Indonesian Journal of Electrical Engineering and Computer Science
Vol. 35, No. 1, July 2024, pp. 263~273
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v35.i1.pp263-273

Unveiling deep learning powers: LSTM, BiLSTM, GRU, BiGRU, RNN comparison

Zakir Mujeeb Shaikh, Suguna Ramadass
Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

Article Info

Article history:

Received Nov 19, 2023
Revised Feb 16, 2024
Accepted Mar 16, 2024

Keywords:

Bidirectional GRU
Bidirectional LSTM
Deep learning algorithms
Gated recurrent unit
Long short-term memory
Recurrent neural network

ABSTRACT

Deep learning algorithms have revolutionized various fields by achieving remarkable results in time series analysis. Among the different architectures, recurrent neural networks (RNNs) have played a significant role in sequential data processing. This study presents a comprehensive comparison of prominent RNN variants: long short-term memory (LSTM), Bidirectional LSTM (BiLSTM), gated recurrent unit (GRU), bidirectional GRU (BiGRU), and RNN, to analyze their respective strengths and weaknesses of national stock exchange India (NSEI). The Python application developed for this research aims to evaluate and determine the most effective algorithm among the variants. To conduct the evaluation, data from the public domain covering the period from 1/1/2004 to 30/06/2023 is collected. The dataset considers significant events such as demonetization, market crashes, the COVID-19 pandemic, downturns in the automobile sector, and rises in unemployment. Stocks from various sectors including banking, automobile, oil and gas, metal, and Pharma are selected for analysis. Finally, the results reveal that algorithm performance varies across different stocks. Specifically, in certain cases, BiLSTM outperforms, while in others, both BiGRU and LSTM are surpassed. Notably, the overall performance of simple RNN is consistently the lowest across all stocks.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Zakir Mujeeb Shaikh
Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai-600062, India
Email: shaikhzakir03@yahoo.com

1. INTRODUCTION

The rapid evolution of deep learning algorithms has revolutionized numerous domains, notably transforming the landscape of time series analysis. In this context, recurrent neural networks (RNNs) stand as pivotal architectures, showcasing their prowess in sequential data processing. This research endeavors to delve into their application within the dynamic realm of the national stock exchange India (NSEI). Focusing on a spectrum of RNN variants-long short-term memory (LSTM), bidirectional LSTM (BiLSTM), gated recurrent unit (GRU), bidirectional GRU (BiGRU), and the fundamental RNN. This study embarks on an expansive comparative analysis. The crux of this investigation lies in elucidating the distinct strengths and weaknesses inherent in each RNN variant. Through meticulous evaluation and benchmarking against NSEI data, the primary ambition is to discern the most adept algorithm tailored for sequential data processing in this specific financial domain. Integral to this pursuit is a purpose-built Python application, meticulously designed to serve as an analytical bedrock, facilitating an exhaustive assessment and comparison among these prominent RNN models. Ultimately, the overarching aim is to uncover and endorse the most proficient RNN algorithm, poised to optimize time series analysis within the intricate framework of the NSEI.

The efficient market hypothesis (EMH) in finance asserts that financial markets swiftly and effectively integrate all pertinent information concerning the value of investments, thereby implying market efficiency. According to the EMH, consistently obtaining above-average returns through the analysis of past price data or other accessible information is unfeasible, given that the market has already factored in all known information into asset prices. This hypothesis further posits that market prices follow a random walk. The EMH delineates three forms of market efficiency: weak, semi-strong, and strong [1]. In the semi-strong form of efficiency, neither fundamental nor technical analysis can reliably predict future movements [2]. The strong form of efficiency does not allow for profits above those of the average investor, even if they have access to new insider information [3]. The LSTM model, utilizing sentiment analysis derived from Twitter, yields promising results. However, it also suggests that an increased volume of sentiment from the population can lead to additional noise in the results [4]. LSTM demonstrates promising accuracy in predicting oil company stock prices, even amidst system volatility [5]. LSTM outperforms RNN when evaluating the Dow Jones and Shanghai composite index [6]. Furthermore, the effectiveness of the BiLSTM-GRU forecast pattern is validated, demonstrating accurate predictions with a maximum training error of 0.029, thereby establishing it as a feasible and effective approach for stock price forecasting [7]. The utilization of attention-based BiLSTM (Att. BiLSTM) in devising trading strategies has substantiated the effectiveness of various technical indicators (TIs) like stochastic oscillator, relative strength index (RSI), BIAS, W%R, and moving average convergence/divergence (MACD). Moreover, the proposal of two trading strategies that integrate deep neural networks (DNN) with TIs has showcased notable effectiveness [8]. DNN-reinforcement learning (RL), LSTM-RL, and BiLSTM-RL were compared to test the buy-and-hold strategy on Dow Jones, S&P500, and NASDAQ. During this experiment, BiLSTM-RL outperformed the others [9]. Convolutional neural network (CNN)-GRU and hybrid models were employed to compare stock price predictions. Comprehensive experiments and comparative analysis revealed that the CNN-GRU model demonstrates superior performance [10]. The CNN-GRU attention model used for stock price forecasting notably yielded MAPE at 11.23, RMSE at 5.17, and an R2 score of 0.41% [11]. A hybrid LSTM-GRU network, integrating 25 features encompassing technical indicators, is utilized to predict the adjusted closing price of the standard and poor 500 index accurately. Comparative analysis using performance indicators (return ratio, R2, MSE, optimism, and pessimism ratios) showcases the superiority of the proposed model over standalone LSTM, GRU, and multiple layer perceptron (MLP) models in forecasting stock market prices [12]. LSTM and GRU were employed for stock market prediction using LASSO, and the results were compared with PCA. During this experiment, LSTM outperformed the others in terms of accuracy [13]. The LSTM and GRU models, overcoming the vanishing gradient problem of traditional RNNs, are investigated as prediction models for stock prices. Extensive analysis and comparison highlight these models as promising solutions for traders, offering accurate price predictions to facilitate efficient decision-making [14]. The performance of BiGRU and BiLSTM is compared in terms of stock prediction with a longer window, and the BiLSTM outperforms the BiGRU [15]. Bidirectional encoder representations from transformers (BERT) with generative adversarial networks were utilized to predict stock prices based on non-structural news. The results were compared with ensemble methods including random forests, extreme gradient boosting, and light gradient boosting machines. The experiment found that BERT outperformed the others [16]. The performance of BiGRU significantly increased in stock price prediction when combined with a genetic algorithm [17]. The comparison between the improved complete ensemble EMD (ICEEMDAN)-BiGRU and multi-objective optimization (MOWOA)-BiGRU revealed that MOWOA-BiGRU outperforms ICEEDAN-BiGRU, showing an improvement of 14.4% in results [18]. Three deep learning algorithms, namely vanilla RNN, LSTM, and GRU, were compared for stock prediction on the Nepal stock exchange (NEPSE). The results revealed that LSTM and GRU outperformed RNN in terms of performance [19]. The RNN model trained on Apple’s stock was utilized for predicting future prices, achieving an accuracy of 95% with a loss of 0.1% [20]. ARIMA, RNN, and ARIMA-RNN models were employed to predict stock price movement. The results revealed that ARIMA-RNN outperformed the others [21]. LSTM demonstrates proficiency in predicting future stock market values when provided with historical data, indicating its competence in forecasting based on past information [22]. LSTM models were applied in two phases: one with technical indicators and another without. The results showed significant enhancement by using technical indicators, resulting in a remarkable 2.42 reduction in error [23].

