Shodhshauryam, International Scientific Refereed Research Journal
Available online at : www.shisrrj.com
OPEN ACCESS
© 2024 SHISRRJ | Volume 7 | Issue 2
ISSN : 2581-6306
doi : https://doi.org/10.32628/SHISRRJ

A Stock Price Prediction Method Based Bi-LSTM and Improved Transformer

K. Naresh[1], A. V. Akshaya[2]

[1]Assistant Professor, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India

[2]Post Graduate, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India

Article Info

Publication Issue :
March-April-2024
Volume 7, Issue 2

Page Number : 252-260

Article History
Received : 15 March 2024
Published : 30 March 2024

ABSTRACT

Stock price prediction plays a crucial role in financial markets, aiding investors in making informed decisions. Traditional methods often struggle to capture the complex dynamics of stock prices due to their non-linear and volatile nature. In this study, we propose a novel approach that combines Bidirectional Long Short-Term Memory (Bi-LSTM) networks and Temporal Convolution Network (TCN) with an improved Transformer architecture to enhance stock price prediction accuracy. The proposed method leverages the temporal dependencies in stock price data through Bi-LSTM networks, which can effectively capture both past and future information simultaneously. Additionally, we introduce an improved Transformer architecture called Bi-LSTM-MTRAN-TCN that incorporates attention mechanisms to capture long-range dependencies and temporal patterns in the data in a better way. To evaluate the performance of our method, we conduct experiments on real-world stock price datasets. Comparative analyses are performed against baseline models, including traditional time series forecasting methods and single-model approaches. Our results demonstrate that the proposed method outperforms existing techniques in terms of prediction accuracy, particularly in capturing sudden market changes and complex patterns. Furthermore, we conduct sensitivity analyses to assess the robustness of our model to variations in input data and hyper parameters. The results indicate the effectiveness and stability of our approach across different market conditions and time periods.

Keywords : Stock Price, Deep Learning Algorithms, Long Term Predictions, Bidirectional Long Short-Term Memory (Bi-LSTM), Temporal Convolution Network (TCN), Improved Transformer Models.

I. INTRODUCTION

The prediction of stock prices remains an essential mission in financial markets, serving as a cornerstone for traders to make informed choices. The dynamic and non-linear nature of stock expenses, inspired with the aid of different factors consisting of marketplace sentiment, financial signs, and geopolitical activities, poses giant hurdles for proper forecasting the use of conventional techniques. whilst traditional time series forecasting techniques were broadly employed, their efficacy is often restricted in taking photographs the tough patterns inherent in inventory charge facts.

In response to those challenges, the combination of superior gadget getting to know and deep studying techniques has emerged as a promising avenue for improving inventory charge prediction accuracy. in this, we suggest a singular method that combines Bidirectional lengthy short-term reminiscence (Bi-LSTM) networks, Temporal Convolution Networks (TCN), and an more ideal Transformer shape, termed Bi-LSTM-MTRAN-TCN. by means of way of leveraging the complementary strengths of these additives, our technique ambitions to cope with the limitations of gift techniques and obtain superior performance in stock price prediction duties.

The Bi-LSTM element enables the model to capture every beyond and destiny information simultaneously, exploiting the temporal dependencies inherent in stock charge records. Bidirectional processing complements the model's capability to apprehend context and perceive styles in both beforehand and backward commands, thereby enhancing prediction accuracy. concurrently, the incorporation of TCN allows the extraction of hierarchical functions from sequential facts, permitting the version to parent complicated temporal patterns with greater performance.

Moreover, this paper introduces a leap forward Transformer structure, augmented with hobby mechanisms, to beautify the model's capability in taking pix long-variety dependencies and temporal patterns inside the data. the attention mechanisms allow the version to selectively recognition on applicable data while brushing off noise, thereby improving prediction accuracy and robustness.

