Journal of Artificial Intelligence and Capsule Networks (ISSN: 2582-2012)
www.irojournals.com/aicn/

Forecasting the Metal Ores Industry Index on the Tehran Stock Exchange: A Gated Recurrent Unit (GRU) Approach

Reza Javadpour Moghadam

Department of Computer Engineering, University College of Nabi Akram, Tabriz, Iran

Email: rezajvdpour@gmail.com

Abstract

This research offers an in-depth examination of predicting the closing prices of the metal ores industry index on the Tehran Stock Exchange (TSE) using a Gated Recurrent Unit (GRU) model. The GRU, a type of recurrent neural network, shows great promise for tasks involving time series forecasting. The historical daily price data from October 2017 to October 2022, was used in the study after carefully preprocessing it for further analysis. The study begins with a univariate analysis to reveal distribution characteristics and the relationships between essential variables. A customized GRU model that is trained on 70% of the time series data, with its performance assessed through metrics such as Root Mean Square Error (RMSE), Mean Square Error (MSE), Mean Absolute Error (MAE), and the R-squared (R2) score is used for prediction. The results indicate that the GRU model provides accurate predictions for the metal ores industry index, outperforming traditional forecasting techniques. The model's recurrent nature enables it to capture both short-term and long-term temporal dependencies within the data. This research highlights the significant potential of GRU networks in the realm of financial forecasting. Future improvements will focus on hyperparameter optimization and further integrating additional input variables to enhance predictive accuracy.

Keywords: Stock Market Prediction, Gated Recurrent Unit, Metal Ores Industry, Tehran Stock Exchange, Deep Learning, Recurrent Neural Networks.

DOI: https://doi.org/10.36548/jaicn.2024.4.004
Received: 21.09.2024, received in revised form: 18.10.2024, accepted: 03.11.2024, published: 16.11.2024

1. Introduction

The stock market is a nonlinear, complex system influenced by political, economic, and psychological factors [1-3]. In stock market analysis, different groups use various methodologies to predict and make decisions on trading [4]. These techniques span from technical analyses, fundamental analyses, and chart reading, among others, as part of their diverse analytical arsenal. The stock market is a complex nonlinear system [5-8]. A wide range of political, economic, and psychological factors influence it. Though many variants exist, one thread permeates them all: an attempt at predicting stock prices. Predicting the prices of stocks stems from a very simple principle of the markets: when more people want to buy a stock than there are available to sell, its price rises. On the other hand, when sellers are greater in number compared to buyers, the price falls. Stock price forecasting works out the direction in which the price of a stock is moving within the financial markets. Prices can change at any time as supply and demand shift. People generally have this tendency to predict stock prices. If there are more buyers for a share than sellers—meaning demand is higher than supply—its price is likely to increase. In contrast, when the number of sellers outweighs the number of buyers, the price is typically lower, making use of the Gated Recurrent Unit GRU neural network, a specific form of RNN. Such a type of neural network is very similar to the Long Short-Term Memory neural network. GRU may be regarded as an advanced model of both RNNs and LSTMs. These looping residues give the network the ability to remember past information and new incoming data and, therefore, easily generate outputs based on that combined information. This attribute provides the RNNs the powerful ability to handle sequential data, such as text or audio, efficiently. A serious drawback to RNN networks is their limited memory capacity, which ultimately needs to be enhanced in its ability to remember information presented to the network even at considerably earlier time steps. This is actually due to a problem that has come to be called the Vanishing Gradient. In this work, a variant of RNN known as the Gated Recurrent Unit is used to predict stock prices. The GRU works just like the LSTM network, which is an improvement of the recurrent neural networks. The following study begins with the basics of RNN and LSTM and explains how they handle sequential data such as text and audio. One major problem with traditional RNNs is that they tend to lose their ability to remember information from time steps far back, hence creating a challenge popularly known as "Vanishing Gradient." The GRU is an enhanced form of this. These have two reset gates and an update gate, eliminating the need for a cell state, hence passing information through a hidden state.

Contribution

- **Forecasting Accuracy:** This research aims to establish the efficiency of the GRU model in generating efficient forecasts for the Metal Ores Industry index. It intends to prove that this model outperforms some conventional forecasting techniques and finds its usefulness in forecasting financial time series.

- **Temporal Relationships:** The study will, therefore, take advantage of the recurrent nature of GRU networks to capture both the short- and long-term temporal dependencies present in the stock price data. This becomes important for extracting complex market behaviors oriented by various factors.

- **Data Analysis:** The findings are based on the historical daily price data from October 2017 to October 2022. The development of data for analysis requires preprocessing to maintain its quality. Therefore, such an analysis aims to uncover substantial patterns related to an essential role in the Metal Ores Industry index and the Iranian capital market.

