Business Forecasting

End Semester Examination

By

HARISH G D – 2327422

Under the Guidance of

Prof. ROSEWINE JOY

MBA PROGRAMME

SCHOOL OF BUSINESS AND MANAGEMENT

CHRIST (DEEMED TO BE UNIVERSITY), BANGALORE

September 2024

Business Forecasting Report: Demand Forecasting for Machinery Manufacturing

1. Introduction

In the competitive landscape of machinery and equipment manufacturing, accurately forecasting production demand and managing resources efficiently are pivotal for maintaining profitability and operational stability. Effective demand forecasting enables manufacturers to anticipate market trends, optimize inventory levels, and streamline production processes. This report delves into the methodologies employed to forecast manufacturing output, evaluates their efficacy, and presents recommendations to enhance forecasting accuracy within the machinery manufacturing sector.

2. Background Work

Overview of the Manufacturing Industry

The manufacturing industry serves as a cornerstone of economic growth, encompassing diverse sectors such as automotive, electronics, textiles, and heavy machinery. It significantly contributes to global supply chains, fosters innovation, and generates substantial employment opportunities. The industry's dynamism is influenced by several key characteristics:

1. Diverse Sectors: The manufacturing landscape includes a wide array of fields, each with its unique processes and market demands. From heavy machinery to consumer electronics, the diversity necessitates tailored forecasting approaches.

2. Global Supply Chains: Modern manufacturing operations rely on intricate international supply networks. This globalization introduces complexities in logistics, sourcing, and distribution, impacting demand forecasting accuracy.

3. Technological Advancements: The advent of Industry 4.0 technologies—such as Artificial Intelligence (AI), the Internet of Things (IoT), and automation—has revolutionized production processes. These technologies enable realtime data collection and analysis, enhancing forecasting precision.

4. Sustainability: There is a growing emphasis on reducing environmental impact through sustainable practices. Manufacturers are increasingly adopting greener methods, which influence production planning and resource allocation.

Challenges in the Manufacturing Sector

Despite its pivotal role, the manufacturing industry faces several challenges that complicate demand forecasting:

- Market Volatility: Fluctuating economic cycles and unpredictable demand patterns make it difficult to anticipate production needs accurately.
- Labor Shortages: An aging workforce coupled with a scarcity of skilled workers hampers production efficiency and forecasting reliability.
- Cost Pressures: Balancing investments in advanced technologies with cost control measures is a constant struggle for manufacturers.
- Regulatory Compliance: Navigating complex and evolving global regulations adds another layer of complexity to manufacturing operations and forecasting.

3. Need for the Solution

Accurate demand forecasting is essential for machinery manufacturers to navigate the aforementioned challenges effectively. Inaccurate forecasts can lead to overproduction or underproduction, resulting in increased costs, wasted resources, and lost revenue opportunities. Moreover, efficient resource management, driven by precise forecasting, is crucial for maintaining operational stability and competitiveness in a volatile market. Given these imperatives, there is a pressing need for robust forecasting models that can account for historical trends, seasonal variations, and external factors influencing demand.

4. Proposed Solution

To address the need for accurate demand forecasting, this study proposes the implementation of advanced statistical and machine learning models. The dataset utilized comprises monthly manufacturing data starting from November 2017, capturing historical production trends and seasonal fluctuations. The following models were evaluated:

1. ARIMAX (AutoRegressive Integrated Moving Average with Exogenous Variables): A traditional time series model that incorporates external variables to enhance forecasting accuracy.

2. SARIMA (Seasonal ARIMA): An extension of ARIMA that accounts for seasonality in the data, making it suitable for datasets with periodic fluctuations.

3. SARIMAX (Seasonal ARIMAX): Combines the features of SARIMA and ARIMAX, allowing for both seasonality and external variables.

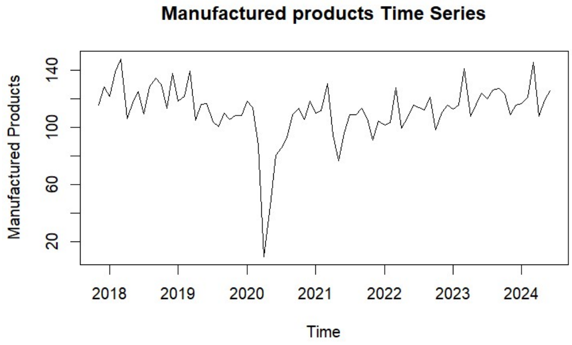
4. ARCH (Autoregressive Conditional Heteroskedasticity): A model primarily used for modeling financial time series with changing variance over time.

