Proceedings of the 5th Faculty of Industrial Technology
International Congress Bandung, Indonesia,
September 17-18, 2025
ISSN 2962 – 1798

A Machine Learning–Based Approach to Modeling Embroidery Machine Effectiveness Using Overall Equipment Effectiveness (OEE)

Faula Huzaifah Nugrahayu [a], *Yuniar* [a,1], and *Alif* Ulfa Afifah[a]

[a] Industrial Engineering Department, Institut Teknologi Nasional Bandung, Indonesia

Abstract. The embroidery process represents the initial stage of production at PT XXX; therefore, evaluating its performance is essential to prevent disruptions in subsequent manufacturing stages. To support decision-making related to embroidery machine performance, the company requires an effective machine performance monitoring system. This study aims to develop a predictive model of Overall Equipment Effectiveness (OEE) using a machine learning approach and to design a performance monitoring dashboard using Power BI. The results show that the average OEE value of the embroidery machines exceeds the World Class OEE standard. Among the OEE components, performance efficiency is the most dominant factor influencing OEE, primarily due to variations in product size, while machine downtime continues to contribute to fluctuations in OEE values. Among the evaluated machine learning models, Linear Regression demonstrates the best predictive performance, achieving an R^2 value of 45%. Furthermore, the developed Power BI dashboard effectively presents machine performance indicators and OEE prediction results in a visual and integrated manner, thereby supporting continuous monitoring and informed decision-making regarding embroidery machine performance.

1 Introduction

The rapid advancement of modern manufacturing industries requires companies to continuously enhance the effectiveness and efficiency of their production processes in order to remain competitive. Machine performance plays a crucial role in manufacturing systems, as it directly affects production capacity, operational speed, and product quality [1]. Poorly managed machine performance can lead to excessive time losses, reduced output, and delays in order fulfilment, ultimately impacting overall company performance.

PT XXX is a doll manufacturing company that implements a Make-to-Order (MTO) production system, in which production activities are initiated based on customer demand. The production process begins with fabric measurement using a measuring table, followed by pattern creation on the fabric using embroidery machines, before proceeding to sewing and stuffing processes. As the initial stage of production, embroidery machines have a critical role in determining the continuity and efficiency of downstream operations. Any

[1] Corresponding author: yuniar@itenas.ac.id

inefficiencies or disruptions at this stage can propagate through subsequent production stages and negatively affect the overall production flow.

Based on operational data, the main issue affecting the embroidery machines at PT XXX is high downtime caused by machine disturbances and operational errors, which reduces effective operating time and leads to fluctuations in production output. In addition, machine performance monitoring is still conducted manually, resulting in limited visibility into machine conditions and the absence of continuous and integrated performance evaluation. This condition makes it difficult to identify performance degradation patterns and to support timely, data-driven decision-making.

These challenges have resulted in a significant decline in daily embroidery machine output, with an average reduction of approximately 20% in 2025. Under the MTO system, such a decline increases the risk of production delays and order fulfillment failures due to discrepancies between planned and actual output. Therefore, the company requires a comprehensive machine performance monitoring system that is capable of accurately representing operational effectiveness, tracking performance changes over time, and presenting concise and visual information to support decision-making.

Overall Equipment Effectiveness (OEE) is widely used as a standard metric to evaluate machine performance by integrating availability, performance efficiency, and quality rate into a single indicator. However, existing studies and industrial practices generally utilize OEE as a descriptive and retrospective evaluation tool, focusing on historical performance analysis rather than future performance prediction. Furthermore, the application of machine learning for OEE prediction in embroidery machine operations—particularly when integrated with performance visualization tools—remains limited. This gap indicates a lack of predictive and proactive approaches for machine performance monitoring in labor-intensive manufacturing environments such as the textile and garment industry.

To address this research gap, this study proposes the development of a machine learning–based model to predict the Overall Equipment Effectiveness (OEE) of embroidery machines using historical production and machine operational data. In addition, this study designs an integrated Power BI dashboard that visualizes machine performance indicators and OEE prediction results to support continuous monitoring and data-driven decision-making at PT XXX. The proposed approach is expected to enable proactive identification of performance issues, reduce the impact of machine downtime, and improve the overall effectiveness of embroidery machine operations within an MTO production system.