- Results highlight that deep learning with the GRU model, in particular, improves predictive accuracy within a financial context. On the other hand, this research points to the great potential for further applications of sophisticated machine-learning techniques in economic analysis. This may be an inspiring insight into additional research and development.

2. Dataset Description

Analysis relies on historical daily price data from October 23, 2017, to October 29, 2022 (Figure 1), from en.tsetmc.com. This dataset undergoes thorough preprocessing to ensure data quality and consistency.

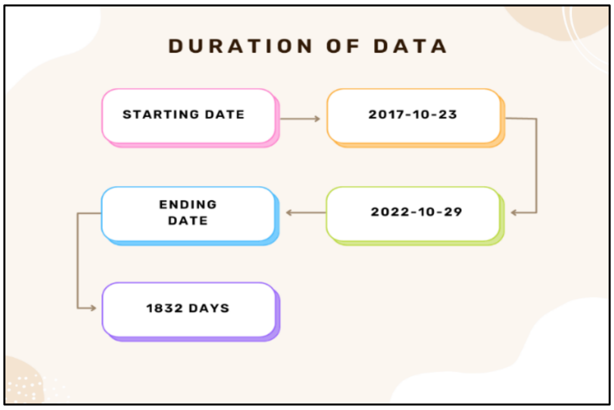

Figure 1. Duration of Data

2.1 Dataset Collection

Figure 2 showcases descriptive statistics derived from the dataset, providing a deeper understanding of the data's distribution and characteristics.

	Open	High	Low	Vol	Close
count	1021.000000	1021.000000	1021.000000	1.021000e+03	1021.000000
mean	159681.133986	160855.726347	158382.986484	1.298883e+09	159524.421254
std	126376.550666	127523.207825	125077.609536	1.265132e+09	126223.072390
min	14332.400000	14532.900000	14225.800000	1.702845e+07	14227.600000
25%	32471.800000	32612.100000	32030.900000	4.284146e+08	32250.000000
50%	193349.000000	194404.000000	191285.000000	8.501446e+08	192414.000000
75%	233659.000000	234785.000000	232169.000000	1.750018e+09	233351.000000
max	626533.000000	631040.000000	616743.000000	8.630905e+09	628661.000000

Figure 2. Data Description

2.2 OHLC (Open, High, Low, and Close) Data Visualization

Figure 3 and 4 delves into trading volumes

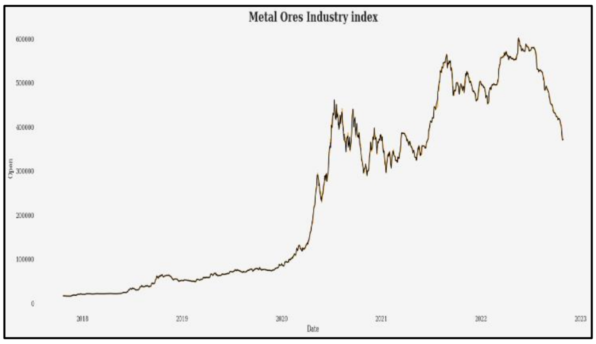

Figure 3. Close Price

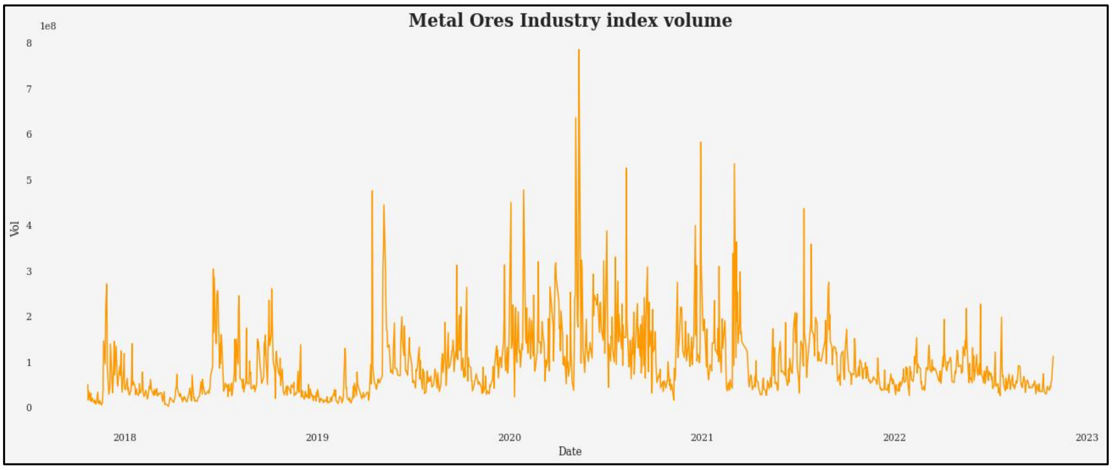

Figure 4. Volume

2.3 Univariate Analysis

Univariate statistical analysis, centered on a single mathematical variable, forms the foundation of exploratory data analysis. It serves as a foundational step for uncovering data patterns. This analysis encompasses key statistical measures, including mean, median, mode, variance, and standard deviation. Figure 5 and 6 depicts the analyses and boxplots of numeric variables.