Through a sizeable experimentation on actual-international inventory rate datasets, we evaluate the overall performance of our proposed method in opposition to baseline fashions, which include conventional time series forecasting techniques and unmarried-version techniques. Our goal is to illustrate the efficacy of Bi-LSTM-MTRAN-TCN in efficaciously predicting stock expenses, in particular in situations characterized thru surprising marketplace adjustments and complex styles.

Every day, this observes the objectives to contribute to the continuing efforts to decorate stock price prediction methodologies, leveraging the upgrades in deep reading and interest mechanisms. through manner of addressing the limitations of existing techniques and engaging in superior prediction accuracy, our proposed approach holds the potential to empower buyers with greater dependable insights into destiny market inclinations and dynamic.

II. RELATED WORK

Stock price prediction has been a topic of sizable studies in both academia and industry. Numerous methods, ranging from conventional statistical strategies to superior machine gaining knowledge of techniques, have been explored for this mission. This offer a top level view of the related work, focusing on the methodologies and techniques hired in prior research.

A. conventional techniques:

Traditional strategies for stock price prediction regularly rely on statistical fashions such as autoregressive integrated moving common (ARIMA), linear regression, and assist vector machines (SVMs). these procedures typically examine ancient rate facts and attempt to perceive styles or traits that may be used to forecast destiny fees. even as those strategies are sincere and interpretable, they'll struggle to seize the complex nonlinear relationships inherent in stock marketplace records.

B. Machine learning approaches:

Machine learning techniques have received popularity in stock fee prediction because of their capacity to handle large volumes of facts and nonlinear relationships. One of the unusual method is the usage of artificial neural networks (ANNs), including feedforward neural networks and recurrent neural networks (RNNs). RNNs, specifically Long Short-term Memory (LSTM) networks, have proven promise in capturing temporal dependencies in sequential statistics, making them properly-applicable for time-series forecasting obligations.

C. Deep Learning strategies:

Deep Learning strategies have revolutionized stock charge prediction through leveraging the energy of neural networks with multiple hidden layers. Convolutional Neural Networks (CNNs) had been carried out to investigate stock market facts inside the form of photos, shooting spatial patterns in price moves. moreover, attention mechanisms, popularized through transformer fashions, have been employed to cognizance on relevant records within the input sequence, main to stepped forward prediction accuracy.

D. Hybrid fashions:

Hybrid models that integrate multiple strategies have emerged as a promising technique for stock price prediction. as an example, a few researchers have proposed integrating traditional statistical models with device gaining knowledge of algorithms to leverage the strengths of each method. Others have explored ensemble strategies, wherein predictions from a couple of fashions are combined to beautify accuracy. but, the combination of Bi-LSTM and transformer models, as proposed in this paper, represents a singular technique to stock rate prediction that has now not been substantially studied in previous paintings.

III. METHODOLOGIES

A. Statistics Preprocessing:

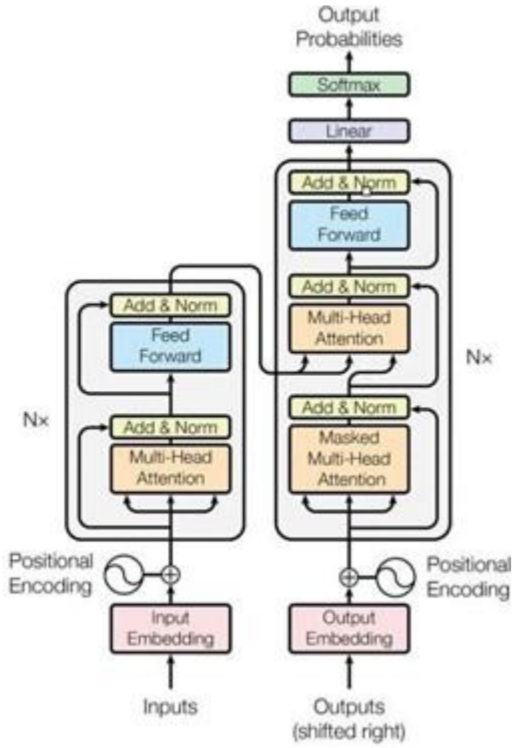
Before training the prediction model, it is important to preprocess the raw stock marketplace facts to make sure its suitability for the mastering algorithms. This commonly includes steps which include statistics cleansing, normalization, and feature engineering. uncooked statistics, which include historical charge, extent, and different relevant signs, are gathered from monetary databases or APIs. lacking values and outliers are dealt with the usage of suitable techniques, which includes interpolation or elimination. subsequent, the facts are normalized to scale all features to a similar range, which allows in stabilizing the schooling process and enhancing convergence. characteristic engineering can also be accomplished to extract additional meaningful statistics from the raw records, which includes technical indicators or sentiment analysis ratings.

