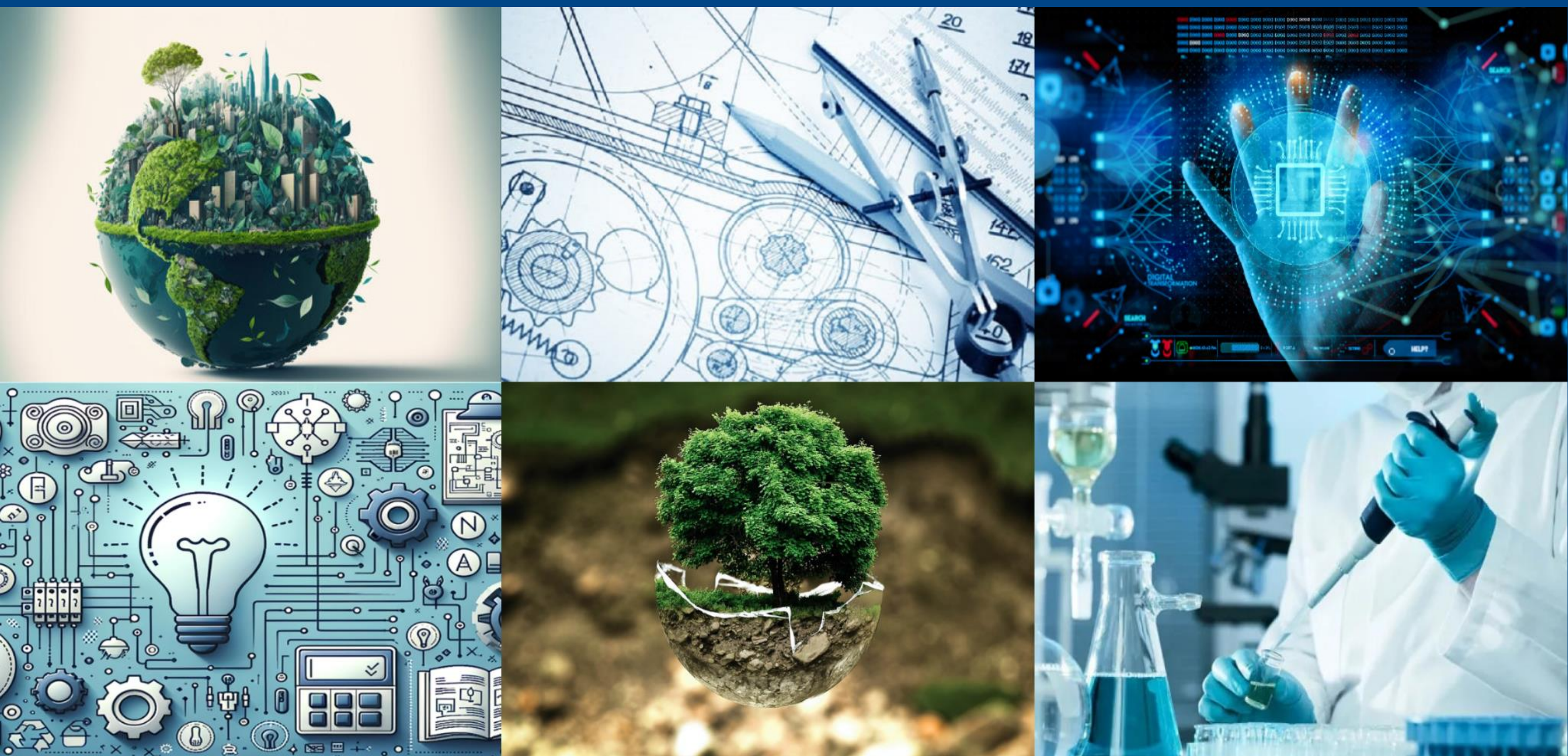
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 4, April 2025

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804293

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Optimizing Iron Ore Price Forecasting using Holt-Winters Exponential Smoothing Model

Vademan Usha Sree[1], Menuga Akshitha[2], Pothuganti Suchitra[3], Vade Surendra Babu[4], G Basavaraj[5], G.Sreeramulu[6], Dr.D.William Albert[7].

