A FORECAST FOR STOCK PRICE PREDICTION

A PROJECT REPORT

Submitted by

MEGHA VARSHINEE S J (210701156)

in partial fulfillment for the award of

the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

ANNA UNIVERSITY, CHENNAI

MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this Thesis titled **“A FORECAST FOR STOCK PRICE PREDICTION**” is the bonafide work of “**MEGHA VARSHINEE S J (2116210701156)”** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Dr . S Senthil Pandi M.E.,Ph.D.,

PROJECT COORDINATOR

Professor

Department of Computer Science and Engineering

Rajalakshmi Engineering College

Chennai - 602 105

Submitted to Project Viva-Voce Examination held on ________________

Internal Examiner **External Exam**

ABSTRACT:

In the financial sector, in order to increase the precision and dependability of stock price forecasts, the study presents a brand-new method called bidirectional long short term memory. By adding TCN (Temporary Revolution Network), this method enhances it and produces a new model called TCN that is especially designed for stock price prediction. The suggested solution completely exploits the capabilities of all three models: Transformer, bidirectional long short-term memory, and TCN by merging TCN with Bi-directional Long Short Term Memory. Transformers are great for gathering full-range distance data; however, TCN takes care of sequence dependencies, which makes the model more broadly applicable. Bidirectional information inside sequences is captured by Bi-directional Long Short Term Memory in the interim. The method offers a comprehensive framework for reliable stock price forecasting by integrating, which captures bidirectional sequence information, which is skilled at handling global relationships within the data. The efficacy of the suggested strategy is demonstrated by the experimental findings, which exhibit enhanced prediction stability and accuracy in comparison to previous techniques. Overall, a viable path forward for the development of stock price prediction methods in financial research and applications is the Bi-directional Long Short Term Memory approach.

Keywords: Bidirectional long short term memory, Temporary Revolution Network, deep learning, combined neural network and stock price prediction.

I. INTRODUCTION:

For a long time, forecasting stock prices has piqued the curiosity of both investors and scholars given its significant impact [1]. He significance of tracking stock price patterns is being increasingly recognized by investors [2]. Foreseeing future changes in stock prices is a significant task for economists [3–4], which is valuable to investors looking to maximize their profits. However, economists rarely succeed in doing so. Yet, the task is made more difficult by the complex and erratic character of stock market patterns, which are made worse by high volatility and the impact of random noise [5]. Predictability in financial time series is still vital despite these significant challenges. From the beginning, the ARIMA moving average autoregressive model has become one of the most well-known techniques [6].

Subsequently, Narendra et al. used the ARIMA model with GARCH, or the conditional heteroskedasticity model, has been used to examine data from the NSE Indian stock market [7]. The Kalman filter model and the Bayesian vector autoregression model have also been used. These approaches perform poorly in long term forecasting and are inappropriate for nonlinear problems, even though they are successful in short term forecasting [8].

As a potential solution to this problem, time series analysis using machine learning has been introduced and has shown to be effective in stock price forecasting [2], [9], and [10]. Machine learning has shown to be useful in managing complex and large-scale datasets.

Machine learning, a subset of artificial intelligence, empowers algorithms to learn from historical data and make predictions or decisions without being explicitly programmed. In the context of stock price prediction, machine learning algorithms analyze vast amounts of historical market data, including price movements, trading volumes, and other relevant factors, to identify patterns and trends that could indicate future price movements.

However, the task of predicting stock prices is inherently challenging due to the dynamic nature of financial markets, which are influenced by a multitude of factors, including economic indicators, geopolitical events, and investor sentiment. Moreover, stock prices exhibit inherent volatility and non-linearity, making accurate predictions elusive.

In this paper, we embark on a journey to explore the application of machine learning techniques in predicting stock prices. Through empirical analysis and experimentation, we seek to evaluate the efficacy of various algorithms, identify key features for prediction, and assess the practical implications for investors and financial professionals. Our quest is to unlock the secrets hidden within the data and illuminate the path towards more accurate and reliable stock price forecasts in the age of machine learning.