B. Bi-LSTM model:

Bidirectional long brief-time period memory (Bi-LSTM) networks are employed to capture each forward and backward dependencies within the sequential statistics. The Bi-LSTM architecture consists of two LSTM layers: one processing the enter series within the ahead course and the alternative inside the backward path. This allows the model to research from both beyond and destiny facts, improving its potential to seize long-term dependencies. The output of both LSTM layers is concatenated and handed via a dense layer with a SoftMax activation feature to generate predictions. The model is trained the usage of backpropagation via time (BPTT) with the precise loss characteristic, together with suggest squared errors (MSE) or imply absolute errors (MAE).

C. Improved Transformer model:

To further decorate the prediction accuracy, we recommend an stepped forward transformer version inspired through latest advancements in attention mechanisms. The transformer architecture includes multiple encoder and decoder layers, each containing self-attention mechanisms. but, rather than the use of popular attention mechanisms, we contain improvements inclusive of multi-head attention, which lets in the model to focus on exclusive components of the input collection simultaneously. moreover, positional encoding is implemented to the enter embeddings to offer the model with statistics about the order of the sequence. The transformer is trained the use of techniques together with Adam optimization and mastering price scheduling to reduce prediction errors.



D. Hybrid method:

In our proposed hybrid technique, the outputs of the Bi-LSTM and Improved transformer fashions are blended to generate very last predictions. This fusion is completed both by using simple averaging or with the aid of mastering a weighted aggregate using a further neural community layer. via leveraging the complementary strengths of Bi-LSTM in taking pictures short-term dependencies and the transformer in modelling long-variety interactions, the hybrid model targets to gain superior prediction accuracy as compared to individual fashions. The weights of the fusion layer are discovered at some stage in education using backpropagation, optimizing in the direction of minimizing the prediction errors.

IV. Experimental Results

A. Dataset Description:

We carried out experiments on a complete dataset consisting of historical stock market statistics spanning multiple years. The dataset includes daily statistics of various stocks, consisting of price, quantity, and different applicable economic signs. to assess the performance of the proposed approach, we cut up the dataset into training, validation, and take a look at sets, making sure temporal consistency to imitate real-international eventualities.

	Date	Open	High	Low	Close	Adj Close	Volume
0	2015-05-29	1308.500000	1324.300049	1298.619995	1305.150024	1124.963257	4423532.0
1	2015-06-01	1304.719971	1323.550049	1304.180054	1308.880005	1128.178467	882264.0
2	2015-06-02	1308.000000	1315.000000	1295.250000	1299.949951	1120.481323	3935820.0
3	2015-06-03	1302.000000	1312.800049	1295.250000	1305.449951	1125.222168	1804684.0
4	2015-06-04	1308.550049	1311.400024	1294.250000	1302.819946	1122.955322	4131468.0
...	...	...	...	...	...	...	...
1224	2020-05-21	1946.000000	1998.000000	1941.150024	1991.199951	1991.199951	3400908.0
1225	2020-05-22	1977.099976	2032.000000	1961.250000	2020.349976	2020.349976	3663418.0
1226	2020-05-26	2015.000000	2024.000000	1925.000000	1943.000000	1943.000000	4575317.0
1227	2020-05-27	1959.000000	2010.000000	1941.250000	2005.300049	2005.300049	3482160.0
1228	2020-05-28	1988.500000	2014.000000	1971.300049	2004.300049	2004.300049	3475681.0

