Time Series Forecasting Using ARMA, LSTM, and GRU Models

Aishwarya Sukumaran (22058088)

GitHub Link

1. Introduction

Time series forecasting plays a critical role in various fields like finance, healthcare, and retail. Traditional methods like ARMA (Autoregressive Moving Average) have long been used for forecasting, particularly in stationary data. However, with the advancement of machine learning, models such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Units) have gained popularity for their ability to capture complex patterns and long-term dependencies (ArunKumar et al., 2021; Zaman et al., 2021). In this report, we compare the performance of ARMA, LSTM, and GRU models using Johnson & Johnson (JJ) sales data from 1960 to 1980 and Amazon stock prices from 2018, to assess the suitability of each method in forecasting these time series datasets.

2. Background

Traditional methods like ARMA are effective for stationary data and linear patterns. These methods assume that the future values in the series depend linearly on past values and errors. Machine learning models like LSTM and GRU, on the other hand, are well-suited to capture more complex, non-linear relationships, especially when data exhibits long-term dependencies (Fauzi et al., 2021; Liu et al., 2021). LSTM addresses the vanishing gradient problem, allowing it to maintain long-term memory, while GRU simplifies the LSTM architecture, making it computationally less expensive and faster to train (Fang et al., 2021). Research indicates that machine learning models often outperform traditional methods, particularly when the data involves long-term dependencies and non-linearities (He et al., 2023). The choice of model, however, depends on the characteristics of the data being analysed.

3. Methodology

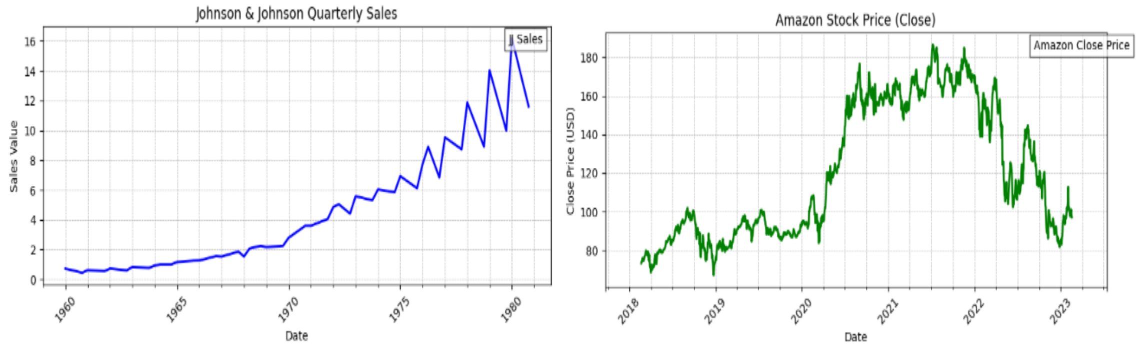
The forecasting process involved several key steps:

3.1. Data Preprocessing:

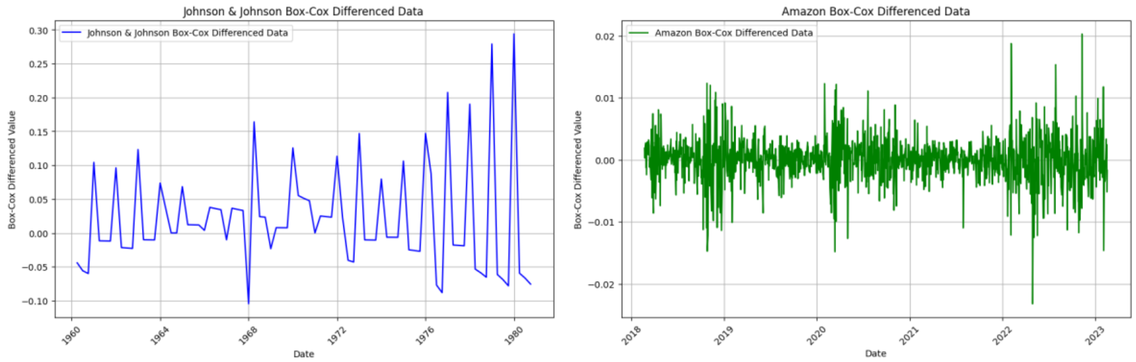
- **ADF Test Results :** The Augmented Dickey-Fuller (ADF) test was performed to check for stationarity. Both datasets were non-stationary after the Box-Cox transformation, but stationarity was achieved after differencing (Lawi et al., 2022).