Existing studies have focused on specific datasets or time periods, lacking a broader understanding of their strengths and weaknesses. Further research is necessary to analyze these algorithms using diverse evaluation metrics, datasets, and market conditions. Exploring the generalizability of deep learning algorithms across different market conditions remains a research gap, necessitating investigation. Enhancing the interpretability and explainability of these algorithms is crucial for their practical application.

2. METHOD

The research methodology employed in this study involves a systematic comparison of RNN architectures RNN, LSTM, BiLSTM, GRU, and BiGRU aimed at a specific task. To conduct this comparison, the top five companies listed in NSEI, representing distinct sectors including finance service, automobile and auto component, oil gas and consumer fuels, metal and mining, and healthcare, were selected within the timeframe of 01/01/2001 to 03/11/2023. The initial dataset underwent preprocessing using standard techniques and was subsequently divided into distinct partitions for training, validation, and testing purposes. The implementation phase leveraged the Tensorflow framework executed on hardware comprising an Intel Core i7 12^{th} Generation CPU, Intel-7 16 GB graphics card, and 32 GB RAM with NVME 2 configurations. Specific indicators such as the bollinger band (BB), MACD, RSI, and money flow index (MFI) were chosen from various indicator categories. Each model's architecture was meticulously outlined, incorporating common regularization techniques and convergence criteria to ensure robustness and accuracy. The evaluation process encompassed diverse look-back values and metrics like mean square error (MSE), RMSE, mean absolute error (MAE), mean absolute percentage error (MAPE), and R2, where lower metric values typically signify higher precision in forecasting, thus reflecting the comprehensive accuracy of the analysis.

Figure 1 illustrated the system architecture of the research framework. The system architecture entails the utilization of a stock market dataset sourced from reputable financial repositories. Data preprocessing involves cleaning, normalization, and feature engineering, incorporating technical indicators essential for market analysis. The prepared dataset undergoes segmentation into training, validation, and test sets. Employing diverse RNN architectures RNN, LSTM, BiLSTM, GRU, and BiGRU each model undergoes hypertuning parameter optimization through cross-validation techniques. The final system presents a detailed comparison of these models' performance, offering insights into their effectiveness in analyzing stock market behavior.

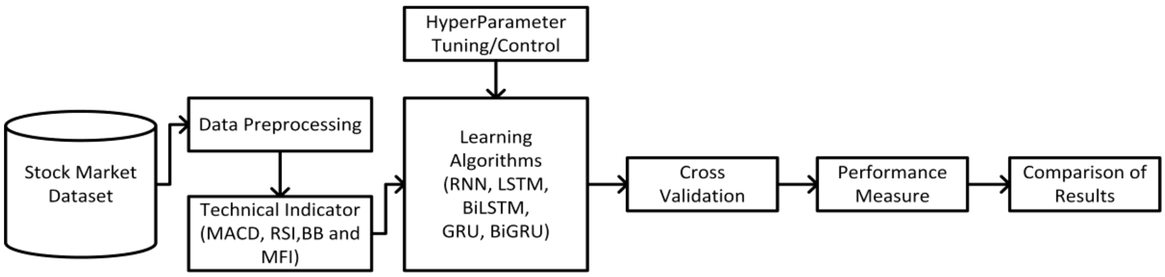

Figure 1. System architecture

2.1. LSTM

An LSTM, known as a long short-term memory network, stands out among RNNs for its specialized capability in recognizing and learning patterns within sequential data while handling prolonged dependencies. It consists of memory cells, gates, and connections that enable the retention and controlled access to data across varying time spans. Diverging from conventional RNNs, LSTMs address the challenge of vanishing or exploding gradients through a more intricate architecture, empowering them to effectively capture and retain crucial information over extensive sequences. Through the management of information flow via distinct gates such as input, forget, and output gates, LSTMs possess the ability to selectively preserve or discard data, showcasing their proficiency in modeling and predicting sequences across diverse fields like natural language processing, time series analysis, and speech recognition. The mathematical expression are shown in (1) to (6) [24], [25].

$$Forgot\ gate: f_t = \sigma(W_f * [h_{t-1}, x_f] + b_f) \quad (1)$$

$$Input\ gate: i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (2)$$

$$Cell\ state\ update: \widetilde{c_t} = tanh(W_c * [h_{t-1}, x_t] + b_c) \quad (3)$$

$$Cell\ state\ update\ through\ input\ gate: c_t = f_i * c_{t-1} + i_t * \widetilde{c_t} \quad (4)$$

$$Output\ gate: o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (5)$$

$$Hidden\ state\ update: h_t = o_t * tanh\ (c_t) \tag{6}$$

Where, f=forgot gate, t=time step, $\sigma =$ sigmoid function, $W =$ weight, b=bias, h=hidden state, i=input state, $\tilde{c}$ =cell state, c = current cell state, $o =$ output gate.

