

"OPTIMIZING CRUDE OIL PRICE PREDICTIONS WITH DEEP LEARNING: A COMPARATIVE PERSPECTIVE"

Prathibha P.H

Sree Sankara Vidyapeetom College, Valayanchirangara, Kerala , India

Abhay Krishna Santosh

Amrita Vishwa Vidypeetham, India

Abstract

This research investigates the performance of various recurrent neural network (RNN) models—namely Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Units (GRU), and Bidirectional GRU (Bi-GRU)—in the context of time series forecasting for crude oil price prediction. Accurate price forecasting is crucial in financial markets, offering potential benefits for investors and traders by providing valuable insights for decision-making. The study utilizes historical crude oil price data, incorporating open, close, high, and low prices, along with trading volume and other relevant indicators. After preprocessing and splitting the data into training and testing sets, each RNN model is trained to capture the temporal patterns and dependencies inherent in the price movements. The performance of each model is evaluated using key metrics, including mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE), to assess their predictive accuracy on the testing set. The results are compared to determine which RNN architecture provides the most reliable and accurate predictions, contributing valuable insights into the application of deep learning for financial forecasting tasks.

Keywords: Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Units (GRU), Bidirectional GRU (Bi-GRU), Time Series Forecasting, Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE)

1. Introduction

Time series analysis and forecasting play a crucial role in numerous fields, including finance and economics. Accurate prediction of future prices in financial markets can provide valuable insights for investors, traders, and decision-makers. In this research, we focus on the time series forecasting of crude oil prices, a key commodity with significant implications for the global economy.

To tackle this prediction task, we explore the application of recurrent neural network (RNN) models. Specifically, we compare the performance of four different RNN architectures: Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Gated Recurrent Unit (GRU), and Bidirectional GRU (Bi-GRU). These models have proven to be successful in capturing temporal dependencies and patterns in sequential data.

LSTM, a variant of RNN, introduced memory cells and gating mechanisms to overcome the vanishing gradient problem. This architecture enables the model to capture long-term dependencies effectively. Bi-LSTM extends LSTM by incorporating bidirectional processing, allowing it to consider past and future information simultaneously. GRU, on the other hand, simplifies the LSTM structure, making it more computationally efficient while retaining the ability to capture long-term dependencies. Bi-GRU combines the benefits of bidirectional processing with the simplicity of GRU, offering a balance between accuracy and efficiency. By comparing these models on the task of crude oil price prediction, we aim to identify the most suitable architecture for this specific time series dataset.

2. Methodology

This research, utilize the help of various forms of RNN, namely, LSTM, Bi-LSTM, GRU and Bi-GRU to help create and understand dependencies that are short and long term in the time series crude oil data.

3. DATA COLLECTION AND PREPROCESSING

- Gather historical crude oil price data, including open, close, high, low prices, and volume.
- Ensure the data is in a consistent time series format.
- Handle missing values, outliers, and perform any necessary data transformations.
- Split the dataset into training and testing sets.
- Reserve a portion of the data for model evaluation.

4. MODEL IMPLEMENTATION

4.1 LSTM

LSTM stands for long short-term memory networks, used in the field of Deep Learning. It is a variety of recurrent neural networks (RNNs) that are capable of learning long-term dependencies, especially in sequence prediction problems.

- Configure the number of LSTM layers and nodes per layer.
- Set appropriate activation functions, regularization techniques, and optimizers.
- Train the LSTM model using the training set.