Test Result	ADF Statistic	p-value	Stationarity
JJ Box-Cox Transformed	-0.3689	0.9152	Non-stationary
JJ Box-Cox Differenced	-2.9955	0.0353	Stationary
Amazon Box-Cox Transformed	-1.8461	0.3578	Non-stationary
Amazon Box-Cox Differenced	-36.7938	0.0000	Stationary

Table 1: ADF Box-cox Test results



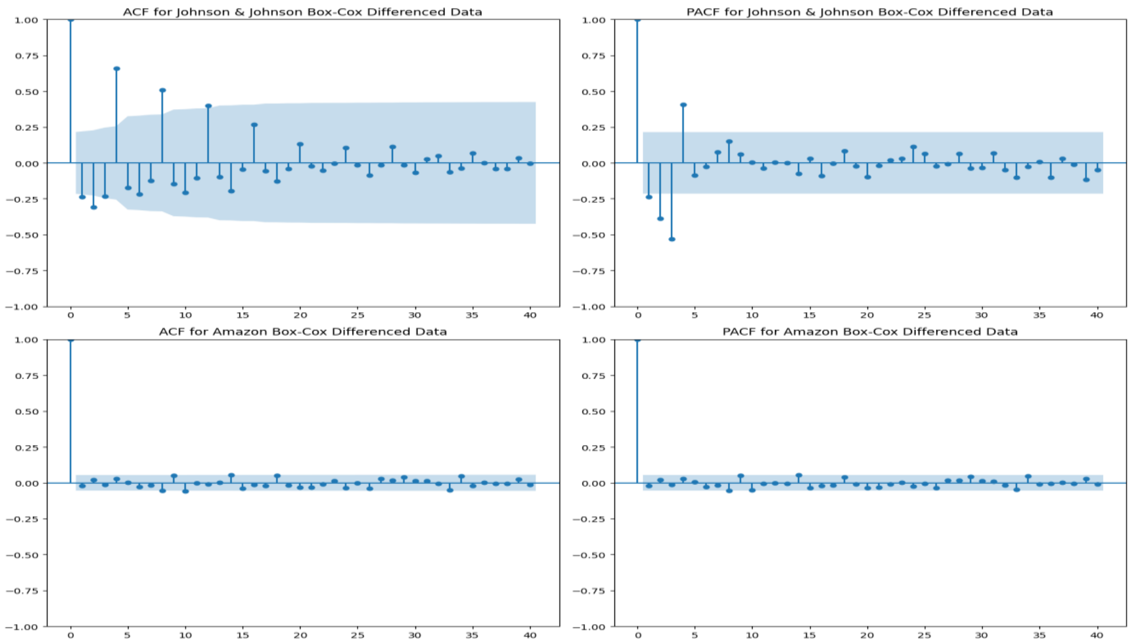
Graph 1: Original Time Series Data – Raw data for both Johnson & Johnson and Amazon stock prices.



Graph 2: Box cox Differenced Data – This graph shows the transformed data after applying Box-Cox differencing to both Johnson & Johnson sales data and Amazon stock prices, demonstrating the achieved stationarity

3.2. ARMA Model :

- The ARMA model was fitted to the stationary data. ACF and PACF plots were used for parameter selection (Di Mauro et al., 2024). After preprocessing, ARMA models were fitted to both datasets. The best manual ARMA models were selected based on AIC values.



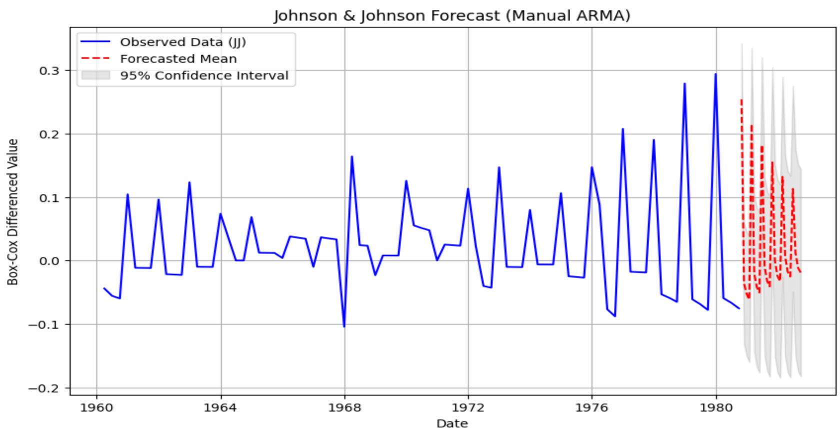
Graph 3: ACF and PACF Plots – These plots were used to select the optimal parameters for the ARMA model.

