University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)

ISSN: 2714-3627

A Journal of the Faculty of Computing, University of Ibadan, Ibadan, Nigeria

Volume 16 No. 1, January 2026

journals.ui.edu.ng/uijslictr
http://uijslictr.org.ng/
uijslictr@gmail.com

University of Ibadan
Journal of Science and Logics in ICT Research

Cargo Revenue Prediction Model Using Machine Learning Approach

[1]✉ Olorunda S. P., [2]Ayodele E.O., [3]Adeyemi O. J., [4]Waheed A. A. and [5]Idowu P. A.

[1,4]Department of Computer Science, Lead City University, Ibadan, Oyo State
[2]Computer Science Program, Bowen University, Iwo, Osun State, Nigeria.
[3]Department of Computer Science, Tai Solarin Federal University of Education, Ijagun. Ogun State.
[5]Department of Computer Science and Cybersecurity, Obafemi Awolowo University, Nigeria.

[1]seyilux@gmail.com, [2*]emmanuel.ayodele@bowen.edu.ng, [3]adeyemioj@tasued.edu.ng,
[4]waheed.azeez@lcu.edu.ng, [5]paidowu@oauife.edu.ng

Abstract

This study develops a machine learning (ML) predictive model tailored to Nigeria's cargo ecosystem, aiming to enhance revenue forecasting, strategic planning, and operational efficiency. Data encompassing 1,133 records from Nigerian logistics firms (detailing variables such as cargo weight, shipping rates, and transaction dates) was preprocessed using one-hot encoding, normalization, and median imputation. Four primary regression models (Decision Tree [DTR], Random Forest [RFR], Gradient Boosting [GBR], and a Stacked Adaptive Multi-Input Regression Algorithm [SAMIRA]) were deployed via Google Colab. Exploratory Data Analysis (EDA) revealed right-skewed revenue distributions and seasonal operational peaks in January, August, and September. Model evaluation demonstrated that GBR outperformed the others, achieving an R^2 of 0.9989, Mean Squared Error (MSE) of ₦2.74 billion, Root Mean Squared Error (RMSE) of ₦52,350.29, and Mean Absolute Error (MAE) of ₦11,213.09. This superior performance was validated through 10-fold cross-validation (mean R^2 = 0.9969) and further visualized via a normalized error heatmap. Subsequently, the optimal model was prototyped into a Kotlin-based Android application for real-time forecasting. The findings demonstrate that GBR can achieve >99% forecasting accuracy, presenting a robust alternative to traditional methods and offering actionable insights for dynamic pricing and resource optimization in emerging markets.

Keywords: *Cargo Revenue Prediction, Gradient Boosting Regression, Machine Learning, Predictive Modeling.*

1. Introduction

Within the dynamic landscape of global commerce, cargo operations serve as the backbone of international trade, facilitating the movement of goods across borders and sustaining global supply chains. The economic significance of the cargo industry is immense; the global freight transport market, valued at approximately $19.34 trillion in 2023, is projected to reach $28.65 trillion by 2030 [1]. Furthermore, efficient cargo logistics directly stimulate economic growth. According to the World Bank's Logistics Performance Index, a 10% improvement in logistics efficiency correlates with a 0.5% GDP increase in developing economies [9]. As noted in recent literature, this evolution is characterized by increasing integration, digitalization, and operational complexity.

Predicting revenue is a vital part of doing business in this complex industry. It helps companies plan for the future, decide how to use their resources, and make sure they can keep going in the long run. In the past, cargo companies used traditional ways of forecasting, like looking at what happened in the past, adjusting for seasonal changes, and using fixed numbers. These methods relied a lot on human intuition. Jin *et.al.* [2] show that these old ways of doing things are only about 65-75% accurate when things are stable. And when the market gets crazy, they get a lot worse. We've seen some big surprises lately, like the COVID-19 pandemic and the blockage of the [3]. These

Olorunda S. P., Ayodele E.O., Adeyemi O. J., Waheed A. A. and Idowu P. A.. (2026). Cargo Revenue Prediction Model Using Machine Learning Approach. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 16 No. 1, pp. 180 – 191

events showed just how weak traditional forecasting models can be, leading to big mistakes - sometimes 30-40% off the mark.

The growing recognition of data as a strategic asset and the advent of Machine Learning (ML) have introduced a revolutionary shift in revenue forecasting. As a subset of artificial intelligence, ML algorithms are uniquely capable of identifying intricate patterns and dependencies within large datasets without relying on explicit programming. Today, ML has transitioned from an experimental concept to an indispensable industry tool. It is reported that 65% of large logistics companies now utilize ML for financial forecasting, yielding up to a 22% improvement in accuracy compared to traditional statistical models [5]. These improvements translate into tangible business benefits, including an overall profitability increase of up to 5% through optimized resource allocation and pricing strategies.

This study looks at how machine learning can be used to predict cargo revenue in Nigeria. It combines machine learning with local knowledge to help make better decisions about revenue management, especially when there's a lot of data to consider. By doing this, it adds to the ongoing conversation about using data to make informed decisions and provides useful tools for managing revenue in complex environments. The goal is to make revenue management more effective and efficient, which is important for businesses and organizations in Nigeria.

2. Related Works

Recent scholarly efforts have increasingly focused on applying intelligent algorithms to financial forecasting. Zhang [10] introduced an innovative methodology for predicting sustainable production revenue by combining correlation-based feature selection with a Modified Bat Algorithm (MBA). Analyzing production data encompassing material costs and efficiency metrics, the MBA achieved 94.72% accuracy. However, this study predominantly utilized datasets from large-scale manufacturers, highlighting a gap in localized, small-to-medium enterprise contexts where distinct market constraints apply.

