Expert Systems with Applications, vol. 42, no. 6, pp. 3234-3241, 2015.

[25] Sepp Hochreiter and Schmidhuber Jürgen, "Long short-term memory", Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[26] Johan Bollen, Mao Huina and Zeng Xiaojun, "Twitter mood predicts the stock market", Journal of Computational Science, vol. 2, no. 1, pp. 1-8, 2011.

[27] Robert P. Schumaker and Chen Hsinchun, "Textual analysis of stock market prediction using breaking financial news: The AZFin text system", ACM Transactions on Information Systems (TOIS), vol. 27, no. 2, pp. 12, 2009.

[28] S. Lefèvre, D. Vasquez and C. Laugier, "A survey on motion prediction and risk assessment for intelligent vehicles", *ROBOMECH Journal*, vol. 1, no. 1, pp. 1, dec 2014.

[29] C. Tay, K. Mekhnacha and C. Laugier, "Probabilistic Vehicle Motion Modeling and Risk Estimation" in Handbook of Intelligent Vehicles, London:Springer, pp. 1479-1516, 2012.

[30] D. J. Phillips, T. A. Wheeler and M. J. Kochenderfer, "Generalizable Intention Prediction of Human Drivers at Intersections", 2017 IEEE Intelligent Vehicles Symposium (IV), pp. 1665-1670, 2017.

[31] Barik, Smitaranee, ManasaMoharana, Santosini Bhutia, and Nrusingha Tripathy. "Advances in data science, trends, and applications of artificial intelligence within the interaction between natural and artificial computation,2022."