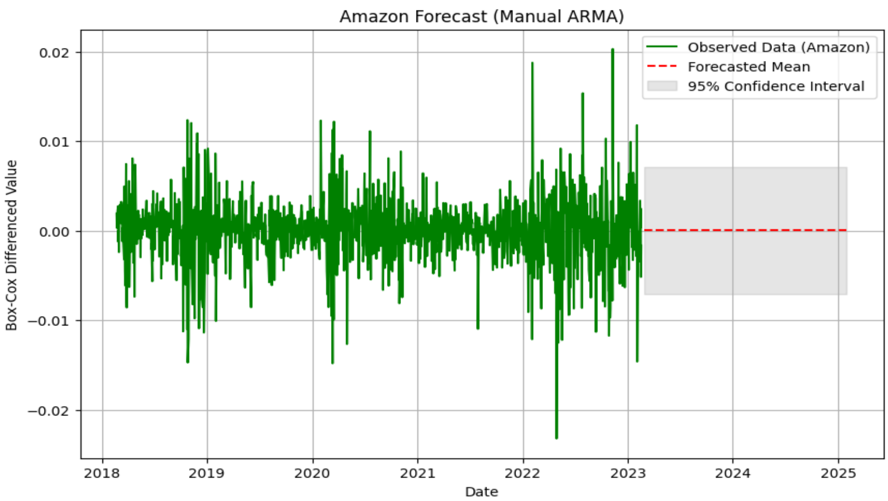
Dataset	Best ARMA Model (Order)	AIC Value (ARMA)	Best Auto ARIMA Model (Order)	AIC Value (Auto ARIMA)
Johnson & Johnson	(4, 0)	-263.611	(4, 0, 0)	-263.611
Amazon Stock	(0, 0)	-10936.959	(0, 0, 0)	-10936.959

Table 2: The table shows the best ARMA and Auto-ARIMA models selected for Johnson & Johnson and Amazon stock data, with orders based on AIC values.

ARIMA Model Results

The ARIMA models for both datasets were evaluated for their fit to the transformed and differenced data. For Johnson & Johnson sales, the ARIMA(4, 0, 0) model provided the best fit with the lowest AIC value of -263.611. For Amazon stock, the ARIMA(0, 0, 0) model was selected, but it had a relatively simple structure with a higher AIC value of -10936.959. These ARIMA models showed that simpler models could be effective for stationary data like Amazon stock, while more complex models were required for the Johnson & Johnson data.



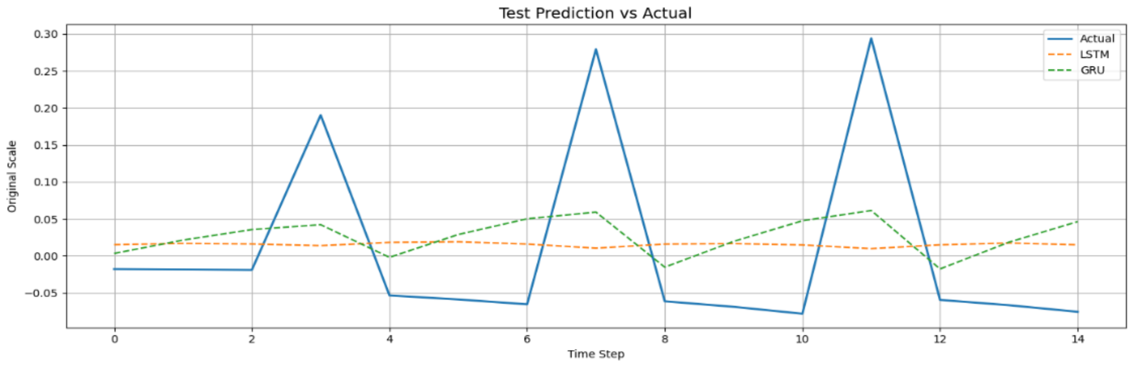


Graph 4: JJ and Amazon Forecast Plots – Forecasted values using ARMA.