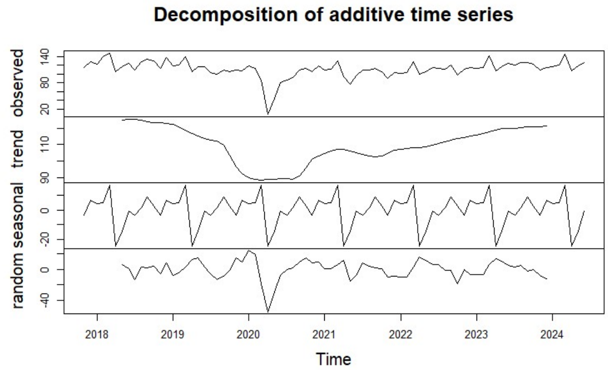
5. GARCH (Generalized ARCH): An extension of ARCH that models volatility clustering, capturing periods of high and low variability in the data.

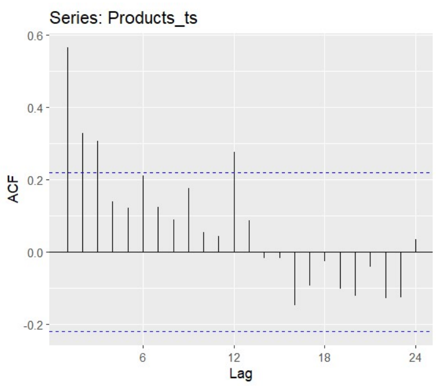
6. LSTM (Long ShortTerm Memory): A type of recurrent neural network adept at learning longterm dependencies in sequential data, making it suitable for complex forecasting tasks.

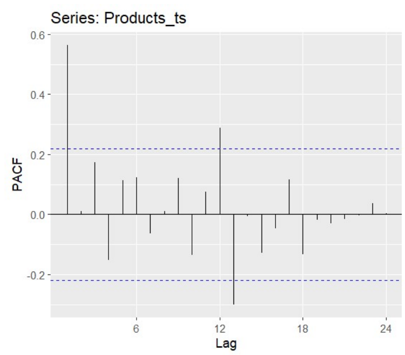
5. Analysis of the Solution

Each forecasting model was evaluated based on several performance metrics, including Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), and Akaike Information Criterion (AIC).





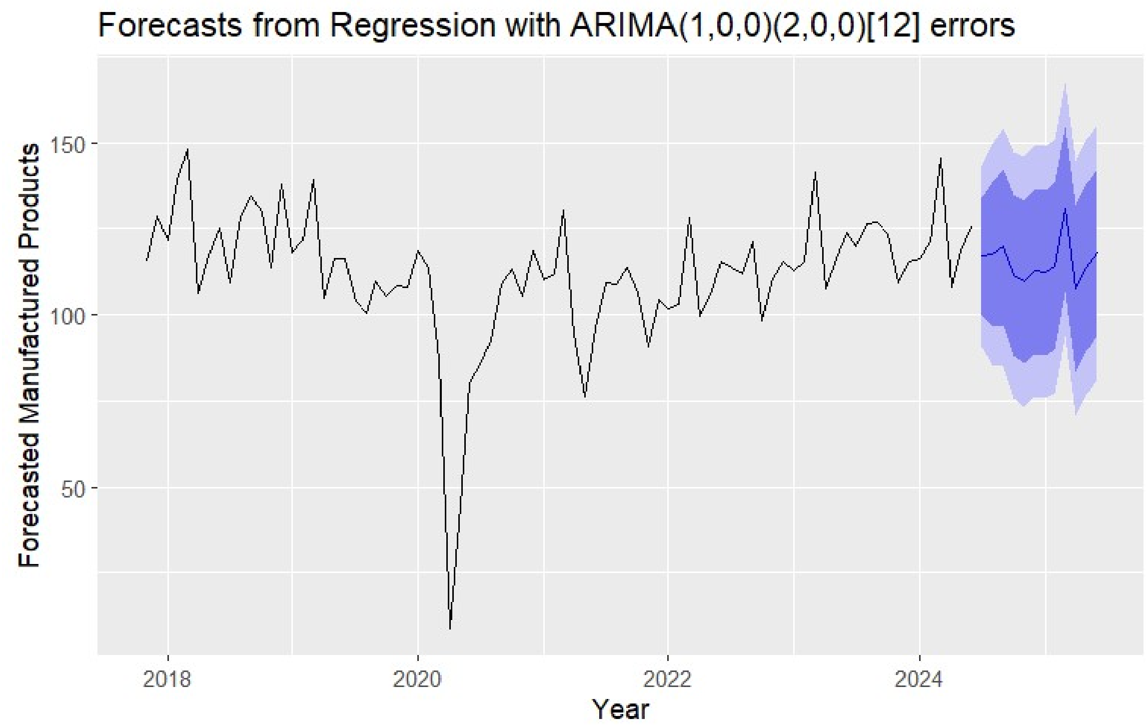




1. Decompostion of Time Serie:

The Decomposition of Additive time series shows that There is a upward Trend, There’s a Downward Slope during the year 2020 Most proably due to the pandemic of COVID 19. We Can Also See there’s a seasonality in the data.