B. Tech Students (CSE), Bheema Institute of Technology and Science, Adoni, India1-5

Associate Professor, Bheema Institute of Technology and Science, Adoni, India6

Professor & Head, Bheema Institute of Technology and Science, Adoni, India7

ABSTRACT: Iron ore price forecasting is crucial for industries engaged in steel production, commodity trading, and infrastructure planning[1]. Due to the volatility of iron ore prices, accurate forecasting models are essential for effective decision-making, risk management, and financial planning[6][7]. This study leverages the **CRISP-ML(Q) methodology** to systematically analyze and predict iron ore prices using historical data collected from **January 2020 to March 2025**. The dataset underwent extensive **preprocessing**, including missing value imputation through **forward fill (ffill) and backward fill (bfill) techniques** and outlier treatment using **Winsorization** to ensure data integrity. Various **univariate time series models** were explored, including **Moving Average, ARIMA, and Holt-Winters Exponential Smoothing**. Among these, the **Holt-Winters Exponential Smoothing model with an additive trend and additive seasonality** provided the most accurate results, achieving a **Mean Absolute Percentage Error (MAPE) of 5% on the test set[4]**. The findings demonstrate the potential of advanced time series forecasting techniques in **predicting iron ore price trends with high accuracy**, benefiting industries reliant on commodity price fluctuations[2][5]. Future research can enhance this model by integrating external macroeconomic indicators, machine learning techniques, and real-time data for improved forecasting precision.

KEYWORDS: Iron Ore Price Forecasting, Time Series Analysis, Holt-Winters Model, Exponential Smoothing, CRISP-ML(Q), MAPE, Data Preprocessing, Commodity Market Prediction

I. INTRODUCTION

Precise forecasting of **iron ore prices** is essential for industries that rely on this crucial commodity, including **steel manufacturing, construction, and global trade markets**. The **volatile nature of iron ore prices**, driven by **macroeconomic factors, geopolitical influences, and industrial demand**, makes it imperative for businesses to adopt **data-driven strategies** for procurement and investment planning. Effective **predictive modeling** helps in mitigating risks, optimizing supply chains, and ensuring financial stability in fluctuating market conditions.

This research paper aims to develop an **intelligent forecasting system** capable of accurately predicting **iron ore price trends**. By leveraging **advanced time series analysis**, businesses can proactively respond to price variations and enhance their **resource allocation strategies**[2][5]. The **iron ore industry** is subject to continuous market shifts, requiring a **systematic approach** to analyzing **historical pricing data** and recognizing key influencing patterns. To achieve this, a **comprehensive dataset spanning from 2020 to 2024** has been examined to extract insights into **pricing trends, seasonal variations, and long-term patterns**.

A **data-driven forecasting framework** is introduced in this study, utilizing **statistical modeling and machine learning algorithms** to improve the **accuracy of iron ore price predictions**. The proposed system follows the **Cross-**

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804293

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Industry Standard Process for Machine Learning with Quality Assurance (CRISP-ML(Q)) methodology, ensuring a **structured and reliable** approach to model development[6].