In the transportation industry, a team of researchers led by PwC [6] created a system to help airlines make more money from cargo. They used a huge amount of data - over 140,000 bookings - to test their idea. The results showed that their method, called Gradient Boosting Machines, was very good at predicting demand, getting it right 87.3% of the time. However, the system was mainly tested on big international routes, and didn't take into account places like Nigeria, which have their own set of problems with infrastructure and seasonal changes that affect cargo transport. This means the system might not work as well in these emerging markets.

Similarly, Mane *et. al.* [4] evaluated sequential ML algorithms (specifically Long Short-Term Memory (LSTM) and Prophet) for revenue forecasting using a generalized dataset. While LSTM achieved a superior mean absolute percentage error (mAPE) of 12.49%, the reliance on generic data limits the real-world applicability of their findings. Furthermore, evaluating models solely on mAPE can prove problematic in datasets containing near-zero values. Topaloğlu *et. al* [8] tackled e-commerce revenue forecasting using a hybrid intelligent method involving Random Forest and DBSCAN clustering. While effective, the study inadequately addressed temporal dynamics in revenue patterns, impacting long-term reliability.

This research aims to fill in the gaps that have been identified by using a real-world dataset from the cargo sector in Nigeria. It uses advanced machine learning models and strict testing methods to make sure the results are reliable and useful for logistics in emerging markets.

3. Methods

The dataset used for this study was a total of 1,133 obtained from Evergreen Logistics, one of Nigerian top-tier cargo firm. The dataset was downloaded as csv file and contains transaction-level records relating to cargo shipments, service types, financial payments, client categories, and realized revenue. The dataset consists of several features describing each cargo operation from waybill numbers, transaction dates, destinations, service types, amount paid, outstanding balances, payment methods, and revenue realized, among others. The target label identifying the revenue outcome is labeled as "Revenue Realized", while others serve as independent variables for the regression modeling task.

3.1 Method of Pre-Processing of Collected Data

The information used in this study came from 1,133 records of transactions from Evergreen Logistics, a well-known cargo company in Nigeria. This data includes things like waybill numbers, when the transactions happened, where the cargo was going, what kind of clients were involved, how payments were made, and how much money was actually made from these transactions. All of these details are broken down in Table 1.

To get the data ready for modeling, we broke down the transaction dates into smaller parts like month and day of the week. This helped the algorithms find patterns that happen at the same time every year or week. We also had to deal with missing financial numbers, and used a method called median imputation. This method is good at handling high or low numbers that can throw off the results. The categorical features were then changed, like words or categories, into numbers using special codes. All the numerical features were on the same scale, so some numbers wouldn't be too big or small compared to others. Finally, the data were separated into two groups: one for training the model, which was 80% of the data, and the remaining 20% for testing.

3.2 Method of Formulation of Machine Learning Models

A supervised machine learning approach was adopted, deploying the following regression models:

1. Decision Tree Regressor (DTR): Selected for its simplicity and capability to map non-linear relationships via recursive dataset splitting.
2. Random Forest Regressor (RFR): An ensemble approach that aggregates multiple decision trees to mitigate overfitting and enhance accuracy.
3. Gradient Boosting Regressor (GBR): Builds additive models iteratively, minimizing residual errors at each stage for highly refined predictions.
4. Stacked Adaptive Multi-Input Regression Algorithm (SAMIRA): A customized hybrid ensemble method leveraging stacking and boosting for optimal generalization.

Model development, simulation, and testing were conducted within the Google Colaboratory (Colab) environment utilizing Python libraries such as *scikit-learn*, *pandas*, *NumPy*, *matplotlib*, and *seaborn*.

3.3 Evaluation Metrics

Model performance was systematically examined using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2).

Table 1: Description of Identified Variables for Revenue Prediction

Feature	Description	Data Type
Waybill Number	Unique identifier for each shipment	Object
Date	Date of transaction	Date/Time
Month	Month extracted from the date	Integer
Day of Week	Day of the week from transaction date	Categorical
Client Category	Type of client (e.g., Corporate, Retail)	Categorical
Destination	Delivery destination city or location	Categorical
Branch	Origin branch responsible for the shipment	Categorical
Amount Paid	Amount paid by the customer for the service	Float64
Balance	Outstanding amount yet to be paid	Float64
Payment Method	Mode of transaction payment (e.g., Cash, POS, Transfer)	Categorical
Service Type	Nature of cargo service (e.g., Local, Waybill)	Categorical
Number of Items	Quantity of packages shipped	Integer
Revenue Realized	Final revenue earned from the transaction (Target Variable)	Float64

To ensure that results were not biased by specific data splits, a 10-fold cross-validation approach was applied.

4. Results and Discussion

4.1 Exploration Data Analysis (EDA)

The dataset contains 7 columns with 1133 records, aimed at predicting revenue (Total) in the cargo business, with variables including both numerical and categorical types. Key numeric variables like rate, kilo, month, year, and total(revenue) distance show a right-skewed distribution, while categorical fields such as name and load provide insight into operational diversity. A small proportion of missing values was found in the load and total fields, which were imputed using the median and mode, respectively. This exercise makes the dataset fit for model building. These variable datatypes were converted to their best state which is numerical so that the ML models can process, analyze, and generate more accurate and reliable predictions.

^{72} is statistically.

From the descriptive statistical summary, Table 1 shows the distribution of kilo values is highly slanted to the right. With its mean value is approximately 294 KG, the median is significantly lower at 65 KG, indicating that most vendors deal in relatively small quantities, The maximum value of 5,487 KG suggests the presence of large outliers, which may represent bulk purchases. The Rate feature exhibits low dispersion with a mean of ₦1,992.25 and a standard deviation of ₦546.27. The 25th, 50th, and 75th percentiles are all ₦2,000, indicating a high concentration around a fixed price point, possibly a standard pricing policy. However, the minimum (₦1,600) and particularly the maximum at the same fixed price. (₦2,000).