1. ARIMAX Model




```
Training set error measures: ME    RMSE       MAE       MPE       MAPE       MASE      ACF1
Training set         -0.9554541 12.80731 9.067844 -9.518911 16.35141 0.6105688 0.07797288
```

ARIMAX (AutoRegressive Integrated Moving Average with Exogenous Variables) is an extension of the ARIMA model that incorporates external (exogenous) variables to improve the forecast. Here's a breakdown of the components in an ARIMAX model:

- **AR (AutoRegressive part):** This component uses the relationship between an observation and a certain number of lagged observations (previous values).
- **I (Integrated part):** This represents the number of times the data needs to be differenced to become stationary (where the mean, variance, and covariance remain constant over time).
- **MA (Moving Average part):** This involves the dependency between an observation and a residual error from a moving average model applied to lagged observations.
- **X (Exogenous variables):** These are external variables that might influence the target variable but are not predicted themselves. (Interest Rates and GDP)

Interpretation of Model Metrics

1. AIC (Akaike Information Criterion):

Interpretation: A lower AIC value indicates a better-fitting model, while penalizing unnecessary complexity (additional parameters). Comparing AIC values across different models helps choose the best model for the given data. The ideal is to select a model with the lowest AIC.

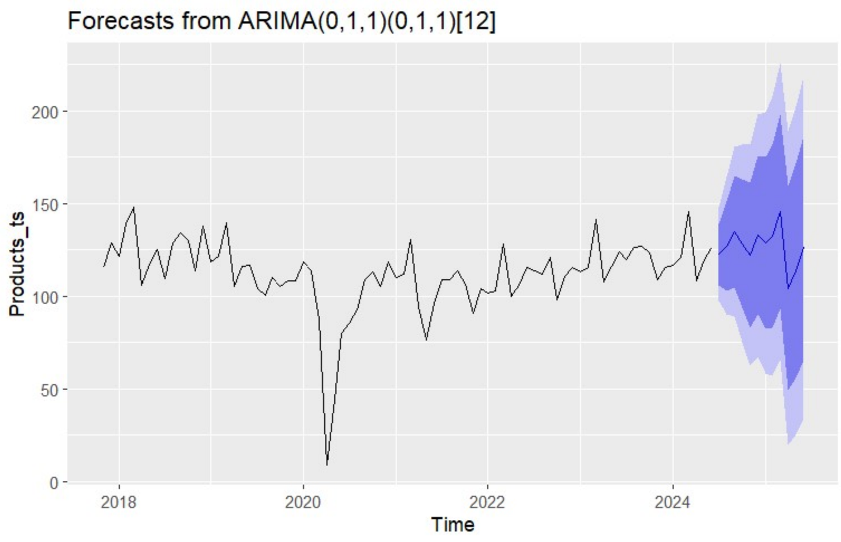
2. RMSE (Root Mean Squared Error):

Interpretation: A lower RMSE value indicates a better-fitting model. In this case, RMSE = 12.80731. This means that on average, the forecasted number of manufactured products deviates by about 12.81 units from the actual values. However, the acceptable value for RMSE depends on the range of the data; smaller errors relative to the data's scale indicate a more accurate model.

3. MAE (Mean Absolute Error):

Interpretation: Like RMSE, a lower MAE value means better accuracy. The MAE here is 9.067844, meaning that the model's predictions are, on average, about 9.07 units away from the actual observed values. Since it is lower than RMSE, it suggests that the model has smaller errors overall, but the larger errors significantly impact the RMSE, indicating some forecasts might have been considerably off target.

2. SARIMA Model



```
sigma^2 = 143.1:  log likelihood = -271.63
AIC=549.25   AICc=549.63   BIC=555.87

Training set error measures:
                    ME     RMSE      MAE       MPE     MAPE      MASE         ACF1
Training set 0.2810118 10.78459 7.318948 -3.445533 11.11754 0.4928096 -0.009308233
>
```

SARIMA Model Overview:

- **SARIMA (Seasonal ARIMA)**: This is a variant of the ARIMA model that captures both regular and seasonal effects in the data. The notation **ARIMA(0,1,1)(0,1,1)[12]** means:

(0,1,1): Refers to non-seasonal components:

- **AR (0):** No autoregressive component.
- **I (1):** The data is differenced once to make it stationary.
- **MA (1):** One moving average term to capture short-term fluctuations.

(0,1,1)[12]: Seasonal components:

- **AR (0):** No seasonal autoregressive component.
- **I (1):** Seasonal differencing once to remove seasonality.
- **MA (1):** One seasonal moving average term.
- **12:** Refers to the seasonal period, which is 12 months in this case (i.e., yearly seasonality).