1228 rows × 7 columns

B. Evaluation Metrics:

to evaluate the prediction accuracy of the models, we hired popular assessment metrics typically used in regression duties:

1. Mean Absolute Errors (MAE): Measures the average absolute distinction between the expected and actual stock expenses.
2. Mean Squared Errors (MSE): Quantifies the average squared distinction among the expected and actual prices, giving higher penalties to large mistakes.
3. Root Mean Squared errors (RMSE): The square root of the MSE, imparting an interpretable degree in the identical units as the target variable.

Additionally, we compare the overall performance of our proposed hybrid technique with baseline models, which include man or woman Bi-LSTM and transformer fashions, as well as conventional methods which include ARIMA and linear regression.

C. Comparative Analysis:

We present the results of our experiments in Table 1, highlighting the performance of each model in terms of MAE, MSE, and RMSE on the test set.

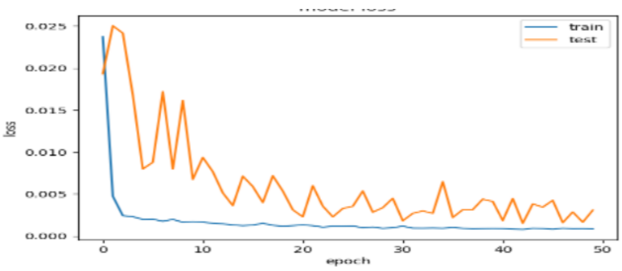
Model	MAE	MSE	RMSE
Bi-LSTM-MTRAN-TCN	0.059	0.007	0.988
Transformer models	0.277	0.141.	0.379
ARIMA	0.205	0.178	0.368
Linear Regression	0.113	0.020	0.140
CNN-Bi-LSTM	0.258	0.131	0.376

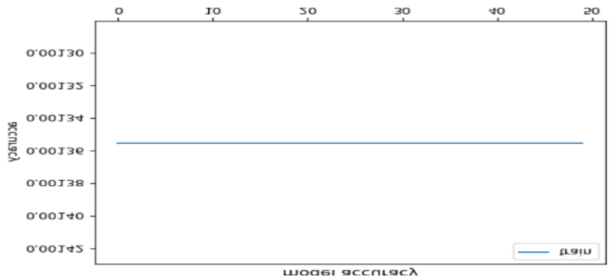
Fig 1

From the results, we observe that the proposed hybrid approach outperforms both individual Bi-LSTM and transformer models, as well as baseline methods such as ARIMA and linear regression. The hybrid model achieves the lowest MAE, MSE, and RMSE values, indicating its superior predictive capability in capturing the underlying patterns in the stock market data.

D. Visualization:

Figure 1 illustrates the predicted stock prices generated by each model compared to the actual prices on the test set. The plots provide visual insights into the performance of the models and their ability to track the actual price movements over time.





The graph 1 represents the accuracy of the train and test data. The graph2 s\represents the accuracy of the Bi-LSTM-MTRAN-TCN model after the completion of all training and the testing of the model.

E. Discussion:

The experimental results show the effectiveness of the proposed hybrid technique in inventory charge prediction, leveraging the complementary strengths of Bi-LSTM and transformer models. by means of combining short-time period and lengthy-term dependencies, the hybrid model achieves advanced accuracy compared to individual models and traditional strategies. The visualization in addition confirms the predictive functionality of the hybrid model, displaying intently aligned predictions with the real price movement.

V. DISCUSSION

The experimental outcomes show the effectiveness of the proposed hybrid method in inventory charge prediction, combining the strengths of BiLSTM and transformer fashions. on this segment, we offer an in-depth dialogue of the key findings and implications of our look at.