2.2. BiLSTM

A BiLSTM, short for bidirectional long short-term memory network, belongs to the family of RNNs and enhances the standard LSTM by simultaneously handling input sequences in both forward and backward directions. Comprising memory cells, gates, and connections, the BiLSTM includes two LSTM layers. One layer processes the input sequence forward, while the other processes it in reverse. This unique bidirectional approach enables the network to comprehend patterns and relationships from past and future contexts, resulting in a more comprehensive grasp of sequential data. By combining information from both directions, the BiLSTM effectively overcomes the limitations of one-way LSTMs, making it particularly suitable for tasks involving natural language processing, speech recognition, and various sequence modeling challenges. The BiLSTM expression show in (7) [26].

$$\mathrm{h_t} = [\overrightarrow{\mathrm{h_t}} : \overleftarrow{\mathrm{h_t}}] \tag{7}$$

Where, h=hidden state, t=time step, $\overrightarrow{\mathrm{h_t}}$ and $\overleftarrow{\mathrm{h_t}}$=forward and backward state.

2.3. GRU

The GRU stands as a specialized version within RNNs, specifically tailored for handling sequential data by understanding and storing complex patterns. With its components consisting of gates and memory units, the GRU architecture empowers the network to grasp connections over different time points, offering enhanced computational efficiency when compared to the conventional LSTM networks. Differing from LSTMs, GRUs combine the forget and input gates into a single "update gate," enabling precise control over information flow to efficiently manage crucial data within sequences. This simplified design of GRUs expedites training processes and renders them particularly suitable for tasks involving natural language processing, machine translation, and time series analysis. The GRU is expressed using in (8) to (11) [27].

$$Update\ gate: z_t = \sigma(W_z * [h_{t-1}, x_t]) \tag{8}$$

$$Reset\ gate: r_t = \sigma(W_f * [h_{t-1}, x_t]) \tag{9}$$

$$\tilde{\mathrm{h}}_\mathrm{t} = \mathrm{tanh}(\mathrm{W_f} * [\mathrm{r_t} * \mathrm{h_{t-1}}, \mathrm{x_t}]) \tag{10}$$

$$\mathrm{h_t} = (1 - \mathrm{z_t}) * \mathrm{h_{t-1}} + \mathrm{r_t} * \tilde{\mathrm{h}}_\mathrm{t}) \tag{11}$$

Where: Z=memory keep or forget, t=time step, W=weight matrix, $\tilde{\mathrm{h}}_\mathrm{t}$=candidate status output, $\sigma =$ sigmoid function, x=previous state, r=reset gate.

2.4. BiGRU

BiGRU stands as a specialized type of RNN design crafted explicitly to handle sequential data intricacies by capturing intricate patterns in both forward and backward sequences. Employing the GRU model as its foundation, the BiGRU comprises dual GRU layers: one scrutinizing the input sequence in a forward trajectory, while the other examines it in reverse. This bidirectional method enables the network to grasp dependencies and patterns from both past and future contexts, enriching its comprehension of sequential data. By amalgamating insights from both directions, the BiGRU effectively overcomes the constraints of unidirectional models, proving its suitability for diverse research tasks such as natural language processing, sentiment analysis, and time series forecasting, especially in scenarios demanding the capture of bidirectional relationships. The BiGRU mathematically expressed using as (12) to (20) [27]. Forward GRU,

$$\mathrm{z_t^{(f)}} = \sigma\left(\mathrm{W_z^{(f)}} \cdot \mathrm{x_t} + \mathrm{U_z^{(f)}} \mathrm{h_{t-1}^{(f)}}\right) \tag{12}$$

$$\mathrm{r_t^{(f)}} = \sigma\left(\mathrm{W_z^{(f)}} \cdot \mathrm{x_t} + \mathrm{U_r^{(f)}} \mathrm{h_{t-1}^{(f)}}\right) \tag{13}$$

$$\tilde{\mathrm{h}}_\mathrm{t}^{(\mathrm{f})} = \mathrm{tanh}\left(\mathrm{W_h^{(f)}} \cdot \mathrm{x_t} + \mathrm{r_h^{(f)}}(\mathrm{U_z^{(f)}} \mathrm{h_{t-1}^{(f)}})\right) \tag{14}$$

$$h_t^{(f)} = \left(1 - z_t^{(f)}\right) \cdot \tilde{h}_t^{(f)} + z_t^{(f)} \cdot h_{t-1}^{(f)} \tag{15}$$

and the backward GRU.

$$z_t^{(b)} = \sigma\left(W_z^{(b)} \cdot x_t + U_z^{(b)} h_{t+1}^{(b)}\right) \tag{16}$$

$$r_t^{(b)} = \sigma\left(W_z^{(b)} \cdot x_t + U_r^{(b)} h_{t+1}^{(b)}\right) \tag{17}$$

$$\tilde{h}_t^{(b)} = \tanh\left(W_h^{(b)} \cdot x_t + r_h^{(b)}(U_z^{(b)} h_{t+1}^{(b)})\right) \tag{18}$$

$$h_t^{(b)} = \left(1 - z_t^{(b)}\right) \cdot \tilde{h}_t^{(b)} + z_t^{(b)} \cdot h_{t+1}^{(b)} \tag{19}$$

The final hidden state is obtained by concatenating the forward and backward hidden states:

$$h_t = \left[h_t^{(f)}, h_t^{(b)}\right] \tag{20}$$

where, h =hidden state, Z=memory keep or forget, t=time step, f=forward GRU, x=input, W and U weight matrix and σ as the sigmoid.