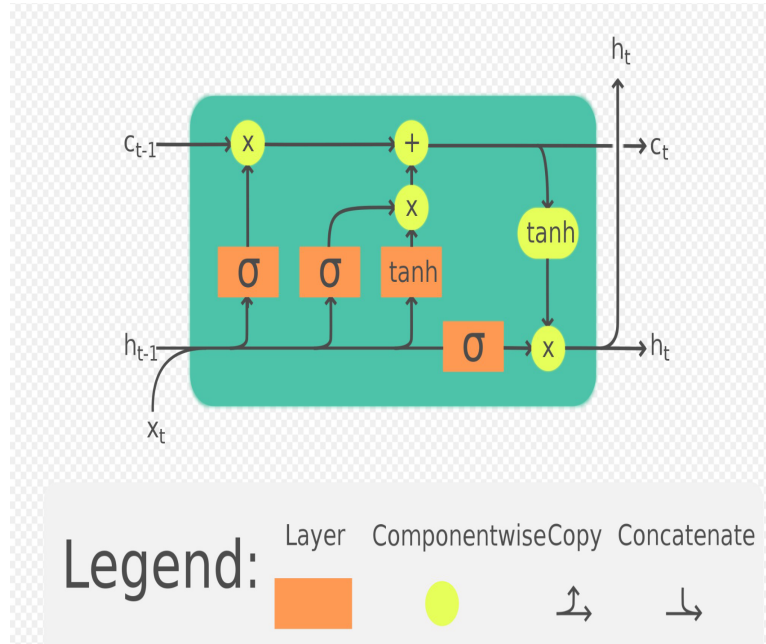


Figure 1: LSTM Architecture

4.2 Bi-LSTM

Bi-LSTM stands for Bidirectional Long Short-Term Memory, which is a type of recurrent neural network (RNN) architecture commonly used in sequence modeling tasks such as natural language processing and time series analysis.

The "bidirectional" aspect of Bi-LSTM refers to the fact that it processes input sequences in both forward and backward directions. This allows the network to effectively capture information from both past and future time steps, enabling it to learn dependencies in the entire sequence.

Bi-LSTMs are well-suited for price prediction because they can capture complex temporal patterns and dependencies in historical price data. By considering past and future context simultaneously, Bi-LSTMs can model relationships and trends that might be missed by traditional unidirectional models.

- Construct a bidirectional architecture by adding a layer of LSTM cells processing the input sequence in the reverse direction.
- Configure the remaining parameters similar to the LSTM model.
- Train the Bi-LSTM model using the training set.

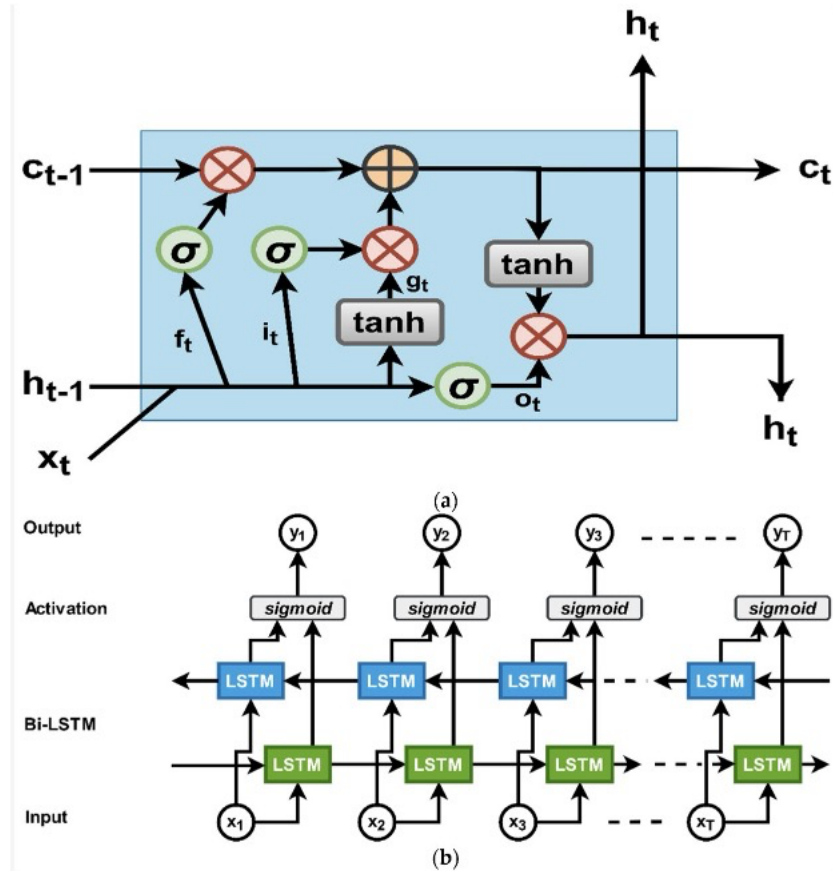


Figure 2: Bi-LSTM Architecture

4.3 GRU

GRU stands for Gated Recurrent Unit, which is another type of recurrent neural network (RNN) architecture. Similar to LSTM, GRU is designed to address the vanishing gradient problem and capture long-term dependencies in sequential data.

