Jurnal Inovasi Teknologi dan Edukasi Teknik, 5(2), 2025

EVALUATING MACHINE LEARNING MODELS FOR OPTIMIZING OVERALL EQUIPMENT EFFECTIVENESS IN FOOD MANUFACTURING SMES

Igbokwe Nkemakonam Chidiebube, Mgbojikwe Godson Uzochukwu, Charles Onyeka Nwamekwe and Aguh Patrick Sunday*

Nnamdi Azikiwe University, Awka, Nigeria, Along Enugu-Onitsha Expressway, Ifite Road, Awka 420110, Nigeria

*Corresponding author, email: ps.aguh@unizik.edu.ng

doi: 10.17977/um068.v5.i2.2025.5

Keywords
OEE Optimization
Machine Learning
Food Manufacturing
Linear Regression
Smart Data
Driven Manufacturing

Abstract

The optimization of Overall Equipment Effectiveness (OEE) is crucial for enhancing productivity and operational efficiency in food manufacturing Small and Medium Enterprises (SMEs). This study evaluates the application of machine learning models, including Linear Regression, Random Forest, Support Vector Machine (SVM), and Gradient Boosting, to optimize OEE based on key production factors such as Availability, Performance, and Quality. The data was analyzed using these models, with performance evaluated based on Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R-squared (R^2) scores. The results indicate that Linear Regression outperformed other models, achieving the highest accuracy with an R^2 score of 0.9546, demonstrating its ability to capture complex relationships within the data. These findings underscore the potential of machine learning-driven OEE optimization, enabling predictive maintenance, real-time efficiency monitoring, and proactive decision-making in food manufacturing SMEs. The study highlights how data-driven approaches can minimize downtime, enhance production efficiency, and contribute to smart manufacturing transformation. Future research can explore real-time implementation and integration with IoT-enabled systems to further enhance predictive accuracy and automation.

1. Pendahuluan

Production managers are driven to accomplish significantly higher levels of operational performance due to the difficulty of today’s dynamic production environment (Igbokwe & Godwin, 2022). Small and medium enterprises (SMEs) are critical to ensuring food security and economic development (Chukwutoo & Nkemakonam, 2018)in today's highly competitive food manufacturing industry. The National Bureau of Statistics in Nigeria defines a small and medium enterprise as an independent and distinct organization, which may include cooperative societies and non-governmental organizations, operated by one or more individuals and has branches or subsidiaries. Specifically, a small enterprise employs between 10 and 49 employees. It has assets (excluding land and buildings) worth between ₦5 million to ₦50 million, while a medium enterprise employs 50 to 199 people and possesses assets between ₦50 million to ₦500 million (Igbokwe & Mba, 2019). They contribute significantly to the economy and account for about 96% of businesses and 84% of employment in the country. Within the manufacturing sector, food processing SMEs represent

A major segment, supporting food security, value addition, and local raw material utilization (Nwamekwe & Nwabunwanne, 2025). However, these enterprises often face significant challenges in optimizing production efficiency due to limited resources, inconsistent equipment performance, and inadequate maintenance strategies (Igbokwe & Mba, 2019). One of the key performance metrics used to assess manufacturing efficiency is Overall Equipment Effectiveness (OEE), which provides insights into equipment availability, performance, and quality. Achieving optimal OEE is crucial for minimizing downtime, reducing waste, and improving productivity (Igbokwe & Godwin, 2021).

Traditional methods for OEE optimization depend on experience-based decision-making, static statistical models, and manual data collection, all of which are frequently ineffective and prone to human error. With the increasing adoption of Industry 4.0 and innovative manufacturing concepts, integrating Machine Learning (ML) models presents a promising approach to enhancing OEE in food manufacturing SMEs (Nwamekwe et al., 2024). Machine Learning techniques can analyze large volumes of real-time operational data, detect patterns, and provide predictive insights that facilitate proactive maintenance and process optimization. High-quality requirements, narrow profit margins, and restricted access to cutting-edge technologies particularly challenge SMEs in the manufacturing industry (Igbokwe et al., 2021, Nwamekwe et al., 2024). Studies show that Machine Learning models tailored for these environments can solve challenges such as production variability, defect detection, and compliance with regulatory standards (Chikwendu et al., 2023).

Despite the potential of Machine Learning in manufacturing optimization, there is limited research on the comparative evaluation of different Machine Learning models for OEE improvement, particularly within the context of SMEs in the food industry (Nwamekwe et al., 2024). Most existing studies focus on large-scale manufacturing plants with advanced automation infrastructure, leaving a gap in understanding how Machine Learning can be effectively applied in resource-constrained environments. Therefore, this study aims to evaluate and compare various Machine Learning models for OEE optimization in a food manufacturing SME, assessing their predictive accuracy, efficiency, and practical applicability (Nwamekwe & Okpala, 2025). The findings will provide valuable insights into the most suitable Machine Learning approaches for enhancing production efficiency, reducing losses, and driving sustainable growth in food manufacturing SMEs (Nwamekwe et al., 2024).

The optimization of Overall Equipment Effectiveness (OEE) has been a key area of research in manufacturing industries, as it directly influences productivity, cost efficiency, and competitiveness. OEE is traditionally evaluated based on three key components: Availability, Performance, and Quality (Chidiebube et al., 2025; Igbokwe & Godwin, 2021). While conventional statistical and heuristic methods have been widely used to improve OEE, recent advancements in Machine Learning (ML) offer more dynamic, data-driven approaches that can enhance predictive maintenance, fault detection, and real-time process optimization.

