International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

Robust Time Series Forecasting Using Transformer-Based Models for Volatile Market Conditions

Suresh Sankara Palli

Independent Researcher, USA.

Abstract

Accurately predicting Realised Downward Semi-Variance (RDS), a crucial metric for assessing the downside risk of prices for assets, may help traders steer their investment strategy and mitigate the effects of price decline-driven market volatility. The availability of large data resources in several domains has made it feasible to use deep learning models with exceptionally high dimensionality that may represent long-term temporal and geographical context. In the analysis of time series data, conventional techniques including Autoregressive Integrated Moving Averages (ARIMA), Long Short-Term Memory Networks (LSTM), Gated Recurrent Units (GRUs), and Recurrent Neural Networks (RNN) have offered reliable frameworks. Due to their reliance on one-dimensional macroeconomic data, traditional prediction systems are difficult to adjust to complicated market developments. This study examines the use of transformer models to energy market time-series forecasting, with a special emphasis on inter-provincial spot pricing. The suggested approach outperforms more conventional models like ARIMA and LSTM in capturing long-range relationships by using the transformer's self-attention mechanism. For stakeholders, such as utilities, legislators, and energy dealers, our study emphasises the need of increased forecasting accuracy. To highlight its resilience, the suggested strategy is tested in a range of market circumstances, including times when the market is steady and times when it is turbulent. The model's potential to optimise buying strategies and promote energy market stability is shown by the results, which show significant performance gains, especially in turbulent markets.

Keywords: - Realized Downward Semi-Variance (RDS), Autoregressive Integrated Moving Average (ARIMA), Traditional Methods, Single-Dimensional, Time Series Data, Energy Market, Long Short-Term Memory Networks (LSTM), Performance Improvements, Market Stability.

I. INTRODUCTION

High levels of volatility and intricate interdependence are characteristics of energy markets, which are influenced by supply-demand dynamics, weather variations, and geopolitical events [1, 2]. Despite their widespread usage, traditional forecasting models like ARIMA are constrained by their incapacity to identify long-term relationships and non-linear patterns in the data. Even while LSTM and other recurrent neural networks (RNNs) are better than ARIMA, they still struggle to retain computational efficiency when working with lengthy time-series data sequences [3, 4].

Understanding how the financial market works, enhancing the financial theoretical framework, assessing and controlling investment risks, and other related topics all benefit greatly from research on financial market risk [3]. Indicating the range and regularity of market price fluctuations across time, volatility captures the unpredictability of financial asset values and the risk associated with asset investments [3, 5]. Despite being a widely used risk indicator, volatility captures both rising and falling asset values.

Nonetheless, investors often focus more on downside risk, or the possible losses brought on by declining asset values [4]. Thus, research on downside risk is crucial to helping investors avoid severe losses and to help financial institutions assess risk management strategies [5, 6]. This is especially true of the prediction accuracy of the risk on the downside measurement index, which is additionally known as realised downward semi-variance (RDS) [5].

The models of multiple regression and autoregressive models are the two categories of volatility prediction models based on whether or not different market elements are taken into account [6, 8]. By reflecting the characteristics of volatility itself, the autoregressive

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

model achieves good prediction. Early researchers used traditional econometric models, such as the stochastic volatility (SV) model, the generalised autoregressive conditional heteroskedasticity (GARCH) model, and the autoregressive conditional heteroskedasticity (ARCH) models to estimate and forecast volatility under low-frequency data [8, 9]. In the residual sequence, these models are able to preserve the uncertain information while extracting the linear and predictable information [9, 10].

An essential metric for assessing the soundness and resilience to risk of financial institutions is the bank stability index. The proper functioning of national and local economies depends on the stability of the banking industry, particularly given the high level of volatility in the world's financial markets [10]. The banking sector now faces a wider range of external threats due to the rapid growth of financial markets and information technology. Accurately forecasting bank stability has emerged as a crucial concern in risk management and financial supervision [10]. Conventional prediction techniques often have issues including a single data dimension and the inability to promptly reflect market changes, since they primarily depend on the macroeconomic situation data and expert judgement [11].

