2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)
21-22 February, 2024, Mazandaran University of Science and Technology, Babol, Iran

2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP) | 979-8-3503-8394-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/AISP61396.2024.10475263

A CNN-LSTM deep neural network with technical indicators and sentiment analysis for stock price forecastings

Fatemeh Moodi
Department of Computer Engineering
Yazd University
Yazd, Iran
0000-0003-3240-5910

Amir Jahangard-Rafsanjani
Department of Computer Engineering
Yazd University
Yazd, Iran
0000-0003-2638-5722

Sajad Zarifzadeh
Department of Computer Engineering
Yazd University
Yazd, Iran
0000-0001-5627-0542

Abstract— **The utilization of social media platforms, such as Twitter, has experienced a notable surge. Consequently, the sentiments expressed by both influential figures and the general public have come to exert a significant influence on public opinion. This influence extends to various domains, including the stock market, where traders are swayed by the ideas and posts disseminated through these channels. Hence, the present study seeks to improving accuracy in stock price for castings with sentiment analysis and utilization of technical indicators, thus an CNN-LSTM framework is introduced to forecast stock market with sentiment analysis and technical indicators. Three deep learning models along with technical indicators and sentiment analysis are used to forecast stock prices. The impact of time windows of different sizes; 1, 7, 14, 21, 30, 45 and 60 was compared on the performance of the model, and the best time window was selected to forecast stock prices. 4 error-based evaluation criteria have calculated. The best model to forecast the stock price was the CNN-LSTM model. The MSE of the model was 0.0011.**

Keywords— ***LSTM; CNN; deep neural networks; gated recurrent unit; technical indicators; sentiment analysis.***

I. Introduction

Stock price forecasting helps to greatly reduce the risks of an investment with analysis. A common method that most traders use to forecast the stock market is to use technical indicators. Considering the long history of stock forecasting, most studies attempts have been made to utilize indicators as a means of assessing the accuracy of stock forecasting [1]. A comprehensive investigation, which encompassed 57 articles that were published in prestigious scholarly journals, revealed that the application of machine learning methodologies to forecast stock prices remains a subject of great contention within the realm of academic research [2]. The present paper aims to elucidate the profound influence of social media on financial markets. In recent times, several deep learning methodologies have surfaced across multiple domains due to the rapid progress of technology and computational capabilities. These methodologies have been widely implemented in the field of finance, specifically in the prediction of stock market trends, optimization of investment portfolios, management of risks, and development of trading strategies. The task of stock indices prediction with noisy data is both intricate and challenging, yet it holds significant importance in determining the opportune moments for buying or selling stocks. This influence extends to various domains, including the stock market, where traders are swayed by the ideas and posts disseminated through these channels. Hence, the present study seeks to improving accuracy in stock price for castings with sentiment analysis and utilization of technical indicators, thus an CNN-LSTM framework is introduced to predict stock market with sentiment analysis and technical indicators. Undoubtedly, this area of finance is highly sought-after and invaluable.

II. Related Work

A. *Long short-term memory (LSTM)*

These RNN models incorporate gates that retain information from previous inputs, enabling them to effectively manage sequential data [3]. However, despite their theoretical capability to handle past sequential data, RNNs encounter practical challenges when it comes to long-term memory. They are limited in their ability to look back only a few steps, rendering them inadequate for processing lengthy sequential data. RNNs face issues such as the vanishing and exploding gradient. The vanishing gradient problem arises when standard RNNs struggle to retain information over extended time intervals during retraining. On the other hand, the exploding gradient problem occurs when the training algorithm assigns excessively high values to matrix weights without any logical basis. To overcome these challenges, researchers have developed RNN models such as the LSTM model, which possesses the ability to maintain both long and short-term memory capabilities.

B. *Gated Recurent Neural Network (GRU)*

Gated Neural Networks have been successfully applied for consecutive or time data [4]. To detect speech, processing natural language, and more appropriate machine translation, along with LSTM in the areas of long sequences, have done well. The gate was intended in the LSTM topic and includes a gate network that produces signals. The gates themselves are weighed and selectively updated according to an algorithm during the learning stage. Gate networks add the computing cost to increased complexity and consequently parametering.