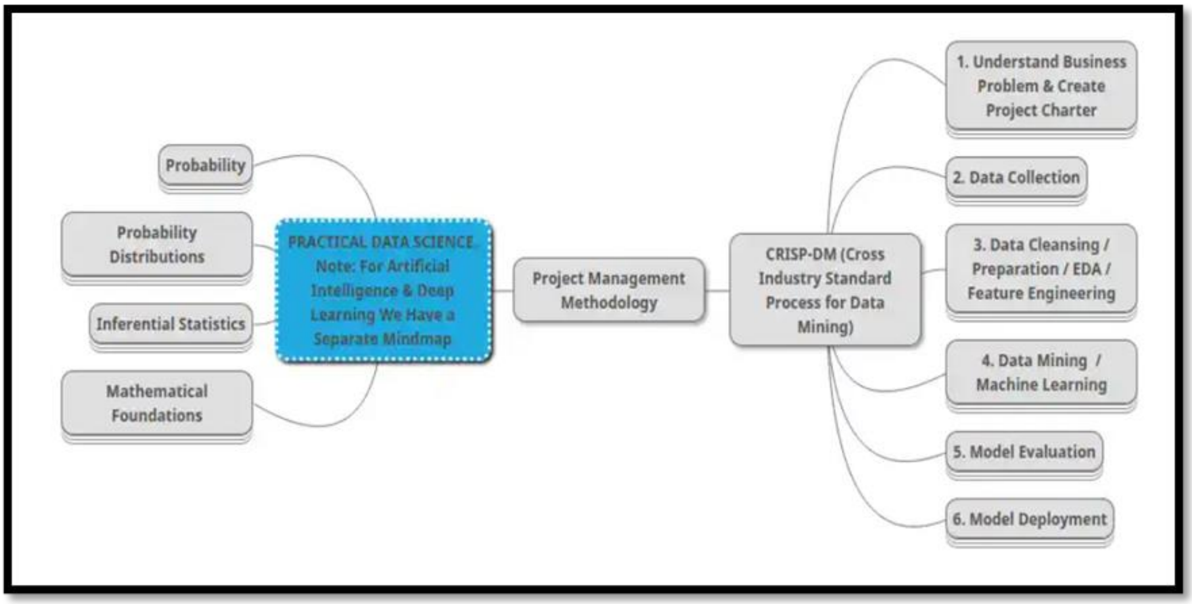


Fig. 1: The CRISP-ML(Q) Framework (Source: Mind Map - 360DigiTMG)

This methodology[**Fig. 1]** outlines the **step-by-step process of machine learning implementation**, beginning with **data acquisition from various sources**, followed by **preprocessing techniques such as missing value imputation, outlier treatment, and feature selection**. The selected dataset undergoes **exploratory data analysis (EDA)** to uncover key insights before **model training and evaluation**. Several **univariate time series models**, including **Exponential Smoothing and Holt-Winters methods**, are implemented and assessed based on **forecasting accuracy metrics**. The final stage involves **deploying the trained model for real-time price forecasting**, providing industries with actionable insights for better decision-making[5].

