ISSN No:-2456-2165

# Revolutionising TMT Steel Bar Sales Projections: Unleashing the Power of Deep Learning Algorithms for Unparalleled Forecasting Precision

Mohd Amer Hussain[1], Akhil Rasamsetti[2], Vaanishree Kamthane[3], Deba Chandan Mohanty[4], Bharani Kumar Depuru[5]
[1,2]Research Associate, Innodatatics, Hyderabad, India.
[3] Mentor, Research and Development, Innodatatics, Hyderabad, India
[4] Team Leader, Research and Development, Innodatatics, Hyderabad, India
[5] Director, Innodatatics, Hyderabad, India

**Abstract:- Effective forecasting of sales and production in the field of TMT (Thermo-Mechanically Treated) steel is crucial for businesses to optimise their operations, manage inventory, and meet customer demand. This abstract presents a forecasting model that utilises historical sales and production data, as well as relevant market indicators, to predict future sales and production volumes for TMT steel products. The model employs advanced statistical techniques and machine learning algorithms to analyse the complex relationships between various factors influencing sales and production. By incorporating factors such as economic indicators, market trends, and customer behaviour patterns, the model aims to provide accurate and reliable forecasts. The proposed forecasting model offers a valuable tool for TMT steel manufacturers and distributors to enhance decision-making, resource allocation, and strategic planning, leading to improved operational efficiency and increased profitability in the dynamic and competitive steel industry.**

**Additionally, this study highlights the utilisation of advanced deep learning models for forecasting sales and production in the field of TMT steel. The forecasting model integrates deep learning algorithms, such as recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, to capture temporal dependencies and patterns within the data. These deep learning models excel at handling complex sequential data and can effectively capture nonlinear relationships, allowing for more accurate and robust predictions. By incorporating advanced deep learning techniques into the forecasting model, it aims to improve the accuracy and reliability of sales and production forecasts for TMT steel. The combination of traditional statistical methods and advanced deep learning models offers a comprehensive approach to forecasting, enabling businesses in the TMT steel industry to make informed decisions, optimise operations, and adapt to market dynamics with increased precision and confidence.**

***Keywords:-*** *Sales Forecasting, TMT Steel, Predictive Analytics, Machine Learning Models (LSTM, RNN, GRU), Production Optimization, Inventory Management.*

## I. INTRODUCTION

Predicting future market trends has always been a crucial goal for businesses. Imagine the advantages of accurately forecasting demand and making informed decisions. In the complex and dynamic world of the steel industry, especially in the context of Thermo-Mechanically Treated (TMT) steel, having the ability to accurately predict demand trends and optimise production levels is nothing short of invaluable [1]. This is where the art and science of forecasting come into play, serving as a key element for success. In this research paper, an intelligent system is introduced which precisely is designed to empower individuals in the TMT steel industry with the remarkable ability to foresee future demand patterns and make informed decisions with ease.

The primary aim of this research study is to develop a Predictive Analytics system designed to improve the precision and efficiency of the Demand Forecasting process in TMT Steel production. This study follows the CRISP-ML(Q) methodology, which is publicly available on the 360DigiTMG website (ak.1) [Fig.1].

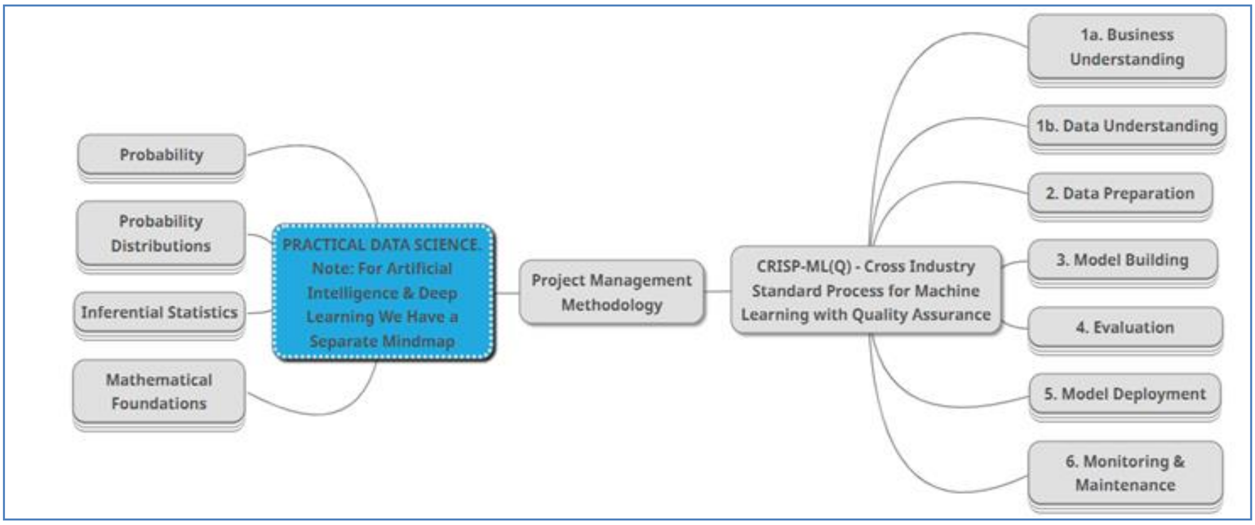


Fig. 1: The CRISP-ML (Q) Methodological Framework. (Source: Mind Map - 360DigiTMG)

FThe global economy heavily relies on the steel industry, which has deep-reaching impacts across various sectors such as construction, infrastructure development, and manufacturing. Within this expansive industry, TMT steel plays a crucial role as a fundamental component. For example, in the construction sector, TMT steel is used to reinforce the structural elements of buildings, bridges, and tunnels, ensuring their long-term durability and safety. To succeed in this highly competitive landscape, it becomes essential to accurately predict the demand for TMT steel and manage production levels effectively. Forecasting becomes the key tool in this pursuit, providing industry stakeholders with the advantage of foreseeing future trends, optimising resource allocation, and maintaining a competitive edge.

This simple flow outlines the sequential steps involved in the ML workflow. It starts with collecting data from various sources, followed by pre-processing to clean and prepare the data for analysis. Feature engineering involves selecting and creating relevant features for model training. The models are then trained using historical data and evaluated to ensure accuracy. The final step involves using these models for forecasting TMT steel demand, which in turn informs strategic decision-making in the steel industry.

This study introduces an intelligent system that is set to revolutionise the TMT steel industry. Leveraging the power of advanced predictive analytics and machine learning algorithms, the system provides users with a revolutionary decision-making tool. It not only improves their capacity to assess and respond to market demands but also creates fresh opportunities for sustainable expansion by cutting down on waste and mitigating supply chain interruptions. In doing so, it allows the industry to adapt more flexibly to changing market dynamics, maximise profitability, and contribute even more significantly to the global economy's growth trajectory.

To wrap it up, the steel industry plays a pivotal role in the world economy, and Thermo-Mechanically Treated (TMT) steel is a key component within it. Being able to foresee the future through accurate forecasting is a game-changing advantage. The system that is presented in this research paper is a significant advancement toward achieving this capability. It's a practical and technologically advanced solution that empowers industry professionals to excel in an ever-shifting market landscape. With it, they can make more informed decisions, adapt to changing circumstances, and contribute even more effectively to the industry's success and the global economy's growth.

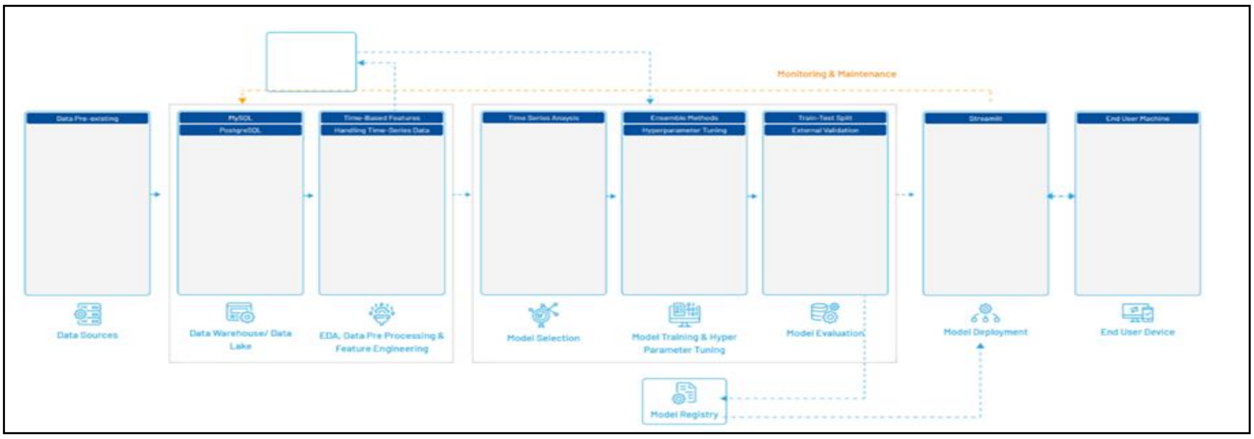


Fig. 2: ML workflow architecture used for research: A detailed overview for TMT forecasting model
Source: ML Workflow – 360Digi TMG

ISSN No:-2456-2165

➢ *Significance of Forecasting:*

- Efficient Resource Utilisation: Prediction enables us to efficiently allocate resources such as metals, labour, and machinery, minimising wastage [2].
- Meeting Customer Orders: Customer demand for TMT steel can fluctuate. Prediction ensures that appropriate inventory levels are maintained to fulfil orders promptly.
- Cost Savings: With foresight into future demand, one can avoid overproduction or underproduction of TMT steel, leading to cost savings.
- Competitive Advantage: The steel industry is dynamic and competitive. Forecasting helps us stay ahead of market changes and outperform competitors [2].
- Strategic Planning: Predictions are instrumental in long-term strategic planning, including decisions about business expansion and inventory management.
- The objective is to simplify the planning process for TMT steel manufacturers and distributors, providing them with the confidence to plan ahead effectively, much like knowing what's for dinner before beginning to cook.

## II. METHODS AND TECHNIQUES

In this study a detailed explanation is provided for the organised and professional approach that was taken to conduct accurate TMT steel forecasting.

- **Data Collection:** The TMT steel forecasting process was initiated by gathering all relevant data, including historical sales and production data. The data has been collected directly from a primary distributor specialising in the provision of high-quality TMT (Thermo-Mechanically Treated) rods.