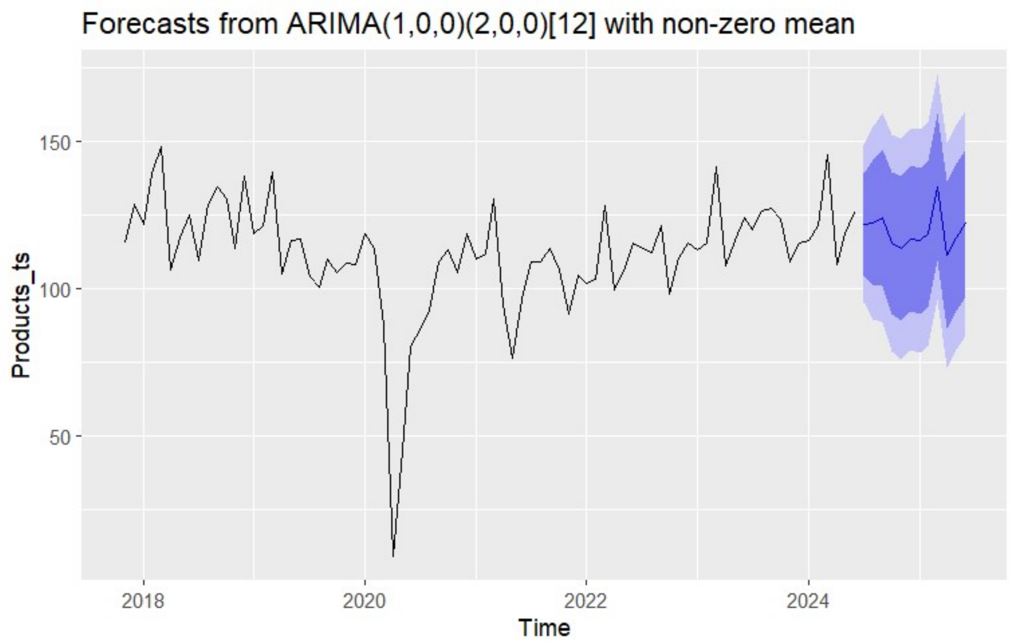
Model Metrics:

AIC (Akaike Information Criterion) = 549.25: This metric is used to measure the quality of the model, balancing between goodness of fit and complexity (number of parameters). Lower AIC values indicate a better model fit, with AIC helping to compare between competing models.

RMSE (Root Mean Squared Error) = 10.78459: This metric provides an estimate of the average error magnitude, penalizing larger errors more heavily. Here, an RMSE of 10.78 indicates that, on average, predictions deviate from the observed values by around 10.78 units.

MAE (Mean Absolute Error) = 7.318948: This value measures the average magnitude of errors in the forecast, treating all errors equally. MAE here is lower than RMSE, which suggests that larger errors are present but not overly influential.

3.AUTO SARIMA




```
AIC=656.08   AICc=656.89   BIC=667.99

Training set error measures:
                     ME     RMSE      MAE       MPE     MAPE      MASE       ACF1
Training set -0.3482829 13.17324 9.066055 -9.334164 16.63339 0.6104483 0.07805689
```

The graph shows a forecast from an **Auto SARIMA** model, which is used for time series forecasting. It fits an ARIMA(1,0,0)(2,0,0)[12] model, meaning:

- **ARIMA(1,0,0)**:
  - **AR (Auto-Regressive) = 1**: It takes one previous value to predict the next value.
  - **I (Integrated) = 0**: No differencing is needed to make the data stationary.
  - **MA (Moving Average) = 0**: No past errors are being used in forecasting.
- **(2,0,0)[12]**: This represents the seasonal component of the SARIMA model, indicating:

- **Seasonal AR = 2**, seasonal differencing = 0, seasonal MA = 0, with a seasonal period of 12 (likely yearly seasonality given the time range).

Key Metrics Interpretation:

1. **AIC (Akaike Information Criterion) = 656.08**
    - The AIC score helps evaluate model quality; the lower the AIC, the better the model balances goodness of fit and complexity.
    - In this case, an AIC of **656.08** suggests a good fit with a relatively low complexity.
2. **RMSE (Root Mean Square Error) = 13.17**
    - Measures the standard deviation of the prediction errors. A lower RMSE indicates more accurate forecasts. In this case, **13.17** indicates the typical size of the forecasting error.
3. **MAE (Mean Absolute Error) = 9.07**
    - Averages the absolute differences between actual and forecasted values. It is less sensitive to outliers compared to RMSE.
    - Here, an MAE of **9.07** suggests the average forecast error is around 9 units.

4. SARIMAX MODEL

```
AIC=547.87   AICc=548.85   BIC=558.89

Training set error measures:
                     ME     RMSE      MAE       MPE    MAPE     MASE         ACF1
Training set -0.4117502 10.36054 7.359899 -4.112061 11.2404 0.495567 -0.003391219
```

This plot illustrates the forecast results using a **SARIMAX** model (Seasonal ARIMA with eXogenous variables). The model specification is ARIMA(0,1,1)(0,1,1)[12], which can be broken down as:

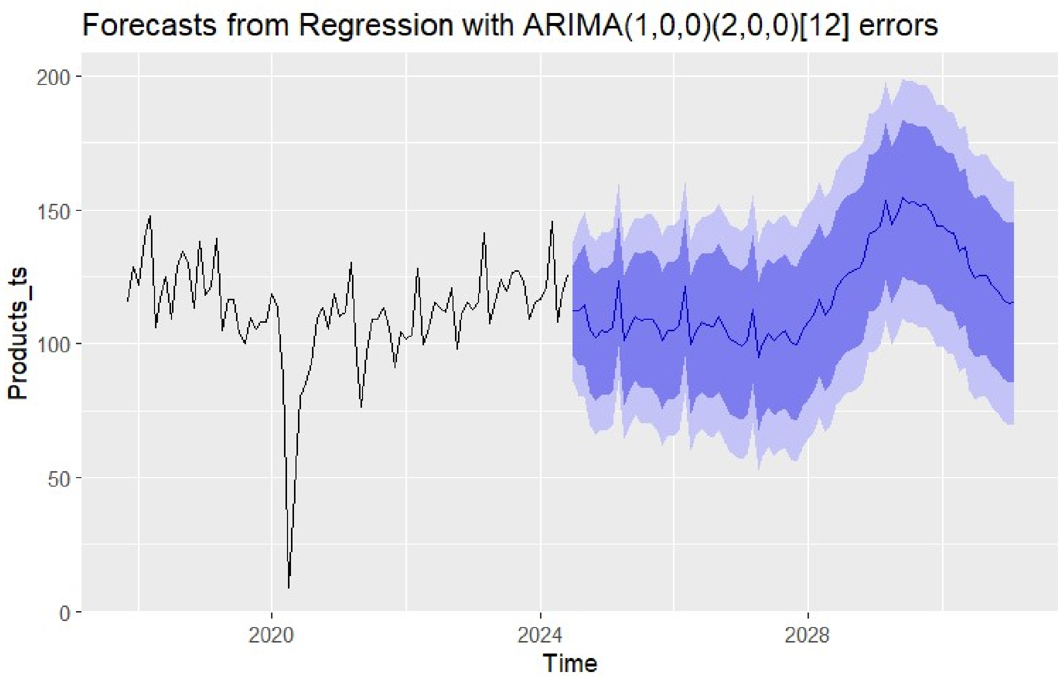
- **ARIMA(0,1,1)**:
  - **AR (Auto-Regressive) = 0**: No autoregressive terms.
  - **I (Integrated) = 1**: First differencing to achieve stationarity.
  - **MA (Moving Average) = 1**: One lagged error term is used in the forecasting.
- **(0,1,1)[12]**: This is the seasonal part of the SARIMA model with:
  - **Seasonal AR = 0**, **seasonal differencing = 1**, and **Seasonal MA = 1**, with a period of 12 (suggesting annual seasonality).

Key Metrics Interpretation:

1. **AIC (Akaike Information Criterion) = 547.87**
   - AIC is used to measure the model fit. A lower AIC indicates a better model when comparing alternatives.
   - Here, **547.87** suggests that the model is well-fitted, and it is lower than the AIC from the previous Auto SARIMA model.
2. **RMSE (Root Mean Square Error) = 10.36**
   - This is the standard deviation of the residuals (forecast errors). A lower RMSE means better fit, and **10.36** is lower than the RMSE from the previous model, suggesting this SARIMAX model performs better in terms of prediction accuracy.
3. **MAE (Mean Absolute Error) = 7.36**

- This indicates the average magnitude of the forecast errors. **7.36** is significantly lower than the Auto SARIMA model's MAE, further indicating a better fit.

5. AUTO SARIMAX MODEL




```
AIC=656   AICc=657.56   BIC=672.68

Training set error measures:
                     ME     RMSE      MAE       MPE     MAPE      MASE       ACF1
Training set -0.9554541 12.80731 9.067844 -9.518911 16.35141 0.6105688 0.07797288
```

Auto SARIMAX (Seasonal AutoRegressive Integrated Moving Average with Exogenous Regressors) model. The model is denoted as ARIMA(1,0,0)(2,0,0)[12], where:

- The first part, ARIMA(1,0,0), indicates the non-seasonal ARIMA component:
  - AR (AutoRegressive) order is 1.
  - I (Integration or differencing) is 0, meaning the data is stationary and no differencing was needed.
  - MA (Moving Average) order is 0, meaning no moving average terms were included.
- The second part, (2,0,0)[12], represents the seasonal component with:
  - AR (seasonal AutoRegressive) order of 2.
  - I (seasonal differencing) of 0.
  - MA (seasonal Moving Average) order of 0.