While the Total payment amounts reflect the greatest variability. The mean total payment is ₦580,629, but the median is just ₦130,000, with a standard deviation of ₦1.39 million. This indicates an extremely right-skewed distribution. While high payments may be expected for large transactions, the maximum value of ₦1.0974 × 10

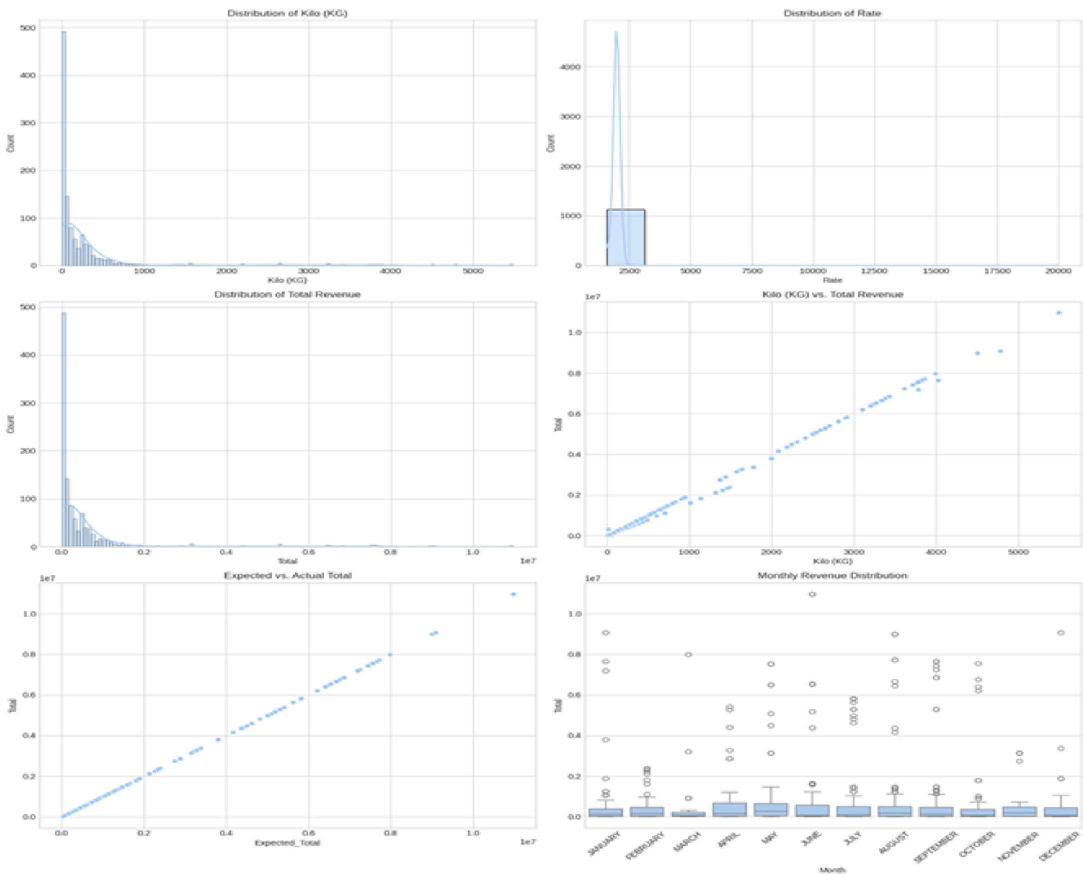

Figure 1: Exploratory Data Visualizations

Table 1 Dataset Descriptive Statistics Summary

Metric	Kilo (KG)	Rate (₦)	Total (₦)
Count	1135	1123	1134
Mean	294.01 KG	₦1,992.25	₦580,629.50
Std Dev	700.83 KG	₦546.27	₦1,389,341.00
Min	0 KG	₦1,600.00	₦0.00
25th %	19 KG	₦2,000.00	₦38,000.00
Median	65 KG	₦2,000.00	₦130,000.00
75th %	256 KG	₦2,000.00	₦504,500.00
Max	5,487 KG	₦2,000.00	₦1.0974×10^{72}

Figure 1 displays the distribution of key variables (Kilo, Rate, and Total), their inter-relationships, and monthly revenue patterns. Visualizations reveal highly skewed distributions for Kilo (KG) and Total revenue, highlighting a few high-volume transactions amid a majority of smaller shipments, a common dynamic in Nigeria's multifaceted business environment. The Rate variable peaks sharply around ₦2,000, suggesting consistent pricing, while a strong linear relationship between Kilo and Total confirms that revenue is primarily volume-driven. The Expected vs. Actual Total plot shows near-perfect alignment, implying accurate, formula-based data recording. Although monthly revenue remains generally stable, box plots highlight extreme outliers in July and August, hinting at seasonal bulk sales. These anomalies were carefully addressed during the data preprocessing stages.

Figure 2 visualizes load and monthly transaction counts. The first load is the most frequent (approximately 360 occurrences), likely representing the primary operational stage. A consistent decline follows, with the second (320), third (270), and fourth (180) loads demonstrating diminishing frequency as the load sequence increases. Regarding monthly distributions, September peaks at 165 transactions, followed closely by January and August (150–155 counts), indicating distinct seasonal activity spikes. Conversely, December, February, and the spring months (March–May) reflect low to moderate activity, with March dipping to a minimum of roughly 40 counts. The stark contrast between high January activity and low December/February activity implies strong year-beginning operational shifts.