2 Methodology

This study adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, which provides a structured and systematic approach for developing data-driven solutions. The research methodology consists of several sequential stages, as illustrated in

ISSN 2962 – 1798

Figure 1, namely business understanding, data collection, data understanding, data preparation, modeling, evaluation, and deployment.

2.1 Business Understanding

At PT XXX, the performance of embroidery machines has not yet been evaluated using a measurable and standardized approach, as machine monitoring is still conducted manually and Overall Equipment Effectiveness (OEE) has not been formally implemented. This limitation results in the absence of an integrated system capable of evaluating machine effectiveness and supporting data-driven decision-making. Therefore, this study focuses on implementing OEE measurement supported by machine learning–based predictive modeling and the development of a Power BI dashboard as a comprehensive machine performance monitoring system.

2.2 Data Collection

The data used in this study were collected from August to October 2025 and consist of embroidery machine production data, machine performance data, and cycle time data. These datasets were obtained from daily operational records and serve as the primary input for OEE calculation and machine learning modelling.

2.3 OEE Calculation and Data Processing

Overall Equipment Effectiveness (OEE) is a widely used performance indicator within Total Productive Maintenance (TPM) to measure how effectively equipment supports the production process. In this study, OEE is calculated based on three main components: Availability Ratio (AR), Performance Efficiency (PE), and Rate of Quality Product (ROQP) [2].

a. Availability Ratio (AR) represents the proportion of effective operating time relative to the available production time [3] and is calculated as:

$$AR = \frac{Operation\ Time}{Loading\ Time} \times 100\% \tag{1}$$

b. Performance Efficiency (PE) reflects the ability of a machine to operate at its designed speed and produce output efficiently [4], calculated as:

$$PE = \frac{Total\ Production \times Ideal\ Cycle\ Time}{Operation\ Time} \times 100\% \tag{2}$$

c. Rate of Quality Product (ROQP) measures the proportion of good-quality products produced [5] and is calculated as:

$$ROPQ = \frac{Total\ Production - Total\ Reject}{Total\ Production} \times 100\% \tag{3}$$

d. OEE Calculation

$$OEE = AR \times PE \times ROPQ \tag{4}$$

2.4 Data Understanding

This stage aims to explore and understand the characteristics of the dataset derived from OEE calculations and operational variables used for machine learning modeling. Descriptive statistical analysis, including minimum, maximum, mean, median, and standard deviation, is employed to identify data distribution patterns, variability, and potential anomalies.

2.5 Data Preparation

Data preparation involves data cleaning and transformation to ensure data quality and modeling readiness. This process includes handling missing values, removing duplicate records, and detecting outliers. Subsequently, data normalization is performed using Z-score standardization to ensure that variables are on a comparable scale, particularly to support Support Vector Regression (SVR) modeling.

2.6 Modeling

Machine learning modeling is conducted using a supervised learning approach with regression techniques. Several algorithms are evaluated, including Linear Regression, Decision Tree Regression, Random Forest Regression, and Support Vector Regression (SVR). The dataset is divided into training and testing subsets to build and validate the predictive models.

2.7 Evaluation

Model performance is evaluated using k-fold cross-validation to ensure robustness and generalization capability. The evaluation metrics used include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). The best-performing model is selected based on the lowest error values and the highest R^2 score.

2.8 Deployment

The deployment stage involves saving the selected best-performing model and implementing a prediction script using Google Colab, enabling OEE prediction without model retraining. In addition, machine performance monitoring results are deployed through a Power BI dashboard to provide an interactive visualization of OEE values and predictions.

2.9 Power BI Dashboard Development

The Power BI dashboard is designed to visualize historical OEE analysis and predicted OEE values for embroidery machines in an integrated and intuitive manner. The dashboard provides key performance indicators, trend analysis, and predictive insights to support

ISSN 2962 – 1798

continuous machine performance monitoring. The overall development scheme of the dashboard is presented in Figure 1.

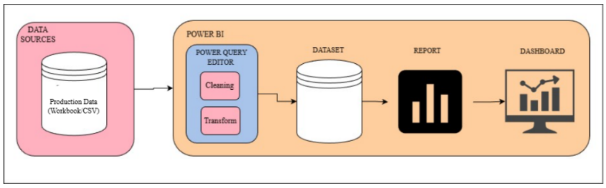

Fig 1. Power BI Dashboard Development Scheme

2.10 Conclusion Stage

At this stage, the results of OEE calculation, machine learning model evaluation, and Power BI dashboard development are synthesized to identify the most effective predictive model and the dominant factors influencing OEE. Based on these findings, recommendations are formulated to support improvements in embroidery machine performance at PT XXX.