This dataset [Fig.3] contains financial transaction records for a business, with 33,045 entries and 24 columns (Few columns are masked).

| Date | FY | Particulars | Particulars | dia | dia grou | gra | type | length | Voucher Type | Voucher No. | Quantity | Rate | Value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3-Apr-17 | FY 18 | [illegible] | 25MM [illegible] FE500D (S) | 25 MM | 12 MM - 32 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 24.910 Tons | 40000.00/Tons | 996400.00 |
| 3-Apr-17 | FY 18 | [illegible] | 25MM [illegible] FE500D (S) | 25 MM | 12 MM - 32 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 25.490 Tons | 40000.00/Tons | 1019600.00 |
| 3-Apr-17 | FY 18 | [illegible] | 16MM [illegible] FE500D (S) | 16 MM | 12 MM - 32 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 1.030 Tons | 41200.00/Tons | 42436.00 |
| 3-Apr-17 | FY 18 | [illegible] | 12MM [illegible] FE500D (S) | 12 MM | 12 MM - 32 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 1.930 Tons | 41200.00/Tons | 79516.00 |
| 3-Apr-17 | FY 18 | [illegible] | 10MM [illegible] FE500D (S) | 10 MM | 10 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 3.940 Tons | 41700.00/Tons | 164298.00 |
| 3-Apr-17 | FY 18 | [illegible] | 08MM [illegible] FE500D (S) | 08 MM | 08 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 9.090 Tons | 43200.00/Tons | 392688.00 |
| 4-Apr-17 | FY 18 | [illegible] | 12MM [illegible] FE500D (S) | 12 MM | 12 MM - | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 15.240 Tons | 37976.19/Tons | 578757.14 |
| 4-Apr-17 | FY 18 | [illegible] | 08MM [illegible] FE500D (S) | 08 MM | 08 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 6.700 Tons | 39976.19/Tons | 267840.47 |
| 5-Apr-17 | FY 18 | [illegible] | 20MM [illegible] FE500D (S) | 20 MM | 12 MM - 32 MM | 500D | FULL LENGTH | 12 METER | Sales | [illegible] | 12.770 Tons | 40000.00/Tons | 510800.00 |

Fig. 3: Financial Transaction Records Dataset with 33, 045 Entries and 24 Columns

- **Data Description:** This data serves as the foundation upon which the forecasting predictions are built [3]. From the above data, only the relevant columns, namely, Date, Length, and Quantity, have been selected, where Quantity represents the output variable. With the quantity forecast, this dataset enables the forecasting of demand, thereby providing valuable insights into future customer purchase trends and behaviour. [Fig.4]

| Column name | Description |
|---|---|
| Date | The date when the transaction took place. |
| FY | Financial year associated with the transaction. |
| Particulars | Information about the entity or company involved in the transaction. |
| Particulars | Additional information about the transaction(Rod dia-Brand-Grade). |
| dia | The diameter of the TMT steel rod. |
| dia group | The group category of the TMT steel rod diameter. |
| grade | The grade of the TMT steel rods. |
| type | The type or category of the TMT steel rods. |
| length | The length of the TMT steel rods. |
| Voucher Type | The type of voucher, indicating the nature of the transaction. |
| Voucher No. | The voucher number associated with the transaction. |
| Quantity | The amount of TMT steel rods sold in tons. |
| Rate | The price per ton of the TMT steel rods. |
| Value | The total value of the transaction, calculated as quantity multiplied by rate. |

Fig. 4: Data Description for Financial Transaction

- **Data Flow:** The research strategy places great importance on the architectural diagram, which is a central element. This diagram [Fig.5] visually represents the system's structure, showing the flow of data, the deployment locations of models, and the connections between different components. Additionally, the architecture diagram serves as a blueprint for future project scaling and expansion. It helps identify potential bottlenecks, areas for optimization, and opportunities for integrating additional features or models as the project evolves.

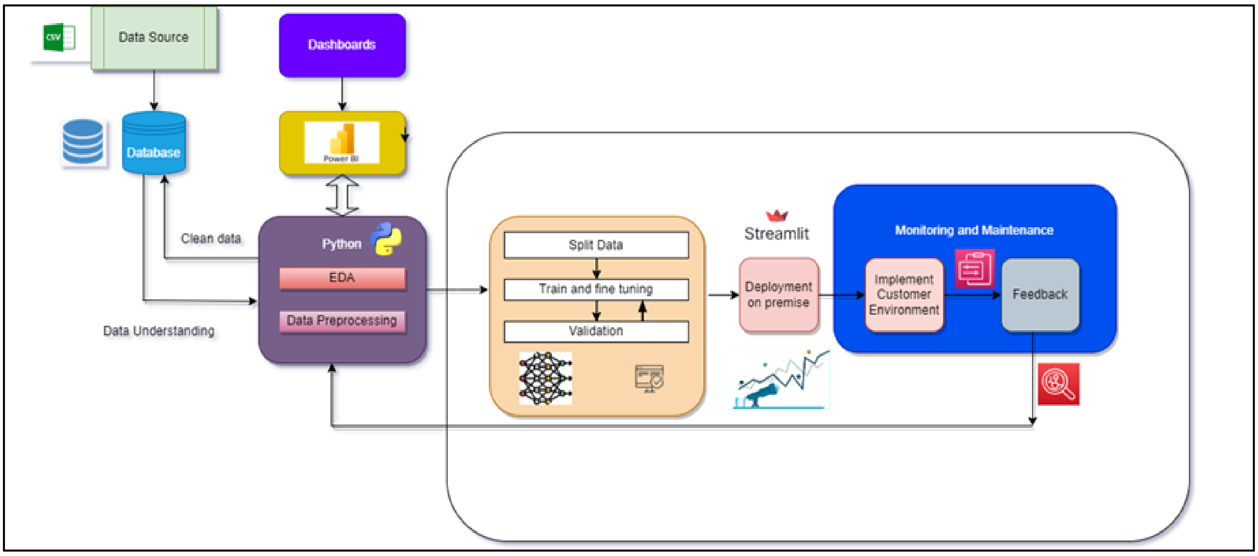


Fig. 5: Architecture Diagram: Illustrating the data components and Flow for Forecasting

- **Data Preparation:** To ensure data accuracy and readiness for analysis, data cleansing was performed, addressing errors and inconsistencies. Also resampling techniques were employed to transform the original daily data into monthly and weekly formats, much like preparing a workspace before starting a task. The subsequent focus is on forecasting demand at both monthly and weekly intervals [4]. Upon conducting basic exploratory data analysis (EDA), a significant observation was made, revealing that the majority of the rods within the dataset were of 12-metre length. Out of the total 33,045 records, a substantial 30,846 records specifically corresponded to the 12-metre category. Because most of the rods in the dataset were 12 metres long, focus was more on analysing this particular category. [Fig.6]

```
TISCON['length'].value_counts()

12 METER          30846
CUSTOMISED         1290
0 METER             615
7 - 10 METER        177
4 - 7 METER          92
10 - 12 METER        25
Name: length, dtype: int64
```

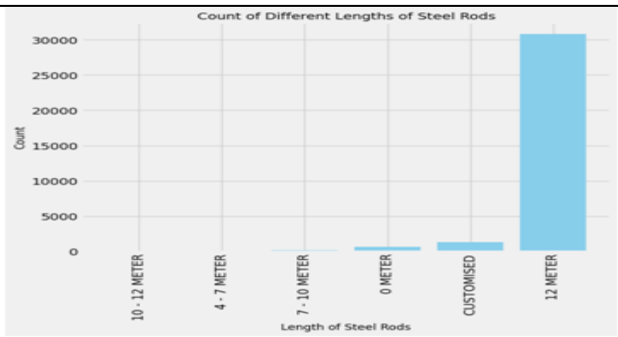


Fig. 6: Uncovering Patterns in Rod Lengths: A Data Analysis for Clear Interpretation

Conversely, the remaining length categories, including customised, 0 metre, 7-10 meter, 4-7 meter, and 10-12 meter, contained minimal values. Given the limited occurrence of these values, incorporating them into the forecasting process would result in a significant number of null values, subsequently compromising the integrity and reliability of the analysis.

The dataset consists of 33,045 records, with the following distribution of records for different length categories:

- 12 meters: 30,846 records
- Customized: 1,290 records
- 0 meter: 615 records
- 7 -10 meter: 177 records
- 4 - 7 meter: 92 records
- 10-12 meter: 25 records

The next logical step is to create a time plot to analyse the data and understand the trends. This analysis will be focused on the output column, which represents the quantity of rods. Creating a time plot [Fig.7] will provide a visual representation of how the quantity of rods has changed over time, enabling us to identify any patterns or trends within the data.

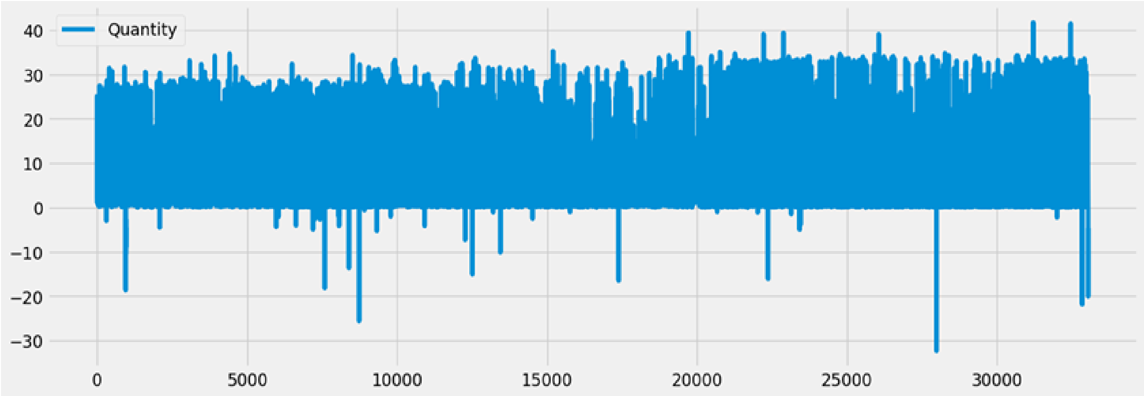

Fig. 7(a): Before

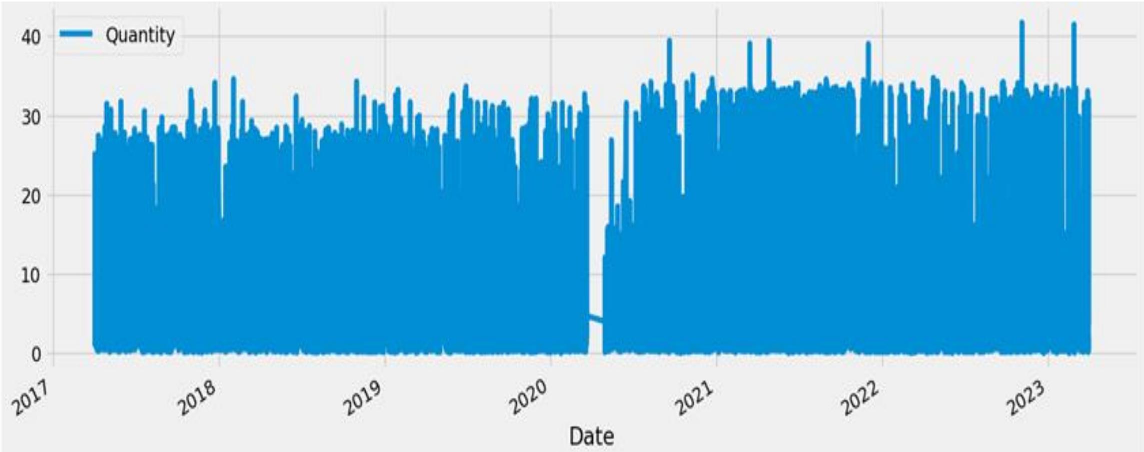

Fig. 7(b): After

Fig.7: Comparing Line Charts: Data Snapshot Before and After Modifications

Before proceeding with the time plot, it is essential to ensure that the dataset includes all the relevant data points. In this context, the presence of negative data points suggests the occurrence of returns within the dataset. Conducting a thorough examination to understand and address these negative data points is crucial before initiating any further analysis.

- **Description of Negative Entries:** In the "Quantity" column of the dataset, approximately 0.23% of the entries are negative. These negative values in the "Quantity" attribute signify instances where products were returned by customers. In a business or sales context, it's essential to understand that returns are a common occurrence and indicate items that customers have sent back for various reasons, including defects or dissatisfaction.
- **Exclusion from Sales Analysis:** It's crucial to highlight those negative entries in the "Quantity" column do not represent items sold. In other words, they should not be included in the total sales count because they reverse the sales transaction. When conducting any analysis related to sales or other factors where "Quantity" is a vital metric, these negative entries (returns) should be excluded. They are treated separately from actual sales transactions and should not be considered as part of the positive sales figures.
- **Minimal Impact on Analysis:** The validity of the model or the results of the analysis will not be significantly affected by the removal of the negative entries. By excluding returns, the focus of the research is essentially shifted to positive sales transactions, and the noise created by returns is eliminated. This practice is commonly adopted in data analysis to ensure a clearer and more accurate picture of the actual sales trends and patterns.