Several studies have explored the application of Machine Learning in optimizing manufacturing operations. For instance, Welte et al., (2023) demonstrated how predictive analytics using ML models could effectively identify machine failures and reduce downtime. Similarly, Maataoui et al., (2023) Random Forest techniques were implemented to enhance predictive maintenance, leading to a measurable improvement in OEE within an industrial setting. A study Biswaranjan et al., (2024) focuses on utilizing machine learning (ML) to improve Overall Equipment Effectiveness (OEE) in industrial manufacturing settings. The aim is to develop a machine learning model that predicts OEE scores, which are crucial for assessing the efficiency and productivity of manufacturing processes.

. However, the application of ML in Small and Medium Enterprises (SMEs) remains underexplored (Le Dinh et al., 2025; Oldemeyer et al., 2024; Peretz-Andersson et al., 2024; Phir & Mbale, 2024). SMEs often operate with limited automation, less structured data, and resource constraints, making it challenging to implement complex ML models (Bauer et al., 2020). Some researchers suggest that lightweight ML models, such as Decision Trees (DT) and Gradient Boosting Machines (GBM), may be more suitable for SMEs due to their lower computational demands and interpretability (Burggräf et al., 2024). In the context of food manufacturing SMEs, adopting ML models is particularly valuable due to the sector's unique challenges, such as stringent quality control requirements and resource constraints. Research highlights the role of ML in enhancing food quality management and safety assessment, leveraging techniques like computer vision and deep learning for defect detection and consistency evaluation (Khan, 2021, Nwamekwe & Okpala, 2025). These advancements improve operational efficiency and ensure compliance with regulatory standards.

Despite the promising results of ML-driven OEE optimization, key challenges remain. A major limitation is SMEs' lack of high-quality, real-time data, which affects model accuracy and reliability (Ghobakhloo, 2020). Additionally, the comparative effectiveness of different ML models in SME manufacturing environments has not been extensively studied, leaving a gap in identifying the most suitable techniques for improving OEE in food production facilities.

While previous studies have Chen et al., 2023; Jin, 2025; Malashin et al., 2024; Vitalis, Nwamekwe, et al., 2024; Vitalis, Onyeka, et al., 2024 highlighted the potential of ML for manufacturing optimization, there is a need for more focused research on the evaluation of ML models specifically for OEE optimization in food manufacturing SMEs. This study aims to bridge this gap by assessing the performance of various ML models and their applicability in optimizing OEE within a resource-constrained SME environment.

2. Method

2.1. Research Design

This study adopts a quantitative experimental research design leveraging historical production and equipment data from a selected food manufacturing SME. Quantitative research is a systematic method of inquiry that focuses on collecting and analyzing numerical data to identify patterns, test hypotheses, and make predictions. It is characterized by objectivity, replicability, and the use of statistical tools to analyze measurable variables. This approach is commonly used in engineering and manufacturing studies, as it allows researchers to quantify relationships between variables and assess the impact of different factors on outcomes. This research design in this study aims to evaluate and compare the performance of multiple machine learning models in optimizing Overall Equipment Effectiveness (OEE). The approach involves data preprocessing, feature engineering, model training and testing, and performance evaluation.

2.2. Study Area

This study was conducted in an SME specializing in food manufacturing in Nnewi, Anambra State, Nigeria. The SME specializes in the processing and packaging consumable goods and operates with key equipment, including mixers, sealers, conveyors, and automated packaging lines. These assets are monitored for downtime, production output, and quality, which are key indicators of OEE.

2.3. Sources of Data

The primary data source for this present study is historical operational records from the company's manufacturing process. It contains monthly records of OEE metrics for three years. The data shown in Table 1 consists of:

- Availability (A%): Measures machine uptime.
- Performance Rate (P%): Evaluates operational efficiency.
- Quality Rate (Q%): Assesses defect rates in production.
- QEE (%): The computed overall OEE value.

Table 1. OEE Metrics

S/No	Category	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	Months 1st Year												
1	Availability (A) %	73.10	72.70	76.60	75.60	73.50	45.20	26.70	70.40	70.30	68.60	52.30	67.90
2	Performance rate (P) %	73.40	60.40	55.80	56.20	66.30	49.30	46.80	83.40	73.20	72.90	72.90	73.60
3	Quality rate (Q) %	93.7	94.10	94.80	92.00	91.90	90.20	90.00	88.90	88.20	88.00	72.50	83.40
4	QEE %	50.32	40.01	40.53	54.53	54.64	21.15	10.62	54.18	45.82	44.64	27.64	41.67
	Months 2nd Year												
1	Availability (A) %	67.90	75.78	77.20	76.00	73.80	66.87	77.87	72.46	72.76	78.43	77.56	76.00
2	Performance rate (P) %	73.60	79.40	80.70	78.30	83.90	73.64	73.74	79.60	74.37	79.80	76.22	79.60
3	Quality rate (Q) %	83.70	95.70	95.40	93.91	90.10	90.00	90.00	94.05	94.02	94.05	91.80	93.52
4	QEE %	41.82	57.58	59.08	56.63	70.12	48.16	56.00	48.76	48.11	53.99	47.26	56.06

S/No	Category	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	Months 3rdYear												
1	Availability (A) %	75.86	73.08	70.88	72.60	78.00	76.40	69.60	69.80	72.00	77.90	77.37	78.50
2	Performance rate (P) %	84.30	81.78	80.88	78.30	85.66	85.22	74.84	74.92	81.73	84.00	84.50	85.52
3	Quality rate (Q) %	90.73	83.99	88.57	86.00	90.10	92.27	83.73	87.22	91.49	92.00	91.49	92.00
4	QEE %	58.02	53.34	48.77	51.72	55.88	61.25	59.11	44.40	44.83	51.97	59.44	62.00

2.4. Data Collection Techniques

Data was collected from the SME's computerized maintenance management (CMMS) and production monitoring systems. The researchers worked closely with plant supervisors and engineers to ensure the accuracy and completeness of the collected datasets.