2.5. Recurrent neural network

A RNN is an artificial neural network specialized in managing sequential data by retaining memory of prior information. Diverging from standard feedforward neural networks, RNNs integrate loops in their structure, enabling them to retain data and apply it for forecasting future elements in a sequence. This looping system empowers RNNs to factor in not only the present input but also the context derived from earlier inputs, rendering them apt for tasks encompassing time series forecasting, natural language processing, speech recognition, and various analyses of sequential data where the sequence's order and context play a critical role in generating precise predictions or classifications. It is expressed using as (21) [28].

$$h_t = f(W_{ih} x_t + W_{hh} h_{t-1} + b_h) \tag{21}$$

Where: h =hidden state, t=time step t, f=activation function, W =wight matrix, i, x=input, b =bias term.

3. RESULTS AND DISCUSSION

3.1. Basis statistics

The provided Table 1 exploring stock price trends: descriptive statistics for ten prominent companies offers a comprehensive analysis of key statistical metrics for five major financial entities, including HDFC Bank, Maruti, Reliance, Sun Pharma and Tata Steel, based on their closing prices. The dataset, comprising 5,956 data points and Maruti is 5,039 for each entity. Mean values reveal that MARUTHI has highest 3,433.36, while Sun Pharma has the lowest at 343.24. Reliance has the highest standard deviation, indicating higher price variability, while Tata Steel has the lowest, suggesting more stability. Sun Pharma has the lowest minimum price, while HDFC Bank has the highest minimum. HDFC Bank, Maruti, and Sun Pharma have relatively low values at the 25th percentile. Maruti has the highest median price, while Sun Pharma has the lowest. HDFC Bank has the highest 75^{th} percentile value. MARUTI has the highest maximum stock price, while Sun Pharma has the lowest.

Table 1. Exploring stock price trends: descriptive statistics for ten prominent Companies

	HDFC Bank	Maruti	Relience	Sun Pharma	Tata Steel
Count	5,956	5.039	5.953	5.956	5.956
Mean	464.5581	3433.363	676.7863	343.24	35.37712
Std	510.1792	3091.651	746.5207	317.8713	29.78737
Min	13.5817	143.868	23.88001	6.380887	1.997227
25%	58.7858	795.9037	103.574	51.87845	17.52377
50%	218.4493	1452.645	426.1733	230.5557	29.05627
75%	864.2262	6685.938	859.4939	575.3492	42.08877
Max	1728.2	10788.45	2831.847	1160.8	139.6047

3.2. Correlation between data

Figure 2 illustrated the correlation between selected stocks where 1 indicating the positive correlation whereas -1 indicate the negative correlation. It's evident that there are strong positive correlations between several companies. For instance: HDFC Bank and Maruti have a very high correlation coefficient of approximately 0.952380. Reliance and Tata Steel also exhibit a strong positive correlation of around 0.892565. Additionally, there are moderate positive correlations between other pairs like HDFC Bank-Reliance, HDFC Bank-Sun Pharma, and Maruti-Reliance, ranging from approximately 0.73 to 0.95. This data suggests that there is a significant degree of correlation between the stock prices of these companies, indicating that movements in the stock prices of one company are likely to correspond with movements in the stock prices of others in the given dataset.

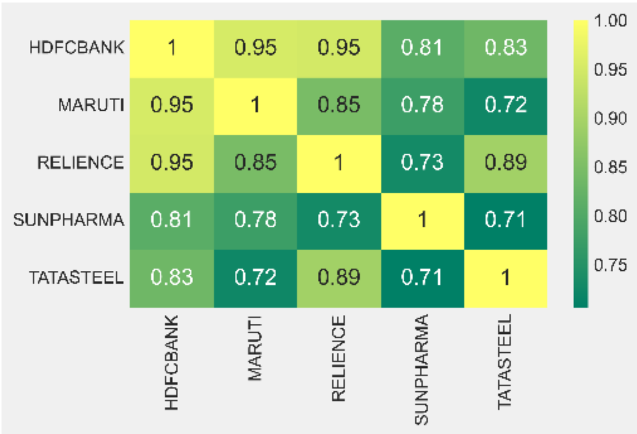

Figure 2. Analyzing the stock market: correlations among companies

3.3. LSTM

Figure 3 illustrated the performance of companies concerning stock price forecasting using the LSTM algorithm. Sun Pharma emerges as the top performer among the evaluated companies, showcasing a highly accurate stock price forecasting capability with the LSTM algorithm. Following closely behind is Tata Steel, which demonstrates a strong performance, albeit slightly lower than Sun Pharma. Reliance secures a mid-tier position, indicating a respectable but comparatively lower accuracy in stock price predictions compared to Sun Pharma and Tata Steel. HDFC Bank falls within a similar mid-tier range, slightly behind reliance in terms of performance. Maruti, while maintaining a commendable stance, falls into the lower tier among the evaluated companies in terms of stock price forecasting accuracy using the LSTM algorithm.