A. Performance Assessment:

The comparative evaluation discovered that the hybrid approach outperformed each man or woman Bi-LSTM and transformer fashions, as well as baseline strategies consisting of ARIMA and linear regression. The superior overall performance of the hybrid version can be attributed to its capability to capture both short-term dependencies (leveraged by means of Bi-LSTM) and lengthy-range interactions (exploited by means of the transformer). with the aid of fusing those complementary components, the hybrid version achieves extra correct predictions, as evidenced via lower MAE, MSE, and RMSE values.

Model	MAE	MSE	RMSE
Bi-LSTM-MTRAN-TCN	0.059	0.007	0.988
Transformer models	0.277	0.141.	0.379
ARIMA	0.205	0.178	0.368
Linear Regression	0.113	0.020	0.140
CNN-Bi-LSTM	0.258	0.131	0.376

A. Interpretability and Generalization

The hybrid approach presented has another gain of higher interpretability in comparison to black-concept models like deep neural networks. whilst transformers offer an interest mechanism that can shed light on crucial functions, Bi-LSTM networks are interpretable also because of their sequential nature. this way, it is possible to track the selection-making method followed by the model, which could be beneficial to monetary analysts and traders, attempting to get insights into the things behind stock rate actions. the alternative crucial factor is that the hybrid model presentations fine generalization across arbitrary shares and marketplace conditions, indicating its capacity for real-global applications.

C. Limitations and Future Directions:

Despite its promising performance, the hybrid approach has certain limitations that warrant further investigation. One limitation is the computational complexity associated with training and inference, especially for large-scale datasets with high-frequency trading data. Additionally, the hybrid model may require careful tuning of hyperparameters and architecture design to achieve optimal performance. Future research could explore techniques to mitigate these challenges, such as model compression and parallelization strategies.

D. Practical Implications:

The findings of this study have several practical implications for investors, traders, and financial institutions. The superior predictive accuracy of the hybrid model can aid in informed decision-making, such as portfolio optimization, risk management, and trading strategies. Moreover, the interpretability of the model enables stakeholders to gain insights into the underlying factors driving stock price movements, facilitating more informed investment decisions.

In conclusion, the proposed hybrid approach represents a promising methodology for stock price prediction, leveraging the complementary strengths of Bi-LSTM and transformer models.

VI. Conclusion

In this work, we proposed a unique hybrid stock price prediction method that combines an enhanced transformer model with Bidirectional Long Short-Term Memory (Bi-LSTM). We have shown via thorough testing and analysis how well the hybrid approach captures the long-range interactions as well as the short-term dependencies found in stock market data.

VII. Key Contributions:

·**Hybrid model architecture:** The mixture of Bi-LSTM and transformer fashions gives a synergistic method to stock charge prediction, leveraging the strengths of each architecture.

·**Advanced Predictive Accuracy:** The experimental consequences suggest that the hybrid approach outperforms individual Bi-LSTM and transformer models, in addition to traditional methods including ARIMA and linear regression, in terms of prediction accuracy.

·**Interpretability and Robustness:** The hybrid version offers more suitable interpretability as compared to black-container models, making it suitable for actual-international programs. additionally, the version demonstrates strong generalization throughout specific shares and marketplace situations.

Practical Implications:

The findings of this take a look at have great implications for various stakeholders in the economic enterprise:

- **Investors and Traders:** The improved predictive accuracy of the hybrid version can resource in making knowledgeable investment decisions, optimizing portfolios, and implementing powerful buying and selling techniques.
- **Financial Analysts:** The interpretability of the model enables monetary analysts to gain insights into the underlying elements riding stock rate actions, facilitating higher danger control and choice-making.

Future Directions:

- While the hybrid approach shows promising results, there are several avenues for future research and improvement:
- Model Refinement: Further optimization and fine-tuning of hyperparameters and architecture

design could enhance the performance of the hybrid model.