Scholars and practitioners have long been interested in stock price prediction as a significant area of study in the financial industry [12]. Data on the stock market is dynamic and complicated. The majority of traditional stock market forecasting techniques depend on time series models, such as GARCH and ARIMA models; however, these techniques have some drawbacks when handling high-dimensional and nonlinear data.

Neural network-based models have steadily taken the lead in stock price prediction as deep learning technology has advanced. In particular, deep learning models like convolutional neural networks (CNN) and long short-term memory networks (LSTM) have produced positive outcomes because of their benefits in processing sequence data.

Nevertheless, classic deep learning models often fall short in leveraging the interplay between global and local variables throughout the modelling process [1, 4], which restricts their ability to interpret intricate stock market data [4, 9].

Important financial metrics that show a bank's resilience to risk and overall health make up the bank stability index [1, 8]. These consist of the liquidity coverage ratio, which evaluates short-term liquidity resilience; the non-performing loan ratio, which quantifies troublesome loans; and the capital adequacy ratio, which gauges a bank's available capital against risks [11]. We used the Transformer model's attention weights to assess these attributes' significance even further. During times of financial instability, for example, the attention heatmaps show a preponderance of emphasis on the liquidity coverage ratio [4, 9], but long-term trends prefer the capital adequacy ratio as the main indicator [1, 12].

This sophisticated comprehension highlights the model's capacity to dynamically modify its emphasis in response to market circumstances. These metrics show banks' risk appetite and overall financial health from a variety of angles [11]. Financial markets have seen sharp volatility as a result of the escalation of global economic instability, particularly the effects of the COVID-19 pandemic and geopolitical disputes, and the strain on the banking system has grown considerably [12]. Conventional statistical model-based prediction techniques have progressively become more challenging to use. The stationary nature of historical data is often assumed by statistical models [14].

The financial system may, however, experience abrupt or nonlinear shifts under very volatile market circumstances, and it is challenging for conventional approaches to adequately represent these intricate dynamics. In light of this, the advent of deep learning models, particularly the Transformer model, [13], opens up new avenues for bank stability index prediction.

Analysing time series data entails looking at datasets made up of time-ordered elements. In many domains, this research is essential for understanding previous behaviours, forecasting future patterns, and making well-informed choices [12]. An essential component of statistical research and data science, time series data analysis is used in a wide range of industries, from engineering and climate science to healthcare and finance. Understanding, modelling, and forecasting temporal data is the fundamental concept of time series analysis. Time series data is essentially sequential as the values are captured at consecutive moments in time, often at evenly spaced intervals [16].

Even though the traditional econometric model can explain the characteristics of volatility aggregation and has the advantages of a strong an explanation, stable parameters, and a clear statistical principle, it has strict requirements on assumptions like data distribution that make it difficult to fully capture the nonlinear relationship in a time series. Additionally, the model has a high processing cost and limited prediction accuracy [12]. As a result, machine learning models are often employed to forecast RV in financial markets with non-stationary, nonlinear, aggregated, and leaping characteristics because they are adept at extracting high-dimensional data elements and capturing complicated nonlinear correlations [13].

Crude oil and natural gas futures contracts are examples of energy commodities whose volatility may be predicted using support vector regression (SVR). The ensemble machine learning model, which combines many weak learners, produces a more robust prediction model than the single machine learning model, which has poor generalisation capacity and outlier sensitivity [14, 15].

Crude oil and stock index RVs were predicted using RF models, respectively. But there is still a lot of feature engineering to be done on the ensemble learning model, and it has to be better able to handle highly dimensional complex data with structures [16]. When it comes to automated feature learning and generalisation, the deep learning model outperforms the ensemble learning model. It uses multi-layer nonlinear transformation to mimic complicated nonlinear features indefinitely. To forecast the volatility of financial stock markets, traditional deep learning algorithms like the gated recurrent unit (GRU) and long short-term memory (LSTM) have been utilised extensively [18].