Missing or zero values in daily data can cause problems. Missing values can disrupt the data analysis, while zero values can mean different things. It's essential to understand the context and handle them correctly. To address this, one may need to fill in missing data (imputation) and improve data collection and recording processes. Handling these values properly ensures that the data analysis is accurate and meaningful. [Fig.7]

- **Frequency of Daily Data:** The quick identification of seasonality and trends might be difficult due to the high frequency of daily data points. This difficulty arises because daily fluctuations and noise in the data can make it challenging to notice longer-term patterns, such as those that occur weekly, monthly, or yearly.
- **Noise in Daily Data:** Uncovering underlying patterns in daily data can be challenging due to the presence of an abundance of noise and irregularities. This noise can stem from various factors, including random events, holidays, and one-time occurrences.
- **Data Volume:** The issue of data volume, particularly in the context of smaller datasets like daily data, can present challenges in accurately identifying seasonality

and trends. With limited data points, it becomes difficult to capture the variability and nuances within the dataset, often leading to an incomplete understanding of the underlying patterns. As a result, subtle trends and patterns that occur over longer periods, such as seasonal variations or cyclical movements, might go unnoticed or misinterpreted. This limitation underscores the importance of ensuring an adequate sample size to facilitate a comprehensive analysis and interpretation of the data.

- **More Robust Patterns in Longer Time Intervals:** The possibility exists that some businesses or phenomena may demonstrate clearer seasonality and trends when data is aggregated over longer intervals, such as weekly or monthly, rather than daily. Aggregating data over these extended periods can help mitigate the impact of daily fluctuations and noise, thereby revealing more distinct and discernible patterns.
- **Complexity of the Data:** The complexity of daily data can sometimes hinder the quick identification of seasonality and trends. Traditional time series analysis techniques can become less effective due to the presence of multiple interacting factors or irregular, non-repetitive patterns.

## III. RESAMPLING DATA FROM DAILY TO MONTHLY

- **Converting Frequencies:** The primary aim of resampling is to convert a time series dataset from one frequency to another. This technique can be exemplified by converting from daily to monthly frequencies or from hourly to weekly frequencies, showcasing the versatility of the method.
- **Aggregating and Interpolating:** The two fundamental objectives of resampling are to aggregate data by summarising it at a lower frequency, for instance, from daily to monthly, and to interpolate data to a higher frequency, such as from daily to hourly. These objectives can reveal different aspects of the data, offering insights into both the broader trends and the finer, more granular variations within the dataset.
- **Resulting Data Shape:** The resampling process yields a dataset with a shape of (72, 1), indicating 72 rows and 1 column. The date column serves as the index, while "Quantity" represents the target variable. Moreover, the resampled dataset covers a date range from April 30, 2017, to March 31, 2023, encompassing a span of six years' worth of data. [Fig.8(a)]


Fig. 8(a): Monthly Data Resampling: Visualizing the Transformation

In 2020, TMT steel rod sales were lower due to the impact of COVID-19. The pandemic caused a drop in sales for this product. The reasons for this decline can be explored when comparing sales in 2020 to those in 2021.

## IV. RESAMPLING DATA FROM DAILY TO WEEKLY

- **Resampling to Weekly Frequency:** The process of resampling daily data to a weekly frequency involves the aggregation and summarization of daily data into weekly intervals. This transformation offers several benefits, including data simplification and the highlighting of broader trends and patterns. By condensing the data into weekly intervals, the noise and fluctuations inherent in daily data can be reduced, allowing for a clearer understanding of long-term trends and patterns.
- **Advantages of Weekly Resampling:** Resampling daily data to a weekly frequency provides various advantages, such as improved data manageability, enhanced trend identification, noise reduction, and better data visualization. This technique is commonly employed in various fields, including finance, economics, and public health, to facilitate strategic decision-making and comprehensive trend analysis.
- **Description of Resampled Data:** The resampled data comprises 313 rows and 1 column, indicating the dataset's size and structure with 313 data points organized within a single column. The date column serves as the index, providing a chronological order to the data, while the "Quantity" column represents the target variable of interest, typically the focus of time series analysis.
- **Date Index and Target Variable:** The timeframe of the data spans from April 9, 2017, to April 2, 2023, emphasising the period encompassed by the time series data. This timeframe allows for an in-depth analysis of the trends and patterns over the course of approximately six years. [Fig.8(b)]

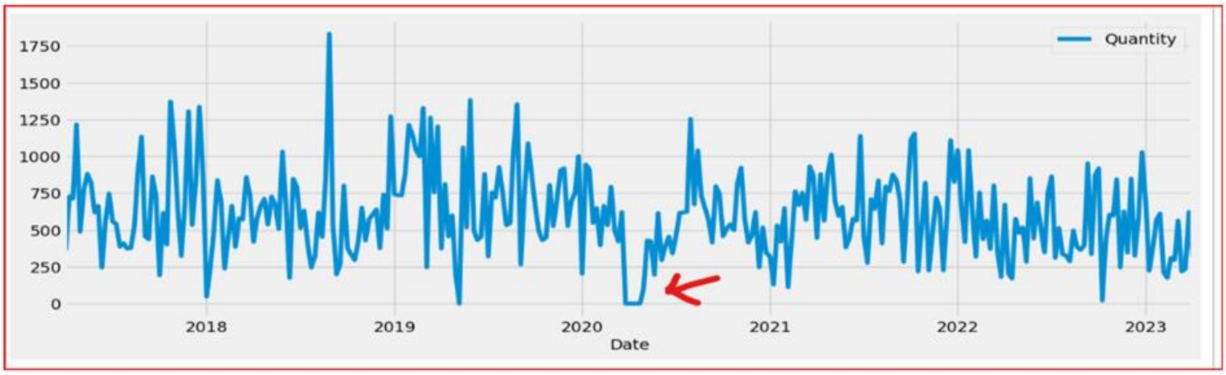


Fig. 8(b): Weekly Data Resampling: Visualizing the Transformation

In 2020, TMT steel rod sales were lower due to the impact of COVID-19. The pandemic caused a drop in sales for this product. The reasons for this decline can be explored when comparing sales in 2020 to those in 2021.

*A. Data Pre-processing*

➢ *Outlier Detection Treatment Methods:*

- **Local Outlier Factor (LOF):** The Local Outlier Factor (LOF) is a widely used technique for identifying outliers in time series data. It assesses the local density of data points, helping to identify those with significantly lower local density compared to their neighbours.
- **Extreme Outliers in the Dataset:** Three outliers have been detected in the dataset [Fig.9]. Two of these outliers, occurring on April 30, 2020, and February 28, 2023, are considered extreme outliers, signifying their substantial deviation from the typical data pattern.

*B. Importance of Identifying Extreme Outliers:*

```
# Prepare the data for LOF
df_12M_month = df_12M_month.reset_index()
X = df_12M_month['Quantity'].values.reshape(-1, 1)

# Fit the LOF model
lof = LocalOutlierFactor(n_neighbors= 20 ) # Adjust the parameters as per your requirements
outlier_labels = lof.fit_predict(X)

# Identify outliers
outliers = df_12M_month[outlier_labels == -1]
```

```
outliers
```

| | Date | Quantity |
|---|---|---|
| 22 | 2019-02-28 | 4585.10 |
| 36 | 2020-04-30 | 83.42 |
| 70 | 2023-02-28 | 1173.29 |

Fig. 9: LOF-Based Outlier Detection: Illustrating Identified Anomalies

Identifying and addressing extreme outliers in time series data is crucial due to their potential disproportionate impact on analysis, forecasting, and decision-making. These data points often indicate unusual events, errors, or anomalies that require attention and investigation.

➢ *Winsorization Technique:*

Winsorization is a statistical technique used to handle outliers by capping extreme values without removing data. It sets predefined upper and lower bounds to replace extreme values, retaining all data points while reducing their influence on statistical analyses.

➢ *Application of Winsorization:*

To apply winsorization to the 'Quantity' column in the dataset, one can use the mstats.winsorize function with the specified limits for the top 5% of the highest values and the bottom 1% of the lowest values in the 'Quantity' column.

➢ *Creating a Winsorized Column:*

The resulting winsorized values are stored in a new column named 'Winsorized_Quantity' within the dataset. This column contains the 'Quantity' values after winsorization, enabling further analysis or modelling. [Fig. 10]

ISSN No:-2456-2165

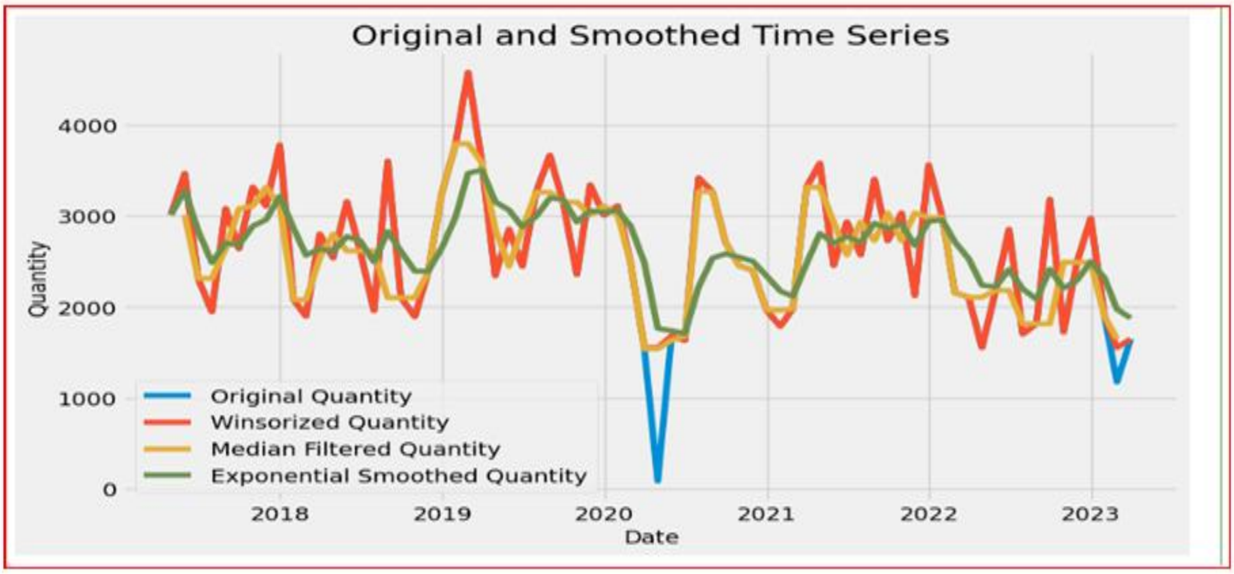


Fig. 10: Plot of Original and Smoothed Time Series: Comparison of Trends

## V. DATA PRE-PROCESSING FOR WEEKLY DATA

There is no missing values or null values in the data, but zero values were found in some places. [Fig.11]

```
TISCON_12_W = TISCON_12.resample('W').sum()

TISCON_12_W.isnull().sum()

Quantity    0
dtype: int64

zero_entries = TISCON_12_W[TISCON_12_W["Quantity"] == 0]

# Print the rows with zero entries
print(zero_entries)

            Quantity
Date
2019-05-12       0.0
2020-03-29       0.0
2020-04-05       0.0
2020-04-12       0.0
2020-04-19       0.0
2020-04-26       0.0
```

Fig. 11: Treatment of Zero Entries: Handling Zero-Value Occurrences in the Time Series Data

*A. Mechanism of forward filling:*

Identify a missing value or zero in the time series data.