2.5. Data Preprocessing

Raw data was cleaned and preprocessed to remove inconsistencies, missing values, and irrelevant attributes. The following steps were undertaken:

- Missing Data Handling: Imputation techniques such as mean substitution or interpolation
- Normalization: Min-max or z-score normalization to scale numerical features
- Categorical Encoding: One-hot encoding for categorical variables
- Feature Engineering: Derivation of OEE components: Availability, Performance, and Quality.

2.6. Machine Learning Models Selection

The study evaluated the performance of the following supervised machine learning models:

Linear Regression (LR): is one of the simplest and most used supervised learning algorithms in machine learning. It is used to model the relationship between a dependent variable (target) and one or more independent variables (features) by fitting a linear equation to the observed data (Ezeanyim et al., 2025). It is represented as:

$$y = \beta \mathrm{o} + \beta 1x + \beta 2x2 + \cdots + \beta \mathrm{n}x\mathrm{n} + \epsilon$$

y Is the predicted dependent variable
x Is the input (independent) variable
β_o Is the intercept (value of y when $x = 0$)
β_1 Is the slope (coefficients)
ϵ Is the error term

Gradient Boosting Regressor: is a supervised ensemble machine learning technique that builds a strong predictive model by combining the outputs of multiple weak learners, typically decision trees, through a stage-wise optimization process. Unlike random forests, which train trees in parallel, gradient boosting builds trees sequentially, where each new tree aims to correct the errors made by the previous ensemble. It is represented as follows:

$F_0(x) = \text{argmin\_}\gamma \ \Sigma \ L\ (y_i, \gamma)$
$r_{im} = -\ [\partial L\ (y_i, F\ (x_i))\ /\ \partial F(x_i)]$ evaluated at $F(x) = F_{m-1}(x)$
Train $h_m(x)$ to approximate r_{im}
$F_m(x) = F_{m-1}(x) + \alpha * h_m(x)$

Random Forest Regressor (RFR) is an ensemble learning method that builds multiple Decision Trees and aggregates their predictions to improve accuracy and reduce overfitting. Each tree is trained on a random subset of the data (with replacement, called bootstrapping) and considers a random subset of features when splitting nodes. It's powerful for regression tasks because it captures non-linear relationships and reduces variance compared to a single Decision Tree (Nwamekwe et al., 2025). It is represented as follows:

$$\acute{y}RFR(x) = \frac{1}{T}\sum_{t=1}^{T}\hat{y}^{(t)}(x)$$

Were
$\acute{y}RFR(x)$ Is the final prediction for the input x
T is the number of decision trees in the forest
$\hat{y}(t)(x)$ Does the t-th decision tree make the Prediction

Support Vector Machines (SVM): Aims to find a function that approximates the data within a certain tolerance margin (ε) while being as flat as possible. Instead of minimizing prediction error like ordinary regression, SVM tries to fit the best line (or curve) that stays within a predefined error margin (epsilon-insensitive zone) and penalizes only predictions that fall outside this margin. It is represented as follows:

$$\hat{y}(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x)$$

were
$\hat{y}(x)$ Is the predicted value for input x
$\alpha_i - \alpha_i^*$ Are the Language multipliers (found during training)
x_l Are the support vectors
$K(x_i, x)$ is the Kernel function (e.g., linear, polynomial, RBF)
b is the bias term

2.7. Model Training and Testing

The dataset was split into training (80%) and testing (20%) sets. This was achieved using the 'train_test_split' function in a Python Library called Sklearn. A 10-fold cross-validation strategy was adopted to reduce bias and ensure robustness of the results. Model hyperparameters were optimized using grid search or randomized search techniques.

2.8. Evaluation Metrics

Model performance was evaluated using the following metrics: Root Mean Squared Error (RMSE), which is a metric used to measure the accuracy of a machine learning model by calculating the average squared differences between actual and predicted values, then taking the square root of that average. It is expressed as:

$$\text{RMSE} = \sqrt{\frac{1}{\text{n}}\sum_{\text{i}=1}^{\text{n}}(\text{yi-ýi})2}$$

Where:
y_i is the actual value
$\acute{y}_i$ is the predicted value from the model
n is the number of observations

Mean Absolute Error (MAE): an evaluation metric measuring the average absolute difference between actual and predicted values. It is expressed as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \acute{y}_i|$$

Where:
y_i is the actual value
$\acute{y}_i$ is the predicted value from the model
n is the number of observations

R-Squared (R^2): an evaluation metric that measures how well the model explains the variability in the data. It is expressed as:

$$R^2 = 1 - \frac{\sum(y_i - \acute{y}_i)^2}{\sum y_i - \hat{y})^2}$$

Where:
y_i is the actual value
$\acute{y}_i$ is the predicted value from the model
$\hat{y}$ Is the mean
These metrics were used to determine each model's accuracy and generalization capability.

3. Results and Discussion

The applied models, Linear Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting Machines (GBM), were evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2). These metrics help assess the accuracy of each model in optimizing overall equipment effectiveness (OEE).

