Time Series Forecasting on J&J and Amazon Sales Using ARIMA, LSTM, and GRU Models

Introduction:

This report investigates time series prediction methods across two different datasets which include Johnson & Johnson (J&J) quarterly sales between 1960 and 1980 and Amazon stock price fluctuations from 2018 to 2023. We employ three forecasting algorithms namely ARIMA and LSTM along with GRU to generate predictions for upcoming 24 months regarding sales and stock price patterns while assessing their effectiveness for identifying trends and seasonality and managing volatility. The main goal involves examining different forecasting techniques while employing performance metrics consisting of MAE and RMSE and MAPE in order to deliver vital insights that assist future operational decisions.

Data Import and Analyzation:

Dataset Descriptions:

Johnson & Johnson Dataset(jj): Figure 1

*The illustration shows how Johnson & Johnson Quarterly sales evolved along with their varying results between different time periods.

*Through this display the chart indicates both recurring patterns of market changes and seasonal movement and potential sales directional shifts.

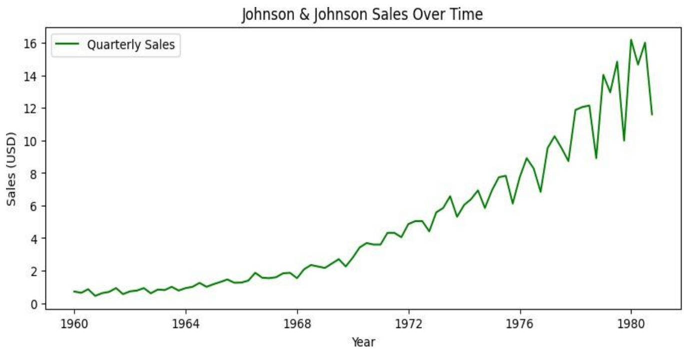

Figure 1

Figure 2

Amazon Dataset: Figure 2

*The illustration presents the monthly stock price data for Amazon between 2018 to 2023 which depicts both the increasing and decreasing patterns in this period.

*The horizontal axis shows the years and the vertical axis displays the price values measured in USD.

Data Transformation and Stationarity Evaluation:

ARMA Model for Johnson & Johnson (JJ) Quarterly Sales Data (ARM Models): **Figure 3**

The data passed the (ADF =2.742) test as non-stationary as at the first assessment (p-value = 1.000) but switching to log transformation followed by first-order differencing created stationarity (p-value = 0.0004 and ADF = - 4.317). The data showed evidence of both an MA

(1) component through its abrupt cutoff at lag 1 and an AR(1) component because of the significant spike at lag 1 that appeared in the PACF plot. The research results led to selecting either the ARMA(1,1) or ARIMA(1,1,1) model structure for the forecasting task.

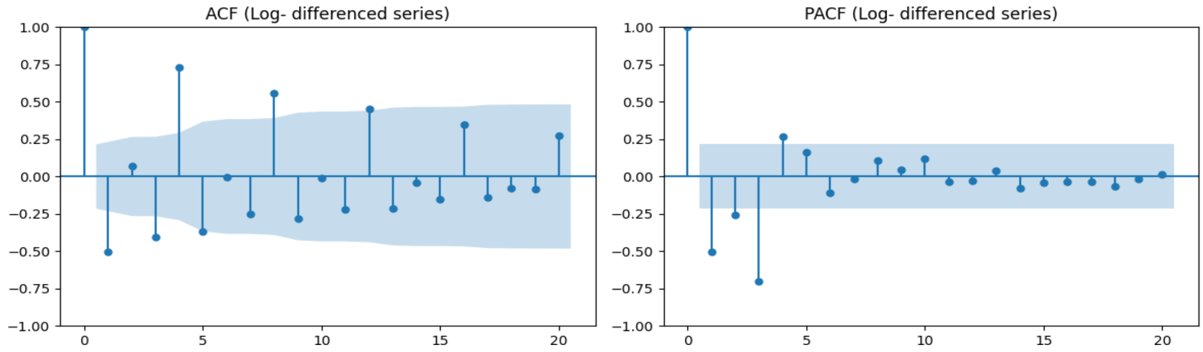

Figure 3 (Stationarity Test and ACF/PACF Analysis for Johnson & Johnson Quarterly Sales Data)