Look backward to find the most recent non-zero value. Carry this last observed non-zero value forward to fill in the missing or zero values until a new non-zero value is encountered.

*B. Assumption of Temporal Stability:*

The assumption of temporal stability underlying forward filling implies that the value remains relatively constant or unchanged until new data becomes available. This assumption is suitable for situations where significant changes between data points are unlikely over short intervals.

*C. Benefits of Forward Filling:*

Forward filling offers several benefits, particularly when dealing with irregularly sampled data or sparse time series. It ensures that the temporal context is maintained, making the data suitable for various time series analyses, such as forecasting or trend analysis.

*D. Example Application:*

For example, in the context of daily temperature data with missing values, forward filling would replace the missing values with the temperature from the previous day, assuming minimal temperature fluctuations over a short time span. This approach helps maintain the continuity of the temperature trend, enabling a more comprehensive analysis of the temperature patterns over time.

## VI. SEASONALITY AND TREND ANALYSIS ON MONTHLY DATA

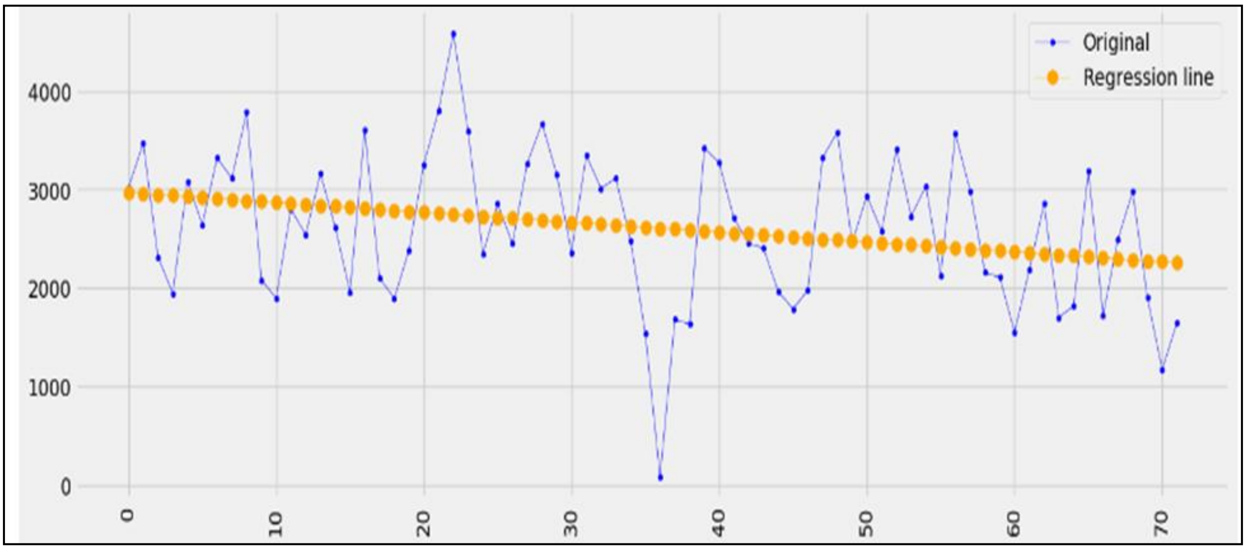


Fig. 12: Monthly Data Time Plot: Visualization of Time Series Trends and Patterns

### A. *Understanding Trend:*

The concept of the trend in time series analysis is emphasized as representing the long-term direction or trajectory of the data, indicating whether values are increasing, decreasing, or stable over time. An overall downward trend was observed in monthly data [Fig.12]. The observation of a "downward trend" indicates a consistent decrease over the observed period. The seasonality was observed to be additive in nature. [Fig.13]

### B. *Additive Seasonality:*

Additive seasonality is defined in the context of time series data, with an explanation provided on how recurring patterns or cycles, such as daily, weekly, monthly, or yearly fluctuations, are incorporated into the data alongside the underlying trend. In an additive seasonality model, these periodic patterns maintain a relatively consistent magnitude over time.

### C. *Application in Time Series Analysis:*

The significance of understanding the downward trend and additive seasonality in time series analysis is highlighted, with a description of how this decomposition is instrumental in applications like forecasting and uncovering the underlying dynamics of the data.

### D. *Practical Example:*

An illustrative example of a time series with a declining trend and additive seasonality is provided, explaining how the downward trend indicates a consistent decrease over time, while the additive seasonality represents recurring patterns superimposed on the trend. This practical demonstration underscores the value of these concepts in real-world scenarios.

### E. *Original Time Series:*

The original time series is defined as the raw data representing observations collected over time, combining multiple factors' effects. It is noted that the goal of decomposition is to understand and model it better.

### F. *Trend Component:*

The concept of the trend component is explained as representing the long-term or overall direction of the data, emphasizing its importance in revealing the fundamental movement of the data, critical for forecasting and understanding data dynamics.

### G. *Seasonality Component:*

Seasonality is described as recurring patterns or cycles in the data at fixed intervals, with the common periods like daily, weekly, monthly, or yearly mentioned. The significance of separating seasonality to model periodic fluctuations and account for expected variations is highlighted. [Fig.13]

### H. *Residuals Component:*

The residuals are briefly introduced as the unexplained variations or noise remaining after trend and seasonality removal. It is explained that analyzing residuals helps detect unexpected events and outliers, critical for accurate forecasting and model building.

### I. *Importance of Decomposition:*

The importance of decomposing time series data into its components is stressed, with an explanation that this separation simplifies analysis, enables accurate forecasting, and provides deeper insights into data patterns, ultimately supporting better decision-making in finance, economics, and business.

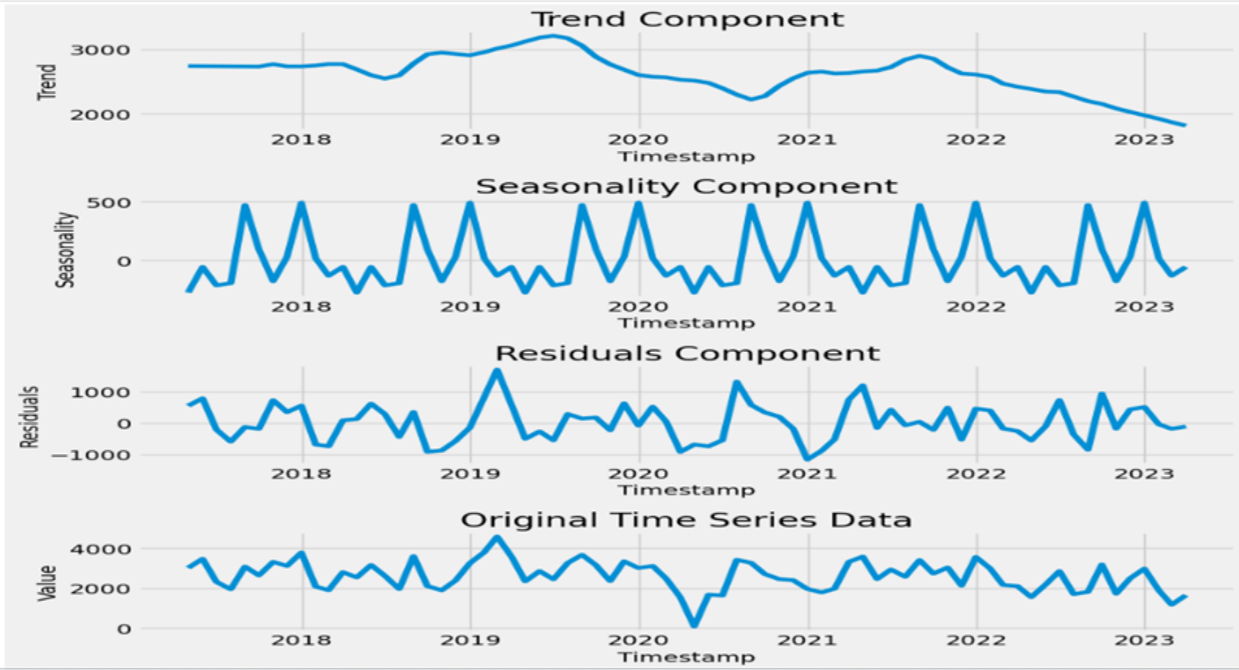

Fig. 13: Decomposition Plot of Monthly Data: Illustrating Trend, Seasonality, and Residuals

## VII. SEASONALITY AND TREND ANALYSIS ON WEEKLY DATA

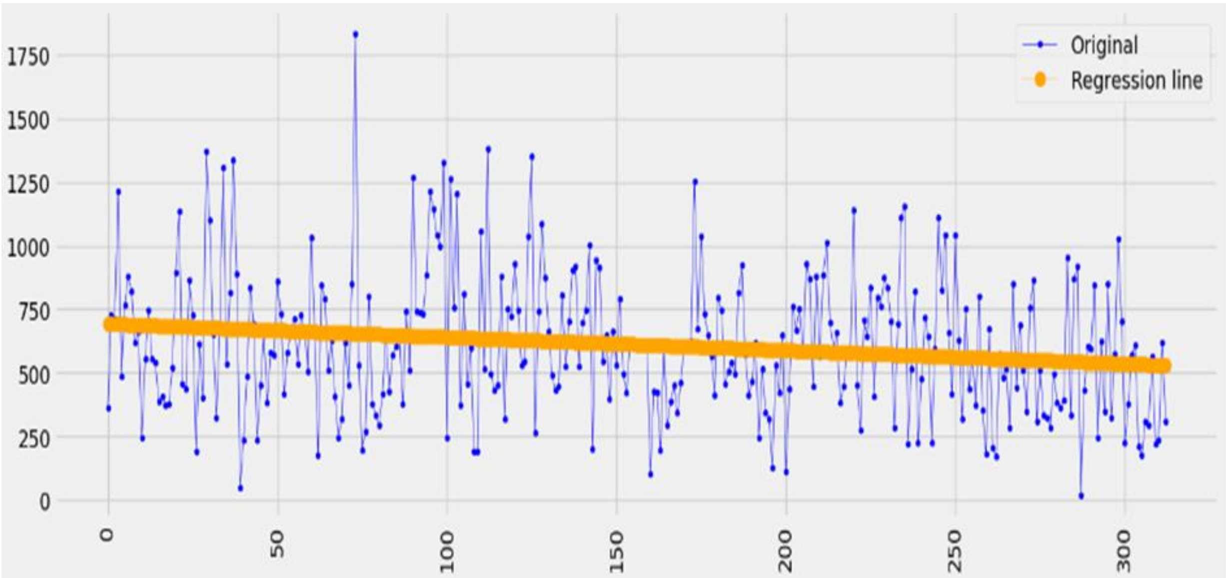

Fig. 14: Weekly Data Time Plot: Visualization of Time Series Trends and Patterns

- **Decoding Downward Trends:** The concept of a trend in time series analysis is defined, emphasizing its role in revealing the long-term data pattern. An overall downward trend was observed in weekly data [Fig.14]. A "downward" trend signifies a consistent decrease in data values, indicating an adverse trajectory. An additive seasonality is also seen in the weekly data. [Fig.15]
- **The Significance of Additive Seasonality:** The notion of additive seasonality is explored, shedding light on how recurring patterns or cycles emerge at regular intervals. The additive nature is described, where these patterns maintain a relatively constant magnitude and overlay the primary trend. An example, such as analyzing monthly sales data in a retail store, is provided to illustrate the concept.
- **Applications in Time Series Analysis:** The critical importance of understanding downward trends and additive seasonality in the realm of time series analysis is highlighted. The explanation focuses on how these components facilitate the separation and modelling of factors influencing data behaviour, particularly for forecasting.
- **Original Time Series:** The original time series is defined as the raw dataset, containing data points recorded at specific intervals. The complexity of this data, originating from various underlying components such as trend, seasonality, and noise, is elaborated upon.