Ultimately, these combined observations outline a business process governed by distinct seasonal cycles and load-dependent operations. Understanding both the load sequence patterns and monthly fluctuations is crucial for accurate optimization and forecasting.

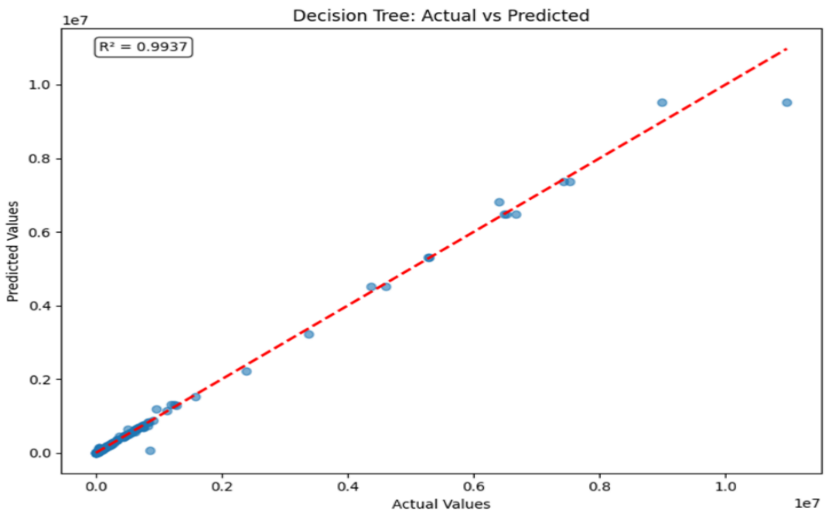

Figure 2: Decision Tree Regression Predicted and Actual Values

4.2 *Results of Formulated Model*

This section presents a detailed results of the performance of each model used in the cargo revenue prediction including Decision Tree Regression, Gradient Boosting Regression, Random Forest Regression and Stacked Adaptive Multi-Input Regression Algorithm (SAMIRA) The evaluation is based on key performance metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), R-square (R^2), and Root Mean Squared (RMS).

a. Result of Decision Tree Regression

As illustrated in Figure 2 and Table 2, the Random Forest model demonstrates high accuracy in forecasting cargo revenue. The scatter plot reveals predicted values clustering tightly around the red dashed line of perfect prediction. This visual alignment is quantitatively validated by an R^2 score of 0.9970, indicating that the model accounts for 99.70% of the variance in actual revenue. Furthermore, the model recorded a Mean Squared Error (MSE) of ₦7,161,241,078, a Root Mean Squared Error (RMSE) of ₦84,624.12, and a Mean Absolute Error (MAE) of ₦16,433.35 (see Table 3). While these metrics reflect strong overall precision, the noticeable gap between the RMSE and MAE indicates a slight sensitivity to outliers, though this does not significantly undermine its predictive strength. Although it performs slightly below Gradient Boosting and SAMIRA in raw metrics, Random Forest provides robust and reliable forecasts, making it a highly practical tool for real-world cargo revenue management.

b. Results of SAMIRA Model

The SAMIRA model exhibits excellent predictive capability, achieving an R^2 of 0.9984, thereby accounting for approximately 99.84% of the variance in total revenue. This exceptionally high coefficient of determination indicates that the model captures underlying data patterns with remarkable accuracy. As detailed in Table 3, the model recorded an MSE of ₦3,922,296,281.73, an RMSE of ₦62,628.24, and an MAE of ₦22,736.50. These relatively modest error metrics confirm that forecasting inaccuracies remain within acceptable financial thresholds. Furthermore, the narrow gap between the RMSE and MAE reflects consistent model stability with minimal sensitivity to extreme outliers.

Figure 3 visually reinforces this performance by mapping actual versus predicted revenue. The individual transaction data points (blue dots) cluster tightly along the red dashed line of ideal prediction. This strong visual alignment demonstrates that the SAMIRA model produces highly precise forecasts with negligible deviation, perfectly corroborating the previously reported quantitative metrics.

c. Results of Gradient Boosting Regression Model

The performance evaluation of the Gradient Boosting model for cargo revenue prediction is illustrated in Figure 4 (Gradient Boosting: Actual vs Predicted). The model achieved an impressive R^2 score of 0.9989, indicating that 99.89% of the variability in actual revenue is accurately captured by the model's predictions.