GRU networks have a simpler structure compared to LSTM, with fewer gates. They combine the input gate and forget gate into a single "update gate" and merge the cell state and hidden state of LSTM into a single "hidden state." This simplification reduces the number of parameters and computations required, making GRU more computationally efficient.

- Set the number of GRU layers and nodes per layer.
- Define activation functions, regularization techniques, and optimizers.
- Train the GRU model using the training set.

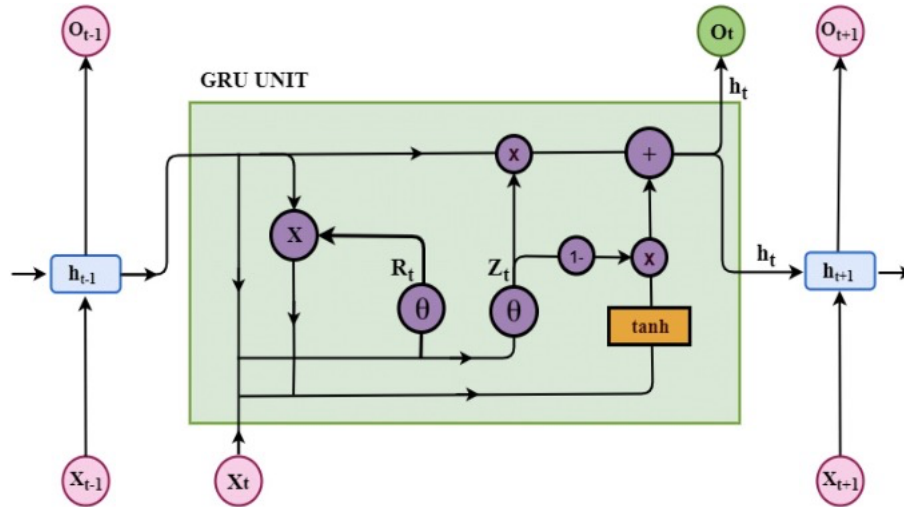


Figure 3: GRU Architecture

4.4 Bi-GRU

Bidirectional GRU (Bi-GRU) is an extension of the GRU architecture that, similar to Bi-LSTM, processes input sequences in both forward and backward directions. It combines two separate GRU layers, one processing the input sequence from the beginning to the end, and the other processing it in reverse.

- Incorporate bidirectional processing by adding a backward GRU layer to the GRU architecture.
- Configure the remaining parameters similar to the GRU model.
- Train the Bi-GRU model using the training set.