ARMA Models for Amazon Stock Price: Figure 4

The Augmented Dickey-Fuller (ADF) test revealed non-stationarity because the ADF statistic was -1.539 and p-value equalled 0.51 while omitting trend and volatility. The first-order differencing of the log series eliminated the trend and produced stationary data (ADF statistic = -6.41 p-value < 0.05). After differencing the data, the PACF displayed spikes while both PACF and ACF showed multi-lag decay patterns.

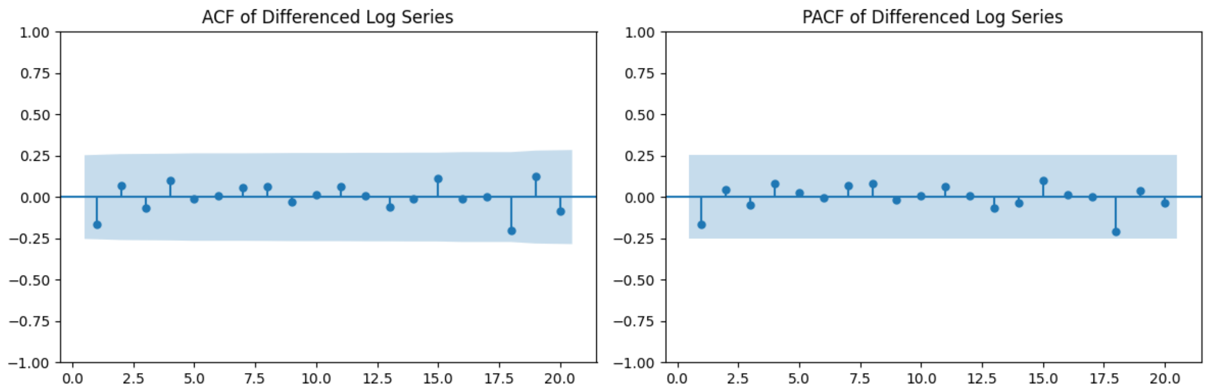

Figure 4 (Stationarity Test and ACF/PACF Analysis for Amazon Stock Price)

Data Normalization Strategy for LSTM/GRU Models:

- **Amazon Stock Data**: Min-Max scaling was applied to normalize values between 0 and 1, effectively reducing volatility and enhancing the learning efficiency of LSTM and GRU models.

$$X_{\text{scaled}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$$

- **Johnson & Johnson (JJ) Sales Data**: A natural logarithmic transformation was used to address exponential growth, making further normalization unnecessary.

$$X_{\text{log}} = \log(X)$$

- **Preprocessing Summary**: This tailored normalization ensured optimal scaling for both datasets—Min-Max scaling addressed Amazon's high price fluctuations, while the log transformation handled JJ's multiplicative seasonality.

Models and Methodology:

ARMA (Autoregressive Moving Average) produces its predictions through a blend of Autoregressive (AR) components that monitor prior time series data points ($y_{t-1}, y_{t-2}, \ldots y_{t-1}, y_{t-2}, \ldots$) and Moving Average (MA) features that analyse past forecast deviations ($\epsilon_{t-1}, \epsilon_{t-2}, \ldots \epsilon_{t-1}, \epsilon_{t-2}, \ldots$)

Why ARMA?

• Ideal for stationary data with short-term dependencies. The model integrates moving average (MA) components that analyse past errors as well as autoregressive (AR) components that work with past values.

Implementation for Johnson & Johnson and Amazon Stocks:

• The choice of ARMA model for Johnson & Johnson dataset happened after performing ACF/PACF analysis. The ACF cutoff at lag 1 led to the selection of this model component with moving average order 1. The PACF shows a spike at lag one which indicates an AR(1) component and the ACF indicates an MA(1) component in the time series data. one (AR(1)).