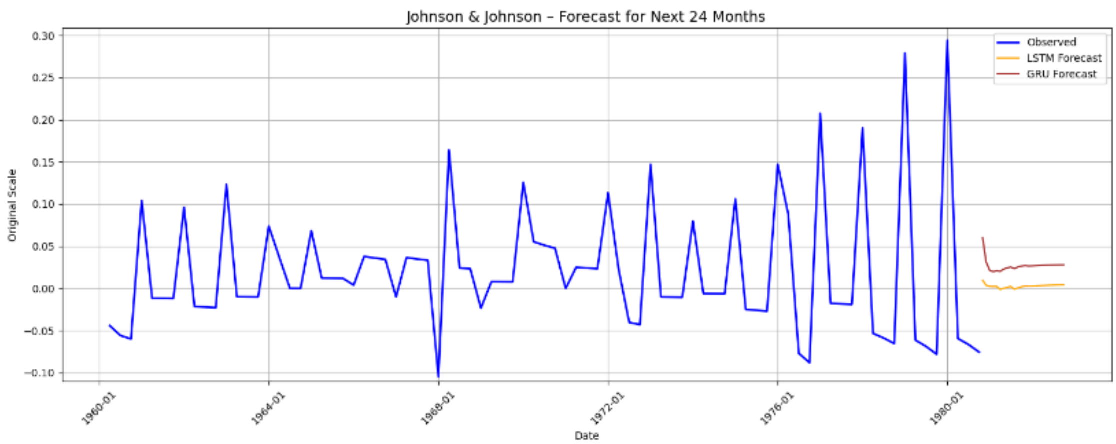
3.3. LSTM and GRU Models: : Both models were trained using the differenced data to capture the long-term dependencies (Zaman et al., 2021).

1. . **Explanation of LSTM and GRU Model Results:**

- **Johnson & Johnson Sales**:
  - **Test Results**: Both LSTM and GRU models showed similar performance on the test set. GRU had a slightly better MAE and RMSE, meaning it captured the patterns in the Johnson & Johnson sales data more effectively than LSTM. However, the difference was small.
  - **Forecast Results**: When forecasting the next 24 months, LSTM outperformed GRU slightly in terms of MAE and RMSE, but GRU showed higher variability in **MAPE**, indicating less stability in its predictions.

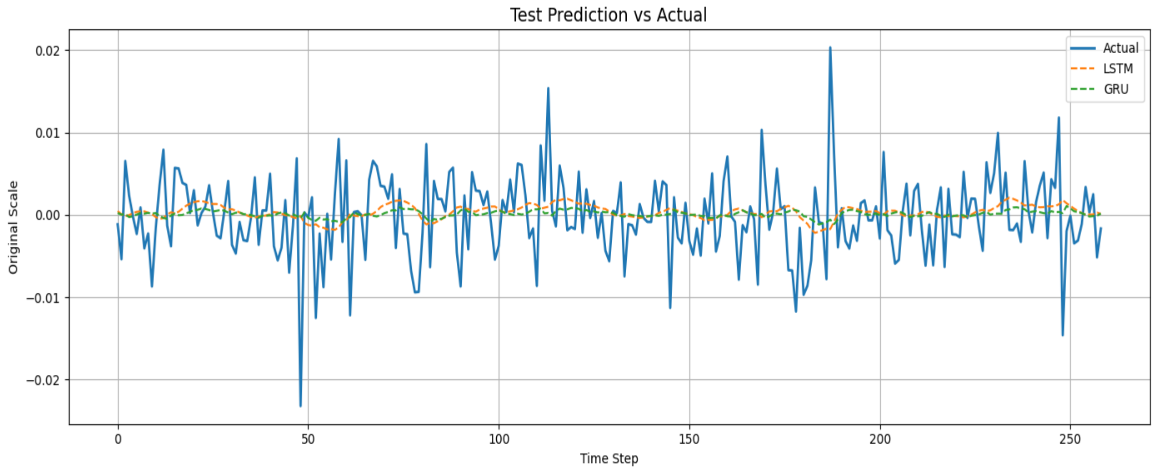


Graph 5: JJ test vs actual graph

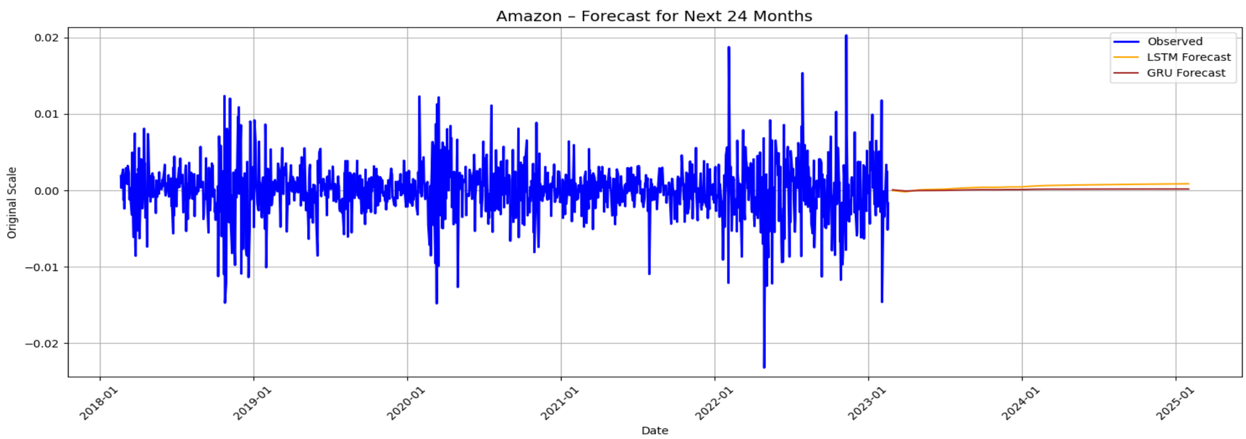


Graph 6: Amazon Forecasted values using LSTM and GRU models.