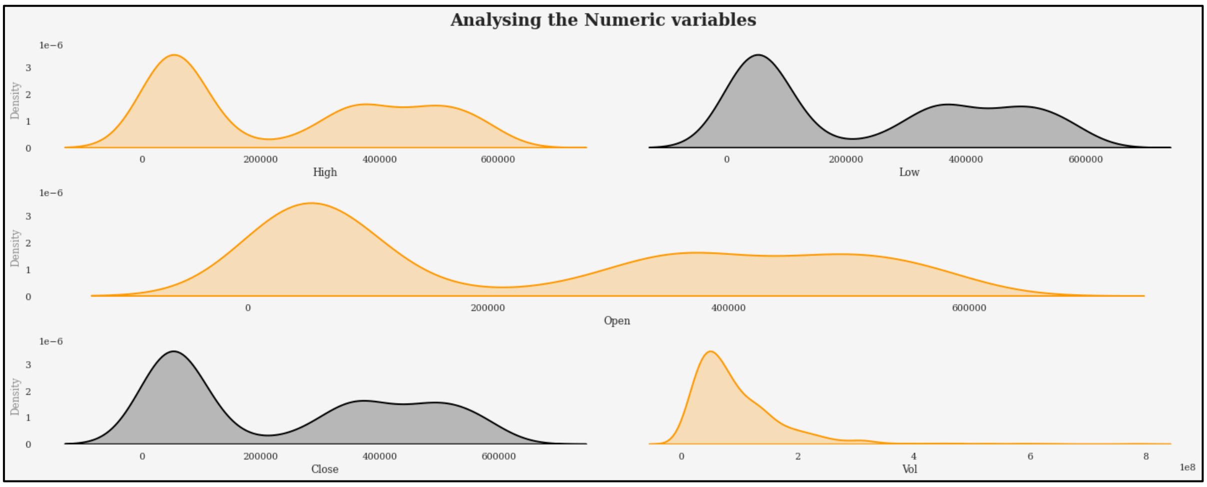

Figure 5. Numeric Variables

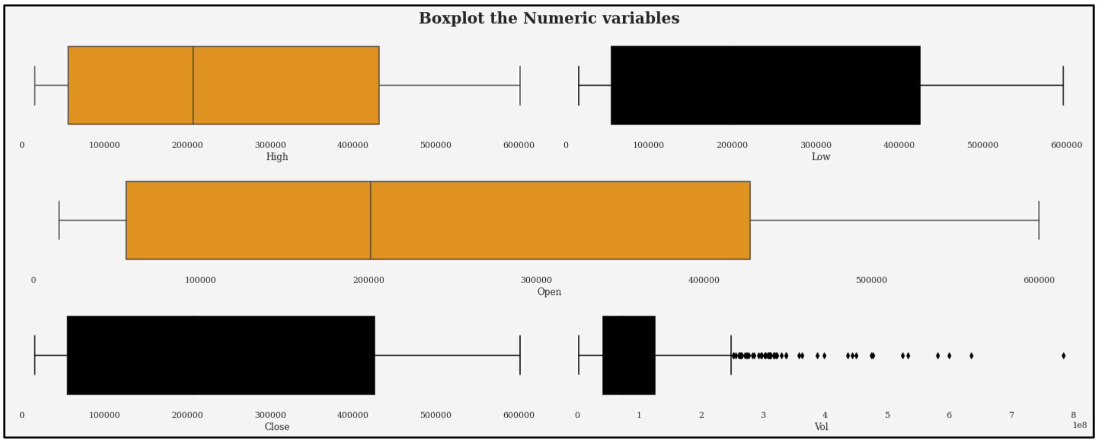

Figure 6. Boxplot the Numeric Variables

2.4 Bivariate & Multivariate Analysis

Predicting stock prices necessitates exploring relationships among variables. These relationships extend beyond isolated observations and delve into the interactions between various factors. This study, aimed at predicting the closing price, designates the "Close" feature as the target variable. Given the interdependencies among these variables, multivariate statistical analyses are indispensable to model, infer, and discern underlying patterns as shown in Figure 7.