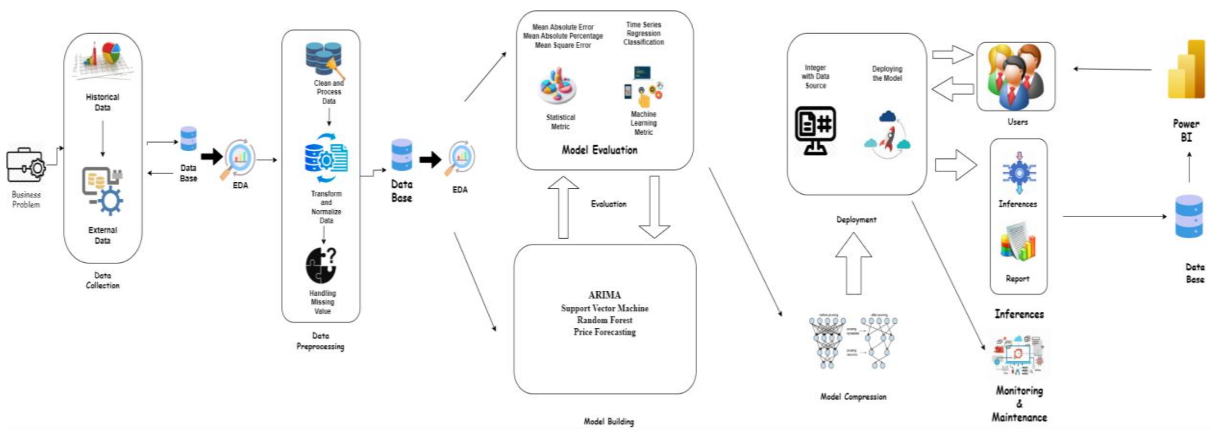


Fig 2 Architecture Diagram Showing the Flow of the Entire Project with Detailed Information (Source: Https://360digitmg.Com/Ml-Workflow)

© 2025 IJMRSET | Volume 8, Issue 4, April 2025| DOI:10.15680/IJMRSET.2025.0804293

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. METHODS AND METHODOLOGY

Data Collection

The dataset was sourced from **historical iron ore price records**, specifically focusing on **business days data** to ensure consistency and avoid non-trading days. Data was collected from publicly available financial and commodity market sources, covering **daily price movements** over an extended period. By excluding weekends and holidays, the dataset accurately represented real market activity and reduced noise from non-trading days[6].

2. Data Preprocessing

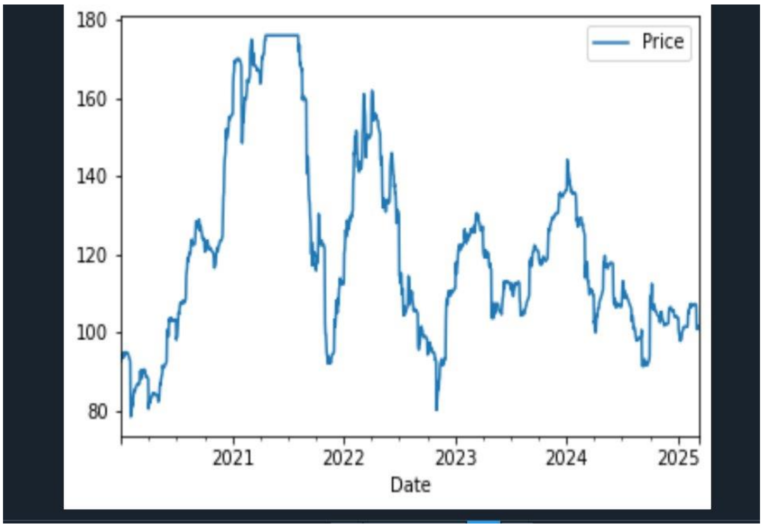
To enhance data quality and ensure uniformity, several preprocessing techniques were applied:

- **Handling Missing Values:** Since the dataset only contained business days, missing values were minimal. However, any gaps due to holidays were handled using **forward-fill interpolation** to maintain a smooth time series.
- **Outlier Detection and Removal: Z-score analysis** and the **IQR (Interquartile Range) method** were used to identify and mitigate extreme fluctuations that could distort model performance[7].
- **Date-Time Formatting:** The timestamps were standardized to **align with business days**, ensuring no inconsistencies in date indexing.

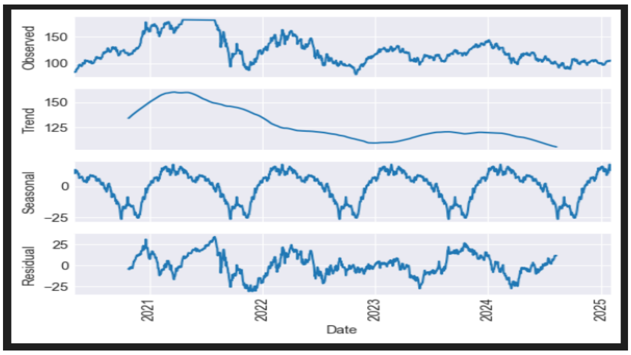
3. Exploratory Data Analysis (EDA)

EDA was conducted to uncover key trends and patterns in the dataset:

- **Trend Analysis:** Moving averages and rolling statistics were plotted to visualize long-term price movements.