II. LITERATURE REVIEW:

Within the financial sector, stock price prediction is a crucial field of research where many different approaches and strategies are investigated to improve forecasting accuracy. The ARIMA (Autoregressive Integrated Moving Average) model, which has served as a platform for further study, is among the oldest and best-known methods. Constructing on this basis, scholars like Narendra et al. have explored the use of models like the ARIMA and autoregressive conditional heteroskedasticity model (GARCH) To forecast stock prices under specific market conditions, like the Indian stock market's NSE.

Though they have demonstrated potential, traditional time series models such as GARCH and ARIMA struggle to capture the complexities of long-term forecasting and nonlinear interactions. As a result, researchers are looking into a number of strategies, including machine learning-based ones The benefit of machine learning techniques is of being flexible in capturing nonlinear patterns and handling vast and complicated datasets with ease.

Interest in recurrent neural networks (RNNs) has grown due to its ability to model sequential data effectively. One variation that has been particularly successful in capturing temporal dependencies within sequences is bi-directional Long Short Term Memory. Furthermore, the Transformer model—which was first created for tasks related to natural language processing—has demonstrated potential in locating global relationships within sequential data because of its attention mechanism.

Recurrent Neural Networks (RNNs) and their specific variant, Long Short-Term Memory (LSTM) networks, have emerged as powerful tools. These models excel at handling the sequential nature of stock price data, allowing them to learn from historical trends and identify patterns within the data series. LSTMs, in particular, shine in capturing long-term dependencies within the data, making them especially well-suited for stock price forecasting [1]

Support Vector Machines (SVMs) also play a role, although their usual focus is classification tasks. In stock price prediction, SVMs can be effective in predicting the overall direction (upward or downward trend) by creating a clear separation between different price categories within a high-dimensional space [2].

Artificial Neural Networks (ANNs) offer another approach. These flexible models can learn complex, non-linear relationships between various features and the target variable, which in this case is the stock price. They can be used to predict both the actual price and the general price trend [2]. However, ANNs require careful tuning to avoid overfitting the data and achieve optimal performance.

III. PROPOSED SYSTEM:

The proposed system makes use of an improved converter and bidirectional long short-term memory technology for stock price prediction while utilizing a variety of machine learning techniques.

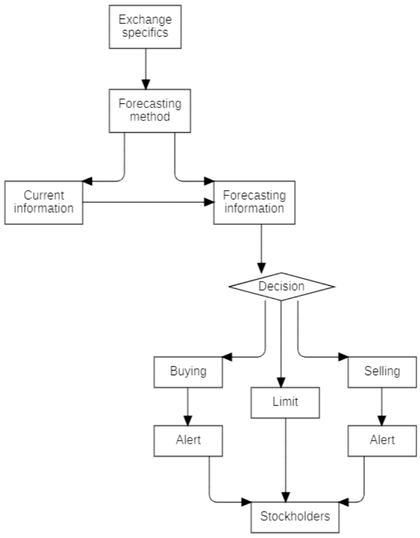

FIGURE 1: Stock price prediction.

The proposed stock price prediction system combines an improved transformer model with Bidirectional Long Short-Term Memory in an effort to improve predictive power via their joint integration. The system uses the modified for improved performance, capitalizing on bi-directional Long Short Term Memory ability to capture temporal dependencies in sequential data. The Z-score normalization technique minimizes feature discrepancies by ensuring standardized input representation during preprocessing. Before the data is subjected to feature extraction by Bi-directional Long Short Term Memory, the Positional Encoding Layer is utilized by the system's painstaking input processing to introduce temporal context into the data. Strong predictions are then produced by MTRAN-TCN, which further refines the sequence characteristics.