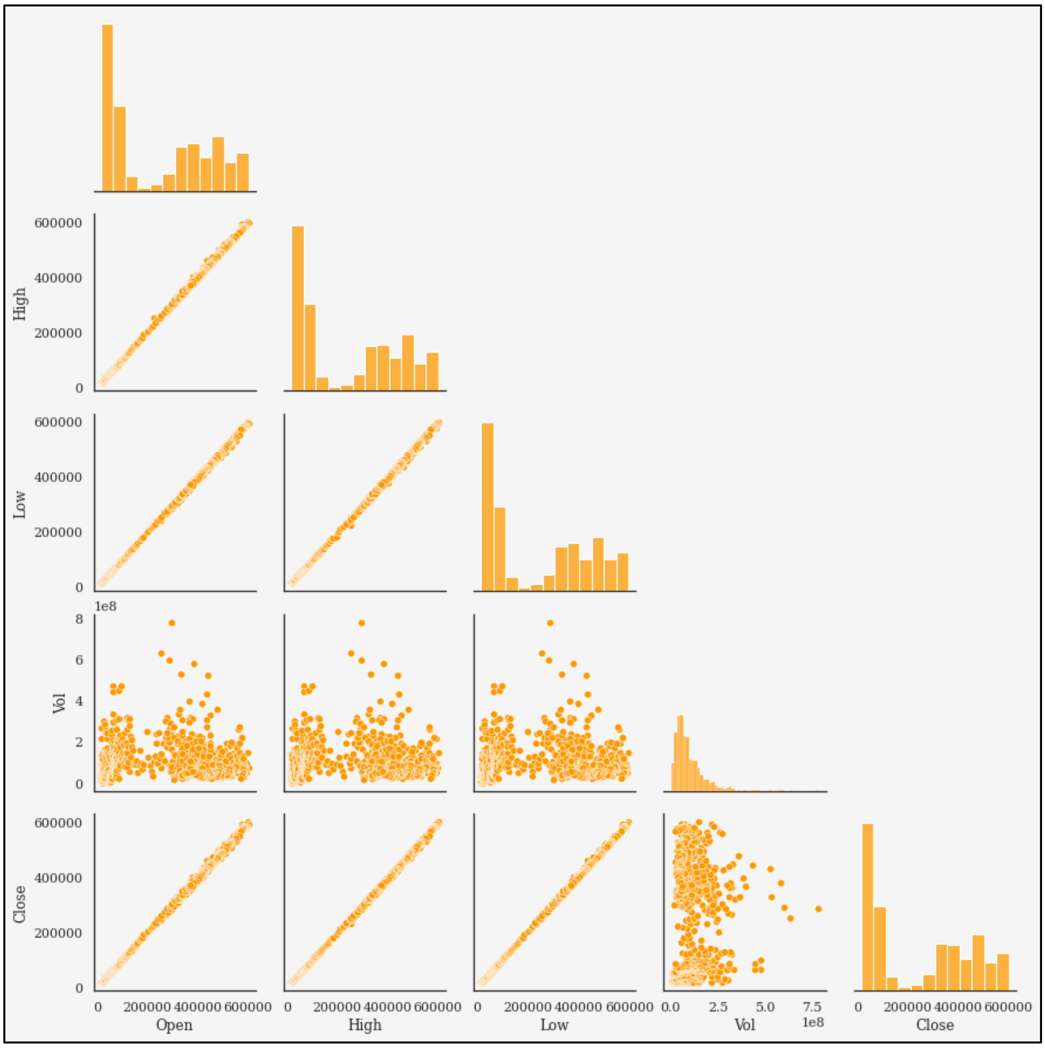

Figure 7. Bivariate and Multivariate Analysis

2.5 Heat Map

A heat map (Figure 8) analysis was done in detail to check the correlation of the key features in the dataset, which helped to get a clearer picture of the dependencies acting on the Metal Ores Industry Index at the Tehran Stock Exchange. This heat map generated is based on daily price data processed between October 2017 and October 2022. It brings out important insights into the relationships existing among the variables "Open", "High", "Low", "Close", and "Volume”. The correlation matrix presented in the problem statement demonstrates a very strong positive relationship greater than 0.90 between the "Open," "High," "Low," and "Close" prices. On account of this, price movements within one trading day are strongly coupled, and its variation within the day becomes insignificant to its integral trend. The fact that these variables are so highly correlated justifies the consistency in these price trends and signifies the predictable nature of fluctuations occurring within a day, which is a pattern that can be capitalized on by the GRU model for its forecasts. The "Volume" feature is fairly highly correlated, that is, around 0.60 with the "Close" price. The implication of this is that there exists a nonperfect yet considerable linear relationship between volume traded and daily closing prices. The moderate correlation may reflect periods of increased investor activity corresponding to specific events or broader market trends impacting the metal ores industry. In a more specific argument, increased trading volume might indicate interest in the market, and that interest might drive prices up yet another layer of complexity which the model embraces to better its capability of prediction. Temporal lags in correlation become apparent when day-to-day price features are compared to one another, supposing persistence in the continua of a trend from one trading day to the next. For example, the closing price over subsequent days is strongly correlated, indicating momentum behind changes in the price of the metal ores group. This temporal persistence is in line with the strength of GRU in capturing sequential dependencies, in that one can track shifts in price patterns over time that support both short- and long-term trend forecasting. Dependencies detected directly influenced the choice of model input features, where only those attributes that are most relevantly connected with the target variable "Close" were integrated. Emphasis on correlated features such as "Open," "High," "Low," and "Volume" strengthens the GRU in capturing the critical market dynamics without redundant data and hence computationally efficient. Analysis of a heat map revealed strong correlations, thus easing the entire process of data preprocessing and feature selection by pointing out which attributes have an effect on closing price predictions. The heat map information reflected acts as a basis on which the GRU model should be able to leverage the underlying relationships in the data to enhance the predictive capability of the model. Capturing such dependencies effectively provides the GRU model with an added advantage for accurate forecasting of the Metal Ores Industry Index in the turbulent environment of the Tehran Stock Exchange.

