Multi-Step Financial Stock Index Forecasting Model Using Convolutional Neural Network with Gated Recurrent Unit Approach

Denis Shakhov[1,*], Inomjon Yusubov[2], Sanat Yakubov[3], Aleksey Ilyin[4], Emil Hajiyev[5], Tatyana Khorolskaya[6]

[1]Department of Economics and Management, Khorezm University of Economics, Urgench, 220100, Uzbekistan

[2]Department of Economics, Urgench State University, Urgench, 220100, Uzbekistan

[3]Department of Economics, Mamun University, Khiva, 220900, Uzbekistan

[4]Kursk Branch, Financial University under the Government of the Russian Federation, Moscow, 125167, Russia

[5]Department of Business Management, Azerbaijan State University of Economics (UNEC), Baku, AZ1001, Republic of Azerbaijan

[6]Department of Money Circulation and Credit, Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, 350044, Russia

Emails: shakhov@mymail.academy; yusubov.inomjon@mail.ru; yakubov_sanatbek1@mamunedu.uz; aeilin@fa.ru, hajiyev.emil@unec.edu.az, tatyana.e.khorolskaya@yandex.ru

Abstract

Prediction of time series is a vital issue related to an extensive array of financial, and social applications, and engineering. The main challenge arises from the intricacy due to the temporal assets of time series and the unavoidable weakening function of analytical systems. Therefore, it is usually problematic to precisely forecast values, particularly in a multi-step ahead situation. Multi-step financial stock price forecast over a lasting perspective is vital for predicting its instability, letting economic organizations charge and evade derivatives, and banks to measure the hazard. Recently, Deep learning systems have been capable to perceive and analyze intricate patterns and connections in the data automatically and haste up the trading procedure. This manuscript designs and develops a Multi-Step Financial Stock Index Forecasting Model Using a Convolutional Neural Network with Gated Recurrent Unit (MFSIFM-CNNGRU) model. The proposed MFSIFM-CNNGRU model relies on enhancing the predicting model for the financial stock index. To accomplish that, the data normalization stage is initially performed by employing z-score normalization to convert input data into a suitable format. Next, the proposed MFSIFM-CNNGRU model designs a hybrid of convolutional neural network and gated recurrent unit (CNN-GRU) technique for the prediction model. Eventually, the hyperparameter selection of the CNN-GRU model can be implemented by the design of the improved whale optimization algorithm (IWOA). The efficiency of the MFSIFM-CNNGRU method has been validated by comprehensive studies using the benchmark dataset. The numerical result shows that the MFSIFM-CNNGRU method has better performance and scalability under various measures over the recent techniques

Keywords: Multi-Step Financial Stock Index; Convolutional Neural Network; Gated Recurrent Unit; Z-score Normalization; Improved Whale Optimization Algorithm

1. Introduction

Economic globalization is one of the 3 major globalization dimensions because the stock market become extremely fluctuating, complex, and polymorphic. Stock price index forecast is essential for economic analysts and investors to develop investment approaches and get sustainable earnings [1]. Consequently, a comprehensive analysis has

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

been done on enhancing economic time series forecasting methods as an effective specialization. Gradually a stock price time- series is a series of succeeding observations gathered from the stock market [2]. The forecast of these time series could be built in multiple- or single-stage forward prediction. Certainly, a long-term prediction that is multiple-stage forward is more effective for stock market investors to reach the finest outcomes and strategies for the future [3]. However, current stock price forecast investigations only address single-stage forward predicting. Thus, there is a necessity to examine multiple-stage ahead stock price prediction [4].

Another concern that needs attention is that economic time series were typically dynamic, highly fluctuated, complex, and non-linear which makes their forecast challenging. In recent times, decomposition-based approaches have been projected [5]. The major concept behind decomposition-based approaches is to deteriorate difficult time series into component counts with a simplest framework that decreases time-series complication [6]. These elements have simpler frameworks and more persistent trends, their forecast is simpler, and the predictable precision is superior [7]. The investigators to forecast and analyse time series data [8] have utilized diverse techniques. To overwhelm these issues, a hybrid method that associates with Support Vector Machine (SVM) to resolve the non-linear regression estimation concern [9]. Another model integrates with Artificial Neural Network (ANN) to forecast the non-linear part. Deep learning (DL) methodologies are also utilized in stock price prediction [10].

This manuscript designs and develops a Multi-Step Financial Stock Index Forecasting Model Using Convolutional Neural Network with Gated Recurrent Unit (MFSIFM-CNNGRU) model. To accomplish that, the data normalization stage is initially performed by employing z-score normalization to convert an input data into a suitable format. Next, the proposed MFSIFM-CNNGRU model designs a hybrid of convolutional neural network and gated recurrent unit (CNN-GRU) technique for the prediction model. Eventually, the hyperparameter selection of CNN-GRU model can be implemented by the design of the improved whale optimization algorithm (IWOA). The efficiency of the MFSIFM-CNNGRU method has been validated by the comprehensive studies using the benchmark dataset.