Although they have been widely used in financial forecasting, traditional statistical models [17], like generalised auto regressive conditional heteroskedasticity (GARCH) and autoregressive integrated moving average (ARIMA), [11], frequently fail to capture intricate dependencies and long-range correlations in financial data [19].

These models' capacity to adjust to abrupt market swings and high-volatility situations is constrained by their assumptions of linearity and stationary behaviour [10]. More complex models that can identify non-linear correlations and sequential dependencies in financial data have been made possible by recent developments in deep learning.

By using memory cells to preserve long-term dependencies, recurrent neural networks (RNNs) and long short-term memory (LSTM) networking have shown significant gains in financial time series forecasting [10]. Nevertheless, these designs are ineffective in managing long-range relationships in high-frequency financial data due to gradient vanishing & exploding issues. Since of its self-attention mechanism, transformer models—which were first developed for natural language processing—have become a potent substitute for time series forecasting since they can interpret quantitative data without being constrained by recurrent architectures.

Transformers have shown significant promise in modelling intricate financial patterns and enhancing forecast accuracy by more accurately capturing long-range relationships. Even with transformer-based forecast models' performance, there is still a significant drawback: the forecasting method does not explicitly include risk [10]. Conventional forecasting models ignore the wider ramifications of financial risk in favour of concentrating only on reducing prediction errors. Financial choices in real-world applications [18] need knowledge of the risks involved in addition to point forecasts of future pricing [10].

To make wise investment choices, market players must evaluate the probability of sharp price swings, unexpected volatility increases, and economic downturns [18]. This calls for a forecasting model that, in addition to predicting price changes, includes dynamic risk indicators to measure possible losses and uncertainty [19].

By addressing the limitations of earlier models, this research focusses on the accuracy of stock market predictions. A variety of often chaotic factors influence stock prices. It is difficult to predict them using conventional methods because of these variables [18].The nonlinear behaviour, multi-noise, and volatility that are inherent in financial markets are difficult for current studies and models to capture. Numerous models that concentrate on local and niche stock markets have already been created [10]. Nevertheless, no single model that works well in both the divergent stock and market for cryptocurrency has been found in the literature [12]. A more robust single model that can handle this complexity and provide accurate forecasts in a variety of financial markets is desperately needed, which is what motivated our study [18]. This work attempts to improve prediction accuracy for time series data by using recent advancements in deep learning, particularly in hybrid models that include several architectures [19].

While some studies utilise categorisation to anticipate trends, others concentrate on utilising regression to forecast stock prices. However, future stock

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

developments are of primary importance to investors and organisations [20]. Stock prices cannot be predicted using historical data, according to random walk theory, and only 50% of academics can predict stock values properly. This theory holds that news and policies have the most effects on stock prices [21].Other scholars, however, contend that experimental findings lend credence to the notion that historical data might provide insightful insight. These endeavours are dangerous because of the unpredictability of equity developments. Furthermore, governments often find it helpful to ascertain the state of the market [22, 23].

This is mainly due to the fact that stock values are usually nonparametric, nonlinear, and dynamic. These traits may thus result in statistical models performing poorly, making it difficult to make precise predictions about values and movements [11]. RNNs are perfect for economic forecasting because they can create connection between individual network units and utilise networks to preserve recent memory events. The LSTM approach is recognised as an improved RNN technique [14]. LSTM has three separate gates in order to accurately remove the drawbacks of RNNs [19]. Both single data points and whole data sequences may be handled by it.

The Texas electric grid crisis of 2021 brought to light the shortcomings of the current forecasting methods. Energy costs increased from an average of $30 per MW to approximately $9,000 per MWh during an extraordinary winter storm [19]. The inability of ARIMA and comparable models to handle this kind of severe volatility highlights the need for more sophisticated models that can handle abrupt changes in market dynamics. In an effort to overcome these obstacles, this research suggests including transformers into energy price predictions [20]. Transformer models, which were first developed for natural language processing problems, have drawn interest recently due to their capacity to use self-attention processes to represent long-range relationships. This makes the model particularly well-suited for time-series forecasting as it enables it to dynamically balance the significance of various time steps.