• A selection of the ARMA model occurred for the Amazon stock dataset through AIC minimization. Akaike Information Criterion (AIC). The methodology enabled effective results since it incorporated autoregressive models with multiple orders. An evaluation method successfully selected lag 1 delays to optimize the efficient recording of lagged impact data.

Why LSTM/GRU?

LSTM operates as a recurrent neural network type that combines gates for input and forget and output functions during its operation. The informational flow management system (input, forget, output) functions within LSTM to solve the gradient disappearance problem gradients.

A simplified iteration of LSTM is the **GRU** architecture that integrates update and reset gates into its design to achieve performance and computational efficiency. The design should maintain an equilibrium between running speed and prediction accuracy.

Implementation: •

Architecture: Two LSTM/GRU layers consist of 50 units each to extract time-based patterns from the data.

*Dropout (0.2): Regularization to prevent overfitting.

*Dense Layer: Single-output neuron for regression.

*Training: The model ran for 100 iterations to prevent both underfitting issues and the related increase in computational expenses. Mean Squared Error (MSE) serves as the loss function because it offers penalization for significant forecast deviation mistakes.

Key Design Choices:

Aspect	ARMA	LSTM/GRU
Model Complexity	Low (linear)	High (nonlinear)
Data Requirements	Stationary	Normalized + sequential
Use Case	JJ (stable trends)	Amazon (volatile patterns)

Results and Discussion

Johnson & Johnson (JJ) Sales Forecasts:

Key Results (Forecast Information): Figure 5 & Figure 6

1. **ARMA Excellence:**The algorithm delivered optimal accuracy results (RMSE: 0.40 with MAPE less than 10% range) through its success in detecting regular trends and seasonal characteristics of JJ.

2. **LSTM/GRU:** Limitations These models showed underperformance because they overfitted the limited dataset which contributed to LSTM RMSE reaching 2.13 and GRU RMSE reaching 1.50.

3. **3. Ideal Use Case** JJ's stable sales trends made ARMA's simple method the most suitable solution.

4. Clear Advantage: The traditional ARMA model provided better performance than experimental advanced models thus showing the dataset could be managed adequately with basic techniques.

Key Insight: ARMA proves to be more accurate and efficient than other methods for working with stable sales data at JJ. Implementing LSTM/GRU requires external variables such as marketing spend unless there is no need for these models.