2. Literature Works

Agarwal et al. [11] developed an innovative hybrid technique that incorporates Successive Variational Mode Decomposition (SVMD) with an LSTM system. SVMD is utilized as a data-decomposition model to assist the division of difficult unique series into its decomposed basic mode functions to decrease non-stationarity. This method is examined for either synthetic signals or financial time series data. Bai et al. [12] focused on dynamically representing the real-time decision-making situations from a finer granularity; the TSIS is developed depending on the information method. Subsequently, a Denoised Neighborhood Rough Set (DNRS) technique depending on the TSIS is projected by local density feature to attain the objective of feature selection (FS) that could poor the sound effect on instance data. Then, the MKELM and multivariate empirical mode decomposition (MEMD) are utilized to forecast and decompose. Eventually, the developed TSIS prediction method is applied to stock prices. Fan et al. [13] introduced a developed stock price prediction model based on the Extended LSTM (xLSTM) model, intended to improve analytical precision through either long or short-term periods. We manage extended investigations by evaluating and building techniques that depend on LSTM and xLSTM structures for several stocks. Our outcomes establish that the xLSTM technique dependably surpasses the LSTM method through each time horizon and stock, with the implementation gap expanding as the forecast period amplifies.

Diane and Brijlal [14] initially focus on dual new non-linear ML models, like the ANN and RF. We develop and then relate the execution of these dual techniques in forecasting stock market accomplished volatility for the JSE Financials Index (JFIN) and the JSE Basic Material Index (JBIND) through a period of 5 years. Depending on project outcome, the RF classifier outperformed the ANN technique for either the JBIND or JFIN index. The authors [15] project an innovative hierarchical FS with local shuffling (HFSLS) and models reweighting (MR) depend upon LSTM, called HFSLSMR-LSTM, for stock price prediction. Particularly, for every layer, local shuffling disturbs every aspect to re-predict, its forecast value is related to the true value to determine the feature significance, and the essential characteristics are chosen and returned to the subsequent layer.

In [16], we aimed to the consequences of novelty like stock prediction and financial risk management. We deliberate dual substantial models that have a prominent role in stock prediction. ANN as absence of few data points does not hamper the function of network. Next, SVM has multiple attributes, and owing to simplest decision limitations, it evades over-fitting. This work initially looks at the diverse technologies utilized in SMP. Fang et al. [17] developed a new long-term prediction FTS movement approach based on an adaptive LSTM technique. The adaptive network primarily contains dual layer of LSTM monitored by a pair of Batch Normalization (BN) layers, binary classifier, and dropout layer. To acquire the significant advantages points, we projected to utilize an adaptive cross-entropy loss function that improves the predictive capability on the sharp variations and diminishes the minor oscillations.

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

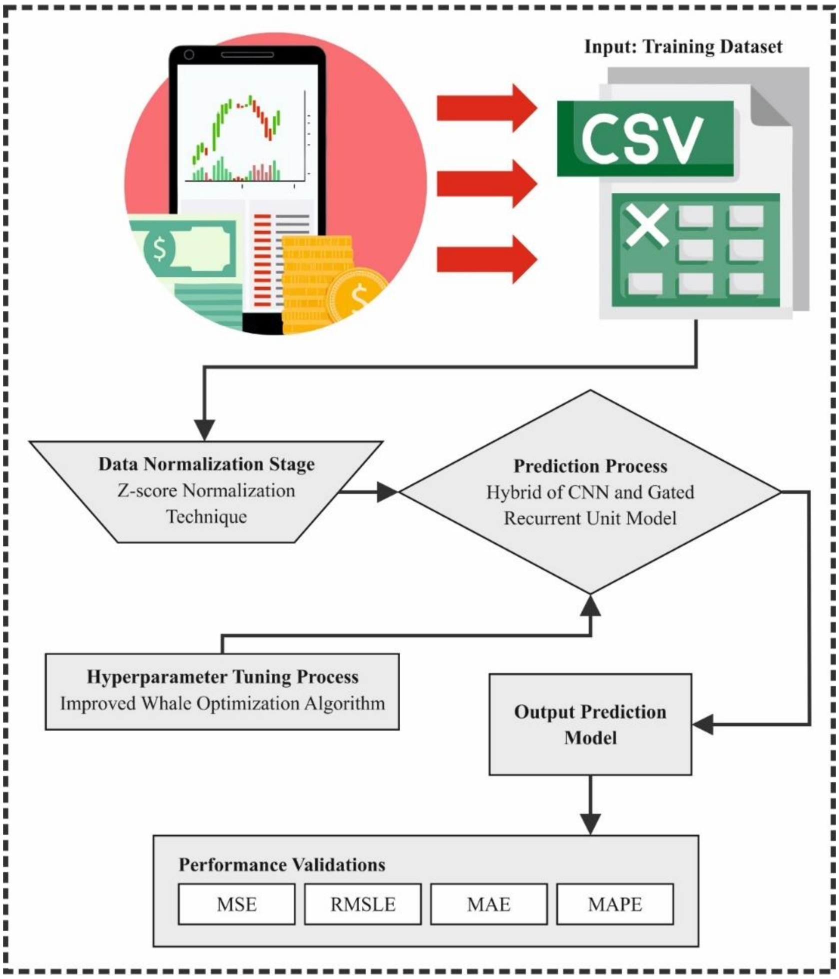

Figure 1. Overall Process of MFSIFM-CNNGRU model

3. The Proposed Method

In this manuscript, we focus on designs and develop a MFSIFM-CNNGRU model. The proposed MFSIFM-CNNGRU model relies on enhancing the predicting model for financial stock index. To accomplish that, the MFSIFM-CNNGRU model has z-score normalization, prediction using CNN-GRU, and parameter optimizer are demonstrated in Fig. 1.