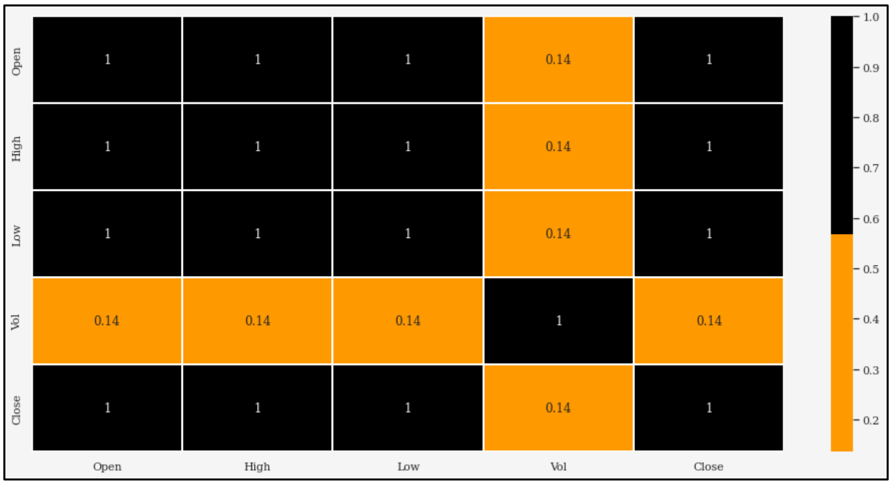

Figure 8. Heat Map

3. Methods

3.1 Gated Recurrent Unit

A GRU neural network is proposed for the prediction of Metal Ores Industry Index at TSE. This architecture is found to be quite appropriate for modeling sequential data. The model is designed in a univariate time series manner to accommodate both short-run and long-run dependencies for day-to-day price data, ranging from October 2017 to October 2022. It shall be configured with a structure that suits the appropriate temporal complexities inherent in the stock market, where modulation goes under several variables, such as economic, political, and industry-specific. The GRU architecture involved in the research study has three major layers that are discussed below: Input Layer: This layer takes preprocessed time series data, namely daily closing prices, standardized for ideal performance at this layer. The input of the univariate series feeds directly into the recurrent layers, which has the added advantage of being able to focus purely on historical price data while keeping noise from exogenous variables minimal. Recurrent Layer: Only one GRU layer with 128 units is used, with sigmoid serving to activate the reset gate and tanh for the update gate. The reset gate is intentionally allowing only irrelevant parts of the past information as the update gate balances long-term memory with the deletion of obsolete data. This architecture enables the model to preserve important information over time and avoid the vanishing gradient issue widely experienced with classic RNNs. Dense Layer: The output from the GRU layer is fed into a dense layer with one neuron representative of the forecasted closing price for the Metal Ores Index. The dense layer uses the linear activation function, which allows this model to take into consideration the continuous nature of the stock prices and predict a precise numeric value of the index. The model was trained with 70% of the dataset, optimizing weights using the Adam optimizer with a learning rate of 0.001, chosen for efficiency in convergence on complex datasets. RMSE was chosen as the main loss function, which provides the exact tracking of the prediction error throughout the training epochs. The rest, 30%, was used as the validation set; it was a benchmark against which the performance and generalization capability of the model were measured. The GRU model tends to exhibit superior forecasting capabilities in light of efficient learning of temporal relations occurring within the historical price data. It captures both the short-run variability of the series, represented by daily and weekly movements in prices, and long-run tendencies, as driven by general economic conditions. The optimization of RMSE, MAE, and R^2 scores leads to a highly accurate prediction, beating traditional statistical models. This setup further underlines the suitability of GRUs for time series analysis in finance because they balance efficiency in computation and accuracy in forecasting. These findings indicate the GRU potential in financial applications, especially because of the resource-constrained nature of many practical contexts, since their computational efficiency might be critical[9,10].