C. CNN-LSTM

The CNN-LSTM model is a specially designed LSTM architecture for solving sequence prediction problems involving spatial inputs, such as images or videos [5]. This innovative CNN-LSTM architecture combines convolutional neural network layers (CNN) to extract the features of the input data, which are then utilized by the LSTM to accurately predict sequences. The development of the CNN-LSTM model was driven by the need to tackle the challenges associated with forecasting visual time series and utilizing textual descriptions derived from a sequence of images [6].

The stock data is characterized by its highly intricate nature and vast volume. Chen introduces a CNN-GRU-attention model for predicting medium to long-term trends in stock prices [7]. The model employs three distinct data decomposition techniques, namely EMD, EEMD, and CEEMDAN, for data preprocessing and selects the most optimal approach. Upon comparative analysis, CEEMDAN was identified as the optimal data preprocessing technique. Ultimately, the CNN-GRU-Attention model exhibited superior alignment with actual values in terms of forecasting and significantly enhanced the quality and effectiveness of predictions. By evaluating performance with MAE, RMSE and R, it was discovered that the CNN-GRU-attention model exhibits the highest level of prediction accuracy.

Lin et al. present three steps for stock investment [8]. The selection of stocks is predicated upon a distinctive ratio, namely the division of forward PE by trailing PE. This ratio affords a superior means of evaluating the growth potential of individual stocks. Two stock market forecasting models, founded upon deep learning, are proposed to forecast future trends: LSTM and GRU models. These models are separately employed to prognosticate the prospective trajectory of stock prices, leveraging historical price data. Empirical evidence demonstrates that the GRU model enhances prediction accuracy and minimizes time lag in comparison to the consequences of the LSTM model.

By employing denoising methods on a range of wavelet basis functions and identifying the most optimal feature from the set of features, Gang Ji and et al. proposed a technique to enhance technical indicators [9]. Within this investigation, denoised stock price data was utilized to more effectively compute the technical indicators. The dataset encompassed four stock markets, namely, the SSEC stock market, Hong Kong, the S&P 500, and the DJI of the US stock market. A total of 18 technical indicators were examined in this study. Time windows of varying sizes were employed, specifically 3, 5, 10, 15, 30, 45, and 60, with the objective of forecasting stock prices three days in advance. The findings demonstrated an enhancement in model accuracy subsequent to the denoising procedure.

Patil enhanced the precision of stock forecasting, here a highly effective technique has been employed [10]. This methodology encompasses the amalgamation of two network classifiers, namely Rider Deep LSTM and Deep RNN, which are trained utilizing a novel algorithm referred to as SCSO. The dataset utilized in this study comprises the stock markets of two distinct companies, acquired between the time span of January 1, 2019 and April 1, 2021. Following the extraction of features, a feature vector is created by combining the extracted attributes. The wrapper approach is employed for the selection of features. The prediction outcome with the least error value is identified as the optimal solution.

Song and Choi put forward three hybrid models based on Recurrent Neural Networks (RNN) [11]. The aim is forecasts for the closing price of three stock market DAX, DOW, and S&P500 indices. The results revealed that the proposed models displayed a remarkable outperformance over the benchmark models. The results is evaluated by MSE and MAE metrics. Interestingly, the proposed ensemble model demonstrated comparable performance to the GRU model, which exhibited superior performance among the benchmark models and outperformed them in numerous instances. Furthermore, they introduced a fresh feature known as "medium". The findings indicated that the incorporation of this novel feature enhanced the overall performance of the models.

This paper presents a concise examination of several existing approaches through which a retail investor may forecast the price of a stock [12]. The price is subject to fluctuations, either upward or downward, contingent upon factors such as quarterly results, the influx of financial news, the technical behavior of the market, or the market sentiment as influenced by the global scenario in recent days. The authors extensively discussed various methodologies for projecting the movement of stock prices and assessed the accuracy proposed by the respective authors of these methodologies in their own publications. The methodologies are founded upon the utilization of LSTM (Long Short-Term Memory) along with other pertinent attributes.