A. Stage I: Z-score Normalization

Initially, the data normalization stage is performed by employing z-score normalization to transform an input data into a suitable format. Z-score normalization is also called standardization, which is a method employed to rescale data by converting it into a distribution with a mean of 0 and a standard deviation of 1 [18]. In the situation of multi-step financial stock index prediction, Z-score normalization aids in managing the fluctuating measures of financial data by regulating the features and creating them similarly. This normalization technique is particularly beneficial when utilizing ML techniques such as regression or neural networks, as it speeds up convergence and enhances prediction accuracy. By eliminating the impact of outliers and ensuring that every feature has equivalent alteration, Z-score normalization improves the model's steadiness and performance in predicting stock index trends over multiple steps.

B. Stage II: Prediction using CNN-GRU

Next, the proposed MFSIFM-CNNGRU model designs a hybrid of CNN-GRU technique for the prediction model. The GRU is a variation of the RNN. The GRU was presented to improve the RNN calculation and, more accurately, to overwhelm the gradient vanishing problem [19]. Stimulated by the new LSTM, such that an individual cell of the memory is lost, the GRU accepts dual gates, such as the reset and the update gates that are only liable for updating and inputting the information and determining which information transfers to the following phase and that must be prevented. The mathematic expression of the GRU is transcribed as shown:

$$R_t = \sigma(W_r x_t \oplus W_r h_{t-1} \oplus b_r) \quad (1)$$

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

$$Z_t = \sigma(W_Z x_t \oplus W_Z h_{t-1} \oplus b_Z) \tag{2}$$

$$C_t = tanh(W_c x_t \oplus W_c(R \odot h_{t-1}) \oplus b_c) \tag{3}$$

$$o_t = (1 - Z_t) \odot C_{t-1} + Z_t \oplus c_t \tag{4}$$

Whereas W_r, W_z, W_c, b_r, b_z, and b_c represents bias and weights trainable parameters in the update and reset gates, while Z_t and R_t signifies outputs of the update and reset gates, correspondingly x_t denotes input vector at time-step t, h_{t-1} symbolize accessible data at the preceding time step, σ stands for the function of sigmoid, $\oplus$means basic addition process, O_t denotes hidden layer (HL) at present time t, and c_t signifies candidate output state.

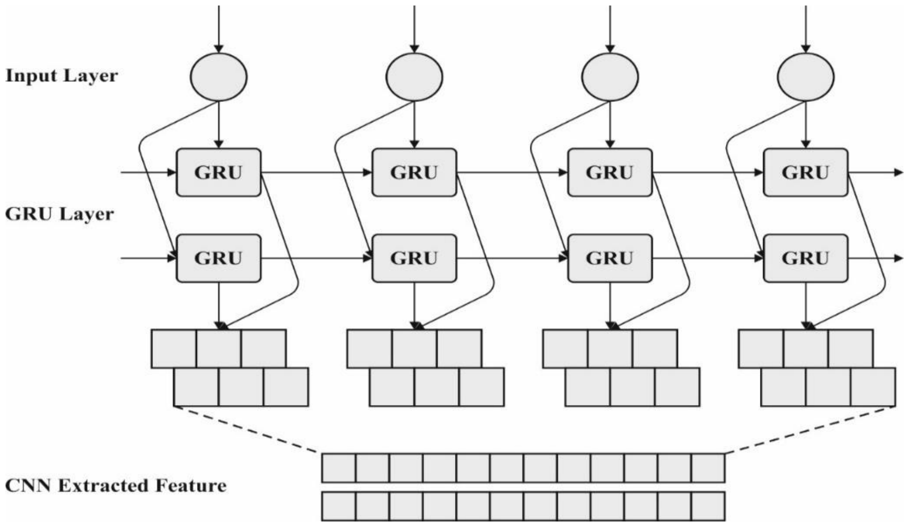

Figure 2. Structure of CNN-GRU

CNN is a type of DL structure mostly applied to resolving different types of time series, audio, text, and, significantly for processing image information. It operates in dual differentiated phases: classification and feature extraction. The CNN structure closely resembles the normal ANN to which it can be established that numerous layers should be required to attain the last task. There is a collaboration of HL, an input layer, and an output layer. Over the output and input layers, the HL is made from a pooling, a convolutional, a fully connected (FC), and an activation layer. The convolutional layers are applied to remove the possible characteristics from the input area. The layer of pooling was presented to improve the computing by reducing the feature count mapping by remembering simply the necessary. During these layers, the max pooling is normally applied. The FC layer is the final stage, and it is finally the CNN method utilized for the concluding prediction of the method.

It is a mixture of the CNN and the GRU. The GRU restores the information nearly the important characteristics offered by the block of CNN; that can be attained by transitioning the fattened layer output value to the gate components for $tracking$ the sequence state. Lastly, the significance of joining the CNN by the GRU is that reaction to either CNN or GRU models has been $concatenated$ to offer the last response. Fig. 2 illustrates the structure of CNN-GRU.