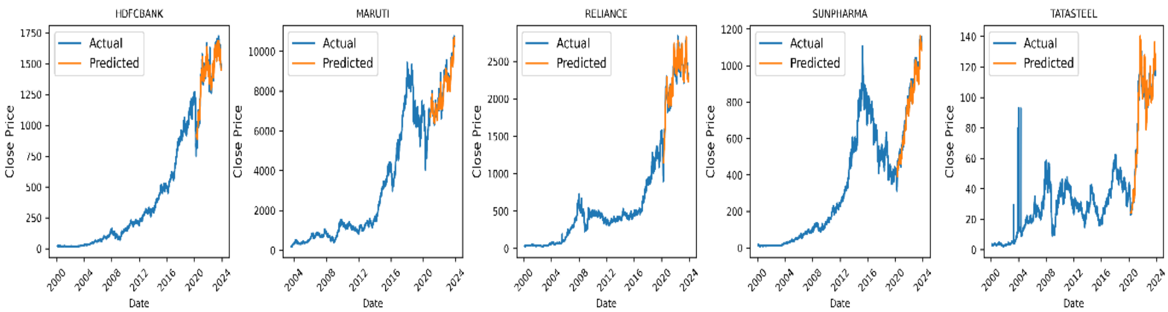

Figure 3. LSTM predictions for HDFC Bank, Maruti, Reliance, Sun Pharma, and Tata Steel stocks

3.4. BiLSTM

Figure 4 depicts the performance of BiLSTM algorithm across different companies' stock price forecasting. Reliance emerges as the top performer among the assessed companies, displaying exceptional accuracy in stock price predictions with the BiLSTM algorithm. Sun Pharma closely follows, demonstrating a highly accurate forecasting capability. HDFC Bank secures a mid-tier position, showcasing respectable but comparatively lower accuracy in stock price predictions compared to Reliance and Sun Pharma. Tata Steel, while maintaining a commendable stance, falls slightly behind HDFC Bank in the assessment of stock price forecasting accuracy using the BiLSTM algorithm. Maruti, although still performing well, ranks lower among the evaluated companies in terms of accuracy in stock price predictions with the BiLSTM algorithm.

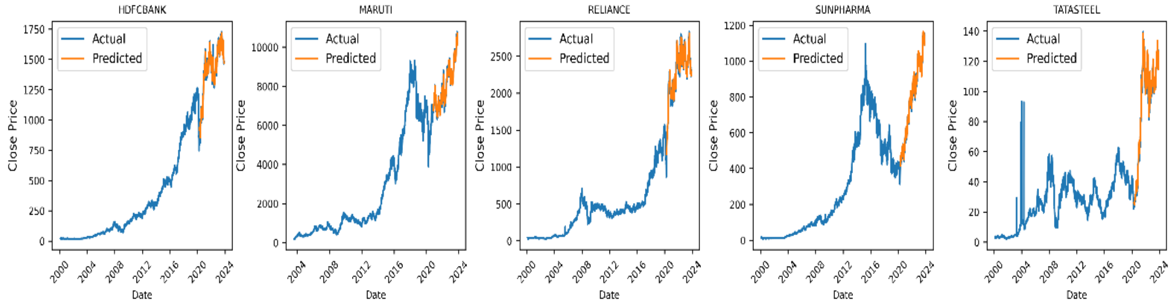

Figure 4. BiLSTM predictions for HDFC Bank, Maruti, Reliance, Sun Pharma, and Tata Steel stocks

3.5. GRU

Figure 5 illustrate that the GRU performance on different stocks with respective actual versus prediction. Reliance and Maruti exhibit comparable performance, both securing a similar level of effectiveness in their stock price predictions using the GRU algorithm. Sun Pharma closely follows, demonstrating high accuracy in its forecasting abilities. HDFC Bank stands in the mid-tier range, showcasing a respectable performance but falling slightly behind Reliance, Maruti, and Sun Pharma. Tata Steel, while still performing reasonably well, ranks comparatively lower among the evaluated companies when using the GRU algorithm for stock price predictions.

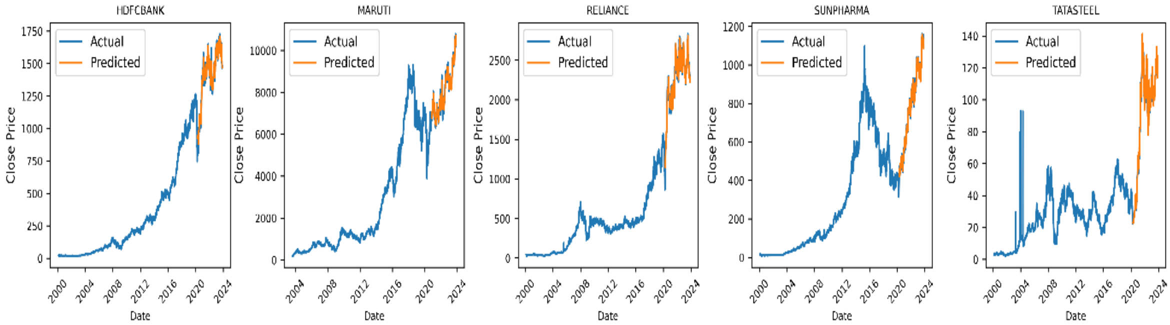

Figure 5. GRU predictions for HDFCBANK, Maruti, Reliance, Sun Pharma, and Tata Steel stocks

3.6. BiGRU

Figure 6 illustrates the performance of different tickers with respect to the date. The sun pharma emerges as the top performer among the evaluated companies, showcasing highly accurate stock price predictions with the BiGRU algorithm. HDFC Bank secures a mid-tier position, displaying respectable accuracy but falling slightly behind sun pharma in forecasting stock prices. Reliance follows closely behind HDFC Bank, demonstrating a respectable but comparatively lower accuracy in stock price predictions with the BiGRU algorithm. Maruti holds a position slightly below Reliance, indicating a slightly lower accuracy in stock price forecasting. Tata Steel, while demonstrating a commendable performance with the BiGRU algorithm, falls into the lower tier among the assessed companies in terms of accuracy in stock price predictions.