Naik and Mohan employed a total of 33 distinct combinations of technical indicators to make predictions regarding stocks for short-term transactions [13]. Initially, relevant features were selected using the Borota feature selection technique and subsequently utilized to forecast stock prices via the ANN Regression forecast (Artificial Neural Network). The employed ANN consisted of three layers and was accompanied by the sigmoid function as its activation function. A threshold value of 0.5 was also taken into account. The parameters of gradient descent momentum were examined to ascertain weight and minimum global reduction. Within this investigation, the evaluation was conducted based on the criteria of the average absolute error (MAE) and the square root of the average squares (RMSE), yielding results that surpass the existing method by 12%.

Moodi et al. had analyses the performance of different regression methods for predicting stock prices based on technical analysis indicators [14]. Specifically, they explored two sequential forward and backward selection methods with 10 estimators using the last 13 years of Apple's dataset, which consisted of 123 technical indicators. The researchers employed 5 error-based evaluation criteria to assess the effectiveness of machine learning models. These evaluation criteria were then used to investigate the results. According to their findings, the proposed method, MLPSF, exhibited a 56/47% superior performance compared to MLP. Additionally, SVRSF showed a 67/42% improvement relative to SVR. Furthermore, LRSF demonstrated a notable 76.7% improvement compared to LR. The researchers also observed that different machine learning methods selected different features based on various evaluation parameters.

Ying Xu and colleagues presented two novel clustering-based prediction models based on two-stage prediction models [15]. The direct forecasting of stock closing prices enables the prediction of stock trend; thus this research aims to forecasting the closing stock price for 1, 5, 10, 20, and 30 days. Subsequently, the K-Means algorithm [16] is employed on a subset of ten technical indicators. Based on the Silhouette coefficient, which assesses clustering coherence and separation, the best and worst clustering results are determined. By dividing the initial data samples, the stocks are allocated to two distinct clusters. The Bagging ensemble learning algorithm [17] is implemented to develop a new prediction model named ensemble learning-support vector regression and random forest (ESVR&RF) in order to predict the stock closing price n days in advance. Finally, the K-Means algorithm and ensemble learning algorithm are merged to enhance the accuracy of stock price forecasting. The dataset employed in this study consists of four Chinese stocks, spanning from 2008 to 2019. Four evaluation criteria, namely MAPE, MAE, RMSE, and MSE, are utilized. The experimental results demonstrate that the combined forecasting model achieves the highest overall forecasting accuracy for financial stock forecasting.

III. METODOLOGY

In this study, 3 deep learning models, LSTM, CNN-LSTM and GRU models, are used to forecast stock prices. The goal here is to forecast stock price with the least error. The proposed flowchart to forecast stock price is in Figure 1.

Initially, Tesla's stock data is collected from Yahoo finance. Then the technical indicators and attributes required are calculated from the tweet data set. Next, the preprocessing of the collected data is done to eliminate the lost values and replace it with the average data available. The data is then normalized by the minimum (min-max); To become a common scale. We use

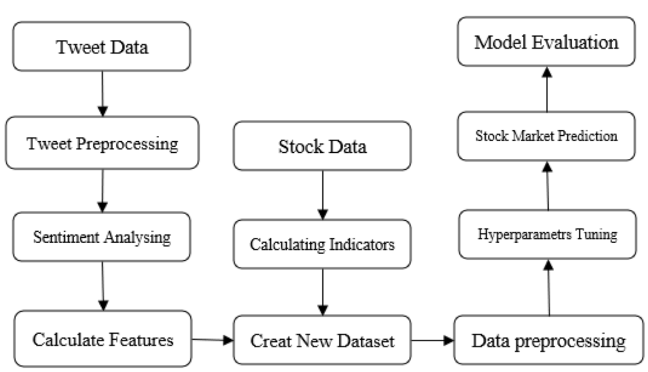

Fig. 1. Flowchart of suggested method for stock price forcasting.

the 8:2 ratio to divide the data set into the training set and testing set, and 20% of the training set for validation. The time window serves as a mechanism to transform a sequence of forecasting challenges into a supervised learning problem. This concept is analogous to the approach presented in article [14], where a window of variable size is employed in the dataset. The impact of time windows of different sizes; 1, 7, 14, 21, 30, 45 and 60 was compared on the performance of the models.