C. Stage III: Parameter Optimizer

Eventually, the hyperparameter selection of CNN-GRU model can be implemented by the design of the IWOA. The WOA is a new model, which differentiates itself in its novel architecture and its smaller controller parameters after comparison with another optimization [20]. Additionally, the WOA converges rapidly is most impressive, and implements superior to other intellectual models in various numeric optimization and engineering difficulties.

Nevertheless, the WOA even required breaking down several obstacles. Once directing complex issues with several variables, it hunts gradually and is liable to become trapped in local ideals using a weaker convergence capability. As a result, this paper improves the filling of the primary population locations of the model over chaotic mapping. Then, it balances the local and global searching competence in the convergence procedure of the model, thus enhancing the speed of convergence.

Chaotic Inverse Learning

Attaining optimum searching performance depends cripplingly on the accurate initialization of swarm methods. As regards the WOA, its arbitrary setup can result in an uneven distribution of whales through the searching region.

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

As a result, when the wealth is not allowed, the majority of the whales will produce in places apart from the best point that might improve the time to hunt for the optimum point and slow rates of convergence.

There is a dual chaotic perturbation equation: The Tent and the Logistic maps that are typically utilized by people. The Logistic mapping tends to allocate points more consistently near the intermediate values. The Tent map is simpler and even offers improved uniform traversal, providing the model with an advanced searching speed. Moreover, when we utilize reverse learning to increase the initialization of the population, they might importantly improve the speed of the convergences. Therefore, to utilize the advantages of either initialization technique, this study makes the primary population using the principles of chaotic mapping and reverse learning.

During population optimization methods, while a greater N improves the global searching ability of the model, it considerably decreases the speed of convergence. The chaotic map application to manage the first whale population to decrease the correlation among sample individuals. The mathematic description of the model is described as shown:

$$x_{i+1} == \begin{cases} 2x_i + rand(0,1) \times \frac{1}{N} & 0 \le x \le \frac{1}{2} \\ 2\,(1 - x_i) + rand(0,1) \times \frac{1}{N} & \frac{1}{2} < x \le 1 \end{cases}, i = 1, \dots, N \quad (5)$$

Whereas N characterizes the individual counts within the population group. x_i signifies the individual whale objects. The randomly generated variable aids stop the iterations into small or unstable periodical designs. The development creates a greater change in increasing its efficiency and robustness.

Subsequently, we apply the chaotic sequence to make the consistent primary population.

$$x_{ij} = x_{\min j} + Z_{kj}\left(x_{\max j} - x_{\min j}\right) \quad (6)$$

$$x_{ij}^{*} = x_{\min j} + x_{\max j} - x_{ij} \quad (7)$$

Whereas x_{ij}^{*} characterizes the inverse population. $x_{\min j}$ and $x_{\max j}$ symbolize the maximum and minimum whale individuals in the $j\ th$ column, individually. Lastly, we associate the reverse population that are composed of the best N individuals selected based on their optimal fitness.

Dynamic Inertia Weight and Non-linear Convergence Factor

In the WOA, it is generally recognized that A and B are the dual essential parameters. A is typically reliant on a, whereas C is frequently affected by r. Due to WOA, A plays an important role in the local exploitation abilities and WOA's exploration. To improve its capabilities of global exploration, we utilize 3 approaches: changing the variation curve of A, presenting dynamical inertia weights, and joining the WOA by the simulated annealing process and adaptive mutation perturbation.

1. Non-linear Convergence Factor

We apply non-linear convergence factors to change the variation curve. Based on Eq. (3), as a improves, A additionally improves. It is detected that as A becomes greater, the iteration step size of the optimizer model rises. Nevertheless, as the step size improves, the model's concentration on the local area reduces. As a result, by decreasing the value of a, and thus minimizing the size of the step, the local exploitation ability is improved. Normally A reduces linearly from (2-0) within the WOA, then in several composite optimizer issues, this might cause the model to untimely converge to local exploitation.

$$\vec{a} = \vec{a}_0 \cdot e^{-\tan\left(\left(1.2 * \frac{t}{T_{\max}}\right)\right)^2} \quad (8)$$

Here t embodies the present iteration count. $T_{\max}$ signifies the maximal iteration amount. $\vec{a}_0$ denotes the primary control factor. Earliest to the iteration, the value of a is greater than the new. During the last phases of iteration, the curve suddenly reduces to a lower value.

2. Dynamic Inertia Weight

Based on this, we can apply dynamical inertia weights to update the locations of the whales'. The equation is as demonstrated:

$$\omega(t) = \omega_{\max} \cdot \left(1 - \sin\left(\frac{\pi t}{2T_{\max}}\right)\right) + \omega_{\min} \cdot \sin\left(\frac{\pi t}{2T_{\max}}\right) \quad (9)$$

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

Where ω characterizes the dynamic inertia weight. $\omega_{\min}$and $\omega_{\max}$ refers to the minimum and the maximum of the dynamical inertia weight. $T_{\max}$ signifies the maximal iteration counts, whereas t represents the present iteration amount. Finally, the equation below characterizes the spiral location adjustment:

$$\overrightarrow{X(t+1)} = \left|\overrightarrow{X_{best}(t)} - \overrightarrow{X(t)}\right| \cdot e^{bl} \cdot \mathrm{coss}(2\pi l) + \overrightarrow{X_{best}(t)} \cdot \omega(t) \quad (10)$$

Whereas l denotes randomly selected value inside the interval [1, 1]. ω is moderately greater at the start of the iteration that creates the model have improved abilities in global exploring, stopping initial convergence to local ideals.