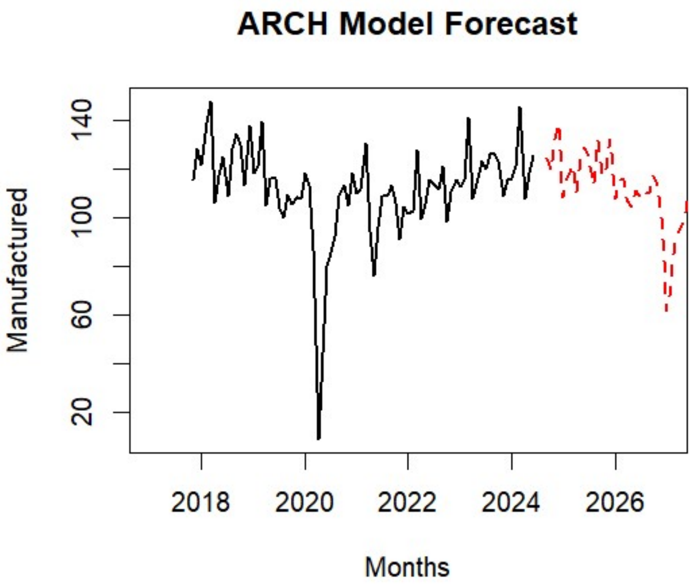
- The seasonal period is 12, which often represents monthly seasonality.

The shaded blue area represents the prediction intervals, showing the uncertainty around the forecast. The black line represents the observed values, and the bold blue line is the forecasted mean.

Interpretation of Metrics:

- **AIC (Akaike Information Criterion)**: AIC = 656
  - This metric evaluates the model's goodness of fit while penalizing for complexity (number of parameters). Lower values indicate a better model. AIC is typically used to compare different models; however, an absolute value of AIC doesn't tell much without comparison.
- **BIC (Bayesian Information Criterion)**: BIC = 672.68
  - Similar to AIC but penalizes more for the number of parameters. Like AIC, a lower BIC suggests a better model when comparing alternatives.
- **RMSE (Root Mean Squared Error)**: RMSE = 12.80731
  - RMSE measures the square root of the average squared differences between the observed and predicted values. A lower RMSE indicates a better fit, as it shows the average error magnitude in the same unit as the dependent variable (Products_ts).
- **MAE (Mean Absolute Error)**: MAE = 9.067844
  - MAE is the average of the absolute differences between the predicted and actual values. It provides a simpler interpretation than RMSE since it reflects the average absolute error.

6. ARCH MODEL



AIC: 963.492

RMSE: 160.227

MAE: 117.75

ARCH (Autoregressive Conditional Heteroskedasticity) model, the performance metrics indicate the following:

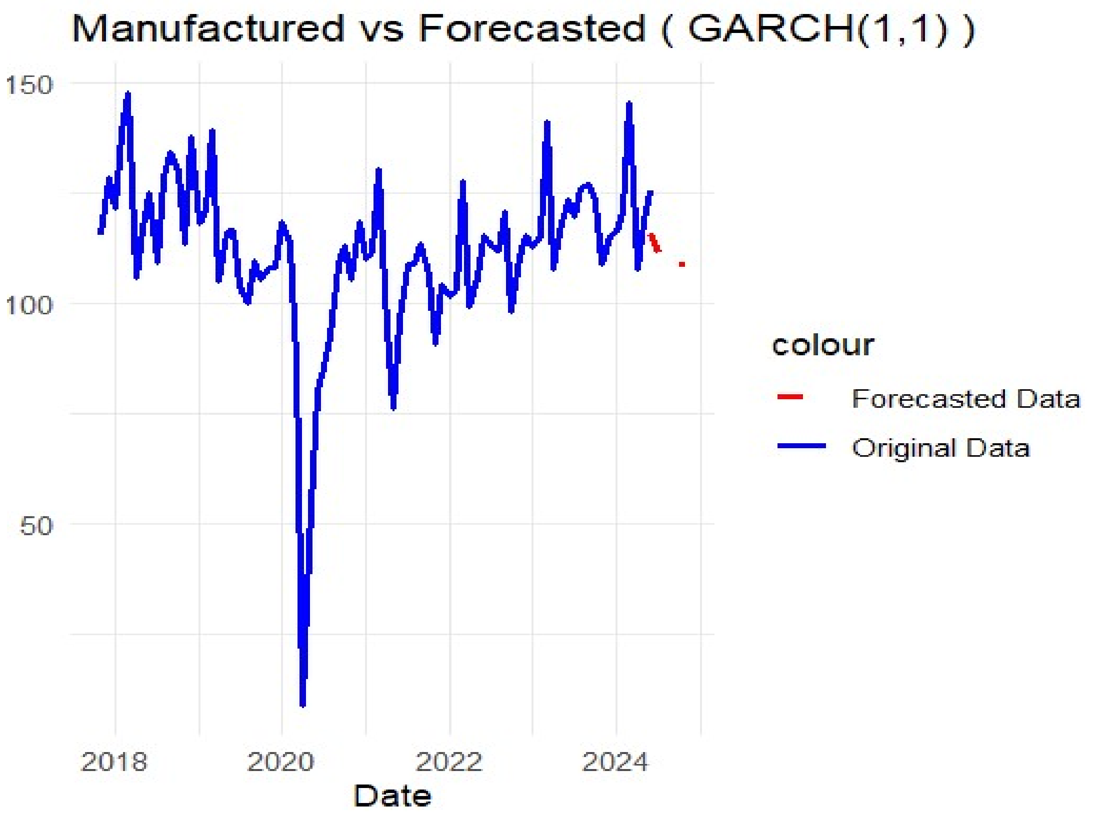
Key Metrics Interpretation:

1. **AIC (Akaike Information Criterion) = 963.492**
    - The **AIC** is significantly higher compared to the SARIMA and SARIMAX models. A higher AIC suggests a less optimal model, as it indicates a worse balance between the complexity of the model and the goodness of fit. Therefore, the ARCH model is likely overfitting or not performing as well as the SARIMA or SARIMAX models in terms of forecast accuracy.