To analyze anticipated accuracy, the system makes use of a strong assessment framework and standard For evaluation, metrics like mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE), and R-squared (R2) are frequently employed. The method not only shows that it can accurately predict stock prices but also highlights areas that require additional research and development, highlighting the ongoing pursuit of improvement. Subsequent efforts might focus on optimizing neural network architecture and using multi-source data integration strategies, promoting a comprehensive strategy to enhance prediction capabilities in financial markets. Improvements in stock price prediction methods are made in an orderly and impartial manner by following a methodical approach.

IV. ALGORITHM INTRODUCTION:

A. TRANSFORMER:

In 2017, the Google team unveiled the Transformer, a well-known NLP model. The Transformer design is also the foundation for BERT, which is still in widespread use today. The Transformer, as opposed to sequential RNN structures, has a self-attention mechanism that makes parallelization easier and allows the model to capture global information. Notable improvements in NLP tasks have resulted from this break from conventional sequential processing paradigms.

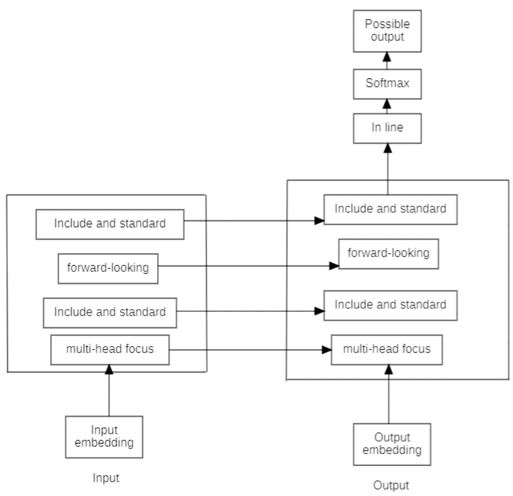

FIGURE 1. General structure of a transformer

B. TCN (TIME CONVOLUTIONAL NETWORK)

The Time Convolutional Network (TCN) is a unique technique for time series prediction that was proposed by Andrew et London in 2013. Remaining connections, dilation convolution, and causal convolution make up the majority of the TCN network topology. A causal convolution with a unidirectional structure is used in each convolution layer of a Time

Convolutional Network (TCN). Because the convolution at time T only works on elements that exist before time T, this design decision guarantees that information will not leak into the model in the future. Additionally, as long as the sequences match when the causal convolution is computed, the TCN may tolerate sequences of different lengths for both input and output.

F (s) = X k−1 i=0 f (i) xs−di

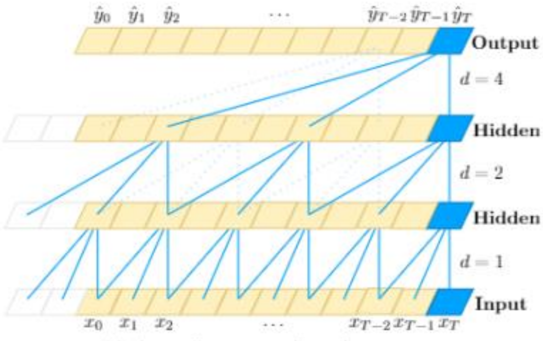

FIGURE 1. Transformer's overall structure was altered.

V. METHODOLOGY:

We present a thorough discussion of the improvements made to the transformer model to better its performance before diving into the recommended method's network design.

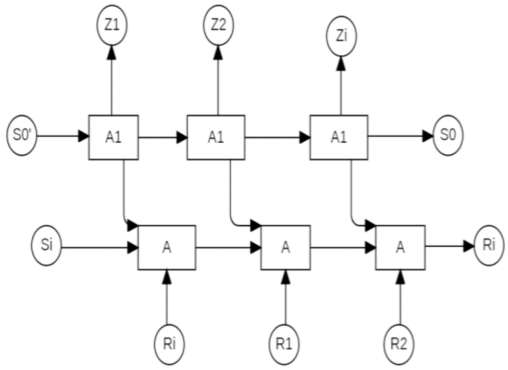

FIGURE 1: Diagram of the bidirectional long short term memory.