3. Adaptive Mutation Perturbation and Simulated Annealing Operation

Consequently, we can combine the simulated annealing principles into the WOA. Additionally, for each individual is possible to converge toward the optimum solution, this optimum individual might even not be the global optimal. So, we utilize the equation.

$$X_{b_{new}}^{t} = \frac{t}{\mathrm{Maxiter}} \cdot X_{b_{gaussian}}^{t} + \left(1 - \frac{t}{\mathrm{Maxiter}}\right) \cdot X_{b_{cauchy}}^{t} \quad (11)$$

At the start of the iteration, we utilize the Cauchy mutation for sharper and larger variations. The t is lower which may assist the WOA get away from the local ideals. Once the iteration continues, t becomes higher. The IWOA is applied to define the hyperparameter contained in the CNN-GRU system. The MSE is measured as function of an objective and its mathematical formulation is given below.

$$MSE = \frac{1}{T}\sum_{j=1}^{L}\sum_{i=1}^{M}\left(y_j^i - d_j^i\right)^2 \quad (12)$$

While L means a resultant value of data; M signifies the resultant value of layer; d_j^i and y_j^i means the proper and achieved scales for jth unit in tth time respectively.

4. Experimental Validation

The experimental evaluation of the MFSIFM-CNNGRU technique can be tested using three datasets [21]. Dataset 1: Dow Jones Composite Average (DJCA), Dataset 2: Dow Jones Industrial Average (DJIA), and Dataset 3: Dow Jones Transportation Average (DJTA).

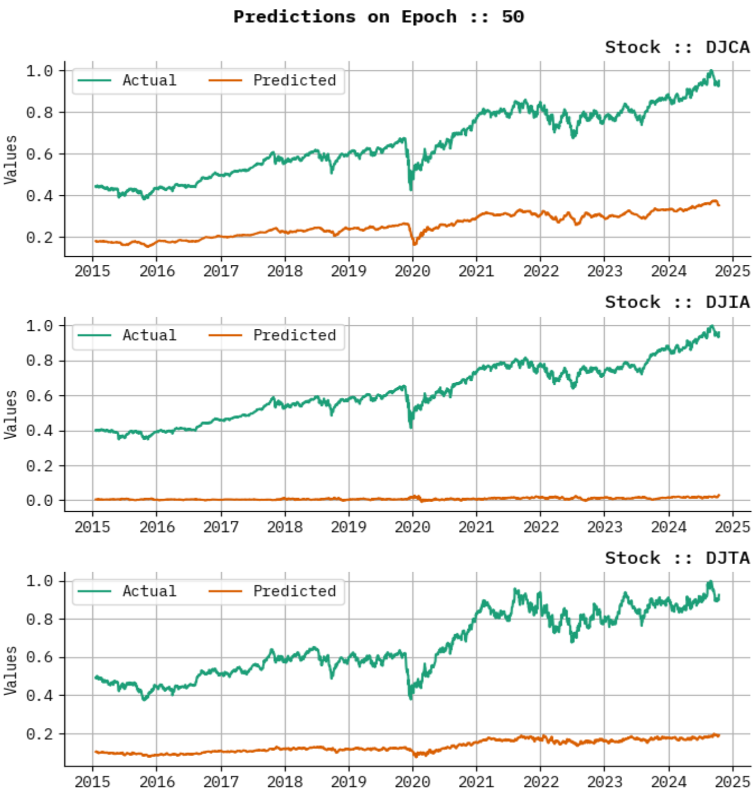

Figure 3. Actual and Prediction Results on 50 Epochs

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

Fig. 3 displays a thorough predictive outcome analysis of the MFSIFM-CNNGRU methodology with 50 epochs. The figure identified that the MFSIFM-CNNGRU algorithm has accomplished effective predictive values below three datasets. It is noted that the difference between the actual and predictive values is significantly low.