In this study, we have employed various specific attributes derived from the analysis of sentiments conducted by the Vader. Initially, we counted the number of positive, negative, and neutral comments present in the daily tweets data set, which were subsequently assigned to the attributes "Count_positive", "Count_neutral" and "Count_negative". Additionally, we used “score_mean_positive”, “score_mean_neutral”, and “score_mean_negative”, "tweet volume" and "positive tweet ratio to negative tweets" features.

A. *Data set*

Tesla's stock data was used in this article from 01/01/2015 to 31/12/2019. The price of closure as the dependent variable and the rest of the information (open price, closed, top, low, volume and adjusted) was removed from the dataset after calculating the technical indicators. This article uses all technical indicators for stock price forecasting [14]. The tweet data set used in total comprises 3,717,964 unique tweets [18]. 1,096,868 tweets are about Tesla.

B. *Estimating the sentiment of each tweet*

Manually analyzing sentiments is a sensitive and sometimes prone to error process. So it is not surprising that many NLP researchers strongly rely on existing dictionaries as primary sources. Stock sentiments can be measured using a company's tweets. Here is a NLTK called Vader [19]. The Vader uses a vocabulary-based approach to assign sentimental privileges to words in a text, taking into account the context and intensity of

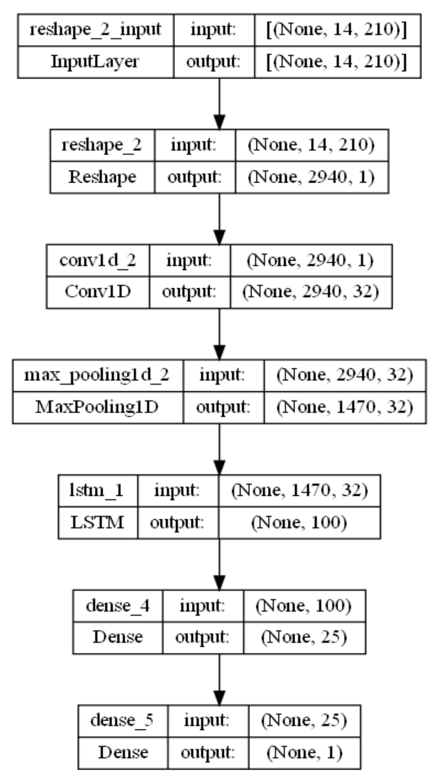

Fig. 2. Structure of the CNN-LSTM model.

feeling. One of the advantages of this method is the ability to determine positive and negative privileges for each tweet. This allows us to get a more accurate understanding of users' feelings in tweets. For example, if a tweet contains positive words such as "great", Vader recognizes these sentiments as positive sentiments. Likewise, if there are negative words such as "bad" or "disappointing" in tweets, it recognizes these sentiments as negative.

The architecture of the LSTM, GRU and CNN-LSTM models are show in Figure 2 to 4.

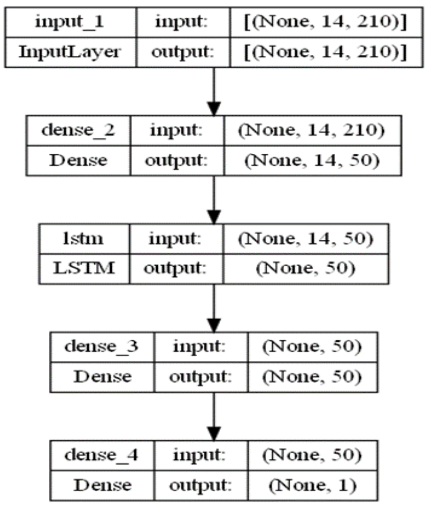

Fig. 3. Structure of the LSTM model.

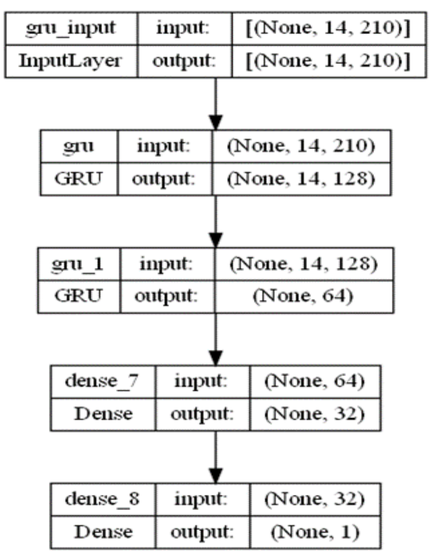

Fig. 4. Structure of the GRU model.