To improve stock price prediction, adjust the transformer model, primarily by changing the transformer model's decoder.

- The first input embedding module is removed because it is not required for processing stock price data. It is in charge of vectorizing language and text. When text needs to be converted into vectors for a task like machine language translation, this module is usually necessary.

- It is recommended that the Position Encoding module be moved to the front of the BiLSTM, adjusted appropriately, and removed from the MTRAN-TCN design.

- The encoder's output is the only input that the decoder uses now that all other decoder inputs have been eliminated.

A. NETWORK STRUCTURE:

The transformer model, which serves as the study's basis, is improved with TCN to better fit the forecast of stock series data. Although the transformer paradigm is very good at catching global signals and facilitating parallel processing, it is not very effective at gathering sequence information, which makes its direct application to stock prediction flawed.

SNO	COMPANY NAME	CODE
1	BUSINESS BANK	JVIEWN
2	SUPREME PETROL	NJIQOS
3	ISDA LTD	OMWDK
4	NAVEEN LTD	MXOWK
5	SHALU MEDICARE	JXNDIW

TABLE 1. A few index stocks.

The dataset part includes data on technical indexes and stock trading, which could be used as input variables. The output line up with the next day's closing price. A 3D tensor format is used to organize the input data, These are features, time intervals, and examples. The transformer's encoder performs additional processing on the sequence characteristics that the bidirectional long short term memory has captured after they have passed through the Positional Encoding layer for input processing. A bidirectional long short term memory ability to capture sequence information and successfully influence each time step's output depends on both the input at that particular time step and the previous memory. However, the wide time range may cause problems for the bidirectional long short term memory during processing when working with lengthy time series data, which is common in stock prediction analysis.

VI. EXPERIMENT ENVIRONMENT:

We present the experimental dataset, evaluation indices, and experimental parameters in this section.

A. DATASET:

Stock price datasets are available from a variety of sources, both for free and for a fee. Here are a few well-known ones: The chosen index stocks are displayed in Table 1. The dataset part includes data on technical indexes and stock trading, which could be used as input variables.

Category	**Large-cap stock**	**Small-cap stock**
Finance	Canara bank(0.01.Cr)	Fink Financial (8.8)
Steel	PI industries (0.01, Cr)	Kiri industries (6.7)
Coal	Hinda industries (0.5. Cr)	SE power (7.5)
Automative	Tata Motors (0.2. Cr)	Piso digital(5.4)
Estate	Poly develop (0.1 Cr)	Urban investment (6.2)
Petrochemical	UPL (0.2 Cr)	IPC (7.2)
Metal	Copper industries (0.2 Cr)	Lithium industries (4.5)

TABLE 2. Chosen stocks from the marketplaces in India.

Since every attribute in the stock dataset varies greatly from the others, standardization is necessary. The Z-score is the normalization method used in this article. The notations used in this equation are y for the normalized value, x for the input data, x for the mean of the input information, and s for the standard variation of the input data.

B. PERFORMANCE EVALUATION:

The methods' assessment criteria include R-square (R2), mean absolute error (MAE), root mean square error (RMSE), and mean square error (MSE). The following procedure is used to calculate these error evaluation indices:

$$\left.\begin{aligned} \mathrm{MSE} &= \frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \\ \mathrm{MAE} &= \frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \\ \mathrm{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \\ \mathrm{R}^2 &= 1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(\bar{y}-\hat{y}_i)^2} \end{aligned}\right\}$$

Parameters	**Value**
Batch size	5
Sequence length of training data	10
Hidden Size	64
No of head	3
No of layer	8
TCN layer	32
Kernal size	7

TABLE 3. The bidirectional long short term memory technique’s parameter settings.