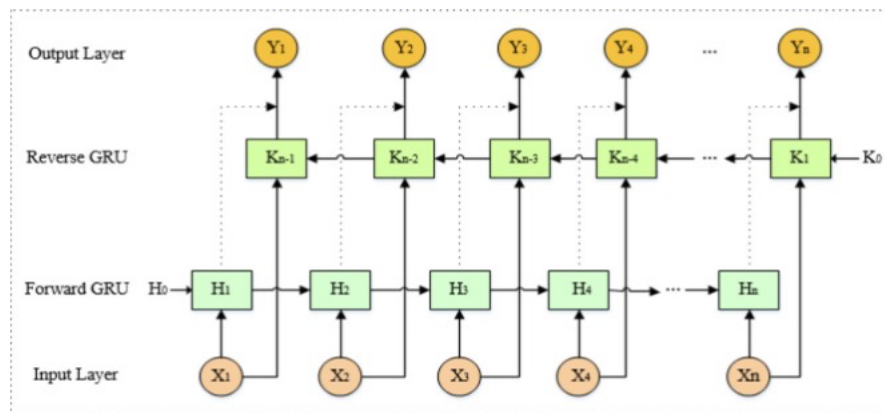


Figure 4: Bi-GRU Architecture

5. COMPARISON OF DL MODELS

5.1 LOSS COMPARISON

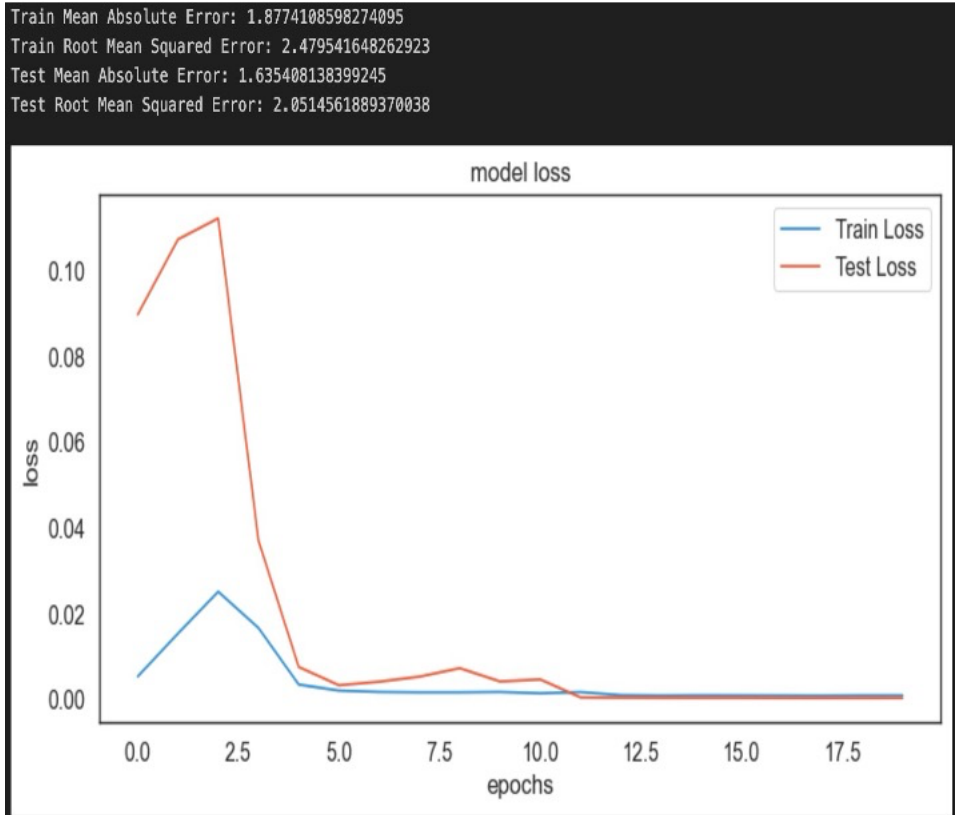


Figure 5: LSTM Train and Test Loss

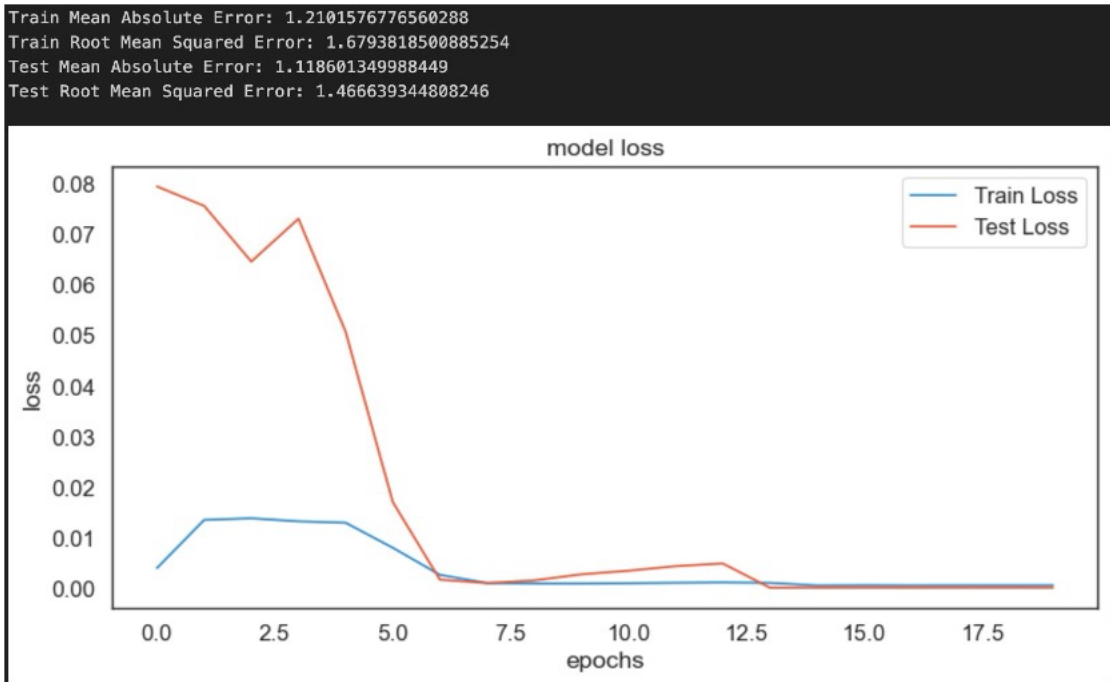


Figure 6: GRU Train and Test Loss

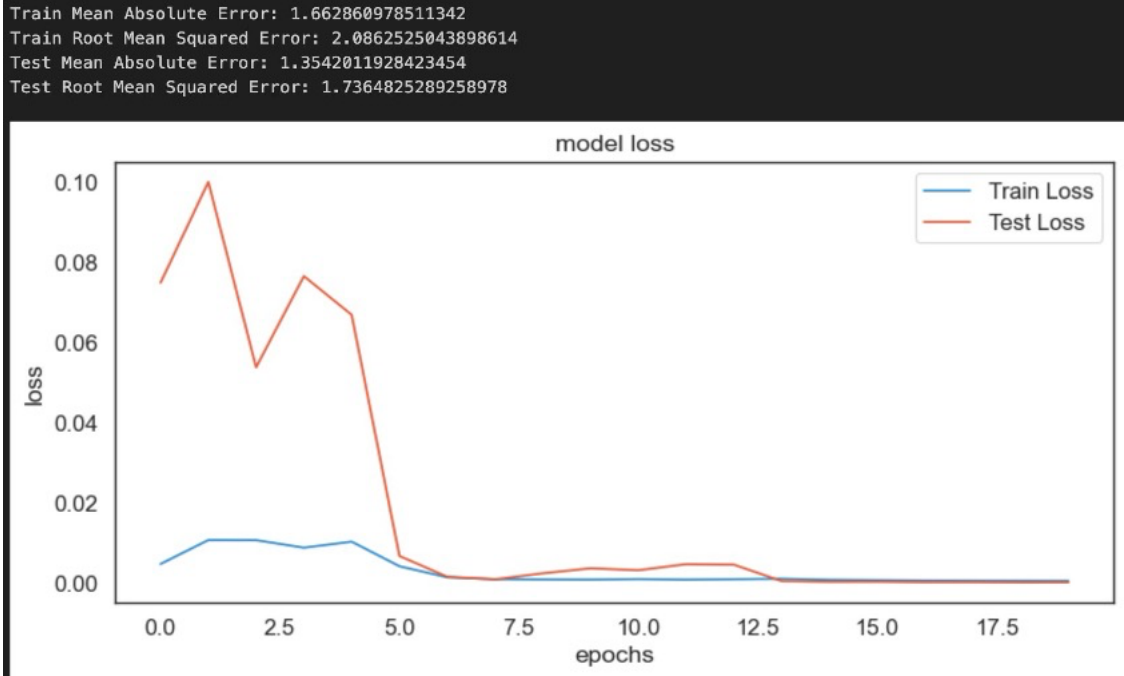


Figure 7: Bi-LSTM Train and Test Loss

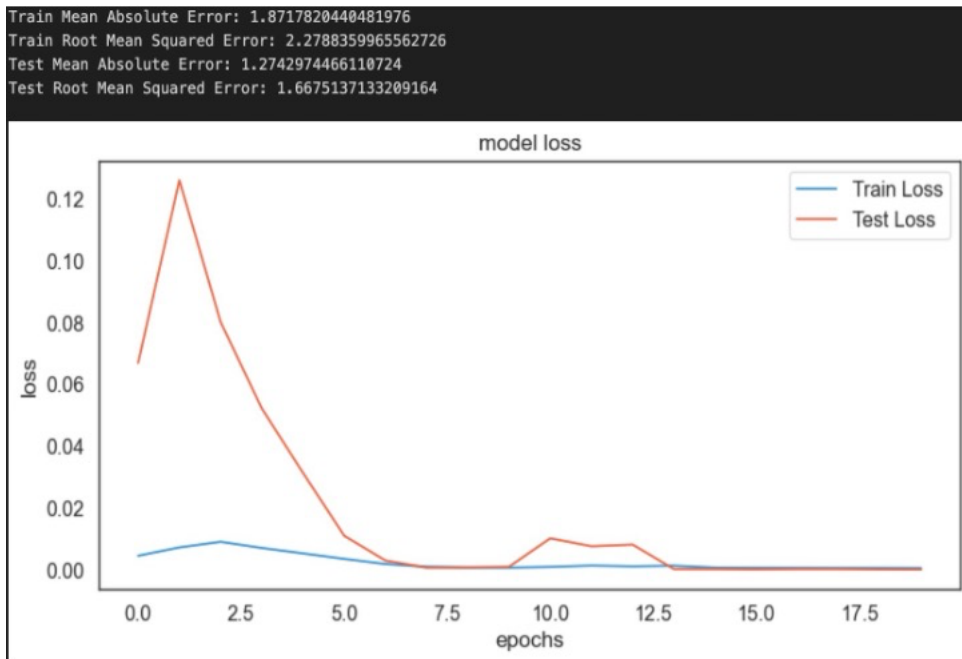


Figure 8: Bi-GRU Train and Test Loss

6. ACTUAL VS PREDICTED GRAPHS

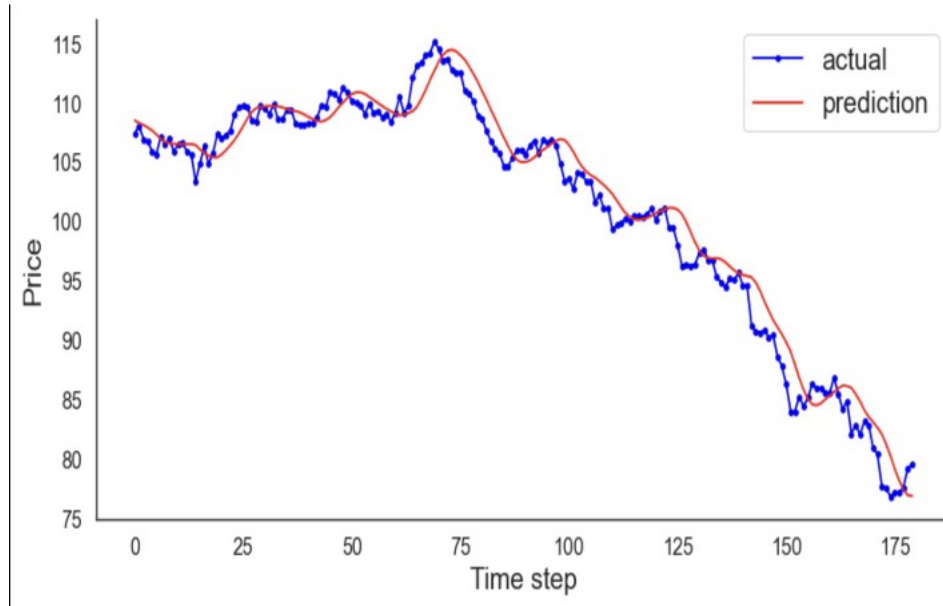


Figure 9: Graph for LSTM Actual vs Predicted

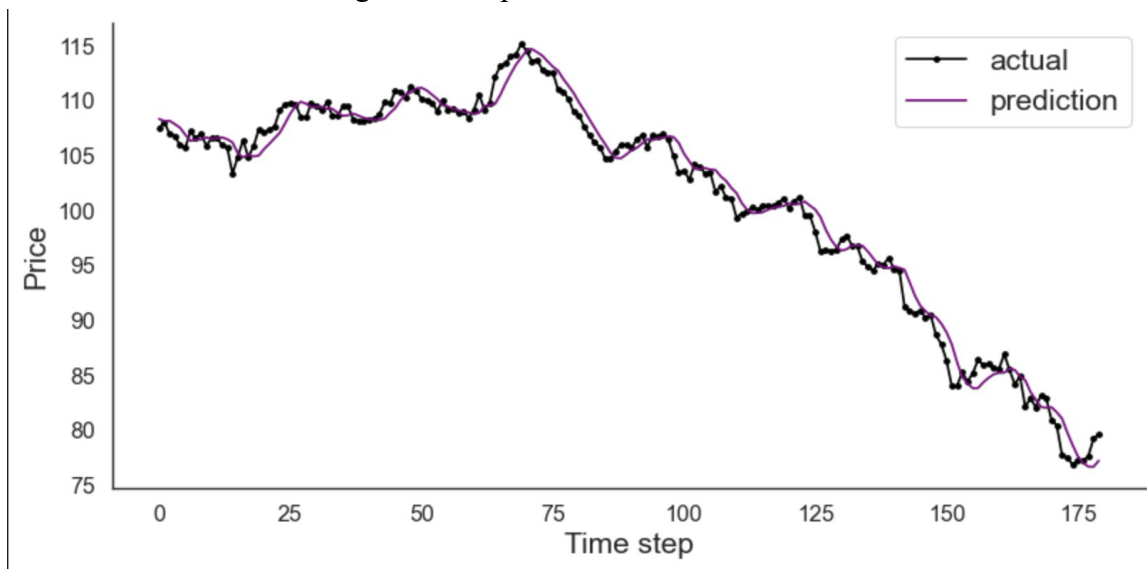


Figure 10: Graph for GRU Actual vs Predicted

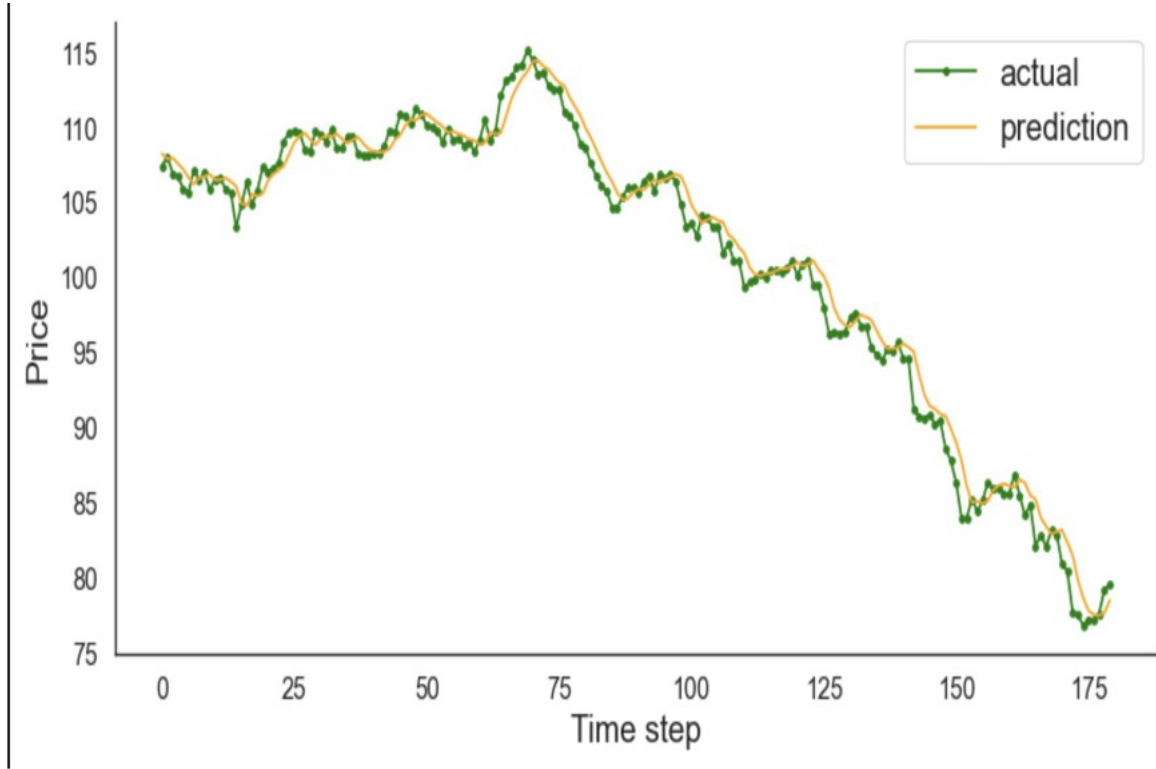