Energy market time-series forecasting has historically depended on linear models such as ARIMA. Because these models rely on stationarity, the underlying data must behave consistently across time. However, because of market fluctuations and external shocks, energy prices are infamously non-stationary [25]. LSTM networks enhanced this by using a gated technique to capture short- and medium-term dependencies, enabling information to endure across time. However, when processing lengthy sequences, LSTMs have trouble becoming computationally efficient. We provide a comparison chart between the transformer, LSTM, and ARIMA models to show how they vary in terms of complexity, managing non-stationary data, and capturing long-term relationships [23]. Recent studies show that by using the self-attention mechanism, transformer-based designs perform better in time-series forecasting than ARIMA and LSTM [24].

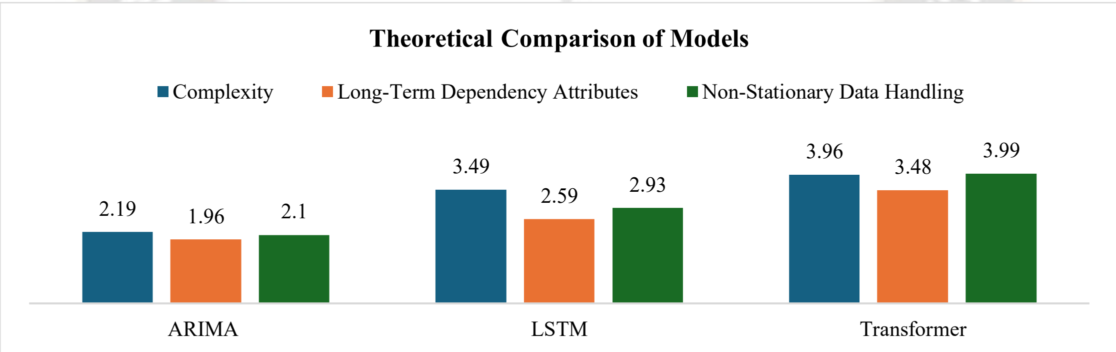

Fig. 1 Theoretical Comparison of Models. [20]

II. METHODOLOGY

2.1 Modeling Techniques

Three models were used for forecasting:

1) **ARIMA:** conventional components that use stationary time-series data, such as moving averages and autoregressive [15]. Because of its widespread use in time-series forecasting, it is used for baseline comparisons.
2) **LSTM:** a kind of RNN that is restricted in its ability to handle dependencies that are long-term but is intended for sequence prediction [18]. Designed to use a 30-day sequence length for input in order to capture temporal relationships [9]. Hyperparameters

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

were adjusted, such as dropout rates & the quantity of hidden units.

3) **Transformer:** The transformation algorithm is quite versatile in capturing long-range relationships in time-series data because it dynamically balances various time steps using a self-attention technique. Used positional encoding and self-attention algorithms to dynamically prioritise important time steps [19]. Feed-forward neural network structures and multi-head attention layers were added to the design [20].

2.2 Data Collection and Preprocessing

Demand projections, meteorological information, historical interprovincial spot prices, and indications of geopolitical events are all included in the dataset [20]. These characteristics were picked in order to capture both internal and external market factors. The following procedures were part of data preprocessing:

(1) **Stationarity Transformation for ARIMA:** The time series was made stationary by using differencing and log transformation [23].
(2) **Normalization for LSTM and Transformer:** The data was normalised into a range of [0,1] using min-max scaled.
(3) **Feature Engineering:** To improve model inputs, external factors such as precipitation, temperatures, and industrial production indices were stored as extra features. [24].