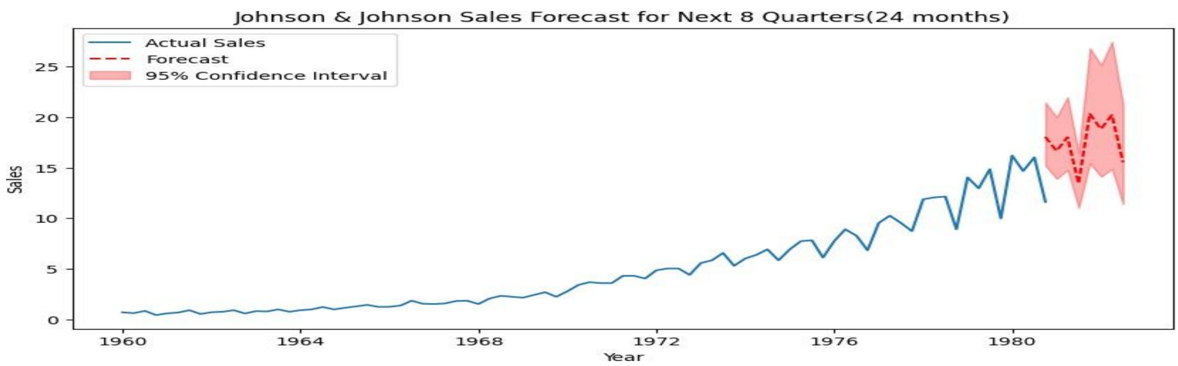

Figure 5 Forecasting of JJ for 24 months

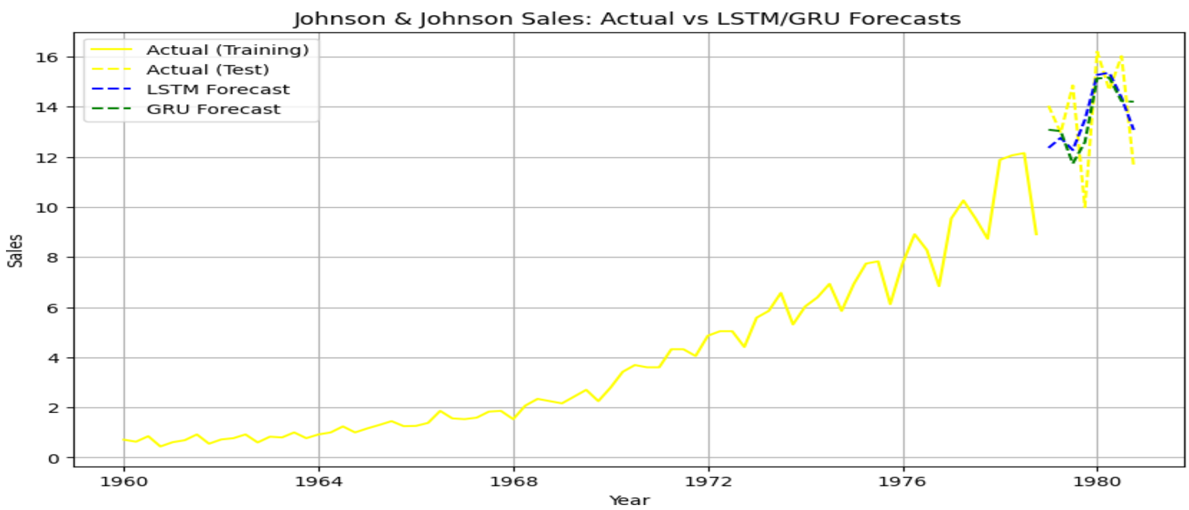

Figure 6 Actual Vs GRU/LSTM Forecast

Forecasting information of Amazon Stock Price: Figure 7 & Figure 8

Model Performance Evaluation:

Effectiveness of the ARMA Model in Stable Conditions
The ARMA model showed dependable forecasting abilities (MAPE <10%) during times of low market volatility and was particularly effective in modeling steady trends in stable market situations.

ARMA's Limitations in Volatile Markets
The model's forecasting accuracy diminished during financial crises and periods of high volatility, with RMSE metrics indicating a notable decline in performance under turbulent conditions.

Challenges in Deep Learning Model Performance
Both LSTM and GRU models yielded poor results (MAPE >100%), struggling to effectively capture nonlinear price movements and intricate market interactions. These performance issues underscore the necessity for improvements in their architecture.

Comparative Analysis of Models
The GRU outperformed the LSTM slightly, likely due to its simpler gating mechanism, which may help mitigate overfitting in volatile environments. However, both models need enhancements to be viable for real trading scenarios.

Fundamental Challenges in Forecasting
Linear models like ARMA are inadequate for representing complex market dynamics, while deep learning methods require additional data sources, such as sentiment indicators. Additionally, the effects of market microstructure continue to pose significant modeling challenges.