Based on Figure 2 to 4, LSTM architecture is made up of 5 layers, GRUs of 5 layers, and CNN-LSTM is composed of 7 layers. To evaluate the performance of a model, a measure called loss is employed. Loss serves to quantify the error generated by the model.

The evaluation of the alignment between a deep learning model and the training data is done using a metric called training loss. Put differently, it measures the extent of error exhibited by the model on the training set. In contrast, the validation loss metric is utilized to gauge the deep learning model's performance on the validation set. The validation set is a subset of the dataset that is specifically reserved for validating the model's performance. Similar to the training loss, the validation loss is calculated by summing up the errors for each example in the validation set.

The chart of the training loss against the validation loss for the data set during the number of epochs is shown for the CNN-LSTM model in Figure 5.

Figure 5 shows good fit for CNN-LSTM model, as the training loss and validation loss are reduced and stabilized at a particular point. The results show that as the number of epochs progresses, the levels of training and the loss of validation are willing to decrease and are reduced to about zero.

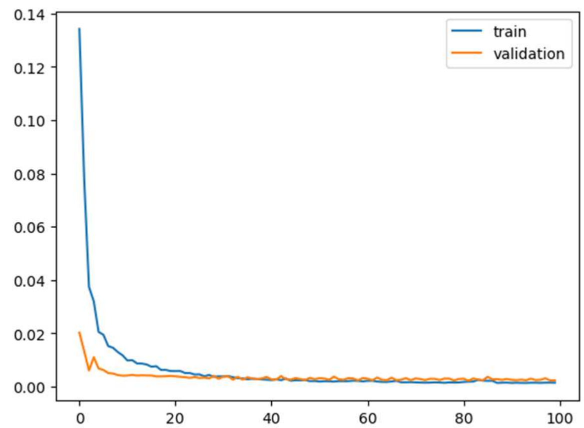

Fig. 5. Comparison of Training Loss and validation loss for CNN-LSTM Model.

C. *Evaluation criteria*

The evaluation criteria utilized in this article include MAE, MAPE, RMSE, and MSE as formula 1 to 4.

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \quad (1)$$

$$RMSE = \sqrt{MSE(y,\hat{y})} \quad (2)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i| \quad (3)$$

$$MAPE = \frac{100}{N}\sum_{i=1}^{N} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (4)$$

IV. Results

The outcomes of the 3 models employed in this investigation are presented in Table 1. Forecasting results were reviewed with different time windows and the best model for forecasting the stock price had the CNN-LSTM with a 45-day window. Also, the LSTM with the 1-day window had the best performance.

TABLE I. THE STOCK PRICE FORCASTING BY 3 MODELS

MODEL	TW	MSE	MAPE	MAE	R2
LSTM	**1**	**0.0016**	**0.0752**	**0.0313**	**0.9527**
	7	0.0024	0.0804	0.0348	0.9283
	14	0.0029	0.1150	0.0447	0.9094
	21	0.0028	0.1166	0.0423	0.9148
	30	0.0020	0.0817	0.0325	0.9390
	45	0.0023	0.1021	0.0381	0.9352
	60	0.0085	0.1798	0.0773	0.7739
CNN-LSTM	1	0.0021	0.0870	0.0338	0.9381
	7	0.0018	0.0741	0.0306	0.9463
	14	0.0022	0.0802	0.0334	0.9317
	21	0.0027	0.0964	0.0383	0.9186
	30	0.0029	0.0860	0.0365	0.9133
	45	**0.0011**	**0.0661**	**0.0252**	**0.9674**
	60	0.0034	0.0913	0.0388	0.9102
GRU	1	0.0027	0.0776	0.0359	0.9174
	7	**0.0021**	**0.0826**	**0.0346**	**0.9362**
	14	0.0025	0.0900	0.0378	0.9209
	21	0.0136	0.2667	0.1041	0.5879
	30	0.0028	0.1013	0.0421	0.9153
	45	0.0035	0.1053	0.0427	0.8995
	60	0.0047	0.1563	0.0593	0.8740

The results of the stock price forecasting by 3 models, on the test data, are show in figures 6,7 and 8, respectively.