Table 2: Result of Decision Tree Regression Performance Metrics

Models	MSE	RMSE	MAE	R2
Decision Tree	152187519901	123364	26536	0.993661

Table 3: Result of SAMIRA Performance Metrics

Models	MSE	RMSE	MAE	R2
SAMIRA	3922296282	62628	22737	0.998366

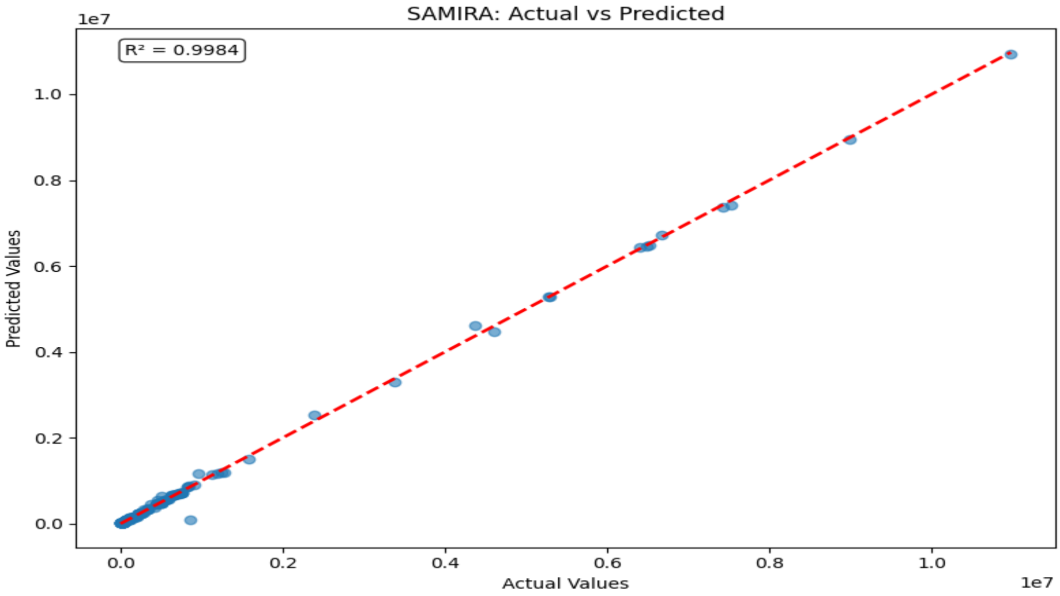

Figure 3: SAMIRA Predicted and Actual Values

Table 4: Result of Gradient Boosting Regression Performance Metrics

Models	MSE	RMSE	MAE	R2
Gradient Boosting	2740552680	52350	11213	0.998858

The plotted data in Figure 4 shows a tight clustering of points around the red dashed line (representing the line of perfect prediction), suggesting minimal deviation between actual and predicted values. Also, the Gradient Boosting model achieves the strongest overall performance with the highest R^2 value of 0.998858, MSE: ₦2,740,552,679.94, RMSE: ₦52,350.29, MAE: ₦11,213.09 and R^2: 0.998858 when measured with performance metrics as shown in Table 4. The model shows exceptional accuracy with the lowest RMSE (₦52,350.29) and remarkably low MAE (₦11,213.09), indicating superior prediction precision. The close relationship between RMSE and MAE suggests stable performance across different prediction scenarios with minimal outlier influence

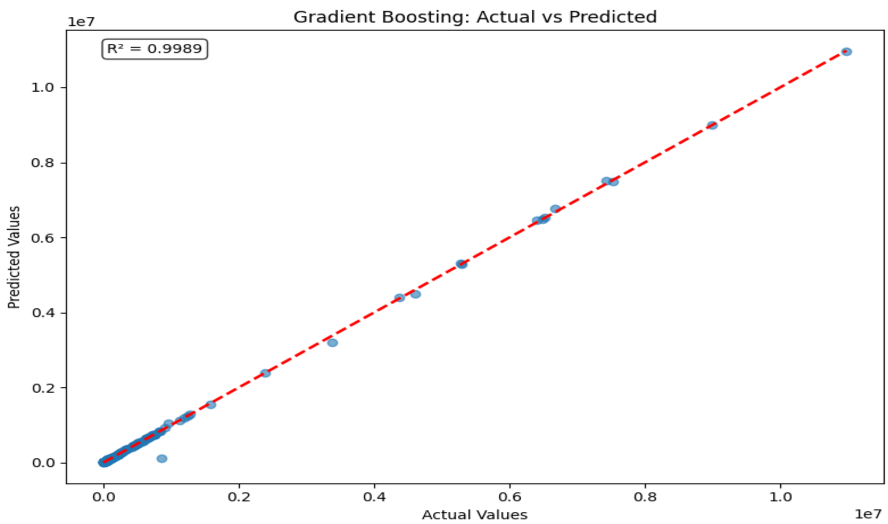

Figure 4: Gradient Boosting Regression Predicted and Actual Values

d. *Results of Random Forest Regression Model*

As illustrated in Figure 5, the Random Forest model delivers a high degree of accuracy in forecasting cargo revenue. The scatter plot mapping actual versus predicted values reveals a tight clustering around the red dashed line, denoting ideal prediction. This visual precision is quantitatively supported by an R^2 score of 0.9970, indicating that the model successfully accounts for 99.70% of the variance in the actual data.

Furthermore, the model yielded a Mean Squared Error (MSE) of ₦7,161,241,078, a Root Mean Squared Error (RMSE) of ₦84,624.12, and a Mean Absolute Error (MAE) of ₦16,433.35. While these metrics reflect strong predictive capabilities, the notable disparity between the RMSE and MAE highlights a moderate sensitivity to extreme outliers. Nevertheless, this characteristic does not diminish the model's overall reliability. Although its raw performance metrics trail slightly behind those of Gradient Boosting and SAMIRA, Random Forest remains a highly practical option for real-world logistics applications. Its excellent generalization, high explanatory power, and inherent resistance to overfitting make it particularly valuable for operations that prioritize model interpretability alongside accuracy.

4.2 Results of Model Validation

To mitigate data-split bias, prevent overfitting, and ensure robust generalization on unseen data, this study employed a 10-fold cross-validation framework. The resulting R^2 scores for the top-performing models are detailed in Table 5 and ranked visually in Figure 6.

Gradient Boosting Regression (GBR) maintained its superior standing, achieving the highest mean R^2 (0.9969) alongside a narrow standard deviation of 0.0027. This highlights both exceptional accuracy and fold-to-fold consistency, solidifying its position as the premier model for cargo revenue forecasting. Random Forest Regression also delivered strong results (mean $R^2 = 0.9961$; SD = 0.0037); however, its slightly higher variability points to modest performance fluctuations across different data subsets.

Conversely, the Decision Tree model (the simplest algorithm tested) recorded a stable mean R^2 of 0.9960 with a low standard deviation of 0.0023.