Table 2. Machine Learning Model Performance Comparison

Model	RMSE	MAE	R^2
Linear Regression	1.4229	1.113	0.9547
Random Forest	1.8003	1.608	0.9275
Support Vector Machine	4.6499	4.0447	0.5164
Gradient Boosting	1.7796	1.095	0.9292

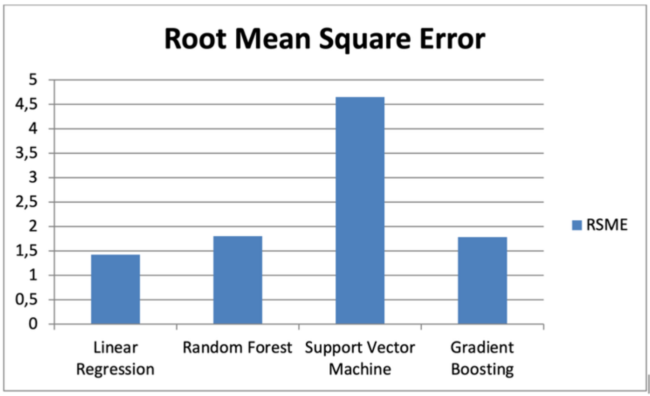

Figure 1. Root Mean Square Error Evaluation metric

Figure 1 represents the average error magnitude in a model's predictions using Root Mean Square Error. A lower RMSE value indicates that the model's predictions are closer to the actual values, meaning better performance. While a higher RMSE implies that the model's predictions are more spread out from the exact values, indicating poor performance. As shown in Table 2, Linear Regression with an RMSE value of 1.4229 indicates that its predictions are closer to the actual values, meaning better performance, accuracy, and smaller prediction errors. Gradient Boosting and Random Forest with RMSE values of 1.7796 and 1.8003 performed slightly poorly compared to the Linear Regression model. Support Vector Machine showed the worst performance with an RMSE value of 4.6499, indicating that the model's predictions are more spread out from the actual values.

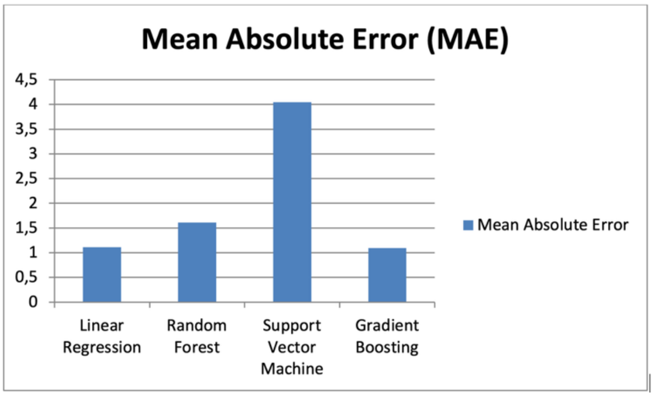

Figure 2. Mean Absolute Error Evaluation Metric

Figure 2 measures the average magnitude of errors in predictions, without considering their direction. It shows that Gradient Boosting has the lowest MAE (1.095), with the lowest average prediction error. Linear Regression (1.113) is close in performance to Gradient Boosting. Random Forest (1.608) performs moderately well but has higher errors, while Support Vector Machine (4.0447) performs the worst, showing high prediction errors.