Method	MAE	MSE	RMSE	R^2
TRAN	0.284	0.224	0.438	0.778
MTRAN	0.175	0.056	0.229	0.945
MTRAN-TCN	0.173	0.055	0.231	0.945
BiLSTM	0.087	0.014	0.118	0.986
BiLSTM-MTRAN	0.136	0.035	0.182	0.965
BiLSTM-TCN	0.104	0.020	0.138	0.981
BiLSTM-MTRAN-TCN	0.147	0.040	0.195	0.960

TABLE 4. Techniques are compared.

VII. RESULT:

A. INDEX STOCKS:

The LSTM strategy is compared and analyzed in the paper with the five previously described strategies using index stocks and Shanghai and Shenzhen equities. This extension of the experiment's scope helps to accentuate how much better the recommended approach is than the others. When compared to alternative techniques, bidirectional long short term memory raises R2 by 1.25% to 11.7% and decrease error by 25.22%.

METHOD	BYD			DONG MOTOR		
	MAE	MSE	R^2	MAE	MSE	R^2
BiLSTM	0.08	0.01	0.98	0.08	0.01	0.98
CNN-BiLSTM	0.27	0.13	0.85	0.25	0.13	0.86
CNN-BiLSTM-AM	0.25	0.13	0.86	0.25	0.13	0.86
BiLSTM-SA-TCN	0.11	0.01	0.98	0.12	0.02	0.97
BiLSTM-MTRAN-TCN	0.06	0.08	0.99	0.04	0.06	0.99

TABLE 5. Comparison of BYD and Dong Motor's evaluation indices using various techniques

B. VALIDATION OF GENERALIZATION ABILITY:

Using the trial data from the 12 Shanghai and Shenzhen stocks mentioned above, thoroughly analyze each method's accuracy and generalizability. Notably, there are no outliers in the R2 values for any of the six techniques, which are all centered between 0.7 and 1.1. The bidirectional long short term memory is the one with the largest concentration of R2 values among them, with the maximum concentration of R2 reaching 0.973.

C. VALIDATION OF TIMELINESS ISSUES:

In order to determine whether there is a timeliness issue, experiments were carried out on the five index stocks listed in Table 1. Stock data from four distinct periods were employed. That is, to confirm the method's stability in the temporal dimension.

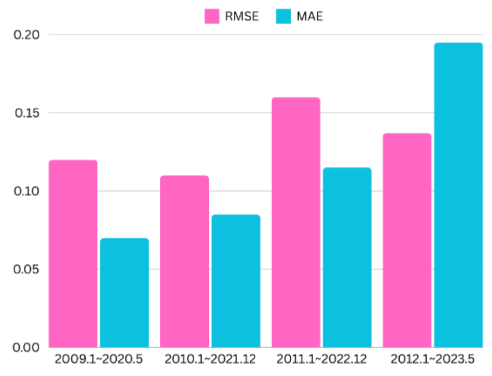

FIGURE 1. Value of the error indicator for every time interval.