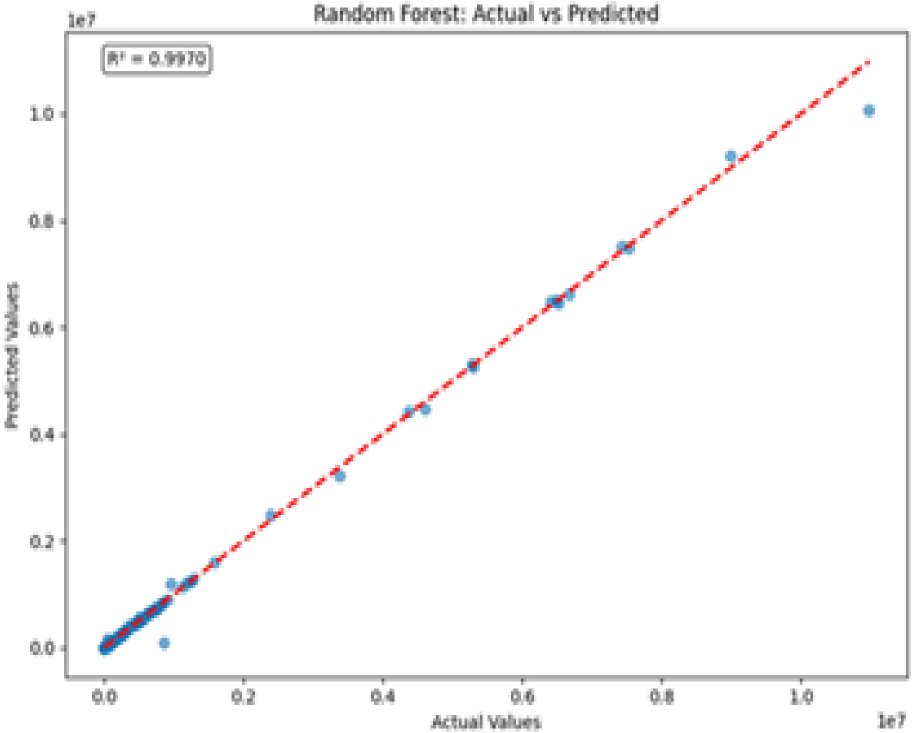

Figure 5: Random Forest Regression Predicted and Actual Values

While it lacks the sheer precision of the ensemble methods, its computational efficiency and high interpretability justify its potential use in resource-constrained environments. Ultimately, these cross-validation metrics confirm that all evaluated models generalize effectively, with GBR retaining a definitive edge in both predictive power and operational stability.

4.3 Results of Model Performance Comparison

In the model comparison as displayed in the Figure 7, Gradient Boosting Regression still achieved optimal results with MSE of 2.74 billion and R^2 of 0.9989, followed by Random Forest Regression, Decision Tree Regression, and SARIMA respectively. SARIMA exhibited 5.5-fold higher MSE compared to Gradient Boosting in error magnitude differences**,** indicating substantial prediction accuracy variations between models. Gradient Boosting and Random Forest Regression significantly outperformed single-model methods, demonstrating improved generalization capabilities due to their ensemble bases.

In model compatibility, SARIMA's poorly performance suggests inappropriate model selection for the given dataset characteristics, highlighting the importance of algorithm-data matching. in this study of cargo revenue prediction, the performance results differences translate to approximately 80% reduction in prediction errors when selecting optimal versus poorest performing models.

4.4 Result of Normalized Error Metrics Heatmap

Figure 8 presents a Normalized Error Metrics Heatmap, utilizing color-coded intensity to visualize model performance for cargo revenue prediction. By scaling the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) onto a 0-to-1 range, this heatmap provides a standardized comparative overview. The resulting color gradients corroborate earlier findings while sharply defining the performance disparities among the evaluated models.

Gradient Boosting Regression (GBR) emerged as the unambiguously optimal choice, achieving flawless normalized scores of 0.000 across all three error metrics. This "perfect green sweep" solidifies its superior predictive reliability. Conversely, Random Forest Regression maintained respectable and consistent results, with error values ranging between 0.34 and 0.45, establishing it as a balanced and reliable alternative without glaring weaknesses.