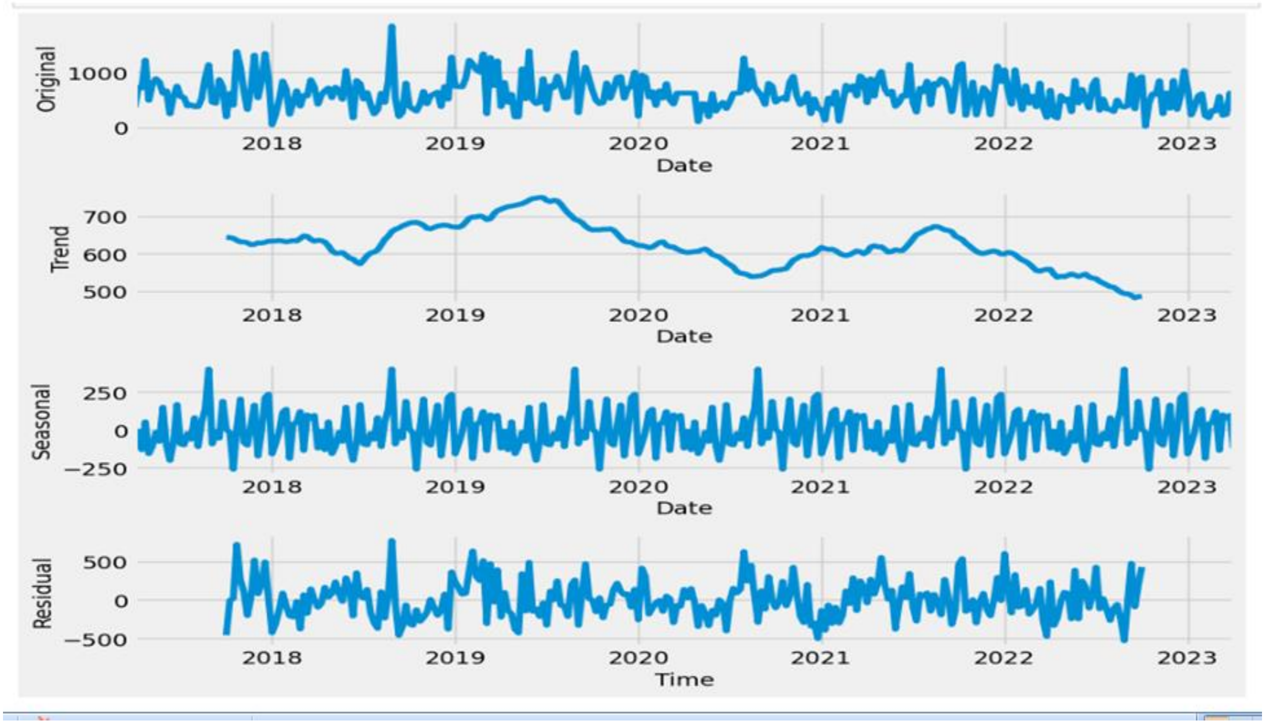


Fig. 15: Decomposition Plot of Weekly Data: Illustrating Trend, Seasonality, and Residuals

- **Understanding Trend:** The concept of the trend is explained as a long-term, underlying direction in the data. The illustration emphasizes how trends can be upward, downward, or stable, and the role of trend decomposition in understanding the data's fundamental movement.
- **Deciphering Seasonality:** Seasonality is described as the regular cyclic patterns in the data occurring at fixed intervals. The emphasis is placed on the external factors influencing these patterns and how seasonality decomposition allows for their isolation and individual modelling.
- **Analyzing Residuals:** The concept of residuals is introduced as the unexplained variations in the data left after removing trend and seasonality. The importance of analyzing residuals in detecting anomalies and outliers, crucial for accurate forecasting and decision-making, is highlighted.
- **Correlation Plots:** In the context of data, ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) [Fig.16 Fig.17] can be incredibly valuable for time series analysis, enabling one to uncover meaningful insights and optimize forecasting models.

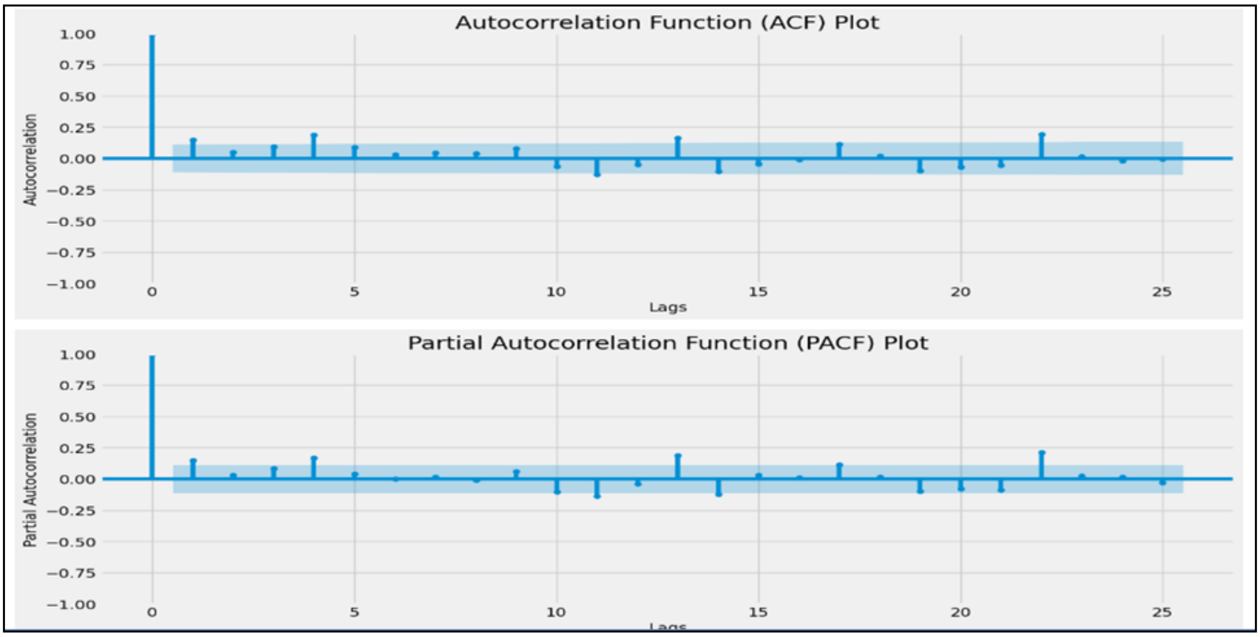


Fig. 16: Decomposing Weekly Trends: Visualizing Data Over Time

ISSN No:-2456-2165

- **ACF (Autocorrelation Function) for Data:** ACF helps reveal the correlation between daily "Quantity" values and their past values (lags). By examining ACF plots, one can identify how each data point is related to its previous data points at various time lags. In the dataset, for weekly data the significant lag values are found to be 22 and for monthly data the significant lag values are found to be 1. For the dataset, ACF can provide insights into the autocorrelation patterns within the "Quantity" attribute, shedding light on whether daily values are correlated with their past values and to what extent. This is crucial for detecting any underlying seasonality or cyclical behaviour in the data.

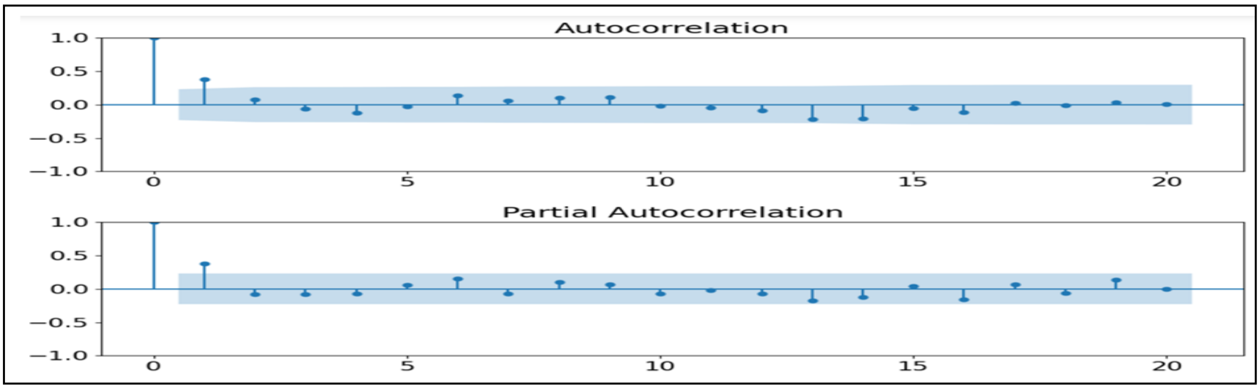


Fig. 17: Decomposing Monthly Trends: Visualizing Data Over Time

- **Identifying Seasonality:** ACF plots can be used to detect any seasonality in daily "Quantity" data. If significant spikes are observed in the ACF plot at regular intervals, it suggests that there might be a recurring pattern or seasonality in the sales data. For instance, if it is seen as autocorrelation at lags of 7 (weekly) or 30 (monthly), it could indicate weekly or monthly seasonality in sales.
- **Lag Selection for Modelling:** ACF is also helpful for determining the appropriate lag order when building Autoregressive (AR) or Moving Average (MA) models. In this case, if an AR model is being developed to forecast daily "Quantity," the ACF plot can guide in selecting the optimal lag order. It helps decide how many past observations should be considered as predictors when forecasting future "Quantity" values. The significant lags identified in the ACF plot can be incorporated into the modelling process.
- **PACF (Partial Autocorrelation Function) for Data:** PACF, on the other hand, takes a more focused approach, accounting for the influence of intermediate lags. In the dataset, for weekly data the significant lag values are found to be 22 and for monthly data the significant lag values are found to be 1. In data, PACF can help pinpoint the specific lags that have a direct impact on the current "Quantity" value. This is particularly valuable for determining the order of an Autoregressive (AR) model, where each lag represents the influence of a specific past value on the current value.