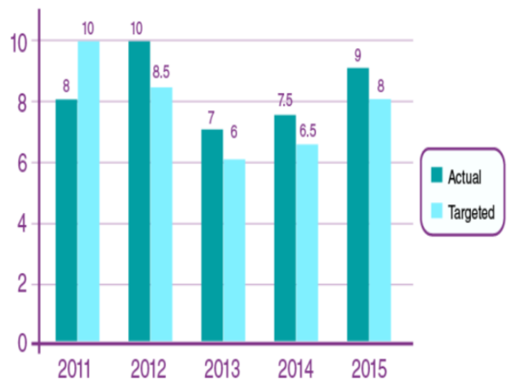

FIGURE 2. The R2-value of several equities over various intervals

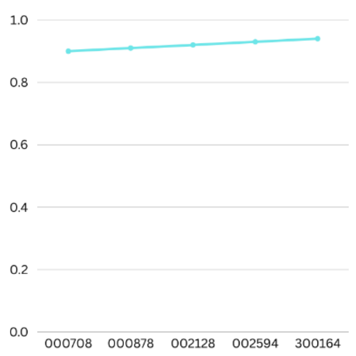

FIGURE 3. The R2 for given stock prices

VIII. CONCLUSION:

This article's recommended bidirectional long short term memory strategy is used to predict stock closing prices. By doing away with By removing unnecessary inputs, replacing the original decoder component with a fully connected layer and TCN layer, and treating the encoder's output as the only input for the decoder, this method modifies the transformer model . The data is first processed by bidirectional long short term memory to extract The Position Encoding Layer processes sequential signals before sending them to the modified transformer for further processing. The effectiveness of incorporating bidirectional long short term memory, the transformer's improvement effect, the method's correctness, its generalizability, and timeliness issues were all carefully examined through testing in this work, Tests have demonstrated that the improved transformer needs to be used in conjunction with bidirectional long short term memory in order to get the best results. Moreover, stock prediction is significantly affected by the addition of bidirectional long short term memory. Additionally, our method tackles a number of important stock prediction issues, such as identifying sequence-dependent signals, improving the transformer's efficiency, and guaranteeing the timeliness, correctness, and generalizability of the approach. We have successfully verified our method's ability to predict stock closing prices through extensive testing. In summary, our findings highlight the significance of utilizing sophisticated deep learning methods like transformer models and bidirectional long short term memory to enhance the precision of stock price forecasting. We present a viable approach to improve predictive capacities in financial markets, assisting analysts and investors in making wise judgments, by merging these techniques and optimizing their structures.

IX. DICUSSION:

Nonetheless, this technique could improve the accuracy of certain stock predictions even more. The following two areas will be the focus of future research:

- Improving the neural network's architecture even more.
- Combining information from several sources, including fundamental, index, and stock price data, to enhance forecasts.

X. REFERENCES:

[1] A. Mu, V. Lin R. Vian, Y. Lin, and D.-F. Lin, ‘‘Using a machine to anticipate stock prices learning and deep learning models based on LSTMs’’ in Proc. 12th ACM Int. Conf.

[2] S. Lin, K. Lang, and P. Ruang, ‘‘A machine learning-based approach to predict customer behavior,’’ Tech Vignesh-Tech. Gazette, vol. 26, 2019.

[3] B. Xiyag, R. Xiya, and L. Siang, ‘‘Forecast of stock price using a mixed model’’ Neural Compute. Apt., vol. 30, no. 12, May 2016.

[4] W. Yug and H. Yang, ‘‘Making stock price predictions with deep neural networks,’’ Neural Compute. Appt., vol. 22, no. 7, pp, Mar. 2020.

[5] Z. Shang, T. Qing, R. Shan, M. Bao, X. Zang, and L. Lin, ‘‘Transformer-based awareness network for forecasting stock movement,’’ Exp. Syst. Appt.,vol. 52, Oct. 2017, Art. no. 1179.

[6] M. K. Kishan and K. M. Rabir, Bi spectral Analysis and Bilinear Time Series Models: An Overview London, USA: Springer, 2010.

[7] R. M. Banu and R. S. Reddy, ‘‘An ARIMA-GARCH model is performance of a chosen Indian stock by partitioning and interpolation.,’’ Appt. Compute. In format., vol. 17, no. 1, pp. 132–133, Jul. 2013.

[8] M. Dong, S. Zhang, and N.Mang, ‘‘NTOCA-AO-SVE model for stock price forecasting,’’ Compute. Eng. Appt., vol.48, no. 8, pp., 2021.

[9] S. Lin, Y. Sui, and M. Kong, ‘‘Temporal hypergraph convolutional neural network-based approach for stock trend prediction,’’ J. Compute. Appt., vol. 44, no. 3, pp. 2021.

[10] L. Zhang, D. Kang, and U. Vu, ‘‘Stock index forecasting using the optimized KEM model,’’ Statist. Decision., vol. 37, no. 13, pp.,2020.