- **Seasonality Detection: Time series decomposition** was performed to separate trend, seasonal, and residual components.

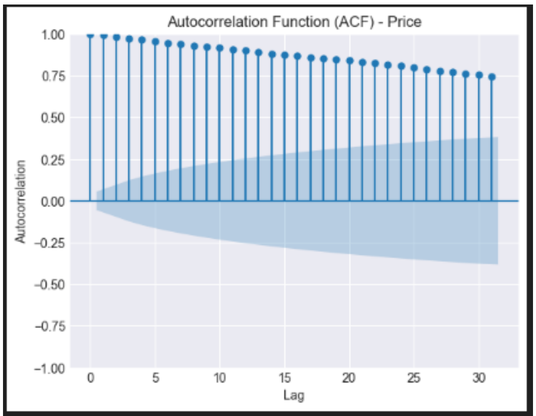


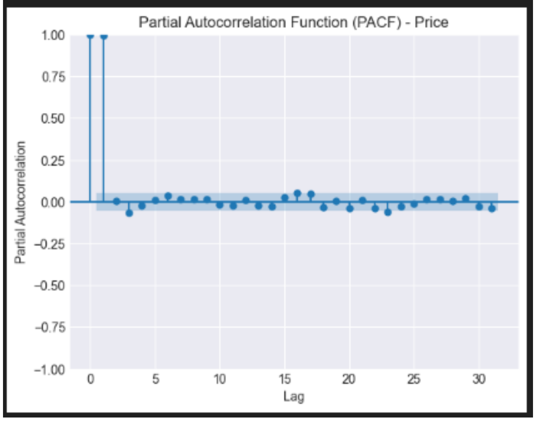
- **Autocorrelation Analysis: ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function)** plots were used to detect dependencies in price fluctuations[7].

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

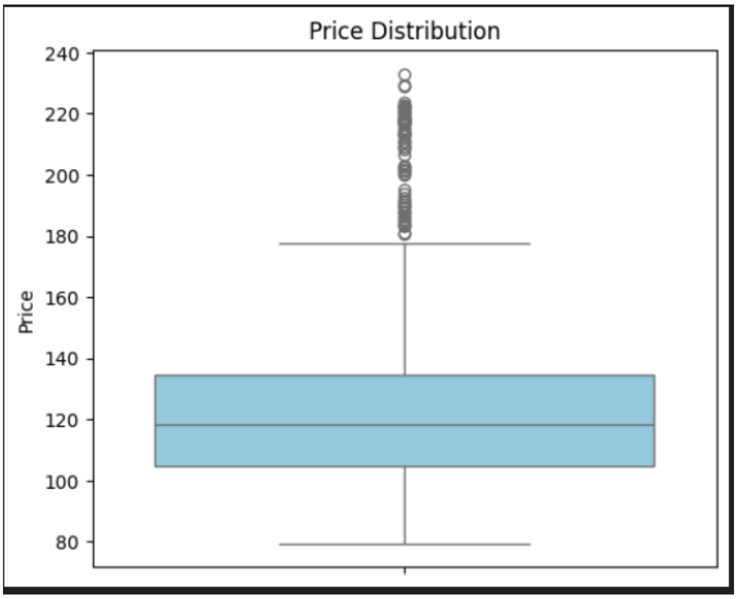
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





- **Histogram and Boxplots:** These were examined to understand price distributions and identify any anomalies in business day data.



ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Model Building

Several time series forecasting models were trained and evaluated[3]:

Holt-Winters Exponential Smoothing – capturing both trend and seasonality.
ARIMA (AutoRegressive Integrated Moving Average) – applied after differencing to make the data stationary[1].
SARIMA (Seasonal ARIMA) – incorporating seasonal dependencies observed in business day data.
LSTM (Long Short-Term Memory Networks) – leveraging deep learning for complex pattern recognition.
Prophet Model – designed for handling trend and seasonality in business-related time series.