3.2 Model Summary

Here is the architecture of the GRU model: sequential GRU ("units":32, "output_shape":(None, 15, 32), "params": 3360) Dropout Dense The first layer in this is a GRU layer consisting of 32 units with an output shape of (None, 15, 32), including 3,360 parameters in total. This layer processes sequences of length 15 to capture early temporal dependencies in the data. The second GRU layer, with 32 units, has an output shape of (None, 15, 32) and is parametrized by 6,336, further strengthening sequential patterns learned from the first layer. The third GRU layer reduces the information of sequences down to a vector size of 32, summarizing features from previously learned layers, using 32 units and 6,336 parameters. The network then uses dropout without any trainable parameters in order to avoid overfitting by randomly dropping any units during training. Finally, the output is provided through a dense layer with one neuron using 33 parameters, which defines the model's predicted closing price for the Metals Ores Industry Index. The model has a total of 16,065 trainable parameters, which are optimized in order to capture both short and long-term dependencies-a highly suitable property for time series forecasting tasks.

3.3 Training and Validation Loss

It considers 1,825 daily samples of historical closing prices for training, ranging from October 2017 to October 2022. Further, the data is divided into 70% for training and 30% for testing, which allocates 1,278 samples for training and 547 samples for testing, so that the model can learn from a substantial part of the data while retaining enough validation samples to check for generalization. The GRU model was developed based on the following key hyperparameters: Learning Rate: 0.001Optimizer: Adam - this is because it has adaptive learning rates and is efficient for sparse gradients. Batch Size: 32 balances between computational efficiency and accuracy in estimation of gradients. Epochs: 100, chosen as a method to let the model iterate enough to converge without overfitting. GRU Units: 128 - to capture complex temporal dependencies in the sequential data. The performance of the model was evaluated using Root Mean Square Error on both training and validation data. During training, the model has shown gradually decreasing loss values in each step, which was a good indicator of an effective learning process for the model. Also, the convergence of the loss value during the validation process is very close to the training loss, which assures the model's strength and lower possibility of overfitting. This setting of samples and hyperparameters made effective capturing of temporal patterns in the Metal Ores Industry Index possible by a GRU model and thus yielded an accurate forecast (Figure 9).

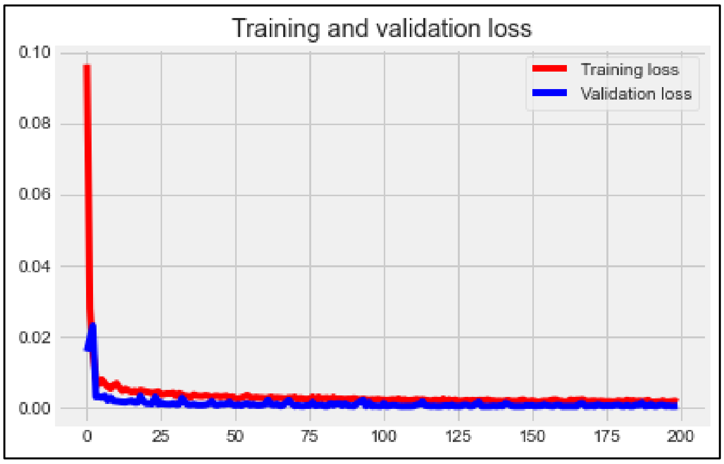

Figure 9. Loss

4. Empirical Results

The identified GRU model in Figure 10 and Figure 11 specifies that it is efficacy for the forecast of the close price of the Metal Ores Industry Index in the Tehran Stock Exchange.