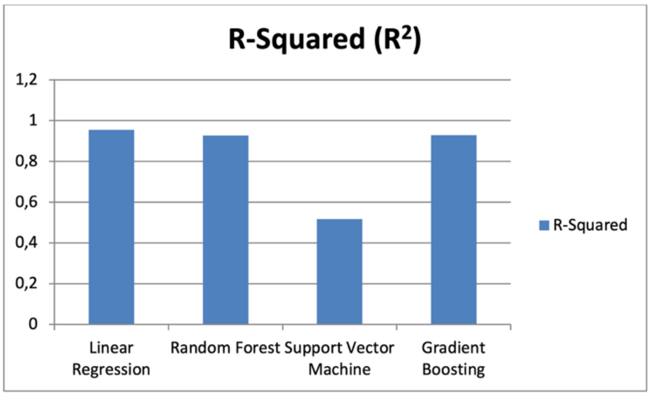

Figure 3. R-Square Evaluation metric

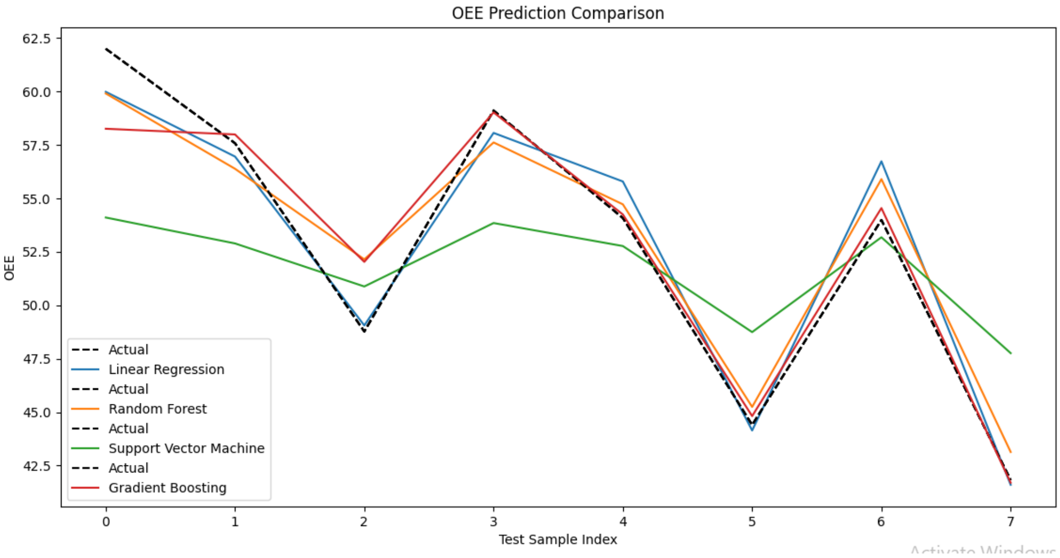

Figure 4. OEE Prediction

Figures 4 and 5 measure how well each model explains the variability in the data. Linear Regression (0.9547) performs best, explaining about 95.47% of the variance. Gradient Boosting (0.9292) and Random Forest (0.9275) are slightly behind, while Support Vector Machine (0.5164) is significantly worse, explaining only 51.64% of the variance. In Figure 4, the Linear Regression model predictions align with the actual data, capturing the underlying trends in OEE. Support Vector Machine model predictions show large deviations from the actual data, indicating poor model performance, where the model struggles to make accurate predictions.

From Table 2, based on the evaluation metrics performance, the best-performing model should have the following:

- Lowest RMSE (to minimize large prediction errors),
- Lowest MAE (to reduce the average error),
- Highest R^2 (to maximize how well the model explains variations in OEE).

The Linear Regression model has the best overall performance with the lowest RMSE (1.4229), Low MAE (1.113), and a high R^2 (0.9547). Gradient Boosting model performed slightly worse than Linear Regression in RMSE (1.7796) but has the lowest MAE (1.0950). At the same time, the Support Vector Machine model has the worst performance overall with the highest RMSE (4.6499) and lowest R^2 (0.5164). Linear Regression is computationally efficient and easier to implement than Gradient Boosting or Random Forest.

If slight improvements in optimization accuracy are needed and computational complexity is not a concern, Gradient Boosting could be a viable alternative.

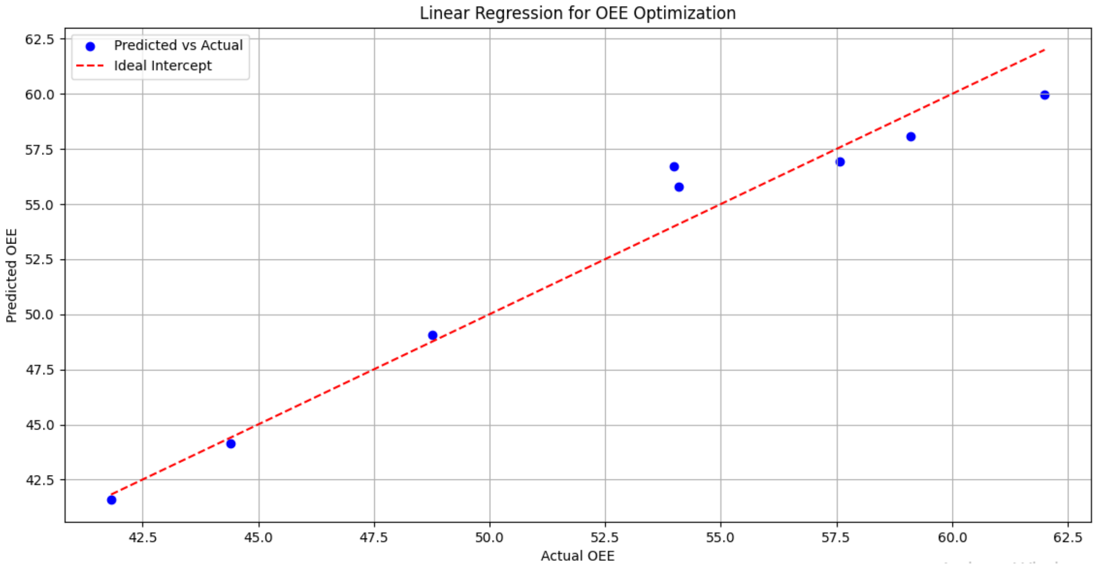

Figure 5. Linear Regression OEE Optimization

Table 3. Linear Regression OEE Optimization Metrics

Metric	Value
Mean Squared Error	2.0246868304162575
R-squared Score (R^2)	0.9547135392245288

Table 4. Model Coefficients

Availability	0.2554543963070465
Performance	0.7820335021434652
Quality	0.7339698020288927
Intercept	-94.73707141435935

Table 3 explains the performance of the linear regression model for optimizing OEE. Mean Squared Error measures the average squared difference between the actual and predicted values. A lower MSE indicates that the model's predictions are close to the exact values. Since 2.0247 is relatively low, the model is making accurate predictions. R-squared explains the proportion of variance in the dependent variable (Overall Equipment Effectiveness, OEE) that can be explained by the independent variables (Availability, Performance, and Quality). An R^2 value of 0.9547 suggests that 95.47% of the variations in OEE can be explained by the three independent factors, which indicates a strong fit of the model as shown in Figure 5.

Table 4 models the relationship between the dependent variable (OEE) and independent variables (Availability, Performance, and Quality) using the following linear regression equation:

$$OEE = (0.2555 \times Avaliability) + (0.7820 \times Performance) + (0.7340 \times Quality) + (-94.7371)$$

From equation 1.3, for every 1unit increase in Availability, OEE is expected to increase by 0.2555, assuming other factors remain constant. Also a 1 A unit increase in Performance increases 0.7820 in OEE, assuming other factors remain constant. In comparison, a 1-unit increase in Quality leads to a 0.7340 increase in OEE, keeping other variables constant. Performance coefficient is the largest coefficient, indicating that Performance has the highest impact on OEE in this case study. At the same time, quality is also a strong determinant of OEE, though slightly less influential than Performance.