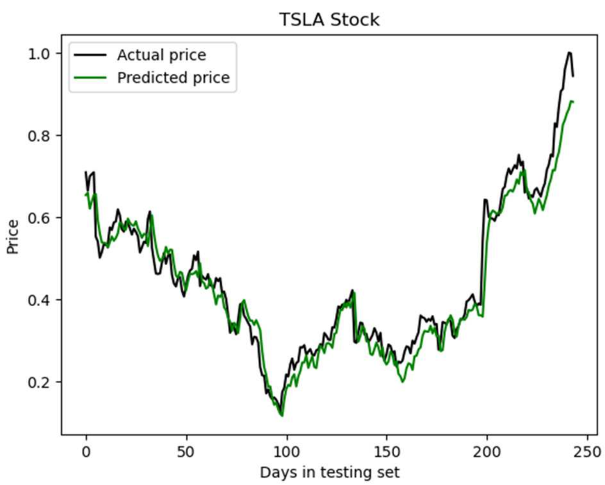

Fig. 6. Stock price forcasting by GRU

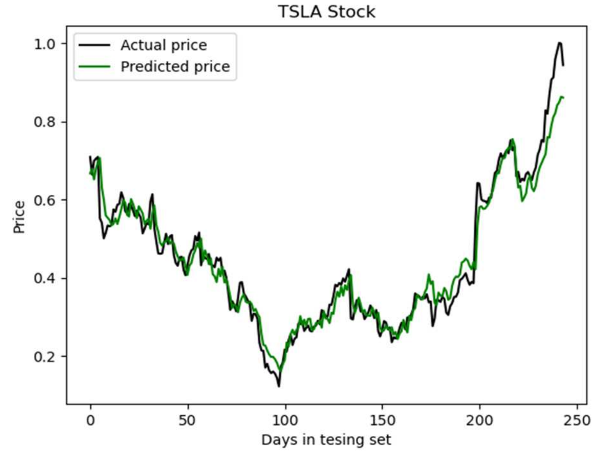

Fig. 7. Stock price forcasting by CNN-LSTM.

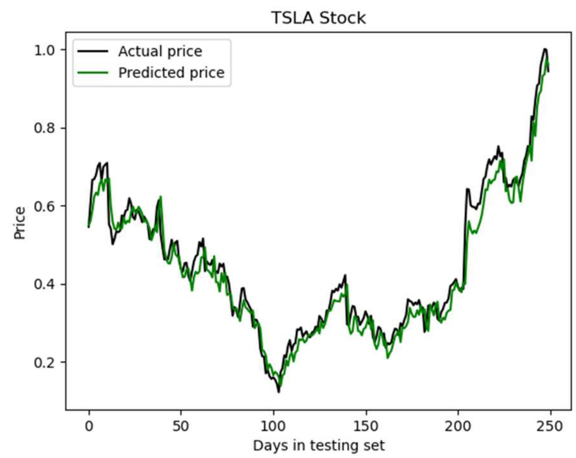

Fig. 8. Stock price forcastiyng by LSTM.

Based on figures 6, 7 and 8, stock price forecasting have been made by 3 models with least error, indicating the impact of sentiment analysis of tweets and technical indicators for stock price forecasting.

Comparison of the models with technical indicators and the sentiment analysis, with the related work is shown in Table 2. According to the results, the optimized models, along with the features used in this study, have reduced the prediction error and have made the stock price forecast more accurately.

TABLE II. COMPARISON OF RESULTS WITH EXISTING WORK

MODEL	TW	MSE	RMSE	MAPE	MAE
CNN-LSTM[11]	21	0.0078	-	-	0.0742
CNN-LSTM	21	**0.0027**	**0.052**	**0.0964**	**0.0383**
ANN [13]	-	-	-	-	15/1221
CNN-LSTM ATTENTION [7]	-	-	1.59	-	1.04

GRU[7]	-	-	1.60	-	1.08
GRU	7	**0.0021**	**0.046**	**0.0826**	**0.0346**
LSTM[7]	-	-	2.30	-	1.67
LSTM	1	**0.0016**	**0.04**	**0.0752**	**0.0313**