- **Amazon Stock**:
  - **Test Results**: Both models performed very similarly on the test set, with GRU outperforming LSTM slightly in terms of **MAE** and **RMSE**.
  - **Forecast Results**: Over the 24-month forecasting period, both LSTM and GRU had very similar performance, but GRU had lower **MAPE**, indicating it was more stable in its forecasts



Graph 7: Amazon test vs actual graph



Graph 8: Amazon Forecasted values using LSTM and GRU models.

4. Discussion

The results of the ARMA, LSTM, and GRU models indicate distinct advantages depending on the nature of the data.

- ARMA: This traditional model proved effective for stationary data, such as Amazon stock prices, where the linear assumptions fit well. The simplicity of ARMA provides an efficient forecasting approach with minimal computational expense.
- LSTM and GRU: Both machine learning models showed competitive performance, with GRU slightly outperforming LSTM for the Johnson & Johnson dataset. GRU's simpler architecture made it computationally faster and less prone to overfitting compared to LSTM, although LSTM showed marginally better accuracy in some cases.

In practical terms, ARMA models are best suited for simpler datasets with stationary trends, while LSTM and GRU models are better for more complex datasets with long-term dependencies, like sales data. GRU's efficiency makes it a better choice when computational resources and training time are limited.

5. Conclusion

In conclusion, the comparison of ARMA, LSTM, and GRU models revealed key insights for forecasting both Johnson & Johnson sales and Amazon stock data:

- ARMA: Ideal for stationary data like Amazon stock prices, providing accurate forecasts with lower AIC values and simpler parameterization.
- LSTM and GRU: Both models performed competitively for more complex datasets, with GRU showing slightly better forecasting performance and stability in Amazon stock data, while LSTM slightly outperformed GRU for Johnson & Johnson sales.

6. Recommendations for Future Work:

- Further tuning of hyperparameters for both LSTM and GRU models could improve forecasting accuracy.
- Exploring hybrid models, combining ARMA with machine learning techniques, could enhance the forecasting performance.
- Additional data preprocessing, such as feature engineering or including external variables, may further improve model performance.

7. References

1. Arokodare, O., 2023. Leveraging Machine Learning for Enhanced Predictive Accuracy in Time Series Forecasting: A Comparative Analysis of LSTM and GRU Models. *ResearchGate*. Available at: https://www.researchgate.net/publication/386461159_Leveraging_Machine_Learning_for_Enhanced_Predictive_Accuracy_in_Time_Series_Forecasting_A_Comparative_Analysis_of_LSTM_and_GRU_Models.
2. ArunKumar, K.E., Kalaga, D.V., Sai Kumar, C.M., Kawaji, M. and Brenza, T.M., 2021. Advances in time series forecasting. *PMC*. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC9453185/.
3. Di Mauro, M., Galatro, G., Postiglione, F., Song, W. and Liotta, A., 2024. Hybrid LSTM-GRU for Time Series Prediction. *ScienceDirect*. Available at: https://www.sciencedirect.com/science/article/pii/S138912862400118X.
4. Fang, Z., Dowe, D.L., Peiris, S. and Rosadi, D., 2021. Hybrid deep learning approaches for time series forecasting. *PMC*. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC8699874/.
5. He, B., Li, L., Bo, Y. and Zhou, J., 2023. Bi-directional LSTM-GRU Based Time Series Forecasting Approach. *ResearchGate*. Available at: https://www.researchgate.net/publication/382418154_Bi-directional_LSTM-GRU_Based_Time_Series_Forecasting_Approach.
6. Lawi, A., Mesra, H. and Amir, S., 2022. Time Series Forecasting Using Deep Learning Models. *Journal of Big Data*. Available at: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00597-0.
7. Zaman, U., Teimourzadeh, H., Hassani Sangani, E., Liang, X. and Chung, C.Y., 2021. Time series forecasting using machine learning algorithms. *IEEE Xplore*. Available at: https://ieeexplore.ieee.org/document/9621650.