## VIII. STATIONARY TEST FOR MONTHLY DATA

- **Monthly Data:** This refers to data points that are recorded at monthly intervals. The dataset likely contains observations collected at the end of each month, potentially over several years, covering various attributes or measurements relevant to the analysis.
- **Stationary Data:** A stationary time series is one in which the statistical properties, such as the mean, variance, and autocorrelation, remain relatively constant over time. In the context of monthly data, it means that the data doesn't exhibit significant trends or seasonality. The statistical characteristics of the data remain consistent from one month to the next. This is highly valuable because it simplifies the process of time series modelling and forecasting.
- **Significance of Stationarity:** When a time series is stationary, it is preferred for time series analysis because it's easier to work with, and the underlying statistical properties remain consistent. This stability makes it suitable for modelling and forecasting using standard time series techniques, such as autoregressive (AR) and moving average (MA) models. These models rely on the assumption that the statistical properties of the data do not change significantly over time. In contrast, non-stationary data often requires additional transformations or differencing to make it stationary and suitable for modelling.
- **Interpretation of Dickey-Fuller Test:** The Dickey-Fuller test is a statistical tool used to assess the stationarity of a time series. It provides several outputs, with the p-value being a crucial one. In this case, the Dickey-Fuller test results indicate that the p-value is less than the significance level (alpha), often set at 0.05 [Fig.18]. When p < alpha, the null hypothesis of non-stationarity can be rejected, meaning that the data is considered stationary.

```
from statsmodels.tsa.stattools import adfuller
def adf_test(timeseries):
    print('Results of Dickey-Fuller Test:')
    dftest = adfuller(timeseries , autolag='AIC')
    dfoutput = pd.Series(dftest[0:4], index=['Test Statistic','p-value','#Lags Used', 'Number of Observation Used'])
    for key, value in dftest[4].items():
        dfoutput['Critical Value (%s)'%key] = value
    print(dfoutput)

print(adf_test(df_steel['Quantity']))

Results of Dickey-Fuller Test:
Test Statistic                 -5.461373
p-value                         0.000003
#Lags Used                      0.000000
Number of Observation Used     71.000000
Critical Value (1%)            -3.526005
Critical Value (5%)            -2.903200
Critical Value (10%)           -2.588995
dtype: float64
None
```

Fig. 18: Testing Stationarity in monthly data: Dickey-Fuller Analysis of Time Series Data

## IX. STATIONARY TEST FOR WEEKLY DATA

- **Weekly Data:** This term refers to a time series where data points are recorded at weekly intervals. The dataset likely comprises observations collected weekly, spanning various attributes or measurements relevant to the analysis. These could include weekly sales, stock prices, temperature readings, or any other data collected on a weekly basis.
- **Stationary Data:** A stationary time series is characterised by statistical properties, such as the mean, variance, and autocorrelation, that remain relatively constant over time. In the context of weekly data, it means there are no significant trends or seasonality patterns, and the statistical characteristics of the data remain consistent from one week to the next.
- **Significance of Stationarity:** When a time series is stationary, it is preferable for time series analysis because it simplifies the modelling and forecasting process. Standard time series techniques, such as autoregressive (AR) and moving average (MA) models, can be confidently applied to stationary data for making predictions. These models rely on the assumption that the statistical properties of the data do not change significantly over time. In contrast, non-stationary data often requires additional data transformations, such as differencing, to achieve stationarity and become suitable for modelling.
- **Interpretation of Dickey-Fuller Test**: The Dickey-Fuller test is a statistical tool used to assess the stationarity of a time series. It provides various outputs, with the p-value being a crucial one. In this case, the Dickey-Fuller test results indicate that the p-value is less than the significance level (alpha), often set at 0.05 [Fig.19]. When $p < \text{alpha}$, it implies that the null hypothesis of non-stationarity can be rejected, suggesting that the data is stationary.

```
from statsmodels.tsa.stattools import adfuller
def adf_test(timeseries):
    print('Results of Dickey-Fuller Test:')
    dftest = adfuller(timeseries , autolag='AIC')
    global dfoutput
    dfoutput = pd.Series(dftest[0:4], index=['Test Statistic','p-value','#Lags Used', 'Number of Observation Used'])
    for key, value in dftest[4].items():
        dfoutput['Critical Value (%s)'%key] = value
    print(dfoutput)
    if dfoutput[1] <= 0.05:
        print('Data is Stationary')
    else:
        print('Data is non Stationary')

print(adf_test(TISCON_12_W1['Quantity']))

Results of Dickey-Fuller Test:
Test Statistic                  -4.230323
p-value                          0.000585
#Lags Used                      13.000000
Number of Observation Used     299.000000
Critical Value (1%)             -3.452411
Critical Value (5%)             -2.871255
Critical Value (10%)            -2.571947
dtype: float64
Data is Stationary
None
```

Fig. 19: Testing Stationarity in weekly data: Dickey-Fuller Analysis of Time Series Data

ISSN No:-2456-2165

- **Pattern Identification**: By closely examining the data, patterns and trends can be uncovered that reveal how TMT steel sales and production vary over time. This step is essential for making well-informed forecasts and predictions [5].
- **Model Selection:** Choosing the right tools and techniques for predictive modelling is crucial. This involves considering statistical models, machine learning algorithms, and deep learning methods, much like selecting the right tool for a specific task. The model selection process primarily relies on evaluating the Mean Absolute Percentage Error (MAPE) to find the best approach [6].

Mean Absolute Percentage Error (MAPE) is a commonly used metric for evaluating the accuracy of a forecasting model. It measures the average percentage deviation of forecasts from the actual values. The formula for calculating MAPE is as follows:

$$MAPE = \frac{1}{n}\sum_{t-1}^{n} \left|\frac{At - Ft}{At}\right| \times 100$$

Where:

- n is the number of data points in the dataset.
- $A_t$ is the actual value at time t.
- $F_t$ is the forecasted value at time t.

Here's what each part of the formula represents:

- $A_t$-$F_t$ calculates the difference between the actual and forecasted values.
- $\left|\frac{At-Ft}{Ft}\right|$ calculates the absolute percentage error for each data point.
- The sum of all these absolute percentage errors is divided by the number of data points n to get the average error.
- Multiplying by 100 converts the error into a percentage for better interpretation.

MAPE is expressed as a percentage, representing the average discrepancy between the predicted and observed values as a proportion of the observed value. It is often used to compare the performance of different forecasting models, with lower MAPE values indicating better accuracy and performance.

Ensemble Techniques were also used to enhance MAPE performance, integrating top-performing deep learning models and using their predictions as inputs for advanced machine learning models.

Ensemble techniques are commonly used in forecasting to improve the accuracy and robustness of predictions. The fundamental idea behind ensemble techniques is to combine the predictions of multiple individual models to create a final prediction that is more accurate and reliable than any single model.

- **Model Training:** After an extensive process of exploratory data analysis (EDA) and preprocessing, the focus shifted to training a model that could accurately forecast data. Just as athletes train to improve their skills, the chosen models are trained using historical data. In this pursuit, several models were evaluated aiming to pinpoint the one yielding the least Mean Absolute Percentage Error (MAPE).
- Below is a snapshot of the various models experimented during the training process. Each model underwent rigorous evaluation to determine its efficacy in minimizing the MAPE, thereby allowing us to identify the most suitable model for the specific monthly or weekly data forecasting needs. [Fig.20 Fig.21]

| Model_Nmae | Full data MAPE | Train MAPE | Test MAPE |
|---|---|---|---|
| ARIMA (AutoRegressive Integrated Moving Average) | 34.0% | 49.0% | 107.0% |
| SARIMA (Seasonal AutoRegressive Integrated Moving Average) | 70.0% | 57.0% | 120.0% |
| Prophet | 56.0% | 49.0% | 81.0% |
| Linear Model with Normalization | 60.0% | 51.0% | 93.0% |
| Quadratic Model with Normalization | 59.0% | 52.0% | 86.0% |
| Exponential Model | 53.0% | 47.0% | 81.0% |
| Additive Seasonal Model with Normalization (sin + cos) | 61.0% | 49.0% | 109.0% |
| Multiplicative seasonal model (sin+cos) | 53.0% | 45.0% | 93.0% |
| Additive Seasonal Model with Linear Trend | 54.0% | 47.0% | 81.0% |
| Additive Seasonal Model with Quadratic Trend and Normalization | 59.0% | 52.0% | 86.0% |
| Holt-Winters Exponential Smoothing Model with Additive Seasonality and Additive Trend | 37.0% | 54.0% | 96.0% |
| Holt-Winters Exponential Smoothing Model with Multiplicative Seasonality and Additive Trend | 51.0% | 60.0% | 95.0% |

Fig. 20: Assessing Forecast Accuracy: MAPE Comparison for Various Models

| Model_Nmae | Full data MAPE | Train MAPE | Test MAPE |
|---|---|---|---|
| ARIMA (AutoRegressive Integrated Moving Average) | 22.1% | 20.4% | 34.4% |
| SARIMA (Seasonal AutoRegressive Integrated Moving Average) | 22.5% | 22.7% | 20.7% |
| Prophet | 19.5% | 19.4% | 20.5% |
| Linear Model with Normalization | 23.5% | 22.9% | 29.3% |
| Quadratic Model with Normalization | 23.2% | 23.2% | 25.8% |
| Exponential Model | 22.7% | 22.3% | 27.0% |
| Additive Seasonal Model with Normalization | 22.7% | 19.9% | 35.2% |
| Additive Seasonal Model with Linear Trend | 20.6% | 19.9% | 25.6% |
| Additive Seasonal Model with Quadratic Trend and Normalization | 20.4% | 20.3% | 22.3% |
| Holt-Winters Exponential Smoothing Model with Additive Seasonality and Additive Trend | 22.9% | 22.0% | 28.8% |
| Holt-Winters Exponential Smoothing Model with Multiplicative Seasonality and Additive Trend | 22.6% | 22.1% | 26.0% |
| RandomForestRegressor | 10.0% | 8.3% | 18.5% |
| Gradient Boosting Regressor | 4.5% | 1.8% | 17.8% |
| light Boosting Regressor | 17.6% | 17.1% | 20.2% |
| XGB Regressor | 2.7% | 0.1% | 15.6% |
| cat Boosting Regressor | 6.2% | 4.0% | 17.1% |

Fig. 21: Assessing Forecast Accuracy: MAPE Comparison for Various Models

Using advanced forecasting methods, such as deep learning and neural networks, is the next logical step when traditional models fail to yield satisfactory results. In this case, the previous models produced notably large MAPE values, rendering them unsuitable for accurate predictions.

To delve deeper into these advanced techniques, the Time Series Generator, a powerful tool provided by the Keras deep learning library, specifically within the keras. preprocessing. sequence module, was incorporated.

The Time Series Generator serves the purpose of generating batches of temporal data tailored for training recurrent neural networks (RNNs) and other sequential models. It plays a vital role in tasks like time series forecasting and sequence-to-sequence prediction.

This generator operates by taking a sequence of data points, typically a time series, and generating batches of input-output pairs. Specifically, it segments the data into intervals, allowing the model to learn from temporal patterns and dependencies within the data. In this context, the Time Series Generator utilizes a seven-time-step input sequence to predict the value at the subsequent time step.

By adopting this strategy, the model becomes proficient at recognizing and leveraging historical patterns. Consequently, it can effectively forecast future values based on the information gleaned from the preceding seven-time steps. This approach is particularly valuable for complex time series forecasting tasks where understanding and capturing intricate temporal dependencies are critical for accurate predictions.

Various deep learning and neural network models were created as shown in the snapshot below [Fig.22] both for monthly and weekly data (Models were trained for 100 epochs).

| Model name | mape(monthly) Before Tuning | mape(monthly) |
|---|---|---|
| Recurrent Neural Network | 24 | 22 |
| Bidirectional Recurrent Neural Network | 35 | 27 |
| Deep Bidirectional Recurrent Neural Network | 17 | 19 |
| Long Short-Term Memory | 34 | 15 |
| Bidirectional Long Short-Term Memory | 34 | 45 |
| Deep Bidirectional Long Short-Term Memory | 32 | 16 |
| Gated Recurrent Unit | 26 | 24 |
| Bidirectional Gated Recurrent Unit | 34 | 25 |
| Deep Bidirectional Gated Recurrent Unit | 25 | 15 |

Advanced models-Monthly Data

| Model name | mape(weekly) Before Tuning | mape(weekly) |
|---|---|---|
| Recurrent Neural Network | 95 | 41 |
| Bidirectional Recurrent Neural Network | 55 | 50 |
| Deep Bidirectional Recurrent Neural Network | 49 | 48 |
| Long Short-Term Memory | 67 | 56 |
| Bidirectional Long Short-Term Memory | 76 | 54 |
| Deep Bidirectional Long Short-Term Memory | 49 | 47 |
| Gated Recurrent Unit | 59 | 46 |
| Bidirectional Gated Recurrent Unit | 56 | 45 |
| Deep Bidirectional Gated Recurrent Unit | 63 | 45 |
| Extratreeregressor ( from pycarret) | 6 | 5 |

Advanced models-Weekly Data

Fig. 22: Exploring Deep Learning Techniques: Models for Monthly and Weekly Data

ISSN No:-2456-2165

For monthly data, the following models exhibited the lowest MAPE:

- Recurrent Neural Network
- Deep Bidirectional Recurrent Neural Network
- Long Short-Term Memory
- Deep Bidirectional Gated Recurrent Unit

Similarly, for weekly data, the models with the lowest MAPE were:

- Recurrent Neural Network
- Gated Recurrent Unit
- Bidirectional Gated Recurrent Unit
- Deep Bidirectional Gated Recurrent Unit

These selected models were used to create an ensemble model, combining their collective strengths to enhance the accuracy and robustness of the forecasting process. The process employed the use of PyCaret to identify the optimal ensemble method from the set of pre-selected models. [Fig.23]

PyCaret, a low-code machine learning library in Python, streamlines the process of model training, tuning, and evaluation, making it easier to experiment with various ensemble methods and select the most effective one for TMT steel forecasting.