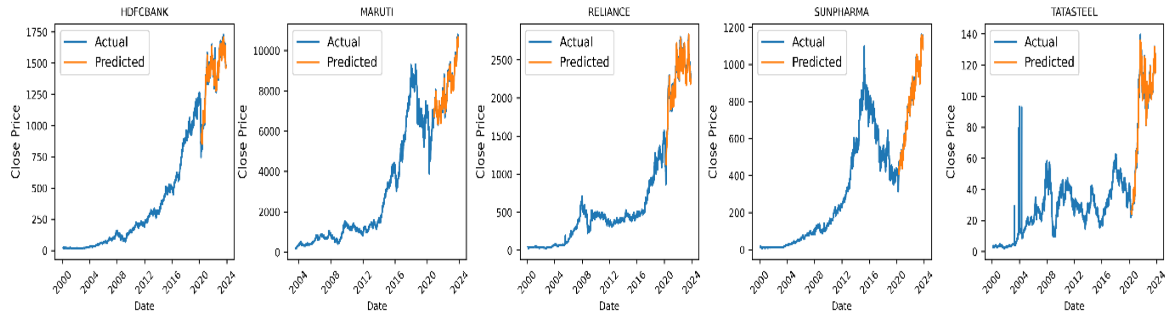

Figure 6. BiGRU predictions for HDFC Bank, Maruti, Reliance, Sun Pharma, and Tata Steel stocks

3.7. RNN

Figure 7 illustrates the actual versus predicted stock prices based on the RNN algorithm. Tata Steel emerges as the top performer among the evaluated companies, displaying highly accurate stock price predictions with the RNN algorithm. Sun pharma secures a position following Tata Steel, showcasing commendable accuracy but slightly lower than Tata Steel in forecasting stock prices. Maruti holds a mid-tier position, demonstrating respectable accuracy but falling behind Sun Pharma in stock price predictions with the RNN algorithm. HDFC Bank follows Maruti, indicating a lower accuracy in stock price forecasting compared to Maruti. Reliance, although displaying a reasonable performance, falls into the lower tier among the assessed companies in terms of accuracy in stock price predictions using the RNN algorithm.

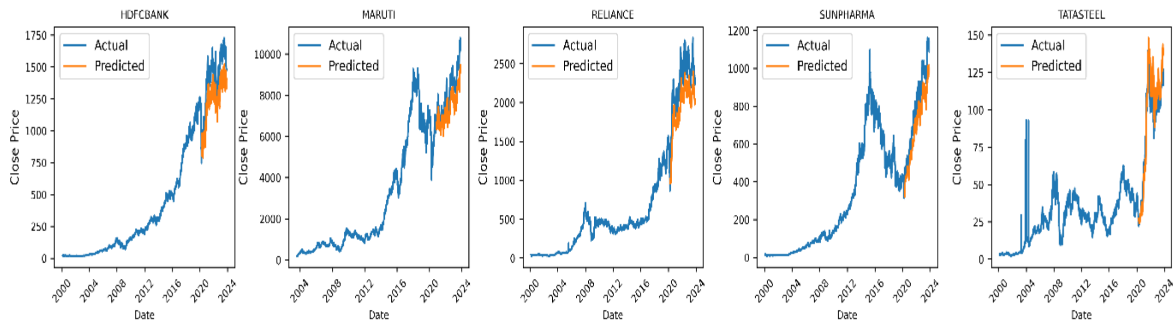

Figure 7. RNN Predictions for HDFC Bank, Maruti, Reliance, Sun Pharma, and Tata Steel stocks

3.8. Comparison of various algorithms

Table 2 illustrate the comparison of different algorithm based on evaluation matrix. These metrics help assess different aspects of a predictive model's accuracy and performance. BiLSTM and BiGRU consistently demonstrate high R2 values, indicating strong predictive capabilities across different stocks. BiGRU generally performs well, with competitive MAE, MSE, MAPE, and RMSE values. GRU and LSTM also show competitive performance, but BiGRU appears to outperform them in several cases. BiLSTM and LSTM exhibit slightly higher errors in certain metrics compared to BiGRU and GRU. RNNs, in contrast, show varied performance across stocks, with significantly lower R2 values, indicating a poorer fit to the data. The RNN algorithm struggles to capture the patterns in the time series data, resulting in higher error metrics.

For HDFC Bank, both BiGRU and LSTM models exhibit strong predictive capabilities with high R2 values of 0.983 and 0.962, respectively. BiGRU outperforms other algorithms with the lowest RMSE of 26.201, suggesting precise predictions. MARUTI's BiLSTM and BiGRU models achieve competitive R2 values of 0.977 and 0.982, demonstrating accurate predictions. RELIANCE's BiLSTM model stands out with an exceptionally high R2 of 0.996, showcasing superior predictive capabilities. Despite a negative R2, the RNN model's performance indicates challenges in accurately forecasting Reliance's stock. Sun Pharma consistently demonstrates high R2 values across all algorithms, with BiGRU achieving the highest at 0.996. The LSTM and BiLSTM models also perform well, showcasing the algorithm's effectiveness for this stock. Tata Steel's BiGRU model outperforms other algorithms with the lowest R2 of 0.994 and minimal error metrics, suggesting accurate predictions.

Table 3 presents the ranking of various algorithms concerning their performance in stock price forecasting for specific companies. LSTM demonstrated superior performance for Sun Pharma, while BiLSTM excelled for Reliance. GRU ranked highest for both Reliance and Maruti, while BiGRU showcased the best predictive abilities for Sun Pharma. Lastly, RNN emerged as the top algorithm for Tata Steel. Each algorithm displayed varying degrees of efficacy across different companies, highlighting their diverse performances in the realm of stock price prediction.