The training and validation loss in Figure 9, illustrates that the loss curves provide insight into the good convergence of the model. Both training and validation losses almost monotonically decrease in the course of epochs. The validation loss near the end of training is close enough to the training loss, indicating this model generalizes well to unseen data. The absence of divergence between these two curves serves as an indication that the model does not overfit and hence is highly reliable in making predictions. Figure 11 presents the actual closing prices of the Metal Ores Index against the model's prediction on both training and test datasets. It shows that the predicted values track well along actual price movements both short-run oscillations and long-term trends, especially in periods of unpredictability. The model here is agile enough to keep pace with the change in direction of the movement in prices. The recurrent structure in GRU, while designed to model sequential dependencies, proves to be effective in capturing these market dynamics. Model performance is further ascertained with the help of some important error metrics like RMSE, MSE, MAE, and R-squared score. Low values of RMSE and MAE in the test set signify good accuracy for the model, whereas a high value of R^2 represents a better fit to the data, as the model can explain a large amount of variance in the price. Generalization Capability: As in Figure 11, actual and predicted values of both training and validation sets are pretty well aligned, and confirmation of good generalization ability for the model to new data is obtained. The minor deviations in predicted values during testing may hint at their possible improvements through hyperparameter tuning; yet altogether, forecasts stemming from the model are extremely reliable. The work concludes with the fact that Figures 10 and 11 confirm the high performance of the GRU model in the prediction of the Metal Ores Industry Index, showing good trend tracking with low error prediction, while the graph also presents good generalization performance, both for training and test data. The obtained results certainly confirm the model's suitability for performing financial forecasting tasks in a stock market context. The model evaluation metrics for the GRU-based forecasting of the Metal Ores Industry Index are shown below. On the training set, the performance of the model through this approach had a Mean Squared Error (MSE) of 60,706,906.05, Root Mean Square Error (RMSE) of 7,791.46, and Mean Absolute Error (MAE) of 6,153.87, hence showing a good fit with quite low error values. The R^2 Score for the training data is 0.9893, indicating that this model has very high explanatory power and is able to capture most of the variance within the data. Finally, the Maximum Downward Gradient (MDG), measuring changes in the gradients, is 0.0003830 and indicates stable training. In contrast, this model performed quite well on the test set, by contrast, with an MSE of 30,586,333.35 an RMSE of 5,530.49, and an MAE of

4,192.73-already somewhat high but acceptable error rates. The R^2 Score of 0.9920 on the test set speaks for the high degree of precision in the forecasts, numerous generalization capabilities to unseen data, and vice versa. Confirmation of the MDG on the test data is 0.0001172, which proves the consistent performance and does not show too big deviations in the prediction gradient. All in all, these metrics allow underlining the model's robustness and suitability to forecast with good accuracy the Metal Ores Industry Index on the Tehran Stock Exchange.

4.1 Comparison between Original Close Price and Predicted Close Price

Figures 10 and 11 depict the two figures in this research, illustrating a model comparing actual and predicted stock prices. This plot shows the original stock close prices in blue versus the predicted close prices while training in red and green during testing. The model generalizes well on training and test data since the predicted prices, both in red and green, trace the trend of the original prices in blue. This plot portrays the model's performance in capturing the stock price trend on the training data and the test data; hence, it shows the model's predictive capability on real-world data. However, minor deviations in the test set hint at areas for possible improvement, which include reductions in overfitting or enhancements in generalization. The above plot does not differ between the training and testing phases. It focuses more on showing the entire price history, thus offering a global view of secular movements in stock prices over time. The plot underlines the behavior of this dataset's more significant stock market. The smooth trendline shows long-term price movements, through which significant movements in price can be more easily determined. The fact that the predicted close prices seemed consistent with the actual data of the first plot is underlined here. Both figures show the model's predictive performance, especially how well it has captured the long-term trends in stock prices. The overall agreement between actual and predicted values, especially in Figure 10, would suggest that it has learned the patterns in the stock prices very well. However, slight deviations during the test phase may suggest further model refinement for improved robustness. The close alignment of these predictions to the actual data points reflects the strength of the model in capturing the trend and fluctuations of the Metal Ores Industry Index and, by extension, its strong forecasting performance. Figure 11 shows the distribution of residual errors, which are the difference between predicted and actual values. Most of the residuals are concentrated around zero, showing very little error in prediction. The nature of the distribution justifies the model's efficiency in providing good and accurate forecasts with little deviation.