Figure11: Graph for Bi-LSTM Actual vs Predicted

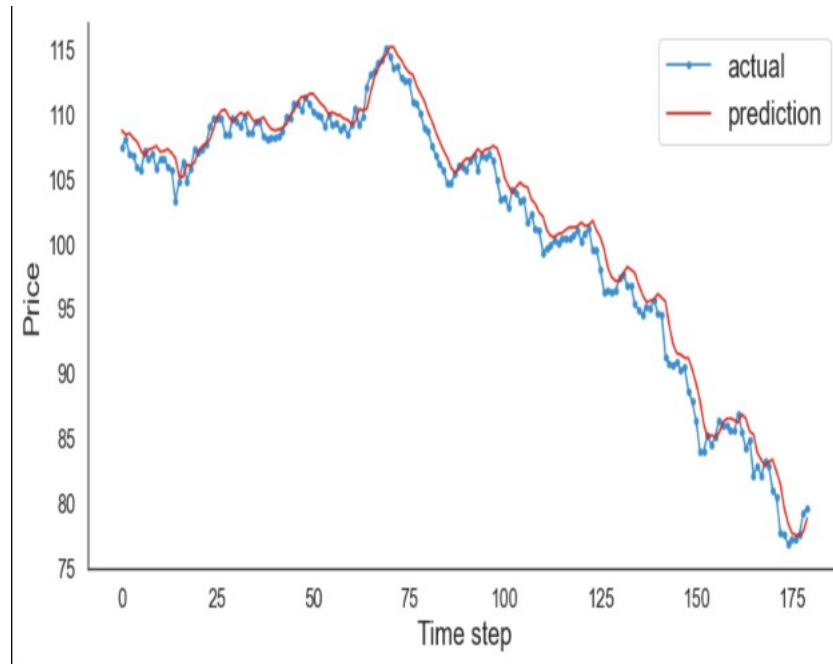


Figure 12: Graph for Bi-GRU Actual vs Predicted

7. RESULT AND CONCLUSION

From the above comparison of loss [Figure 5-8] and actual versus predicted graphs [Figure 9-12], we can infer that for the particular problem of predicting crude oil prices, which is essentially a time series problem, the best deep learning model is GRU and the least effective model is LSTM.

GRU is well-suited for price prediction tasks for several reasons:

- **Efficient Training:** The simplified structure of GRU makes it easier and faster to train compared to LSTM. It has fewer parameters to optimize, which can be beneficial when dealing with large datasets or limited computational resources.