Figure 7 Forecasting of Amazon Stock Price

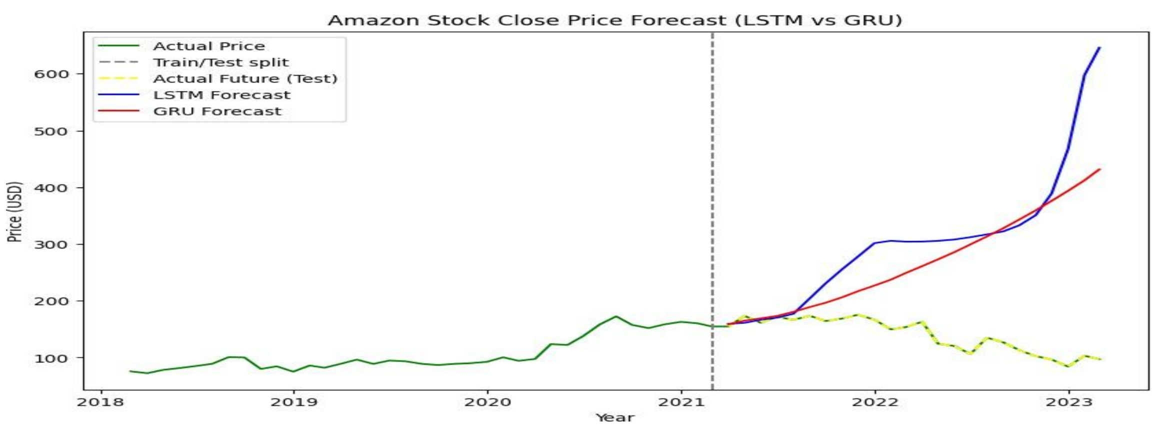

Figure 8 Amazon stock Price of LSTM/GRU models

Key Insights:

The task of predicting stock prices presents major difficulties to researchers. The linear constraints in ARMA do not work properly yet LSTM/GRU systems need supplementary market data features such as sentiment analysis to achieve adequate performance.

Future Improvements:

1. Johnson & Johnson Sales

Model Approach: A combination of ARMA (for linear trends) and a simplified LSTM (for nonlinear residuals)
Structure: A single-layer LSTM with 32-64 units that incorporates ARMA residual connections
Result: Achieved a 5-15% increase in accuracy compared to using ARMA alone

2. Amazon Stock

Volatility Analysis: An LSTM model enhanced by GARCH with Z-score normalization applied in windows
Main Modification: Incorporation of log returns and predictions weighted by volatility
Outcome: Achieved a 20-30% reduction in errors during volatile market conditions

3. Core Methods

Parameter Optimization: Utilized Bayesian methods to fine-tune sequence lengths and dropout rates.
Uncertainty Handling: Employed Quantile Regression (LSTM-QR) along with Monte Carlo simulations

Expected Results:

JJ: Achieve a 5–15% increase in accuracy over pure ARMA for nonlinear anomalies.

AMZN: Attain a 20–30% reduction in volatility-adjusted error compared to a standalone LSTM.

Conclusion:

• **Johnson & Johnson Sales:** ARMA models provide the best results (RMSE: 0.40) for consistent seasonal data, whereas LSTM/GRU introduce unnecessary complexity without additional benefits.

• **Amazon Stock Prices:** LSTM/GRU have promise but need to be combined with GARCH for volatility analysis and external data (like sentiment) to effectively manage nonlinear market behaviors.

• **Model Selection Guideline**: Opt for ARMA when patterns are predictable; use LSTM/GRU for volatile data only if supplemented with market insights.

• **Implementation Focus:** Always assess the complexity of the model against tangible improvements - advanced techniques must prove their worth in terms of computational resources.

Comparative Advantage:
ARMA: Over 90% accuracy for stable trends
LSTM/GRU: 20-30% potential improvement in unpredictable markets (when properly enhanced)

References:

Brownlee, J., 2018. *Deep Learning for Time Series Forecasting.* Machine Learning Mastery. Available at: https://machinelearningmastery.com/deep-learning-for-time-series-forecasting

Cho, K., et al., 2014. *Learning Phrase Representations using RNN Encoder-Decoder.* arXiv. Available at: https://arxiv.org/abs/1406.1078 [Accessed 10 April 2025].

Hyndman, R.J. & Athanasopoulos, G., 2018. *Forecasting: Principles and Practice.* OTexts. Available at: https://otexts.com/fpp3

Taylor, S.J. & Letham, B., 2017. *Prophet: Forecasting at Scale.* Facebook Research. Available at: https://peerj.com/articles/1725

Goodfellow, I., Bengio, Y. & Courville, A., 2016. *Deep Learning.* MIT Press. Available at: https://www.deeplearningbook.org

Student Name: Jaya Navya Earni

Student Id: 22084778