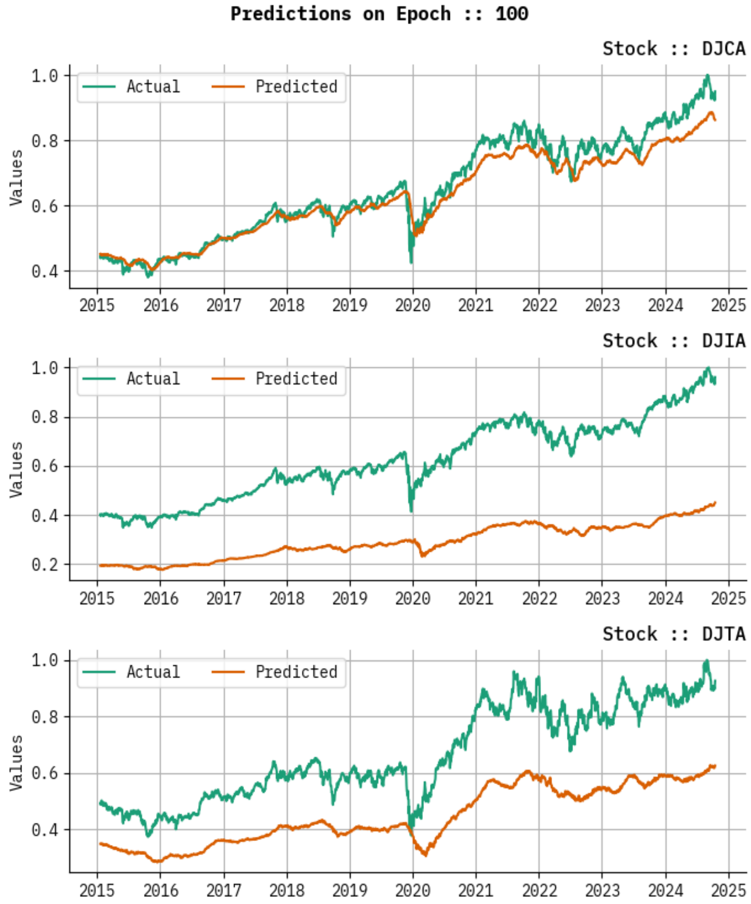

Figure 4. Actual and Prediction Results on 100 Epochs

Fig. 4 displays a detailed predictive outcomes analysis of the MFSIFM-CNNGRU technique with 100 epochs. The figure indicated that the MFSIFM-CNNGRU approach has achieved effectual predictive values below three datasets. It is denoted that the change between the actual and predictive values is considerably low.

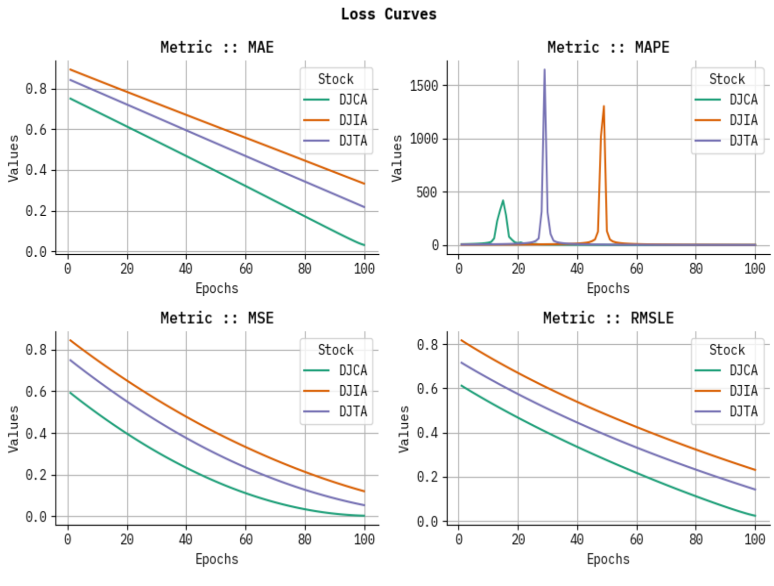

Figure 5. Loss curve of MFSIFM-CNNGRU model under MSE, MAPE, MSE, and RMSLE

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

Table 1: Overall outcomes of MFSIFM-CNNGRU system with various measures and three datasets

Stock Dataset	MSE	RMSLE	MAE	MAPE
DJCA	0.0497	0.1399	0.2138	0.4839
DJIA	0.0458	0.1356	0.204	0.4806
DJTA	0.0527	0.1427	0.2186	0.4859

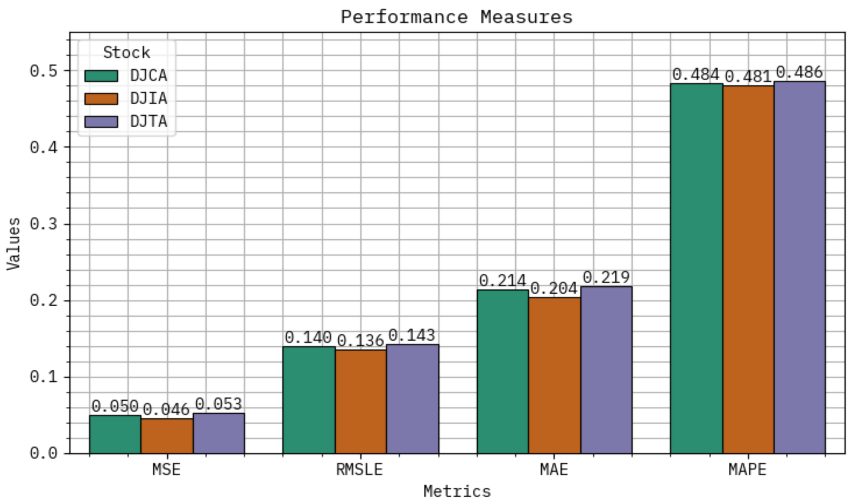

Figure 6. Prediction outcomes of MFSIFM-CNNGRU system with three datasets

The overall outcome of loss analysis of the MFSIFM-CNNGRU approach on the MSE, MAPE, MSE, and RMSLE is shown in Fig. 5. The figure shows that the MFSIFM-CNNGRU algorithm has meant the reduced DJCA loss, DJIA loss across DJTA loss. It is also noted that the values of loss get soaked with the epoch count of 100.