Figure 10. Original vs. Predicted Price

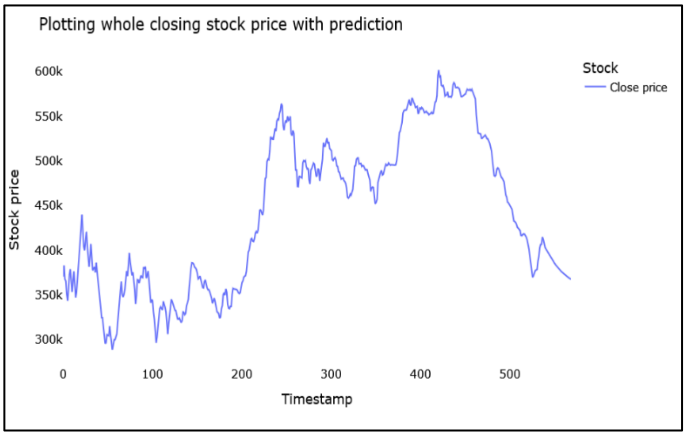

Figure 11. Full Price with Prediction

5. Conclusion and Future Direction

The study is focused on an in-depth analysis of the Metal Ores Industry index on the TSE, using the GRU model to predict closing price estimates. Historical daily price data ranges from October 2017 to October 2022, and the pre-processing for readiness of data analysis has been done very wisely. In this section, the main findings and contributions are presented as follows: The results show how effective the GRU model was in yielding good predictions for verification on the Metal Ores Industry index. This is compared well with more traditional forecasting approaches, such as that shown in the following figure, which underlines the capability of this model in financial time series forecasting. Given the recurrence nature of the GRU, it is allowed to learn short and long-term temporal dependencies in this data. Hence, this could be of considerable advantage in analyzing stock market data, which often exhibits complex and dynamic behaviors. The study provided fundamental insight into the Metal Ores

Industry index of the TSE and its importance within the Iranian market. The research gave insight into the formation of the metal ore mining group within the capital market and its substantial market value. This research pointed out the importance of deep learning representing the GRU model, which will be increasingly important in improving the accuracy of such a forecast. It, therefore, shows that the application area for deep learning methods in financial analysis is enormous.

Future Directions

Building on the insights of this research, future exploration in financial forecasting could focus on enhancing GRU model performance through hyperparameter optimization, including fine-tuning learning rates, batch sizes, and network design, as well as incorporating additional inputs like macroeconomic factors, news sentiment, and social media metrics to improve predictive accuracy. Additionally, exploring ensemble methods such as stacking or bagging, integrating real-time forecasting systems, and incorporating risk assessment and portfolio optimization within the framework could provide more reliable predictions and comprehensive investment strategies.

References

[1] Kumbure, Mahinda Mailagaha, Christoph Lohrmann, Pasi Luukka, and Jari Porras. "Machine learning techniques and data for stock market forecasting: A literature review." Expert Systems with Applications 197 (2022): 116659.

[2] Naghipour, A., A. Salehpour, and B.S. Iranag, Optimizing UPVC profile production using adaptive neuro-fuzzy inference system. International Journal of Information Technology, 2024: p. 1-18.

[3] Salehpour, Arash, Monire Norouzi, Mohammad Ali Balafar, and Karim SamadZamini. "A cloud-based hybrid intrusion detection framework using XGBoost and ADASYN-Augmented random forest for IoMT." IET Communications (2024).

[4] Salehpour, A. and K. Samadzamini, Machine learning applications in algorithmic trading: a comprehensive systematic review. International Journal of Education and Management Engineering, 2023. 13(6): p. 41.

[5] Soflaei, M.R.A.B., A. Salehpour, and K. Samadzamini, Enhancing network intrusion detection: a dual-ensemble approach with CTGAN-balanced data and weak classifiers. The Journal of Supercomputing, 2024: p. 1-33.

[6] Lazib, Lydia, Yanyan Zhao, Bing Qin, and Ting Liu. "Negation scope detection with recurrent neural networks models in review texts." International Journal of High Performance Computing and Networking 13, no. 2 (2019): 211-221.

[7] Almalaq, A. and G. Edwards. Comparison of Recursive and Non-Recursive ANNs in Energy Consumption Forecasting in Buildings. in 2019 IEEE Green Technologies Conference(GreenTech). 2019. pp.1-5

[8] Skrobek, D., et al., Implementation of deep learning methods in prediction of adsorption processes. Advances in Engineering Software, 2022. 173: p. 103190.

[9] Chandola, Deeksha, Akshit Mehta, Shikha Singh, Vinay Anand Tikkiwal, and Himanshu Agrawal. "Forecasting directional movement of stock prices using deep learning." Annals of Data Science 10, no. 5 (2023): 1361-1378.

[10] Omar, Abdullah Bin, Shuai Huang, Anas A. Salameh, Haris Khurram, and Muhammad Fareed. "Stock market forecasting using the random forest and deep neural network models before and during the COVID-19 period." Frontiers in Environmental Science 10 (2022): 917047.