3 Result and Discussion

The collected data were analyzed using Overall Equipment Effectiveness (OEE) calculations and machine learning–based modeling to predict OEE values. The results of these analyses were subsequently used to evaluate the performance of embroidery machines at PT XXX. Furthermore, the analytical and predictive outcomes were visualized through a Power BI dashboard to support comprehensive machine performance monitoring and data-driven decision-making.

3.1 Overall Equipment Effectiveness (OEE) Analysis

Overall Equipment Effectiveness (OEE) was calculated based on three components: Availability Ratio (AR), Performance Efficiency (PE), and Rate of Quality Product (ROQP). The OEE values were computed for each working day during the observation period from August to October 2025. The calculation results are presented in Table 1.

Table 1. OEE Value Calculation

Month	AR (%)	PE (%)	ROQP (%)	OEE (%)
August	95.67	95.00	100	90.24
September	92.98	55.64	100	51.65
October	91.53	96.43	100	88.64
Average	**93.39**	**82.36**	**100**	**76.84**

The Availability Ratio (AR) remained consistently high, exceeding 90% each month, indicating that the embroidery machines operated with minimal downtime and maintained reliable operating conditions. The Rate of Quality Product (ROQP) also consistently reached 100%, showing that all products met quality standards without significant defects or rework.

Variations in OEE were therefore mainly influenced by changes in Performance Efficiency (PE), particularly in September when it decreased to 55.64%. This decline was associated with high-mix, small-batch production that required more frequent setup and adjustments, reducing continuous stitching time even though machine availability and product quality remained stable.

3.2 Data Understanding

This stage aims to explore and understand the characteristics of the dataset derived from OEE calculations and operational variables used for machine learning modeling. Data understanding is an important step in data-driven research to ensure that the dataset is reliable and suitable for subsequent modeling stages [6].

Descriptive statistical analysis, including minimum, maximum, mean, median, and standard deviation, is employed to identify data distribution patterns, variability, and potential anomalies within the dataset [7]. This analysis provides an overview of the central tendency and dispersion of each variable, supporting informed decisions in data preparation and model development.

Table 2. Descriptive Statistical Data Exploration

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Date	57	57	8/1/2025	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Day	57	5	Friday	13	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Setup_Freq	57	NaN	NaN	NaN	2.087719	0.892042	1	1	2	3	5
Total_Qty	57	NaN	NaN	NaN	5296.96491	2706.12789	340	3572	4320	6691	13574
Total_Stitch	57	NaN	NaN	NaN	1489036.23	720815.801	103020	1053293	1253957	2032623	3534570
Weighted_Complexity	57	NaN	NaN	NaN	3.263875	1.885721	1	2	2	6	6
OEE	57	55	78.79%	2	NaN	NaN	NaN	NaN	NaN	NaN	NaN

The results indicate that all variables are suitable for further analysis. The OEE variable, which was initially stored in object format due to the presence of percentage symbols, was converted into numerical (float) values to enable machine learning modeling. Furthermore, the dataset contains no missing values or duplicate records, therefore, no additional data cleansing procedures were required. These findings confirm that the dataset is complete, consistent, and ready for the data preparation and modeling stages.

3.3 Data Preparation

At the data preparation stage, the dataset was divided into input variables (X) and the target variable (Y). The input variables consist of operational and production-related features, while the target variable is the OEE value.

To ensure comparability among features and to prevent certain variables from dominating the modeling process, data normalization was applied. Standardization using the Z-score normalization method was selected because the Support Vector Regression (SVR) model is sensitive to differences in feature scales. This normalization process transforms the data to have a mean of zero and a standard deviation of one, thereby improving model stability and performance.

3.4 Modeling

At the modeling stage, predictive models were developed using supervised machine learning regression algorithms. The dataset was divided into training data (80%) and testing data (20%) to evaluate the generalization capability of the models.

Several regression algorithms were implemented and compared, namely Linear Regression, Decision Tree Regression, Random Forest Regression, and Support Vector Regression (SVR), to assess their predictive performance in estimating OEE values based on the input variables.