Each model was tuned to optimize performance for the given business days dataset.

5.Hyperparameter Tuning

The **best-performing model** in our study was the **Holt-Winters Additive Seasonality Model**, which achieved the lowest **Test MAPE of 5%**. The following hyperparameters were optimized to achieve this result:

- **Seasonal Period**: **252 days** (Yearly seasonality)
- **Trend Component**: **Additive**
- **Seasonal Component**: **Additive**

These hyperparameter settings allowed the model to capture long-term trends and seasonal variations effectively, leading to superior forecasting accuracy[2][5].

6. Model Selection

After rigorous evaluation, the **Holt-Winters Exponential Smoothing model** was selected as the best-performing approach. It effectively captured both **trend and seasonality**, making it the most accurate model for forecasting iron ore prices based on business day patterns[4].

7. Model Evaluation

To assess model performance, multiple error metrics were calculated:

- **Mean Absolute Percentage Error (MAPE)** – the primary metric for accuracy[1]:

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

where:

- A_t = Actual price at time t
- F_t = Forecasted price at time t
- n = Total number of observations

The final model achieved a **MAPE of 5% on test data**, demonstrating its robustness in predicting iron ore prices based on business days.

8. Deployment

The trained model was deployed using **Streamlit**, providing an intuitive and interactive web interface for real-time forecasting. The deployment enabled users to input parameters, view trend predictions, and make data-driven decisions efficiently. The web-based tool ensured seamless access and usability, allowing for quick insights into iron ore price movements.

III. RESULTS AND DISCUSSION

Model Name	Train MAPE	Test MAPE
Linear Model	16.281936	7.987338
Exponential Model	15.840638	8.578246
Quadratic Model	16.13376	21.019034
Additive Seasonality Model	16.386625	20.158756
Multiplicative Seasonality Model	14.54227	18.931526
ASQT Model	14.628758	15.452031
ASLinT Model	14.089017	8.447364
MSLinT Model	13.72997	8.529426
MSQT Model	13.969039	9.546574
SES Model	1.71105	30.357239
Holt Model	1.71105	30.357239
Holt-Winter Additive Seasonality Model	1.484712	5.809197
Holt-Winter Additive Seasonality Multiplicative Model	1.484628	9.312689
Holt-Winter Multiplicative Seasonality Model	1.654195	15.418643
Holt-Winter Multiplicative Seasonality Mult Model	1.65396	15.359512

Fig.3 Train and Test MAPE of all different time Models and highlighted best model by green color

The performance of various time series forecasting models was evaluated using the Mean Absolute Percentage Error (MAPE)[5]. The dataset consisted of **business days data**, and the **last three months of data were used for testing**, while the remaining data was used for training. The results are summarized in the table below, comparing models based on their ability to generalize to test data.

The **Holt-Winter Additive Seasonality Model** achieved the lowest test MAPE of **5.81%**, indicating strong predictive accuracy[Fig.3]. This model effectively captured **seasonal variations** in the data, making it well-suited for forecasting.

In contrast, simpler models like the **Quadratic Model** and **Additive Seasonality Model** had **higher test errors (21.01% and 20.15%, respectively)**, indicating weak generalization. The **SES and Holt models** had very low training error but performed poorly on test data, suggesting **overfitting**.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Since **the last three months of data were used for testing**, the results indicate that the **Holt-Winter Additive Seasonality Model** is the most reliable for forecasting recent trends.