In Table 1 and Fig. 6, the overall prediction outcomes of MFSIFM-CNNGRU system with three datasets. Based on DJCA dataset, the MFSIFM-CNNGRU model attains MSE of 0.0497, RMSLE of 0.1399, MSE of 0.2138, and MAPE of 0.4839. In addition, depending on DJIA dataset, the MFSIFM-CNNGRU approach reaches MSE of 0.0458, RMSLE of 0.1356, MSE of 0.204, and MAPE of 0.4806. Eventually, with respect to DJTA dataset, the MFSIFM-CNNGRU system accomplishes MSE of 0.0527, RMSLE of 0.1427, MSE of 0.2186, and MAPE of 0.4859.

Table 2, an overall comparative outcomes of the MFSIFM-CNNGRU technique with existing approaches in terms of MSE and MAE [2].

Fig. 7 illustrates the comparative MSE inspection of the MFSIFM-CNNGRU system. The table values implicit that the MFSIFM-CNNGRU method outperformed the superior performances. Based on MSE, the MFSIFM-CNNGRU methodology has reached lesser MSE of 0.0458 where the 1D-CapsNet-LSTM, RNN, LSTM, CNN-LSTM, and SVR models have realized better MSE of 0.1320, 0.1156, 0.0976, 0.0797, and 0.0620.

Table 2: MAE and MAE outcomes of MFSIFM-CNNGRU algorithm with existing methods

Methods	MSE	MAE
1D-CapsNet-LSTM	0.1320	0.7651
RNN Model	0.1156	0.7473
LSTM Model	0.0976	0.6317
CNN-LSTM	0.0797	0.6154
SVR Algorithm	0.0620	0.5980
MFSIFM-CNNGRU	0.0458	0.4806

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

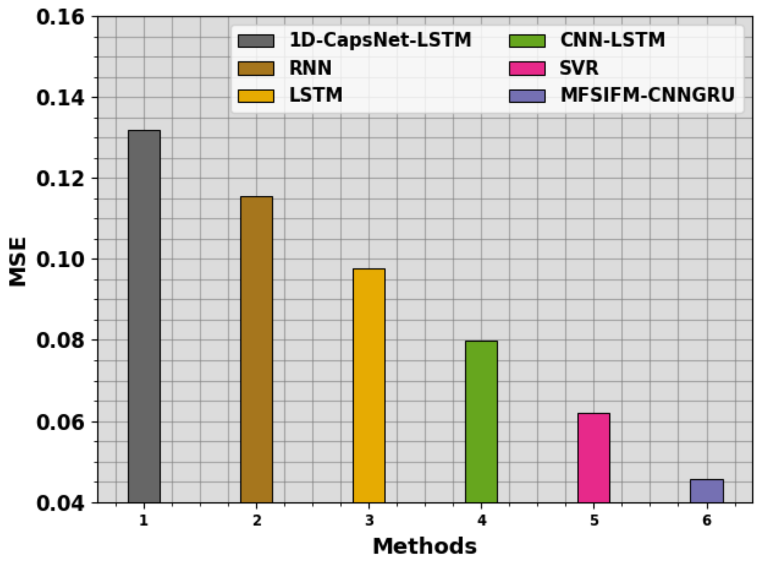

Figure 7. MSE outcome of MFSIFM-CNNGRU algorithm with existing models

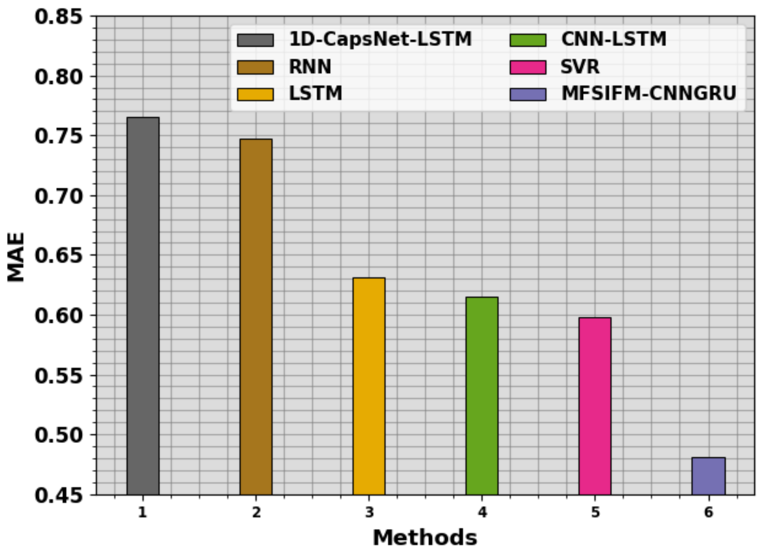

Figure 8. MAE outcome of MFSIFM-CNNGRU algorithm with existing models

Fig. 8 establishes the MAE result of the MFSIFM-CNNGRU methodology is reported. The outcomes confirmed that the 1D-CapsNet-LSTM technique has displayed ineffectual outcomes with greater MAE of 0.7651. In the meantime, the RNN and LSTM methods have shown significant performance with MAE of 0.7473 and 0.6317. Furthermore, the CNN-LSTM and SVR algorithms have proficient reasonable results with MAE of 0.6154 and 0.5980. Finally, the MFSIFM-CNNGRU system establishes maximum performance with lower MAE of 0.4806.