2.3 Model Training

Seventy percent of the dataset was used for training, fifteen percent for validation, and fifteen percent for testing. Because ARIMA is sequential, it was trained repeatedly [23], whereas Transformer and LSTM used parallelised training techniques. An inter-provincial spot price prediction is produced by each of the three models. Single-step forecasts are produced using ARIMA [14]. Both single-step & multi-step predictions are produced by LSTM and Transformer, with the transformer exhibiting superior long-horizon accuracy [20].

III. EXPERIMENTAL DESIGN AND RESULTS

To learn more about the model's performance in severe situations, including sudden price increases brought on by weather abnormalities, further research was done. Transformers have continuously shown precision and durability, indicating their ability to withstand demanding market circumstances [23]. We highlight the alignment of projected values with real prices in our comprehensive graphics that show how well the model performed across various test situations [24].

3.1 Evaluation Metrics

Three primary measures were used to assess the models: Mean Absolute Error (MAE), mean absolute percentage error (MAPE) [18], and Root Mean Squared Error (RMSE) [14]. Each model's success across these measures is shown in the bar charts below:

(1) **ARIMA:** Performed dependably while volatility was low, but since it was unable to capture non-linear relationships, it showed large error rates when volatility was high. [20].
(2) **LSTM:** Performed better than ARIMA in times of moderate volatility but was unable to sustain accuracy over longer sequences, exhibiting higher RMSE values at times of high volatility.
(3) **Transformer:** Shown its resilience in managing both stable and erratic markets by consistently accomplishing the lowest possible MAE, RMSE, [14], and MAPE across all situations.

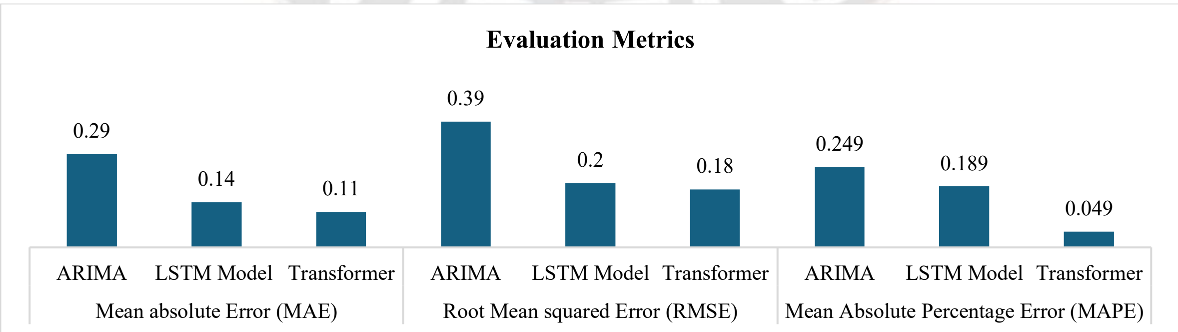

Fig. 2 Evaluation Metrics. [17]

Actual vs. Predicted Prices Diagram During a time of extreme volatility, the line chart contrasts expected and actual prices [23]. The transformer model's ability to handle abrupt changes in the market is shown by the fact

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11s
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

that it follows prices much more closely than ARIMA and LSTM.