- **Capturing Long-Term Dependencies:** GRU is capable of capturing long-term dependencies in sequential data, just like LSTM. It can learn to recognize and remember patterns that are crucial for price prediction, such as trends, seasonality, and recurring patterns.
- **Faster Convergence:** GRU networks often converge faster than LSTM networks, meaning they require fewer training iterations to reach a reasonable level of accuracy. This can be advantageous when training models on large datasets or in time-sensitive applications.
- **Interpretability:** GRU's simplified structure can make it easier to interpret and understand the learned representations. This interpretability can be useful in financial applications, as it allows analysts and researchers to gain insights into the underlying factors driving the price movements.

REFERENCES

- [1] Mehtab, S., Sen, J., & Dutta, A. (2021). Stock price prediction using machine learning and LSTM-based deep learning models. In *Machine Learning and Metaheuristics Algorithms, and Applications: Second Symposium, SoMMA 2020, Chennai, India, October 14–17, 2020, Revised Selected Papers 2* (pp. 88-106). Springer Singapore.
- [2] Seabe, P. L., Moutsinga, C. R. B., & Pindza, E. (2023). Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: A deep learning approach. *Fractal and Fractional*, 7(2), 203.
- [3] Vo, A. H., Nguyen, T., & Le, T. (2020). Brent Oil Price Prediction Using Bi-LSTM Network. *Intelligent Automation & Soft Computing*, 26(6).
- [4] Yamak, P. T., Yujian, L., & Gadosey, P. K. (2019, December). A comparison between ARIMA, LSTM, and GRU for time series forecasting. In *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence* (pp. 49-55).
- [5] Yang, Z., Tian, Y., Zhou, T., Zhu, Y., Zhang, P., Chen, J., & Li, J. (2021). Time-series deep survival prediction for hemodialysis patients using an attention-based Bi-GRU network. *Computer Methods and Programs in Biomedicine*, 212, 106458.