3.5 Evaluation

At the model evaluation stage, the predictive performance of the developed regression models was assessed to determine the most suitable model for OEE prediction. Model validation was carried out using two approaches: evaluation on the testing dataset and k-fold cross-validation to ensure the robustness, accuracy, and reliability of the models. The evaluation metrics employed include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). The performance comparison results of each model are presented in Figures 2 to 5.

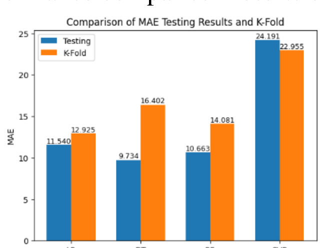

Fig 2. Comparison of MAE

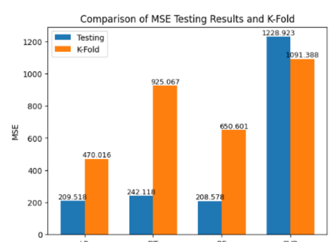

Fig 3. Comparison of MSE

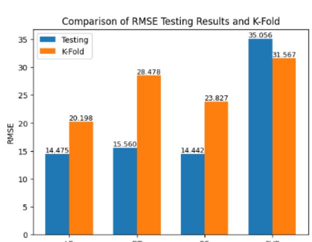

Fig 4. Comparison of RMSE

Fig 5. Comparison of R^2

The initial train–test evaluation showed strong performance for several models, with Linear Regression achieving an RMSE of 14.47 and R^2 of 0.853, slightly outperforming Random Forest (RMSE 14.44; R^2 0.854) and Decision Tree (RMSE 15.56; R^2 0.831). However, this evaluation was highly dependent on a single data split and risked overestimating model capability. To obtain a more reliable assessment, K-Fold Cross Validation was conducted. The results showed that Linear Regression produced the lowest average error (RMSE 20.20;

MAE 12.93) and the highest R^2 (0.452), while Random Forest (RMSE 23.83; R^2 0.246) and Decision Tree (RMSE 28.48; R^2 –0.677) exhibited significant performance drops, indicating overfitting. SVR remained the weakest model with RMSE 31.57 and R^2 0.048.

Based on these cross-validation results, Linear Regression was selected as the final model for predicting embroidery machine OEE because it demonstrated the most consistent and generalizable performance. Despite being simpler, it provides better stability and interpretability compared to more complex models that require larger datasets to perform optimally. These findings indicate that the relationship between operational variables such as total stitch count, setup frequency, and production quantity, and OEE follows a predominantly linear trend, allowing Linear Regression to adequately capture the production behavior without introducing unnecessary model complexity.

3.6 Deployment

At the deployment stage, the selected model, namely Linear Regression, was implemented to predict the OEE values of embroidery machines based on operational variables. This deployment aims to support performance analysis and data-driven decision-making by providing OEE predictions without requiring model retraining.

The OEE prediction results generated by the deployed model for November 2025 are presented in Table 3.

Table 3. Prediction Results

Date	Day	Setup_Freq	Total_Qty	Total_Stitch	Weighted_ Complexity	OEE_Prediction (%)
11/3/2025	Monday	1	12615	2261476	4	100.000
11/4/2025	Tuesday	2	13547	2924729	5	100.000
11/5/2025	Wednesday	4	1489	2888538	1	31.267
11/6/2025	Thursday	4	5613	1028599	1	98.337
11/7/2025	Friday	1	1409	586085	4	41.396
11/10/2025	Monday	1	4477	404845	4	76.523
11/11/2025	Tuesday	1	5101	2853236	4	53.145
11/12/2025	Wednesday	1	4255	1711491	2	74.241
11/13/2025	Thursday	4	12869	1571175	2	100.000
11/14/2025	Friday	2	11052	187105	1	100.000
11/17/2025	Monday	3	9604	2037348	2	100.000
11/18/2025	Tuesday	1	5534	557549	2	100.000
11/19/2025	Wednesday	1	10866	1695710	3	100.000
11/20/2025	Thursday	1	12433	3435729	4	100.000
11/21/2025	Friday	3	10604	2429986	2	100.000
11/24/2025	Monday	2	13091	719077	5	100.000
11/25/2025	Tuesday	3	3795	380781	1	89.106
11/26/2025	Wednesday	1	12468	2713921	5	100.000
11/27/2025	Thursday	1	10751	297665	3	100.000
11/28/2025	Friday	2	11125	1032801	4	100.000

3.7 Development of Power BI Dashboard

The dashboard was developed to present key information related to embroidery machine performance based on the Overall Equipment Effectiveness (OEE) value and its contributing factors. The embroidery machine performance dashboard is shown in Figure 6.