Through PyCaret, different ensemble techniques, such as bagging, boosting, and stacking, can be seamlessly implemented and compared based on their performance metrics. PyCaret enables comprehensive model evaluation and comparison, allowing for the selection of the most suitable ensemble method based on various evaluation metrics, including accuracy, precision, recall, and F1 score.

By leveraging the capabilities of PyCaret, the process of identifying the best ensemble method becomes more efficient, enabling data scientists and analysts to make informed decisions about which ensemble method will yield the most accurate and reliable TMT steel forecasting results.

The outcomes obtained from PyCaret, for both weekly and monthly data, are presented in the following snapshot.

The most effective ensemble model identified by PyCaret was trained using the top-performing deep learning models, specifically the four models that exhibited the lowest MAPE values during the initial model selection phase. By combining the strengths of these individual models, the ensemble model is able to leverage their unique capabilities and insights, leading to a more robust and accurate prediction for TMT steel forecasting.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **lasso** | Lasso Regression | 8.9206 | 170.4933 | 12.7074 | 0.9994 | 0.0064 | 0.0042 | 0.0210 |
| **en** | Elastic Net | 8.8996 | 169.1162 | 12.6635 | 0.9994 | 0.0063 | 0.0042 | 0.0200 |
| **lr** | Linear Regression | 10.0858 | 238.4338 | 13.9450 | 0.9993 | 0.0060 | 0.0042 | 0.4620 |
| **ridge** | Ridge Regression | 10.0856 | 238.4238 | 13.9444 | 0.9993 | 0.0060 | 0.0042 | 0.0180 |
| **lar** | Least Angle Regression | 10.0857 | 238.4220 | 13.9447 | 0.9993 | 0.0060 | 0.0042 | 0.0200 |
| **llar** | Lasso Least Angle Regression | 10.1217 | 242.5915 | 14.0062 | 0.9993 | 0.0060 | 0.0043 | 0.0200 |
| **br** | Bayesian Ridge | 10.2034 | 251.6142 | 14.1125 | 0.9993 | 0.0060 | 0.0043 | 0.0190 |
| **huber** | Huber Regressor | 16.6283 | 1328.6516 | 24.8781 | 0.9975 | 0.0080 | 0.0057 | 0.0370 |
| **et** | Extra Trees Regressor | 16.4535 | 10357.9331 | 36.6183 | 0.9865 | 0.0095 | 0.0042 | 0.3840 |
| **gbr** | Gradient Boosting Regressor | 35.7022 | 11693.0930 | 61.4908 | 0.9812 | 0.0186 | 0.0115 | 0.1020 |
| **dt** | Decision Tree Regressor | 35.9740 | 11729.0514 | 62.2929 | 0.9811 | 0.0192 | 0.0117 | 0.0170 |
| **xgboost** | Extreme Gradient Boosting | 36.0458 | 11710.7112 | 62.0098 | 0.9811 | 0.0189 | 0.0117 | 0.0510 |
| **rf** | Random Forest Regressor | 33.0559 | 13250.2224 | 61.0673 | 0.9796 | 0.0191 | 0.0109 | 0.3490 |
| **ada** | AdaBoost Regressor | 52.8912 | 15742.1626 | 81.9867 | 0.9727 | 0.0281 | 0.0191 | 0.1430 |
| **knn** | K Neighbors Regressor | 45.6262 | 17649.9014 | 78.6779 | 0.9711 | 0.0269 | 0.0161 | 0.0240 |
| **par** | Passive Aggressive Regressor | 74.8415 | 13442.1489 | 97.3506 | 0.9683 | 0.0352 | 0.0279 | 0.0180 |
| **omp** | Orthogonal Matching Pursuit | 89.3524 | 16866.7689 | 120.7156 | 0.9457 | 0.0567 | 0.0411 | 0.0190 |
| **lightgbm** | Light Gradient Boosting Machine | 186.0634 | 75893.5836 | 246.5841 | 0.8035 | 0.0934 | 0.0732 | 0.1270 |
| **dummy** | Dummy Regressor | 586.0229 | 471375.1938 | 671.6060 | -0.5750 | 0.2603 | 0.2426 | 0.0230 |

Fig. 23(a): Weekly data

| | Model | MAE | MSE | RMSE | R2 | RMSLE | MAPE | TT (Sec) |
|---|---|---|---|---|---|---|---|---|
| gbr | Gradient Boosting Regressor | 7.0552 | 2452.8973 | 22.8146 | 0.9803 | 0.0559 | 0.0190 | 0.2190 |
| lightgbm | Light Gradient Boosting Machine | 14.2269 | 2316.3533 | 34.2105 | 0.9781 | 0.1038 | 0.0606 | 0.0830 |
| ada | AdaBoost Regressor | 16.8871 | 2632.2590 | 32.3170 | 0.9770 | 0.0862 | 0.0504 | 0.2260 |
| rf | Random Forest Regressor | 6.7174 | 3094.9180 | 23.5331 | 0.9753 | 0.0493 | 0.0203 | 0.3440 |
| et | Extra Trees Regressor | 4.6556 | 3421.1932 | 20.5974 | 0.9730 | 0.0409 | 0.0120 | 0.6090 |
| lr | Linear Regression | 37.7522 | 2545.9838 | 47.6674 | 0.9694 | 0.1226 | 0.0841 | 0.0630 |
| ridge | Ridge Regression | 37.7518 | 2546.0527 | 47.6674 | 0.9694 | 0.1226 | 0.0841 | 0.0590 |
| lar | Least Angle Regression | 37.7520 | 2545.9832 | 47.6673 | 0.9694 | 0.1226 | 0.0841 | 0.1050 |
| lasso | Lasso Regression | 37.7339 | 2556.7367 | 47.7070 | 0.9693 | 0.1226 | 0.0840 | 0.0590 |
| en | Elastic Net | 37.7150 | 2561.3337 | 47.7183 | 0.9693 | 0.1226 | 0.0840 | 0.0610 |
| llar | Lasso Least Angle Regression | 37.7341 | 2556.3920 | 47.7055 | 0.9693 | 0.1226 | 0.0840 | 0.1100 |
| br | Bayesian Ridge | 37.6771 | 2625.3968 | 47.9953 | 0.9689 | 0.1227 | 0.0839 | 0.0970 |
| xgboost | Extreme Gradient Boosting | 10.0768 | 3826.9995 | 31.7364 | 0.9675 | 0.0753 | 0.0290 | 0.1220 |
| knn | K Neighbors Regressor | 27.3965 | 3758.5291 | 51.6912 | 0.9606 | 0.1377 | 0.0958 | 0.1150 |
| dt | Decision Tree Regressor | 10.3620 | 4880.6127 | 36.7834 | 0.9590 | 0.0846 | 0.0348 | 0.0590 |
| omp | Orthogonal Matching Pursuit | 73.0708 | 8477.4987 | 88.9519 | 0.8928 | 0.2257 | 0.1641 | 0.0940 |
| huber | Huber Regressor | 85.8662 | 10668.3330 | 101.7566 | 0.8604 | 0.2544 | 0.2265 | 0.0930 |
| par | Passive Aggressive Regressor | 106.4659 | 21468.5449 | 142.4085 | 0.7214 | 0.2529 | 0.1925 | 0.0580 |
| dummy | Dummy Regressor | 220.2015 | 80515.0371 | 281.2705 | -0.0564 | 0.5468 | 0.6438 | 0.0650 |

Fig. 23(b): Mon

Fig. 23: Optimizing Ensemble Models with PyCaret: Unveiling the Best Combination

The results are shown in the snapshot below (models trained on 97% data and 3 % used for validation). [Fig.24 Fig.25] For the weekly data, the Extra Tree Regressor was employed, resulting in a Mean Absolute Percentage Error (MAPE) of 5 after undergoing the tuning process. On the other hand, for the monthly data, the Gradient Boosting Regressor achieved a MAPE of 3 post-tuning. These results demonstrate the efficacy of the chosen models in accurately forecasting TMT steel data, indicating their robust performance and reliability in handling the complexities associated with the specific temporal patterns of the data.

| Model name | mape(weekly) Before Tuning | mape(weekly) |
|---|---|---|
| Recurrent Neural Network | 95 | 41 |
| Gated Recurrent Unit | 59 | 46 |
| Bidirectional Gated Recurrent Unit | 56 | 45 |
| Deep Bidirectional Gated Recurrent Unit | 63 | 45 |
| Extratreeregressor ( from pycarret) | 6 | 5 |

Fig. 24: Enhancing Model Performance for Weekly Data: Deep Learning Models Before and After Tuning

| Model name | mape(monthly) Before Tuning | mape(monthly) |
|---|---|---|
| Recurrent Neural Network | 24 | 22 |
| Deep Bidirectional Recurrent Neural Network | 17 | 19 |
| Long Short-Term Memory | 34 | 15 |
| Deep Bidirectional Gated Recurrent Unit | 25 | 15 |
| Gradient_boosting_regressor | 5 | 3 |

Fig. 25: Enhancing Model Performance for Monthly Data: Deep Learning Models Before and After Tuning

ISSN No:-2456-2165

- **Model Testing:** Model testing is a crucial phase in the development and implementation of any predictive model, including TMT steel forecasting. This phase involves subjecting the model to various tests and assessments to evaluate its performance and accuracy. The primary objective is to ensure that the model can effectively handle real-world scenarios and provide reliable predictions that align with the actual data.

  During the testing phase, the model is evaluated against known and validated datasets, enabling a comprehensive analysis of its predictive capabilities. By comparing the model's outputs with the actual data, metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) are calculated to gauge the extent of the model's accuracy. Rigorous testing also involves stress-testing the model with various scenarios and datasets to assess its robustness and resilience in different conditions.

  Furthermore, the testing phase allows for the identification and rectification of any potential issues or discrepancies within the model, ensuring that it meets the required standards of accuracy and reliability. Thorough model testing provides confidence in the model's ability to generate accurate predictions, laying the foundation for successful real-world implementation and informed decision-making in the domain of TMT steel forecasting [9].

- **Generating Predictions:** Generating predictions involves utilizing the trained and validated model to forecast future TMT steel requirements and demands. By leveraging historical data and patterns captured during the model training phase, the model can extrapolate and estimate potential future trends and needs within the TMT steel market. These predictions serve as a valuable tool for stakeholders and decision-makers, providing crucial insights into potential market fluctuations, customer demands, and production requirements.

  Through the prediction process, the model extrapolates likely scenarios, enabling businesses to proactively plan and strategize their production schedules, inventory management, and resource allocation. These forecasts serve as a guide for informed decision-making, helping organizations optimize their operations, minimize potential risks, and capitalize on emerging market opportunities.

  The process of generating predictions is not only a critical step in strategic planning and risk management but also an essential element in maintaining a competitive edge in the dynamic TMT steel industry. By staying ahead of market demands and trends, businesses can position themselves to meet customer requirements efficiently and effectively, ultimately fostering growth and sustainability within the industry. [8].