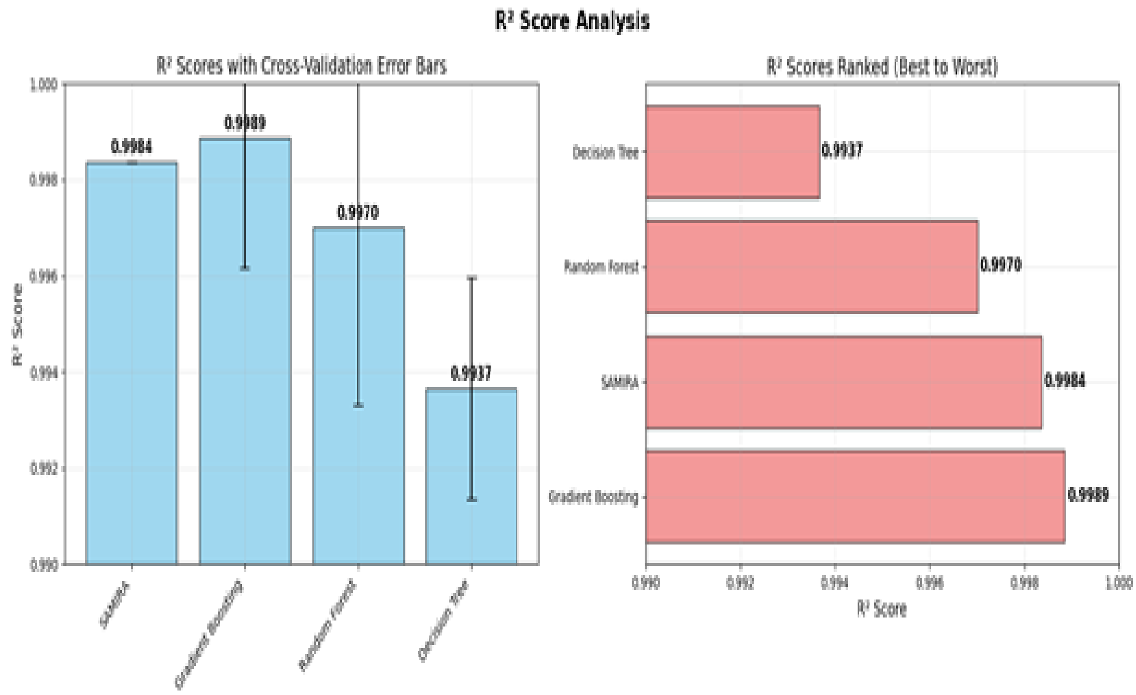

Figure 6: Results of Model Performance Comparison

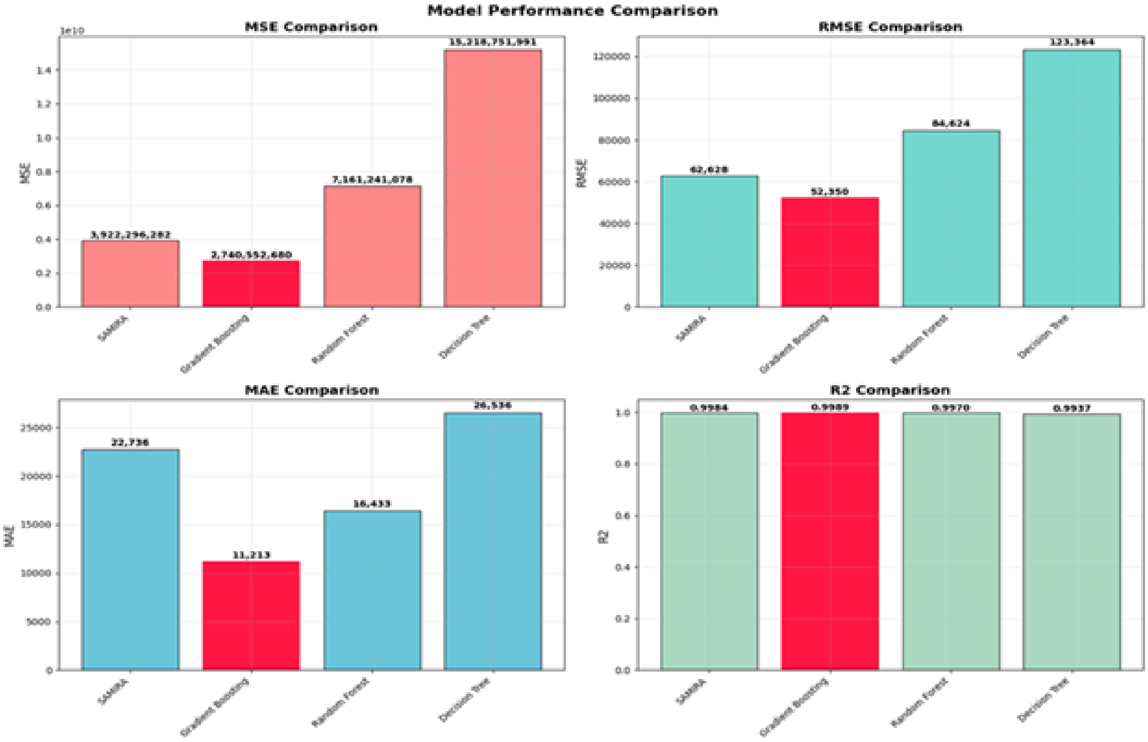

Figure 7: Results of Model Performance Comparison

On the other hand, the Decision Tree model performed poorly across all dimensions. Its normalized values peaked at 1.000 for every metric, indicating the highest error rates and rendering it the least suitable option for real-world deployment. Interestingly, the time-series-specific SARIMA algorithm presented a more nuanced profile; while it delivered relatively strong MSE and RMSE outcomes, it faltered significantly in MAE, revealing performance vulnerabilities that raw numerical tables alone might have obscured

.