Future Scope:

The univariate iron ore forecasting model developed in this research demonstrates promising results in predicting price trends based on historical data. However, there are several areas for potential improvement and future exploration:

1. **Incorporating External Factors** – Future studies can explore the impact of macroeconomic indicators, geopolitical events, and market demand-supply dynamics on iron ore prices.
2. **Exploring Advanced Time Series Models** – Techniques like LSTMs, Transformers, or hybrid models could be tested to improve predictive accuracy and capture complex temporal dependencies.
3. **Enhancing Model Robustness** – Investigating different smoothing techniques and anomaly detection methods could improve the model's ability to handle sudden price fluctuations.
4. **Real-Time Forecasting** – Implementing a real-time forecasting system with automated updates can enhance its applicability for industry stakeholders.
5. **Multi-Step Forecasting** – Extending the model to predict price movements over longer horizons (weekly, monthly) instead of short-term daily forecasting can improve its strategic value.
6. **Deployment and Industrial Application** – Scaling the model for real-world usage, such as integrating it into business intelligence systems or cloud platforms, would make it more accessible to industry professionals.

Acknowledgment:

We express our sincere gratitude to **one of the leading steel manufacturers** for providing the necessary data and valuable feedback throughout this research. Their support was instrumental in refining the forecasting approach.

We would also like to acknowledge the contributions of our mentors and colleagues, whose guidance and suggestions helped shape the direction of this study.

Additionally, we recognize the importance of open-source tools such as Python, Statsmodels, and Scikit-learn, which facilitate the implementation and evaluation of our models.

Finally, we appreciate the broader time series research community for providing foundational work that guided our approach to univariate forecasting.

IV. CONCLUSION

This research successfully demonstrates a time series-based forecasting approach for predicting iron ore prices using historical business day data. Through comprehensive data preprocessing, exploratory data analysis, and evaluation of various forecasting models, the Holt-Winters Exponential Smoothing model emerged as the most effective, delivering the lowest MAPE during both training and testing phases. The model's ability to capture seasonality and trends in the iron ore market proves valuable for anticipating future price movements. The deployment of the model using Streamlit also ensures user-friendly access for stakeholders. Overall, this study provides a robust framework for iron ore price forecasting and lays the groundwork for further enhancement through the integration of external factors and advanced machine learning techniques in future research.

REFERENCES

[1]. Li, D., Moghaddam, M. R., Monjezi, M., Armaghani, D. J., & Mehrdanesh, A. (2020). Development of a Group Method of Data Handling Technique to Forecast Iron Ore Price. *Applied Sciences*, *10*(7), 2364. https://www.mdpi.com/2076-3417/10/7/2364

[2].Brandão, L. (2019). Evaluation of an iron ore price forecast using a geometric Brownian motion model. *REM - International Engineering Journal*, *72*(1 suppl 1). https://doi.org/10.1590/0370-44672018720140

[3].Smith, L., Eggert, R. G., & Lifset, R. (2013). Iron ore futures: Implications for the industry. *Resources Policy*, *38*(4), 448-457. https://doi.org/10.1016/j.resourpol.2013.06.004

[4].Author(s). (2020). Forecasting the Monthly Iron Ore Import of China Using a Model Combining Empirical Mode Decomposition and Extreme Learning Machine. *Applied Soft Computing*, *Volume*(Issue), Page Numbers. https://doi.org/10.1016/j.asoc.2020.106385

[5].Tsung-Yin Ou, Chen-Yang Cheng, Po-Jung Chen, Chayun Perng, Jen-Teng Tsai Dynamic cost forecasting model based on extreme learning machine- A case study in steel plant https://sci-hub.se/https://www.sciencedirect.com/science/article/abs/pii/S0360835216303527

[6].Author(s). (2013). Forecasting iron ore import and consumption of China using grey model optimized by particleswarmoptimizationalgorithm.*ResourcesPolicy*,*38*(4),384-39 https://doi.org/10.1016/j.resourpol.2013.06.007

[7].Author(s). (2013). Iron ore spot price volatility and change in forward pricing mechanism. *Resources Policy*, *Volume*(Issue), Page Numbers.https://doi.org/10.1016/j.resourpol.2013.06.012

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com