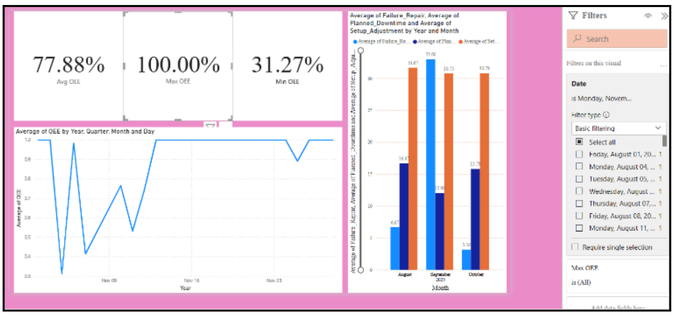

Fig 6. Embroidery Machine Performance Dashboard

The dashboard displays the main OEE indicators, including average, maximum, and minimum values, using card visualizations to provide an overall overview of machine effectiveness. OEE value changes over time are illustrated through a line chart to identify performance trends and fluctuations. In addition, downtime components—namely failure and repair, planned downtime, and setup and adjustment—are visualized using a clustered bar chart to highlight the dominant sources of downtime. A time-period filter is also provided to enable more flexible and focused machine performance analysis based on user-defined time ranges.

4 Conclusion

Based on the research results, embroidery machine number 9 achieves an average OEE value of 76.84%, with an Availability Ratio of 93.39%, Performance Efficiency of 82.36%, and a Rate of Quality Product of 100%. The OEE value from August to October still did not meet the world class OEE standard, which is 85%. The OEE value is predominantly influenced by performance efficiency, which is affected by variations in embroidered product sizes and a relatively low ideal cycle time (ICT) of 0.110 minutes per unit, although machine downtime continues to contribute to performance fluctuations.

The OEE prediction modeling results indicate that the Linear Regression model provides the most reliable predictive performance. Based on k-fold cross-validation evaluation, this model achieved an MAE of 12.93, an MSE of 470.02, an RMSE of 20.20, and an R^2 value of 0.45, demonstrating more consistent error distribution and better generalization capability compared to the other algorithms. These results confirm that the Linear Regression model is the most suitable for representing the relationship between operational variables and OEE. Furthermore, the developed Power BI dashboard effectively visualizes OEE indicators, performance trends, and downtime factors, thereby supporting continuous machine performance monitoring and data-driven operational decision-making.

References

1. P. Muchiri and L. Pintelon, "Performance measurement using overall equipment effectiveness (OEE): Literature review and practical application discussion," Int. J. Prod. Res., vol. 46, no. 13, pp. 3517–3535, (2008). doi: 10.1080/00207540601142645.
2. S. Nakajima, Introduction to TPM: Total Productive Maintenance. Portland, OR, USA: Productivity Press, (1988).
3. L. Turpin Jr., "A note on understanding cycle time," Int. J. Prod. Econ., vol. 205, pp. 113–117, (2018). doi: 10.1016/j.ijpe.2018.09.004.
4. P. G. Yazdi, A. Azizi, and M. Hashemipour, "An empirical investigation of the relationship between overall equipment efficiency (OEE) and manufacturing sustainability in Industry 4.0 with time study approach," Sustainability (MDPI), vol. 10, no. 9, Art. no. 3031, (2018). doi: 10.3390/su10093031.
5. S. Murugesan, S. T., V. G. V. Murugesan, and P. Muthuswamy, "Overall equipment effectiveness (OEE) analysis and improvement in a spinning unit," SSRN Electron. J., (2022). doi: 10.2139/ssrn.4032186.
6. J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Burlington, MA, USA: Morgan Kaufmann, (2012).
7. D. C. Montgomery and G. C. Runger, Applied Statistics and Probability for Engineers, 6th ed. Hoboken, NJ, USA: Wiley, (2014).