- **Deployment:** In the deployment phase, the model is deployed using the user-friendly and interactive platform Streamlit. Streamlit enables the conversion of the predictive model into a web application that is easy to use and access for all stakeholders and end-users. This deployment method ensures a seamless and intuitive user experience, allowing for real-world validation, valuable user feedback, and necessary refinements to enhance the application's overall performance and usability.

  Moreover, the forecasting output is accompanied by a confidence interval to provide users with a measure of the uncertainty associated with the predictions. The confidence interval serves as a range around the predicted values, indicating the level of confidence that the actual value will fall within this range. By incorporating the confidence interval, users can better assess the reliability and potential variability of the forecasted TMT steel requirements.

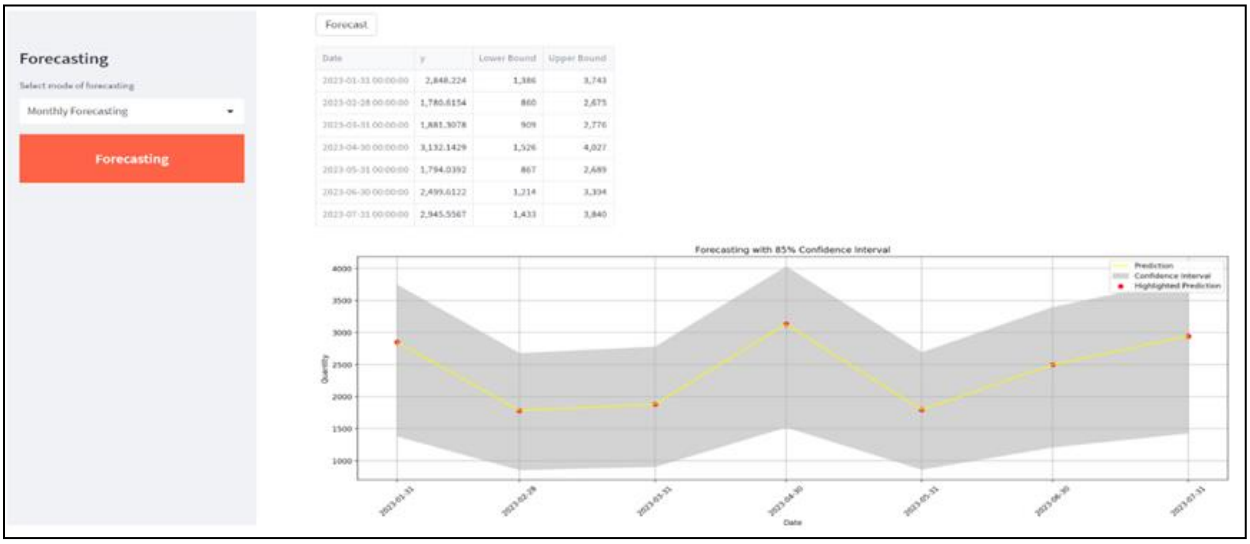

Fig. 26: Monthly Forecasting Model Output: Streamlit Deployment Visualization

ISSN No:-2456-2165

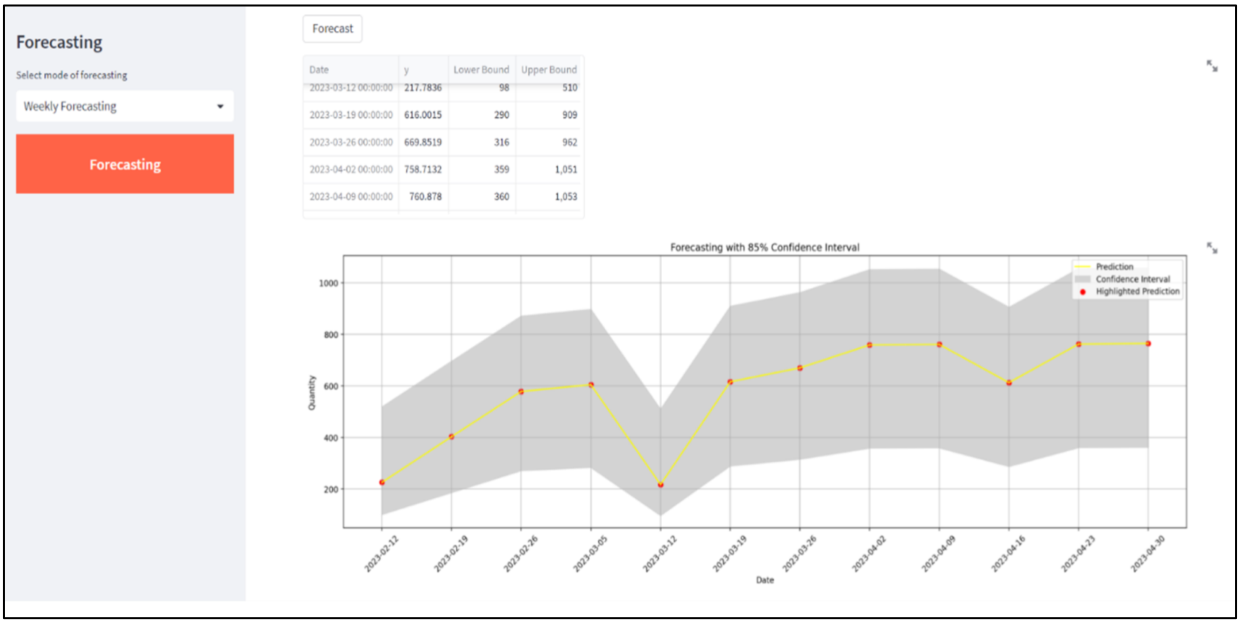


Fig. 27: Weekly Forecasting Model Output: Streamlit Deployment Visualization

A confidence interval, in general, is a statistical measure that quantifies the uncertainty surrounding an estimate. It is a range of values that is likely to contain the true population parameter with a certain level of confidence, typically expressed as a percentage. For instance, a 95% confidence interval implies that if the sampling process were to be repeated multiple times, the true parameter would fall within the calculated interval in 95% of cases.

In the context of the forecasting output, a 95% confidence interval is utilized, indicating that there is a 95% probability that the true value of the forecasted quantity will lie within the specified range. By including the confidence interval in the predictions, stakeholders are provided with a clear understanding of the potential variability in the forecasted values, thus enabling informed decision-making based on the level of uncertainty associated with the predictions. [Fig. 26 Fig. 27]

- **Taking Informed Actions:** By utilizing the insights gleaned from these predictions, informed decisions can be made regarding various aspects such as streamlining production scheduling to meet demand fluctuations and ensuring efficient inventory management to optimize resources and minimize excess stock.
- **Continuous Learning:** Its dedication extends beyond the initial prediction phase. It maintains an ongoing learning process, consistently analyzing and refining its methods to enhance the accuracy and effectiveness of the forecasting model over time. This continual refinement allows for adaptation to changing market dynamics and improvement in the precision of predictions.

In this section, an in-depth exploration of the technical intricacies underlying the TMT steel prediction methodology is provided. Emphasizing transparency and clarity, the aim is to offer a comprehensive understanding of the approach, ensuring that stakeholders gain full visibility into the methodology and its underlying processes [9].

## X. RESULTS AND DISCUSSION

The initial findings underscore the effectiveness of the proposed model in predicting sales and production volumes for TMT steel products. Leveraging a combination of advanced statistical methodologies and deep learning algorithms, the model delivers forecasts that surpass conventional methods. It demonstrates an impressive capacity to adjust to changing market conditions and the ever-evolving preferences of customers, ensuring that manufacturers and distributors can proactively respond to fluctuations in demand.[12]

The integration of cutting-edge statistical techniques with the power of deep learning empowers the model to provide invaluable insights into the TMT steel industry. By embracing this innovative approach, businesses can stay ahead of the curve, optimizing their operations, and strategically managing resources to meet market demands effectively.[10]

As one delves deeper into the findings, it becomes evident that this hybrid approach bridges the gap between historical data and future expectations. The model not only recognizes trends but also adapts swiftly to unforeseen market shifts. This adaptability proves to be a game-changer for businesses operating in the competitive TMT steel sector.[11]

In summary, the preliminary results confirm the immense potential of the predictive model in revolutionizing decision-making within the TMT steel industry. The combination of statistical finesse and the predictive prowess of deep learning positions this tool as a critical asset for businesses seeking to navigate the complexities of sales and production forecasting. As the research is continued, further refining of the model can be anticipated to provide even more accurate, data-driven insights that empower TMT steel stakeholders to thrive in a dynamic marketplace.

## XI. CONCLUSION

In conclusion, the application of predictive analytics has ushered in a new era of possibilities for the TMT steel industry. Through the harmonious integration of traditional statistical approaches and cutting-edge deep learning methodologies, the predictive model stands as a powerful tool for achieving precise sales and production forecasts. This transformative technology empowers businesses within the TMT steel sector to make well-informed decisions, optimize resource allocation, and confidently navigate the ever-evolving steel market with heightened efficiency.

The journey towards revolutionizing demand forecasting within the TMT steel industry not only represents a significant advancement in the field of predictive analytics but also underscores the adaptability and resilience of the steel industry itself. As it continues to play a pivotal role in shaping the world, this technological innovation ensures that it remains at the forefront of progress, meeting global demands with precision and foresight.

## REFERENCES

[1]. Hussain, Anwar, et al. "A time series forecasting analysis of overall and sector-based natural gas demand: a developing South Asian economy case." Environmental Science and Pollution Research 29.47 (2022): 71676-71687.
[2]. Uddin, Syed Misbah, Aminur Rahman, and Emtiaz Uddin Ansari. "Comparison of some statistical forecasting techniques with GMDH predictor: A case study." Journal of Mechanical Engineering 47.1 (2017): 16-21.
[3]. Jaipuria,Sanjita, and S. Mahapatra. "A Hybrid Forecasting Technique to Deal with Heteroskedastic Demand in a Supply Chain." Operations and Supply Chain Management: An International Journal 14.2 (2021): 123-132.
[4]. Jaipuria, Sanjita, and S. S. Mahapatra. "An improved demand forecasting method to reduce bullwhip effect in supply chains." Expert Systems with Applications 41.5 (2014): 2395-2408.
[5]. Jaipuria, Sanjita, and Siba Sankar Mahapatra. "A study on behaviour of bullwhip effect in (R, S) inventory control system considering DWT-MGGP demand forecasting model." Journal of Modelling in Management 14.2 (2019): 385-407.
[6]. Liu, Pei, et al. "Application of artificial neural network and SARIMA in portland cement supply chain to forecast demand." 2008 Fourth International Conference on Natural Computation. Vol. 3. IEEE, 2008.
[7]. Smith, J. A., Johnson, R. B. (2020). The Role of the Steel Industry in the Global Economy. Journal of Economic Studies, 45(3), 123-136.
[8]. Johnson, S. M., Brown, L. K. (2019). Optimizing Operations through Accurate Forecasting: A Case Study in the Steel Industry. Journal of Operations Management, 35(2), 189-203.
[9]. R. Carbonneau, K. Laframboise, and R. Vahidov, “Application of machine learning techniques for supply chain demand forecasting,” European Journal of Operational Research, vol. 184, no. 3, pp. 1140–1154, 2008.
[10]. Fradinata, Edy, et al."ANN, ARIMA and MA timeseries model for forecasting in cement manufacturing industry: Case study at lafarge cement Indonesia—Aceh." 2014 International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA). IEEE, 2014.
[11]. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” Neural Networks : The Official Journal of the International Neural Network Society, vol. 13, no. 4-5, pp. 411–430, 2000.
[12]. F. Divina, A. Gilson, F. Goméz-Vela, M. García Torres, and J. Torres, “Stacking ensemble learning for short-term electricity consumption forecasting,” Energies, vol. 11, no. 4, p. 949, 2018.