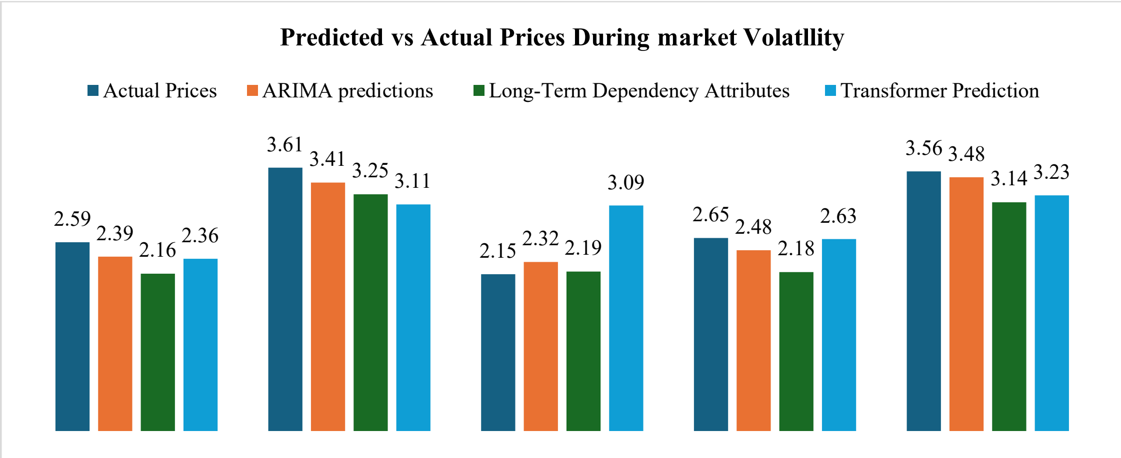

Fig. 3 Predicted vs Actual prices. [17]

IV. ANALYSIS AND DISCUSSION

The findings show that compared to standard models, transformer models provide a significant increase in accuracy. The model's capacity to anticipate in turbulent markets is enhanced by the self-attention mechanism, which specifically enables the model to dynamically allocate significance to various time steps [14].

Transformers have been demonstrated to perform better than more conventional models such as ARIMA in forecasting financial markets because they are able to capture non-linear pattern in stock prices and long-term interdependence. Given that markets are very volatile and impacted by a wide range of external variables, this accomplishment implies that comparable performance may be anticipated in energy price predictions [17]. Incorporating transformers into models for forecasting might assist energy market practitioners minimise the financial effect of unexpected price spikes and optimise their buying strategies [19]. These models may also be used by policymakers to enhance market regulation and guarantee energy security in emergency situations [11].

V. CONCLUSION

This study shows how transformer models may revolutionise time-series forecasting in the electrical energy market, especially for inter-provincial spot pricing. The suggested model outperforms more conventional techniques like ARIMA and LSTM by using self-attention mechanisms to capture intricate relationships and dynamically adjust to market situations. The experimental findings demonstrate transformer models' greater accuracy and resilience, even during times of extreme volatility.

In order to help utilities, dealers, and regulators make well-informed choices, maximise buying strategies, and reduce the financial risks associated with unforeseen price swings, these results highlight their practical relevance for energy market players. Although transformer models have shown promise in predicting energy costs, further studies should concentrate on improving the design to take into account outside variables like weather trends, world energy prices, and geopolitical events. Furthermore, it may be possible to investigate the use of transformers in other time-sensitive fields including anticipating power consumption and producing renewable energy.

VI. REFERENCES

[1] Andersen, T.G.; Bollerslev, T.; Diebold, F.X.; Labys, P. Modeling and forecasting realized volatility. Econometrica 2003, 71, 579–625.

[2] Corsi, F. A simple approximate long-memory model of realized volatility. J. Financ. Econom. 2009, 7, 174–196.

[3] Izzeldin, M.; Hassan, M.K.; Pappas, V.; Tsionas, M. Forecasting realised volatility using ARFIMA and HAR models. Quant. Financ. 2019, 19, 1627–1638.

[4] Liang, C.; Wei, Y.; Lei, L.; Ma, F. Global equity market volatility forecasting: New evidence. Int. J. Financ. Econ. 2022, 27, 594–609.

[5] Korkusuz, B.; Kambouroudis, D.; McMillan, D.G. Do extreme range estimators improve realized volatility forecasts? Evidence from G7 Stock Markets. Financ. Res. Lett. 2023, 55, 103992.