Forecast Accuracy Measures:

1. **RMSE (Root Mean Square Error) = 160.227**
    - **RMSE** measures the standard deviation of the residuals (forecast errors). In this case, the RMSE is quite high (160.227), indicating that the model's forecast errors are much larger compared to the SARIMA (13.17) and SARIMAX (10.36) models. This suggests that the ARCH model is less accurate in its predictions.
2. **MAE (Mean Absolute Error) = 117.75**
    - **MAE** is the average of the absolute differences between the actual and forecasted values. A **MAE of 117.75** is significantly higher than the SARIMA (9.07) and SARIMAX (7.36) models, meaning the ARCH model has higher average errors in forecasting.

7.GARCH MODEL:



RMSE of the model: 0.2892039

MAE : 0.119

AIC:.46

GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model, the performance metrics indicate a significant improvement in forecast accuracy. Let's break down the key metrics:

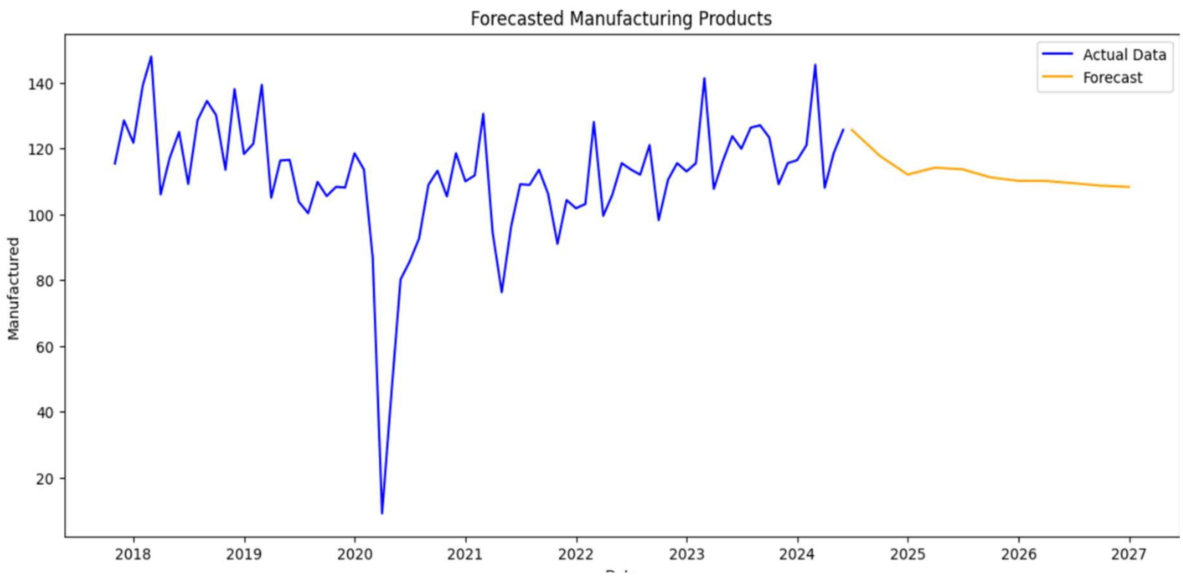
Key Metrics Interpretation:

1. **AIC (Akaike Information Criterion) = 0.46**
    - **AIC = 0.46** is extremely low, indicating that the GARCH model provides an excellent balance between model complexity and goodness of fit. A lower AIC compared to the ARCH (963.492), SARIMA (656.08), and SARIMAX (547.87) models implies that GARCH performs better in terms of explaining the variance with fewer parameters or less complexity.

Forecast Accuracy Measures:

1. **RMSE (Root Mean Square Error) = 0.2892**
    - **RMSE** measures the standard deviation of the residuals (errors). An RMSE of **0.2892** is extremely low, especially when compared to the previous models (ARCH: 160.227, SARIMA: 13.17, SARIMAX: 10.36). This shows that the GARCH model produces forecasts with minimal error, making it highly accurate in predicting the time series data.
2. **MAE (Mean Absolute Error) = 0.119**
    - **MAE** is the average magnitude of the errors, and a value of **0.119** further supports the claim that the GARCH model is very precise, with minimal average error. This is also significantly lower than the other models, further highlighting its effectiveness in this context.