5. Conclusion

In this manuscript, we focus on designs and develop a MFSIFM-CNNGRU model. The proposed MFSIFM-CNNGRU model relies on enhancing the predicting model for the financial stock index. To accomplish that, the data normalization stage is initially performed by employing z-score normalization to convert input data into a suitable format. Next, the proposed MFSIFM-CNNGRU model designs a hybrid of the CNN-GRU technique for the prediction model. Eventually, the hyperparameter selection of the CNN-GRU model can be implemented by the design of the IWOA. The efficiency of the MFSIFM-CNNGRU method has been validated by comprehensive studies using the benchmark dataset. The numerical result shows that the MFSIFM-CNNGRU method has better performance and scalability under various measures over the recent techniques.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025

References

[1] T.H.H. Aldhyani and A. Alzahrani, "Framework for predicting and modeling stock market prices based on deep learning algorithms," *Electronics*, vol. 11, 2022.

[2] C. Zhang, N.N.A. Sjarif, and R. Ibrahim, "1D-CapsNet-LSTM: A deep learning-based model for multi-step stock index forecasting," *Journal of King Saud University-Computer and Information Sciences*, vol. 36, no. 2, p. 101959, 2024.

[3] A. Altan, S. Karasu, and E. Zio, "A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer," *Applied Soft Computing*, vol. 100, 2021.

[4] K.J. Koa, Y. Ma, R. Ng, and T.S. Chua, "Diffusion variational autoencoder for tackling stochasticity in multi-step regression stock price prediction," in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 2023, pp. 1087-1096.

[5] S. Aryal, D. Nadarajah, P.L. Rupasinghe, C. Jayawardena, and D. Kasthurirathna, "Comparative analysis of deep learning models for multi-step prediction of financial time series," *Journal of Computer Science*, vol. 16, 2020.

[6] I.E. Livieris and P. Pintelas, "A novel multi-step forecasting strategy for enhancing deep learning models' performance," *Neural Computing and Applications*, vol. 34, no. 22, pp. 19453-19470, 2022.

[7] H. Nasiri and M.M. Ebadzadeh, "Multi-step-ahead stock price prediction using recurrent fuzzy neural network and variational mode decomposition," Applied Soft Computing, vol. 148, p. 110867, 2023.

[8] D.S. Berman, "DGA CapsNet: 1D application of capsule networks to DGA detection," *Information*, vol. 10, no. 5, 2019.

[9] C. Deng, Y. Huang, N. Hasan, and Y. Bao, "Multi-step-ahead stock price index forecasting using long short-term memory model with multivariate empirical mode decomposition," *Information Sciences*, vol. 607, pp. 297-321, 2022.

[10] M. Mohamed, "Financial Risks Appraisal based on Dynamic Appraisal Framework," *Full Length Article*, vol. 2, no. 1, pp. 50-0, 2023.

[11] S. Agarwal, S. Sharma, K.N. Faisal, and R.R. Sharma, "Time-Series Forecasting Using SVMD-LSTM: A Hybrid Approach for Stock Market Prediction," *Journal of Probability and Statistics*, vol. 2025, no. 1, p. 9464938, 2025.

[12] J. Bai, B. Sun, Y. Guo, and X. Chu, "A new multivariate decomposition-ensemble approach with denoised neighborhood rough set for stock price forecasting over time-series information system," *Applied Intelligence*, vol. 55, no. 4, p. 284, 2025.

[13] X. Fan, C. Tao, and J. Zhao, "Advanced stock price prediction with xlstm-based models: Improving long-term forecasting," *Preprints*, 2024.

[14] D. Diane and P. Brijlal, "Forecasting stock market realized volatility using random forest and artificial neural network in South Africa," *International Journal of Economics and Financial Issues*, vol. 14, no. 2, pp. 5-14, 2024.

[15] Z. An, Y. Wu, F. Hao, Y. Chen, and X. He, "A novel hierarchical feature selection with local shuffling and models reweighting for stock price forecasting," *Expert Systems with Applications*, vol. 249, p. 123482, 2024.

[16] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting," *Annals of Data Science*, vol. 10, no. 1, pp. 183-208, 2023.

[17] Z. Fang, X. Ma, H. Pan, G. Yang, and G.R. Arce, "Movement forecasting of financial time series based on adaptive LSTM-BN network," *Expert Systems with Applications*, vol. 213, p. 119207, 2023.

[18] H. Henderi, T. Wahyuningsih, and E. Rahwanto, "Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer," *International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 13-20, 2021.

[19] R.M. Adnan, W. Mo, O. Kisi, S. Heddam, A.M.S. Al-Janabi, and M. Zounemat-Kermani, "Harnessing Deep Learning and Snow Cover Data for Enhanced Runoff Prediction in Snow-Dominated Watersheds," *Atmosphere*, vol. 15, no. 12, p. 1407, 2024.

[20] H. Yang, S. Yang, D. Meng, C. Hu, C. Wu, B. Yang, P. Nie, Y. Si, and X. Su, "Optimization of Analog Circuit Parameters Using Bidirectional Long Short-Term Memory Coupled with an Enhanced Whale Optimization Algorithm," *Mathematics*, vol. 13, no. 1, p. 121, 2024.

[21] J. Beach, "Dow Jones and S&P500 Indices Daily Update," *Kaggle*, [Online]. Available: https://www.kaggle.com/datasets/joebeachcapital/dow-jones-and-s-and-p500-indices-daily-update.

DOI: https://doi.org/10.54216/FPA.180208
Received: August 03, 2024 Revised October 26, 2024 Accepted: January 18, 2025