[6] Fałdziński, M.; Fiszeder, P.; Orzeszko, W. Forecasting volatility of energy commodities: Comparison of garch models with support vector regression. Energies 2021, 14, 6.
[7] Luong, C.; Dokuchaev, N. Forecasting of Realised Volatility with the Random Forests Algorithm. J. Risk Financ. Manag. 2018, 11, 61.
[8] Demirer, R.; Gkillas, K.; Gupta, R.; Pierdzioch, C. Risk aversion and the predictability of crude oil market volatility: A forecasting experiment with random forests. J. Oper. Res. Soc. 2022, 73, 1755–1767.
[9] S. Lu, Z. Liu, T. Liu and W. Zhou, "Scaling-up Medical Vision-andLanguage Representation Learning with Federated Learning," Engineering Applications of Artificial Intelligence, vol. 126, Article ID 107037, 2023.
[10] J. Cao, R. Xu, X. Lin, F. Qin, Y. Peng and Y. Shao, "Adaptive Receptive Field U-Shaped Temporal Convolutional Network for Vulgar Action Segmentation," Neural Computing and Applications, vol. 35, no. 13, pp. 9593-9606, 2023.
[11]]Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural computation, 31(7), 1235-1270.
[12] C. Saiktishna, N. S. V. Sumanth, M. M. S. Rao, and T. J, “Historical Analysis and Time Series Forecasting of Stock Market using FB Prophet,” 2022 6th Int. Conf. Intell. Comput. Control Syst. ICICCS, pp. 1846–1851, 2022,
[13] Q. Huang, “Forecasting Stock Prices Using Multi-Macroeconomic Regressors Based on the Facebook Prophet Model,” BCP Bus. Manag., 2022.
[14] S. Mahmud, “Bangladesh COVID-19 Daily Cases Time Series Analysis using Facebook Prophet Model,” PSN Dis. Illn. Top., 2020,
[15] O. Mphale, N. Raffing, S. Sheikh, and L. Balasubramanian, “Time Series Forecasting of COVID-19 Mortality in SADC region with Facebook Prophet Model,” 2022 Int. Conf. Smart Appl. Commun. Netw. SmartNets, pp. 1–6, 2022.
[16] Hossain M A, Karim R, Thulasiram R, et al. Hybrid deep learning model for stock price prediction. In 2018IEEESymposium Series on Computational Intelligence (SSCI), IEEE, 2018, 1837-1844.
[17] Vargas M R, De Lima B S, Evsukoff A G. Deep learning for stock market prediction from financial news articles. In 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), IEEE, 2017, 60-65.
[18] Agrawal M, Shukla P K, Nair R, et al. Stock prediction based on technical indicators using deep learningmodel. Computers, Materials & Continua, 2022, 70(1).
[19] Rezaei H, Faaljou H, Mansourfar G. Stock price prediction using deep learning and frequency decomposition. Expert Systems with Applications, 2021, 169: 114332.
[20] Kim, H.Y.; Won, C.H. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Syst. Appl. 2018, 103, 25–37.
[21] Chen, W.; Jiang, M.; Zhang, W.-G.; Chen, Z. A novel graph convolutional feature based convolutional neural network for stock trend prediction. Inf. Sci. 2021, 556, 67–94.
[22] Chen, Q.; Zhang, W.; Lou, Y. Forecasting Stock Prices Using a Hybrid Deep Learning Model Integrating Attention Mechanism, Multi-Layer Perceptron, and Bidirectional Long-Short Term Memory Neural Network. IEEE Access 2020, 8, 117365–117376.
[23] Naeini, M.P.; Taremian, H.; Hashemi, H.B. Stock market value prediction using neural networks. In Proceedings of the 2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM), Krackow, Poland, 8–10 October 2010; IEEE: New York City, NY, USA, 2010; pp. 132–136.
[24] Qian, B.; Rasheed, K. Stock market prediction with multiple classifiers. Appl. Intell. 2007, 26, 25–33.
[25] Guo, T.; Xu, Z.; Yao, X.; Chen, H.; Aberer, K.; Funaya, K. Robust Online Time Series Prediction with Recurrent Neural Networks. In Proceedings of the 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, Canada, 17–19 October 2016; IEEE: New York City, NY, USA, 2016; pp. 816–825.