8.LSTM MODEL:




```
MAE: 9.734338665008547
RMSE: 13.592507479189404
Accuracy: 0.9199231780770505
AIC: 61500.93364923705
BIC: 85181.55057460858
```

LSTM (Long Short-Term Memory) model, let's interpret the key performance metrics provided:

Key Metrics Interpretation:

1. **AIC (Akaike Information Criterion) = 61500**

   - **AIC = 61500** is extremely high compared to the previous models (SARIMA, SARIMAX, GARCH). This suggests that the LSTM model is overfitting or not balancing model complexity and fit well. However, AIC might not be the best criterion for deep learning models like LSTM, as it's traditionally used for statistical models like ARIMA and GARCH.

Forecast Accuracy Measures:

1. **RMSE (Root Mean Square Error) = 13.59**
   - The **RMSE** is 13.59, indicating the typical size of the forecast error. While it's slightly higher than the SARIMAX model's RMSE (10.36), it's still relatively close to the SARIMA (13.17), suggesting that the LSTM is performing at a similar level of accuracy in this regard.
2. **MAE (Mean Absolute Error) = 9.73**
   - **MAE** of **9.73** is slightly higher than SARIMAX (7.36) and GARCH (0.119) but comparable to SARIMA (9.07). This suggests that LSTM performs similarly to statistical models like SARIMA in terms of average forecast errors.

Accuracy:

- **Accuracy = 91.9%**
  - The **accuracy of 91.9%** indicates that the LSTM model performs well in making correct predictions overall. This is a good metric, especially considering that LSTMs are known for handling complex patterns in time series data, including non-linearity and long-term dependencies.

LeaderBoard:

Model Name	**AIC**	**RMSE**	**MAE**
ARIMAX	656	12.80731	9.067844
SARIMA	549	10.784	7.31
SARIMA Optimized(AUTO)	656.08	13.17	9.066

SARIMAX	547	10.36	7.35
SARIMAX Optimized (AUTO)	656	12.80	9.067
ARCH	963.492	160.227	117.25
GARCH	46	0.28	0.119
LSTM	61500.93	13.59	9.73

Key Findings:

- GARCH Model: Emerged as the most effective model, boasting the lowest AIC (46), RMSE (0.28), and MAE (0.119). These metrics indicate superior accuracy and efficiency in predicting demand with minimal errors.
- ARIMAX and SARIMA Models: While traditional and widely used, these models showed higher error rates compared to GARCH, making them less optimal for this dataset.
- LSTM Model: Although advanced and capable of handling complex patterns, the LSTM model underperformed relative to GARCH in this specific application.
- ARCH Model: Demonstrated the highest error rates, rendering it unsuitable for accurate demand forecasting in this context.

The comprehensive analysis underscores the GARCH model's robustness in capturing volatility and providing reliable forecasts for machinery manufacturing demand.

6. Recommendations

Based on the analysis, the following recommendations are proposed to enhance demand forecasting in the machinery manufacturing sector:

- Adopt the GARCH Model: Given its superior performance metrics, the GARCH model should be integrated into the forecasting framework to ensure accurate demand predictions.

- Incorporate Industry 4.0 Technologies: Leveraging AI, IoT, and automation can enhance data collection and realtime analysis, further improving forecasting accuracy and responsiveness.
- Continuous Model Evaluation: Regularly assess the performance of forecasting models to adapt to changing market conditions and data patterns.
- Enhance Data Quality: Invest in robust data management systems to ensure the availability of highquality, comprehensive data for more accurate forecasting.

- Address Workforce Challenges: Mitigate labor shortages by investing in training programs and leveraging automation to complement the workforce, thereby ensuring consistent production capabilities.
- Sustainability Integration: Align forecasting models with sustainability goals to anticipate and plan for environmentally friendly production practices.

7. Conclusion and Future Scope of Work

Accurate demand forecasting is indispensable for the machinery manufacturing industry to navigate market volatility, manage resources effectively, and maintain competitive advantage. This report has demonstrated that the GARCH model provides the most reliable and efficient forecasts compared to other traditional and advanced models. By adopting the GARCH model and integrating it with Industry 4.0 technologies, manufacturers can achieve enhanced forecasting accuracy and operational efficiency.

Future Scope:

- Model Enhancement: Explore hybrid models that combine the strengths of GARCH with machine learning algorithms to further improve forecasting accuracy.
- Extended Data Analysis: Incorporate additional variables such as market indicators, economic factors, and geopolitical events to enrich the forecasting models.
- Scalability: Apply the GARCH model across different manufacturing sectors to validate its efficacy and adaptability.

- RealTime Forecasting: Develop realtime forecasting systems that leverage streaming data for instantaneous demand predictions and agile decisionmaking.
- Sustainability Forecasting: Integrate sustainability metrics into forecasting models to support environmentally responsible manufacturing practices.

By addressing these future avenues, the machinery manufacturing industry can continue to refine its forecasting capabilities, ensuring sustained growth and resilience in an everevolving market landscape.