Table 2. Performance metrics for various algorithms on stock ticker predictions

Sr. No.	Ticker	Algorithm	R^2	MAE	MSE	MAPE	RMSE
1	HDFC Bank	LSTM	0.962	30.341	1531.443	0.022	39.134
		BiLSTM	0.977	23.140	922.961	0.017	30.380
		GRU	0.976	23.489	954.895	0.017	30.901
		BiGRU	0.983	19.828	686.480	0.015	26.201
		RNN	0.082	177.192	37041.059	0.123	192.461
2	Maruti	LSTM	0.951	176.889	55852.484	0.022	236.331
		BiLSTM	0.977	115.575	25746.895	0.015	160.458
		GRU	0.972	126.317	25746.895	0.016	176.793
		BiGRU	0.982	103.913	20698.513	0.013	143.870
		RNN	0.378	724.984	703555.526	0.086	838.782
3	Reliance	LSTM	0.957	48.874	4046.217	0.022	63.610
		BiLSTM	0.996	44.333	3231.285	0.020	56.844
		GRU	0.972	39.642	2668.489	0.018	51.657
		BiGRU	0.971	40.208	2722.787	0.018	52.180
		RNN	-0.232	317.485	115461.447	0.141	339.796
4	Sun Pharma	LSTM	0.986	19.655	637.813	0.032	25.255
		BiLSTM	0.994	13.075	267.800	0.018	16.365
		GRU	0.995	10.961	206.154	0.015	14.358
		BiGRU	0.996	9.988	177.011	0.014	13.305
		RNN	0.744	95.009	11282.049	0.118	106.217
5	Tata Steel	LSTM	0.984	2.994	16.374	0.033	4.047
		BiLSTM	0.991	2.405	9.842	0.030	3.137
		GRU	0.992	2.159	8.579	0.025	2.929
		BiGRU	0.994	1.869	6.775	0.022	2.603
		RNN	0.876	8.711	130.454	0.098	11.422

Table 3. Algorithm rankings for stock price forecasting across companies

Sr. No.	Algorithm	Ticker (1^{st} rank)
1	LSTM	Sunp Harma
2	BiLSTM	Reliance
3	GRU	Reliance, Maruti
4	BiGRU	Sun Pharma
5	RNN	Tata Steel

4. CONCLUSION

The problem statement aims to compare various RNN algorithm variants-LSTM, BiLSTM, GRU, BiGRU, and RNN-specifically focusing on the NSEI for stock market analysis. Historical data spanning from 1/1/2004 to 30/06/2023 was gathered from the public domain to achieve this objective. The stock prediction analysis highlights that BiGRU, LSTM, and RNN exhibit strong predictive capabilities for HDFC Bank, with BiGRU showing superior precision, boasting the lowest RMSE. For Maruti, both BiLSTM and BiGRU models showcase competitive accuracy. However, Reliance's BiLSTM model stands out with an exceptionally high R2 of 0.996. Sun Pharma consistently displays high R2 values across models, with BiGRU achieving the highest at 0.996. Notably, Tata Steel's BiGRU model stands out with the lowest R2 of 0.994 and minimal error metrics, indicating precise predictions. Despite our work's strengths, limitations exist, such as the exclusion of economic and fundamental analysis, focusing solely on MACD, MFI, RSI, BB, and time series data. Moreover, the performance rankings of diverse algorithms in stock price forecasting for individual companies. LSTM stood out as the best performer for SUNPHARMA, whereas BiLSTM exhibited exceptional accuracy for Reliance. GRU emerged as the top-ranking algorithm for both Reliance and Maruti, while BiGRU demonstrated superior predictive capabilities specifically for Sun Pharma. Additionally, RNN was identified as the most effective algorithm for Tata Steel. The varying efficacy showcased by each algorithm across different companies underscores their diverse and distinct performances within the domain of stock price prediction.

REFERENCES

[1] B. G. Malkiel, “Efficient market hypothesis,” in *Finance*, London: Palgrave Macmillan UK, 1989, pp. 127–134, doi: 10.1007/978-1-349-20213-3_13.

[2] M. C. A. Neto, G. D. C. Calvalcanti, and T. I. Ren, “Financial time series prediction using exogenous series and combined neural networks,” in *Proceedings of the 2009 International Joint Conference on Neural Networks*, pp. 2578–2585, 2009.

[3] A. Timmermann and C. W. J. Granger, “Efficient market hypothesis and forecasting,” *International Journal of Forecasting*, vol. 20, no. 1, pp. 15–27, Jan. 2004, doi: 10.1016/S0169-2070(03)00012-8.

[4] Y. Qi, W. Yu, and Y. Deng, “Stock prediction under COVID-19 based on LSTM,” in *Proceedings of IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers, IPEC 2021*, Apr. 2021, pp. 93–98, doi: 10.1109/IPEC51340.2021.9421323.

[5] J. T. Firouzjaee and P. Khaliliyan, “The Interpretability of LSTM Models for Predicting Oil Company Stocks: Impact of Correlated Features,” 2022, [Online]. Available: http://arxiv.org/abs/2201.00350

[6] Q. Jiang, C. Tang, C. Chen, X. Wang, and Q. Huang, “Stock price forecast based on LSTM neural network,” in *Lecture Notes on Multidisciplinary Industrial Engineering*, vol. part F46, 2019, pp. 393–408, doi: 10.1007/978-3-319-93351-1_32.

[7] Y. Yu, “Research on the forecast of stock price index based on BiLSTM-GRU,” in *Proceedings - 2022 Euro-Asia Conference on Frontiers of Computer Science and Information Technology, FCSIT 2022*, Dec. 2022, pp. 81–85, doi: 10.1109/FCSIT57414.2022.00027.

[8] J. T. Firouzjaee and P. Khaliliyan, “LSTM Architecture for Oil Stocks Prices Prediction,” 2020, . *arXiv:2201.00350v1*.

[9] Y. Huang and Y. Song, “A new hybrid method of recurrent reinforcement learning and BiLSTM for algorithmic trading,” *Journal of Intelligent & Fuzzy Systems.*, pp. 1–13, May 2023, doi: 10.3233/jifs-223101.