The intercept (-94.7371) represents the theoretical OEE when all independent variables (Availability, Performance, and Quality) are zero. In practical terms, this value is not meaningful because Availability, Performance, and Quality cannot be zero in a functioning manufacturing setup.

From Figure 5, the blue dots are closely aligned with the red dashed line, meaning the model performs reasonably well at predicting OEE. However, some dots deviate slightly from the red line, indicating prediction errors. The data points are clustered along the diagonal, meaning that most predictions are accurate. The model could be improved by considering other variables such as machine maintenance schedules, operator efficiency, or external disruptions.

To enhance equation 1.3's predictive performance and optimize Overall Equipment Effectiveness (OEE), we explored the interaction effects among the independent variables: Availability, Performance, and Quality.

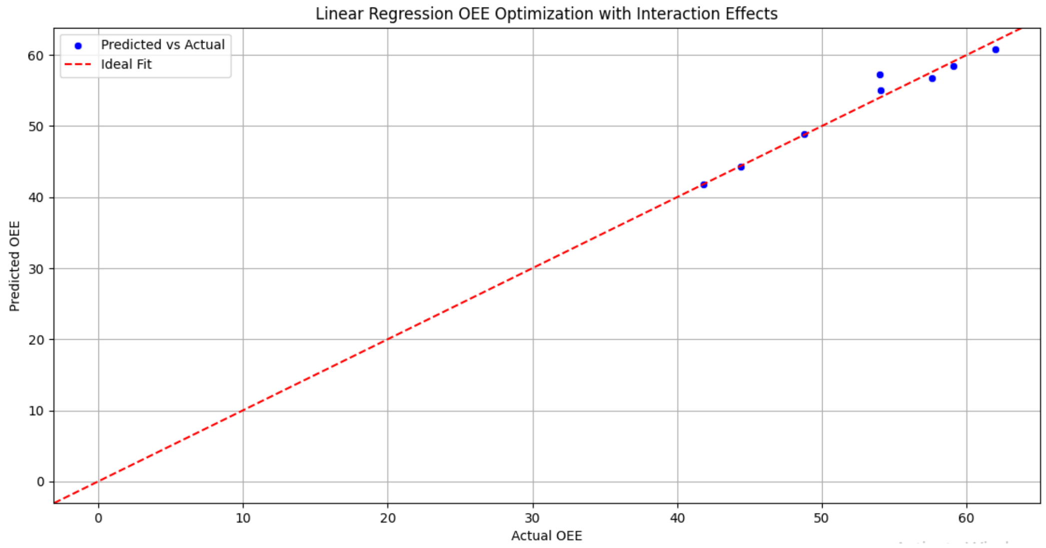

Figure 6. Linear Regression OEE Optimization with Interaction Effects

Table 5. Linear Regression OEE Optimization with Interaction Effects Metrics

Metric	Value
Mean Squared Error	1.7451
R-squared Score (R^2)	0.9610

Table 6. Interaction Effects Model Coefficients

Availability	0.1614
Performance	-0.3787
Quality	-0.0162
Intercept	-12.8699
A*P	0.0047
A*Q	-0.0009
P*Q	0.0088

Table 5 explains the performance of the linear regression model for optimizing OEE with interaction effects. Mean Squared Error (1.7451) is relatively low, suggesting that the model makes precise predictions. R-Squared metric indicates that 96.10% of the variations in OEE can be explained by Availability, Performance, Quality, and their interactions. This high value confirms that the model fits the data well. Table 6 models the relationship between the dependent variable (OEE), independent variables (Availability, Performance, and Quality), and their interaction effects using the following linear regression equation:

$$\begin{aligned} OEE = {} & (0.1614 \times Avaliability) + (-0.3787 \times Performance) + (-0.0162 \times Quality) \\ & + (0.0047 \times A * P) + (-0.0009 \times A * Q) + (0.0088 \times P * Q) + (-94.7371) \end{aligned}$$

From equation 1.4, Availability has a positive but moderate impact on OEE, and for every 1-unit increase in Availability, OEE increases by 0.1614, assuming all other factors remain constant. A 1-unit increase in Performance reduces OEE by 0.3787. This unexpected negative coefficient suggests that, without considering interaction effects, increasing Performance alone may not always improve OEE. If Performance is increased without improving Quality or Availability, inefficiencies may arise (e.g., defective products due to rushed production). A 1-unit increase in Quality decreases OEE slightly by 0.0162. This means that improving Quality alone does not significantly boost OEE unless paired with Performance and Availability improvements. When both Availability and Performance increase together, OEE improves by 0.0047 per unit. This shows that Availability supports Performance in improving efficiency. The interaction effects between availability and quality (A*Q = -0.0009) suggest that Availability and Quality do not reinforce each other significantly. This means that when Availability increases but Quality is inconsistent, OEE may suffer due to increased rework or defects. The most substantial interaction effect occurs when Performance and Quality improve together, thus OEE increases by 0.0088 per unit. This implies that Performance improvements are most effective when coupled with high-quality standards.

Simply increasing Performance without Quality and Availability support can hurt efficiency. Thus, coordinated efforts in all three areas will yield the best OEE improvements. Performance should be carefully managed, since Performance alone negatively impacts OEE; therefore, it is necessary to avoid increasing it at the expense of Quality and Availability. To maximize OEE, ensuring high-quality output while improving Performance is essential. Optimizing specific conditions, such as reducing downtimes while maintaining quality, will help achieve this goal.