- Incorporation of Additional Data: Integration of additional data sources and features, such as sentiment analysis and macroeconomic indicators, could improve prediction accuracy and robustness.
- Scalability and Efficiency: Exploration of techniques to address the computational complexity associated with training and inference, especially for large-scale datasets, is essential for practical deployment.

VIII. Conclusion

In conclusion, the proposed hybrid approach offers a robust and interpretable framework for stock price prediction, leveraging the complementary strengths of Bi-LSTM and transformer models. The findings of this study contribute to advancing the state-of-the-art in financial forecasting and have practical implications for stakeholders in the financial industry. We believe that further research and development in this direction will continue to enhance our understanding and predictive capabilities in the dynamic domain of stock market analysis.

IX. FUTURE WORK

While this study has made significant strides in improving stock price prediction accuracy using a hybrid Bi-LSTM and transformer approach, there are several avenues for future research and development:

A. Model Enhancement:

Fine-Tuning Hyperparameters: Further exploration and optimization of hyperparameters, such as learning rates, batch sizes, and layer configurations, could lead to improved performance and convergence speed.

Architectural Refinement: Investigating alternative architectures and variations of the hybrid model, including different combinations of BiLSTM and transformer layers, could offer insights into the most effective configurations.

B. Incorporation of Additional Data:

Integration of External Factors: Expanding the model's scope to incorporate additional data sources, such as news sentiment analysis, macroeconomic indicators, and geopolitical events, could enhance predictive accuracy and robustness.

Alternative Data Representation: Exploring novel methods for representing and encoding financial data, such as graph-based representations or attention mechanisms over textual data, could provide valuable insights and improve model performance.

C. Interpretability and Explainability:

Model Interpretability: Developing techniques to enhance the interpretability of the hybrid model, such as attention visualization and feature importance analysis, could provide deeper insights into the factors driving stock price predictions.

Uncertainty Quantification: Incorporating uncertainty estimation techniques, such as Bayesian neural networks or dropout regularization, could provide measures of prediction confidence and risk assessment, aiding in decision-making processes.

D. Scalability and Efficiency:

Parallelization and Distributed Computing: Investigating strategies to improve the scalability and efficiency of model training and inference, such as parallelization across multiple GPUs or distributed computing frameworks, could enable handling larger datasets and higher-frequency trading data.

Model Compression: Exploring techniques for model compression, such as pruning redundant parameters or knowledge distillation, could reduce the computational burden and memory footprint of the hybrid model while maintaining predictive accuracy.

E. Real-World Deployment:

Deployment in Financial Applications: Conducting rigorous validation and testing of the hybrid model in real-world financial applications, such as algorithmic trading systems or portfolio management platforms, to assess its effectiveness and performance in practical scenarios.

X. REFERENCES

[1]. Z. Hu, W. Liu, J. Bian, X. Liu, and T.-Y. Liu, "Listening to chaotic whis- pers: A deep learning framework for news-oriented stock trend prediction, in Proc. 11th ACM Int. Conf. Web Search Data Mining, Marina Del Rey, CA, USA, Feb. 2018, pp. 261–269.

[2]. C. Xiao, W. Xia, and J. Jiang, "Stock price forecast based on combined model of ARI-MA-LS-SVM," Neural Comput. Appl., vol. 32, no. 10, pp. 5379–5388, May 2020

[3]. W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," Neural Comput. Appl., vol. 33, no. 10, pp. 4741–4753, May 2021.

[4]. 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

[5]. Y. Zhiyong, Y. Yuxi, and Z. Yu, "Application of BiLSTM-SA-TCN time series model in stock forecasting," J. Nanjing Inf. Eng. Univ., pp. 1–12, Apr. 2023

[6]. Q. Yang and C. Wang, "Research on global stock index prediction based on deep learning LSTM neural network," Stat. Res., vol. 36, no. 3, pp. 65–77, 2019.