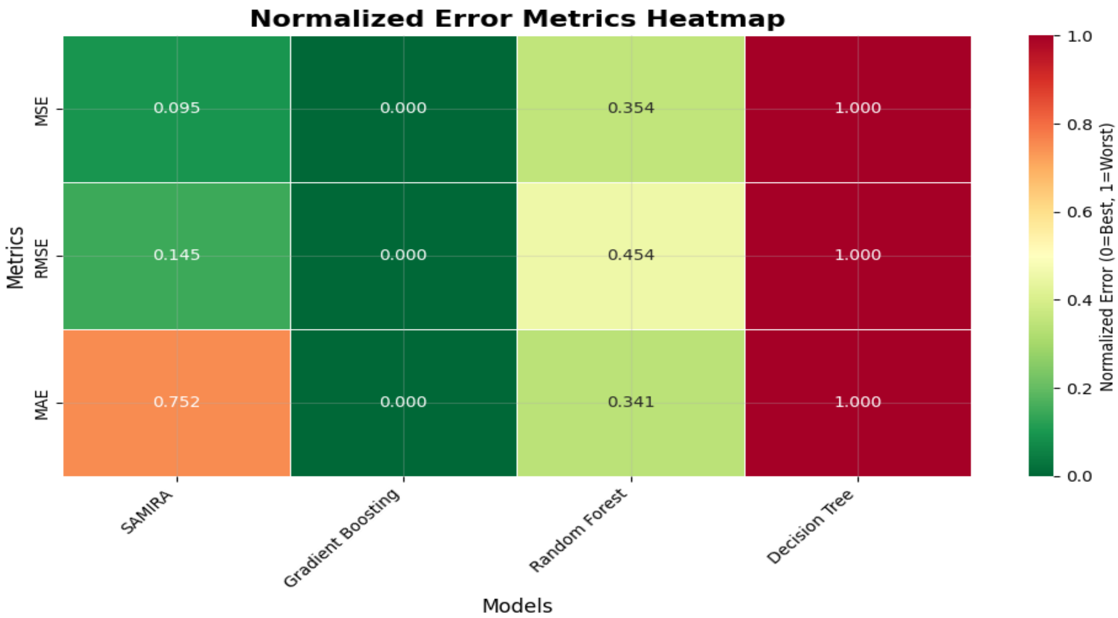

Figure 8: Results of Normalized Error Metrics of the Models

4.5 Discussion of the Results

Exploratory Data Analysis (EDA) of the 1,133-record Nigerian cargo dataset revealed highly right-skewed numerical distributions. The presence of extreme statistical outliers (including maximum cargo weights of 5,487 kg and exponentially high revenue entries) reflects a market characterized by routine small-to-medium shipments punctuated by occasional, massive bulk transactions. Furthermore, categorical variables highlighted distinct operational peaks in January, August, and September, indicative of demand-driven seasonal cycles. Following robust median and mode imputation to resolve missing values, the data demonstrated a strong linear correlation between cargo weight and realized revenue. This linear relationship, interwoven with complex seasonal fluctuations, necessitated the deployment of advanced algorithms capable of mapping multivariate and non-linear interactions.

Consequently, this section synthesizes the performance of the five evaluated predictive models: Decision Tree Regression (DTR), Gradient Boosting Regression (GBR), Random Forest Regression (RFR), SAMIRA, and SARIMA. Model efficacy is assessed using standard error metrics (MSE, RMSE, MAE) and R^2 scores, while robustness is validated through cross-validation and a normalized error heatmap. By contextualizing each model's strengths, limitations, and alignment with the EDA findings, this discussion provides cargo operators with a comprehensive, practical framework for selecting and deploying data-driven forecasting solutions.

5. Conclusion

This study successfully developed and validated predictive machine learning frameworks for revenue forecasting within Nigeria's cargo sector. Analyzing 1,133 transactional records, the research established that ensemble machine learning methodologies (specifically Gradient Boosting Regression) vastly outperform traditional forecasting and basic linear models. GBR achieved an exceptional R^2 of 0.9989, demonstrating robustness against the dataset's right-skewed bulk outliers and seasonal volatility.

By translating raw transactional data into actionable foresight, this localized predictive system addresses crucial operational bottlenecks in emerging markets. It enables logistics managers to optimize resource allocation, formulate dynamic pricing strategies, and secure long-term financial stability. Furthermore, this research enriches the currently limited body of literature concerning advanced predictive analytics in African supply chains, offering a scalable blueprint that regulators and operators can integrate to standardize data practices and elevate industry efficiency.

References

[1] International Transport Forum. (2024). Global Freight Transport Outlook to 2030. OECD Publishing

[2] Jin, G., Feng, W., & Meng, Q. (2022). Prediction of Waterway Cargo Transportation Volume to Support Maritime Transportation Systems Based on GA-BP Neural Network Optimization. *Sustainability*, *14*(21), 13872. https://doi.org/10.3390/su142113872

[3] Kramar, V. and Alchakov, V. (2023). Time-series forecasting of seasonal data using machine learning methods. Algorithms, 16(5), 248. https://doi.org/10.3390/a16050248

[4] Mane, V., Patil, R., Patwardhan, S., & Pethkar, P. (2024). Revenue Prediction using Sequential Machine Learning. 1492–1496. https://doi.org/10.1109/i-smac61858.2024.10714745

[5] Oracle. (2023). CFO Guide to Revenue Forecasting. https://www.oracle.com/ng/erp/cfo/revenue forecasting/

[6] PwC. (2024). Global Digital Operations Study for Logistics Companies. PricewaterhouseCoopers.
Rizzo, S. G., Chen, Y., Pang, L., Lucas, J., Kaoudi, Z., Quiane, J., & Chawla, S. (2020). Prescriptive learning for air-cargo revenue management. In C. Plant, H. Wang, A. Cuzzocrea, C. Zaniolo, & X.
Steinberg Fabian, Peter Burggräf, Johannes Wagner, Benjamin Heinbach, Till Saßmannshausen, and

[7] X. Wu, Ed., "Proceedings - 20th IEEE International Conference on Data Mining, ICDM 2020," pp. 462–471, Art. no. 9338388, Institute of Electrical and

Electronics Engineers Inc., Nov. 2020. doi: 10.1109/ICDM50108.2020.00055.

[8] Topaloğlu, G., Kalaycı, T. A., Pekel, K., & Akay, M. F. (2023). Revenue forecast models using hybrid intelligent methods. International Journal of Mathematics and Computer in Engineering, 2(1), 117–124. https://doi.org/10.2478/ijmce-2024-0009

[9] World Bank. (2023). Connecting to Compete: Trade Logistics in the Global Economy. The World Bank Group.

[10] Zhang, H. (2025). Analysis of Business Factors Influencing Film Revenue. Advances in . Economics, Management and Political Sciences, 166(1), 1–7. https://doi.org/10.54254/2754-1169/2025.20876.