V. DISCUSSION AND FUTURE WORK

Predicting stock prices is a complex task that is affected by several factors. Researchers have been able to accurately predict the direction and amount of price movement by analyzing the feelings of tweets and the use of deep neural networks. The power of tweet sentiments lies in their ability to reflect the market feelings and opinions of the market, but they should not be the only basis for investment strategies. Vader was used to estimate the feelings of tweets. The effects of time windows of different size were compared to the performance of the model, and through the experiments we showed that the appropriate time window size exerts a favorable impact on the performance of the model. In this article, the size window size was set in 1, 7, 14, 21, 30, 45 and 60 days, respectively, and the best time window was selected to predict stock prices. The best model to forecast the stock price was the CNN-LSTM model. In general, the enhanced CNN-LSTM model in this article significantly improves prediction accuracy.

In the future, we shall construct a more resilient framework employing methodologies rooted in deep learning.

REFRENCES

[1] Murphy, J. J. (1999). Technical analysis of the financial markets: A comprehensive guide to trading methods and applications. Penguin.

[2] Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. Expert Systems with Applications. 124. 226-251.

[3] Rezaei, H., Faaljou, H., & Mansourfar, G. (2021). Stock price prediction using deep learning and frequency decomposition. Expert Systems with Applications. 169. 114332.

[4] Dey, R., & Salem, F. M. (2017, August). Gate-variants of gated recurrent unit (GRU) neural networks. In 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS). 1597-1600.

[5] Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., & Darrell, T. (2015). Long-term recurrent convolutional networks for visual recognition and description. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2625-2634.

[6] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555.

[7] Chen, C. (2023, May). Stock Price Prediction Based on the Fusion of CNN-GRU Combined Neural Network and Attention Mechanism. In 2023 6th International Conference on Electronics Technology (ICET). 1166-1170.

[8] Lin, Z., Tian, F., & Zhang, W. (2022, December). Evaluation and Analysis of an LSTM and GRU Based Stock Investment Strategy. In 2022 International Conference on Economics, Smart Finance and Contemporary Trade (ESFCT 2022). 1615-1626.

[9] Ji, G., Yu, J., Hu, K., Xie, J., & Ji, X. (2022). An adaptive feature selection schema using improved technical indicators for predicting stock price movements. Expert Systems with Applications. 200. 116941.

[10] Patil, P. R., Parasar, D., & Charhate, S. (2023). Wrapper-Based Feature Selection and Optimization-Enabled Hybrid Deep Learning Framework for Stock Market Prediction. International Journal of Information Technology & Decision Making. 1-26.

[11] Song, H., & Choi, H. (2023). Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models. Applied Sciences. 13(7). 4644.

[12] Vishwakarma, V. K., & Bhosale, N. (2022, April). Stock Price Prediction using LSTM: An advanced review. In Proceedings of the International Conference on Innovative Computing & Communication (ICICC).

[13] Naik, N., & Mohan, B. R. (2019). Optimal feature selection of technical indicator and stock prediction using machine learning technique. In Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics: Second International Conference, ICETCE 2019. 261-268.

[14] Moodi, F., Jahangard-Rafsanjani, A, Zarifzadeh, S. (2023). Evaluation of feature selection performance for identification of best effective technical indicators on stock market price prediction. *Preprint*. https://doi.org/10.48550/arXiv.2310.09903

[15] Xu, Y., Yang, C., Peng, S., & Nojima, Y. (2020). A hybrid two-stage financial stock forecasting algorithm based on clustering and ensemble learning. Applied Intelligence. 50. 3852-3867.

[16] Moodi, F., Saadatfar, H. An improved K-means algorithm for big data. IET Software. 2021. 16. 48–59.

[17] Liu, H., & Song, B. (2017, November). Stock trends forecasting by multi-layer stochastic ann bagging. In 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI). 322-329.

[18] Doğan, M., Metin, Ö., Tek, E., Yumuşak, S., & Öztoprak, K. (2020, December). Speculator and influencer evaluation in stock market by using social media. In 2020 IEEE International Conference on Big Data (Big Data). 4559-4566.

[19] Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media. 8(1). 216-225.