These models allow manufacturers to predict OEE based on Availability, Performance, Quality, and interaction effects, enabling proactive maintenance and efficiency improvements. Production schedules and workforce planning can be adjusted with accurate predictions to optimize overall performance.

4. Conclusion

This study evaluated multiple machine learning models to optimize Overall Equipment Effectiveness (OEE) in a food manufacturing SME, revealing that machine learning can significantly enhance efficiency by providing predictive insights, identifying inefficiencies, and supporting data-driven decision-making. Among the models tested, Linear Regression emerged as the most effective based on performance metrics, making it a suitable choice for OEE optimization in this context. The findings emphasize the transformative potential of advanced analytics in improving equipment performance, reducing downtime, and enabling proactive maintenance and real-time decision-making, ultimately increasing productivity and cost savings. These insights are valuable for production managers and operational leaders aiming to adopt smart manufacturing strategies aligned with Industry 4.0. Additionally, the study contributes to the literature by providing empirical evidence of machine learning applications in the relatively underexplored food manufacturing SME sector, while highlighting the interpretability and effectiveness of linear models in constrained environments. However, limitations include the use of data from a single SME, lack of real-time implementation, exclusion of advanced algorithms, and omission of external factors like supply chain disruptions. Future research should address these gaps by exploring multi-site studies, real-time deployments, and integration with IoT systems to improve model applicability and resilience.

References

Bauer, M., Van Dinther, C., & Kiefer, D. (2020). Machine learning in SME: An empirical study on enablers and success factors. *Advances in Information Systems Research*, 1–11. https://aisel.aisnet.org/amcis2020/adv_info_systems_research/adv_info_systems_research/3

Biswaranjan, S., Awad Bin, N., & Renato R, M. I. (2024). Machine learning model for improving the overall equipment effectiveness in industrial manufacturing sites. In *Advances in Computational Intelligence and its Application* (pp. 151–160). https://doi.org/10.1201/9781003488682-20

Burggräf, P., Steinberg, F., Sauer, R. C., & Nettesheim, P. (2024). Machine learning implementation in small and medium-sized enterprises: Insights and recommendations from a quantitative study. *Production Engineering, 18*(5), 751–764. https://doi.org/10.1007/s11740-024-01274-2

Chikwendu, O. C., Chidiebube, I. N., & Obiuto, N. C. (2023). Revolutionizing manufacturing: Harnessing the power of artificial intelligence for enhanced efficiency and innovation. *International Journal of Engineering Research and Development, 19*(6), 18–25. www.ijerd.com

Chen, T., Sampath, V., May, M. C., Shan, S., Jorg, O. J., Aguilar Martín, J. J., Stamer, F., Fantoni, G., Tosello, G., & Calaon, M. (2023). Machine learning in manufacturing towards Industry 4.0: From 'for now' to 'four-know.' *Applied Sciences (Switzerland), 13*(3), 1–32. https://doi.org/10.3390/app13031903

Chidiebube, I. N., Nwamekwe, C. O., Harold, G., & Wilfred, M. (2025). Optimization of overall equipment effectiveness factors in a small and medium enterprise. *Journal of Research in Engineering and Applied Science, 10*(1), 836–845.

Chukwutoo, C. I., & Nkemakonam, C. I. (2018). Total productive maintenance (TPM) is a business strategy for manufacturing small and medium enterprises in Nigeria. *Advances in Research, 15*(5), 1–9. https://doi.org/10.9734/AIR/2018/41330

Ezeanyim, O. C., Ewuzie, N. V., Aguh, P. S., Nwabueze, C. V., & Nwamekwe, C. O. (2025). Effective maintenance of industrial 5-stage compressor: A machine learning approach. *Gazi University Journal of Science Part A: Engineering and Innovation, 12*(1), 96–118. https://dergipark.org.tr/en/pub/gujsa/issue/90827/1646993.

Ghobakhloo, M. (2020). Industry 4.0, digitization, and opportunities for sustainability. *Journal of Cleaner Production, 252*, 119869. https://doi.org/10.1016/j.jclepro.2019.119869

Igbokwe, N. C., & Godwin, H. C. (2021). Maintenance performance evaluation and downtime analysis of manufacturing equipment in a food manufacturing company. *Journal of Engineering Research and Reports, 20*(11), 100–107. https://doi.org/10.9734/jerr/2021/v20i1117413

Igbokwe, N., & Godwin, H. (2022). Using an optimal maintenance strategy, cost-effective maintenance and machine reliability for food manufacturing industries. *Archives of Current Research International, 22*(8), 27–34. https://doi.org/10.9734/ACRI/2022/v22i8543

Igbokwe, N., & Mba, W. (2019). Assessing the nature of risk management implementation in manufacturing small and medium enterprises in Nigeria. *Archives of Current Research International, 16*(4), 1–11. https://doi.org/10.9734/ACRI/2019/v16i430098

Igbokwe, N., Chukwuemeka, G. H., & Constance, N. (2021). A GANetXL approach to an optimal maintenance strategy in food manufacturing. *International Journal of Engineering Science and Computing*.