[10] R. Jaiswal and B. Singh, “A Hybrid Convolutional Recurrent (CNN-GRU) Model for Stock Price Prediction,” in *2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT)*, Apr. 2022. doi: 10.1109/csnt54456.2022.9787651.

[11] J. Qi, S. Huang, J. Hu, W. Ni, and H. Chen, “Stock price prediction in Chinese stock markets based on CNN-GRU attention model,” in *International Symposium on Artificial Intelligence and Robotics 2022*, Dec. 2022. doi: 10.1117/12.2663261.

[12] G. R. Patra and M. N. Mohanty, “An LSTM-GRU based hybrid framework for secured stock price prediction,” *J. Stat. Manag. Syst.*, vol. 25, no. 6, pp. 1491–1499, Aug. 2022, doi: 10.1080/09720510.2022.2092263.

[13] Y. Gao, R. Wang, and E. Zhou, “Stock Prediction Based on Optimized LSTM and GRU Models,” *Sci. Program.*, vol. 2021, pp. 1–8, Sep. 2021, doi: 10.1155/2021/4055281.

[14] V. Polepally, N. S. Nandini Reddy, M. Sindhuja, N. Anjali, and K. J. Reddy, “A Deep Learning Approach for Prediction of Stock Price Based on Neural Network Models: LSTM and GRU,” in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Jul. 2021, doi: 10.1109/icccnt51525.2021.9579782.

[15] M. E. Karim and S. Ahmed, “A Deep Learning-Based Approach for Stock Price Prediction Using Bidirectional Gated Recurrent Unit and Bidirectional Long Short Term Memory Model,” in *2021 2nd Global Conference for Advancement in Technology (GCAT)*, 2021, doi: 10.1109/GCAT52182.2021.9587895.

[16] P. Sonkiya, V. Bajpai, and A. Bansal, “Stock price prediction using BERT and GAN,” 2021, *arXiv:2107.09055*.

[17] H. Zhu, A. Yang, and X. Zhang, “Quantitative trading model and price forecasting system based on BiGRU neural network,” in *2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 2022. doi: 10.1109/AEECA55500.2022.9918821.

[18] Y. Duan, Y. Liu, Y. Wang, S. Ren, and Y. Wang, “Improved BIGRU Model and Its Application in Stock Price Forecasting,” *Electronics*, vol. 12, no. 12, 2023, doi: 10.3390/electronics12122718.

[19] M. Agrawal *et al.*, “Stock Prediction and Analysis based on RNN Neural Network,” *J. Phys. Conf. Ser.*, vol. 151, no. 2, pp. 1–19, Mar. 2020, doi: 10.5220/0006749901020108.

[20] Y. Zhu, “Stock price prediction using the RNN model,” *J. Phys. Conf. Ser.*, vol. 1650, p. 32103, 2020, doi: 10.1088/1742-6596/1650/3/032103.

[21] S.-L. YU and Z. Li, “Stock Price Prediction Based on ARIMA RNN Combined Model,” *DEStech Trans. Soc. Sci. Educ. Hum. Sci.*, Mar. 2018, doi: 10.12783/dtssehs/icss2017/19384.

[22] A. Pandey, “Prediction of Stock Price using RNN’s LSTM-Based Deep Learning Model,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 8, pp. 2469–2474, Aug. 2021, doi: 10.22214/ijraset.2021.37791.

[23] S. Mittal and A. Chauhan, “A RNN LSTM Based Predictive Modelling Framework for Stock Market Prediction Using Technical Indicators,” *Int. J. Rough Sets Data Anal.*, vol. 7, no. 1, pp. 1–13, Sep. 2021, doi: 10.4018/ijrsda.288521.

[24] N. M. Rezk, M. Purnaprajna, T. Nordstrom, and Z. Ul-Abdin, “Recurrent Neural Networks: An Embedded Computing Perspective,” *IEEE Access*, vol. 8, pp. 57967–57996, 2020, doi: 10.1109/ACCESS.2020.2982416.

[25] Y. Yu, X. Si, C. Hu, and J. Zhang, “A review of recurrent neural networks: LSTM cells and network architectures,” *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, 2019.

[26] A. A. Sharfuddin, M. N. Tihami, and M. S. Islam, “A Deep Recurrent Neural Network with BiLSTM model for Sentiment Classification,” *2018 Int. Conf. Bangla Speech Lang. Process. ICBSLP 2018*, no. October, 2018, doi: 10.1109/ICBSLP.2018.8554396.

[27] R. Dey and F. M. Salem, “Gate-variants of gated recurrent unit (GRU) neural networks,” in *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)*, 2017, pp. 1597–1600.

[28] A. M. Rather, A. Agarwal, and V. N. Sastry, “Recurrent neural network and a hybrid model for prediction of stock returns,” *Expert Syst. Appl.*, vol. 42, no. 6, pp. 3234–3241, 2015.

BIOGRAPHIES OF AUTHORS

Zakir Mujeeb Shaikh is a research scholar at department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India. He holds M.E. Computer from Bharati Vidyapeeth University Pune, Maharashtra, India. His research areas are time series analysis, data mining, medical image processing, and pattern recognition. He can be contacted at email: shaikhzakir03@yahoo.com.

Suguna Ramadass is a Professor in the Computer Science and Engineering Department at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India. She holds a Ph.D. in CSE from Anna University Chennai, an M. Tech in CSE from IIT Madras, and a B.E. in CSE from Thiagarajar College of Engineering, Tamilnadu, India. With 32 years of academic experience, she is a dedicated educator and a Member of the Board of Studies. Dr. Suguna's research includes over 70 papers published in reputable journals and conferences. She has organized international conferences and workshops, earning her the "Best Supporter to Society" award by Science for the Society organized at IIT Madras. Additionally, she actively participates in academic and technical roles and is a life member of the Computer Society of India (CSI). Her research areas are image processing, data mining, and big data analytics. She can be contacted at email: drsuguna@veltech.edu.in.