Jin, J. (2025). Enhancing manufacturing performance with AI and machine learning: Applications in predictive maintenance and production optimization. In *Proceedings of the 3rd International Conference on Mechatronics and Smart Systems* (Vol. 140, pp. 84–89). https://doi.org/10.54254/2755-2721/140/2025.21393

Khan, R. (2021). Artificial intelligence and machine learning in food industries: A study. *Journal of Food Chemistry and Nanotechnology, 7*(3), 60–67. https://doi.org/10.17756/jfcn.2021-114

Le Dinh, T., Vu, M. C., & Tran, G. T. C. (2025). Artificial intelligence in SMEs: Enhancing business functions through technologies and applications. *Information (Switzerland), 16*(5), 1–22. https://doi.org/10.3390/info16050415

Maataoui, S., Bencheikh, G., & Bencheikh, G. (2023). Elevating manufacturing excellence: A data-driven approach to optimize overall equipment effectiveness (OEE) for a single machine. In *International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD'2023)* (pp. 332–343). https://link.springer.com/chapter/10.1007/978-3-031-54318-0_29

Malashin, I., Martysyuk, D., Tynchenko, V., Gantimurov, A., Semikolenov, A., Nelyub, V., & Borodulin, A. (2024). Machine learning-based process optimization in biopolymer manufacturing: A review. *Polymers, 16*(23), 1–45. https://doi.org/10.3390/polym16233368.

Nwamekwe, C. O., & Nwabunwanne, E. C. (2025). Circular economy and zero-energy factories: A synergistic approach to sustainable manufacturing. *Journal of Research in Engineering and Applied Sciences (JREAS), 10*(1), 829–835. https://qtanalytics.in/journals/index.php/JREAS/article/view/4567

Nwamekwe, C. O., & Igbokwe, N. C. (2024). Supply chain risk management: Leveraging AI for risk identification, mitigation, and resilience planning. *International Journal of Industrial Engineering, Technology & Operations Management, 2*(2), 41–51. https://doi.org/10.62157/ijietom.v2i2.38

Nwamekwe, C. O., & Okpala, C. C. (2025). Circular economy strategies in industrial engineering: From theory to practice. *International Journal of Multidisciplinary Research and Growth Evaluation, 6*(1), 1773–1782. https://www.allmultidisciplinaryjournal.com/uploads/archives/20250212103754_MGE-2025-1-288.1.pdf

Nwamekwe, C. O., & Okpala, C. C. (2025). Machine learning-augmented digital twin systems for predictive maintenance in high-speed rail networks. *International Journal of Multidisciplinary Research and Growth Evaluation, 6*(1), 1783–1795. https://www.allmultidisciplinaryjournal.com/uploads/archives/20250212104201_MGE-2025-1-306.1.pdf

Nwamekwe, C. O., Ewuzie, N. V., Igbokwe, N. C., Okpala, C. C., & U-Dominic, C. M. (2024). Sustainable manufacturing practices in Nigeria: Optimization and implementation appraisal. *Journal of Research in Engineering and Applied Sciences, 9*(3). https://qtanalytics.in/journals/index.php/JREAS/article/view/3967

Nwamekwe, C. O., Ewuzie, N. V., Igbokwe, N. C., U-Dominic, C. M., & Nwabueze, C. V. (2024). Adoption of smart factories in Nigeria: Problems, obstacles, remedies, and opportunities. *International Journal of Industrial and Production Engineering, 2*(2). https://journals.unizik.edu.ng/ijipe/article/view/4167

Nwamekwe, C. O., Ewuzie, N. V., Okpala, C. C., Ezeanyim, C., Nwabueze, C. V., & Nwabunwanne, E. C. (2025). Optimizing machine learning models for soil fertility analysis: Insights from feature engineering and data localization. *Gazi University Journal of Science Part A: Engineering and Innovation, 12*(1), 36–60. https://dergipark.org.tr/en/pub/gujsa/issue/90827/1605587

Nwamekwe, C. O., Okpala, C. C., & Okpala, S. C. (2024). Machine learning-based prediction algorithms for mitigating maternal and fetal mortality in the Nigerian tertiary hospitals. *International Journal of Engineering Inventions, 13*(7), 132–138. https://www.ijeijournal.com/papers/Vol13-Issue7/1307132138.pdf

Oldemeyer, L., Jede, A., & Teuteberg, F. (2024). Investigating artificial intelligence in SMEs: A systematic review of the state of the art and the main implementation challenges. *Management Review Quarterly, 75*(2). https://doi.org/10.1007/s11301-024-00405-4

Peretz-Andersson, E., Tabares, S., Mikalef, P., & Parida, V. (2024). Artificial intelligence implementation in manufacturing SMEs: A resource orchestration approach. *International Journal of Information Management, 77*, 102781. https://doi.org/10.1016/j.ijinfomgt.2024.102781

Phir, R., & Mbale, J. (2024). Leveraging machine learning and artificial intelligence for innovation and sustainability in small and medium-sized enterprises (SMEs): A Kalumbila, Zambia case study. *Sixth International Conference in Information and Communication Technologies, Lusaka, Zambia*, October 56–62.

Vitalis, E. N., Nwamekwe, C. O., Nkemakonam, I., Victoria, N. C., Nwabunwanne, E. C., & Ono, C. G. (2024). Application of machine-learning-based hybrid algorithm for production forecast in a textile company. *Jurnal Inovasi Teknologi dan Edukasi Teknik, 4*(12), 1–9. https://doi.org/10.17977/um068.v4.i12.2024.2

Vitalis, E. N., Onyeka, N. C., Chikwendu, O. C., Chidiebube, N., U-Dominic, C. M., & Victoria, N. C. (2024). Machine learning models application for maintenance operations enhancement in the petrochemical industry. *Faculty of Engineering International (Hybrid) Conference 2024*, 66–76.

Welte, R., Estler, M., & Lucke, D. (2023). A method for implementing machine learning solutions for predictive maintenance in small and medium-sized enterprises. *Procedia CIRP, 93*, 909–914. https://doi.org/10.1016/j.procir.2020.04.052.