International Journal of Computer and Information System (IJCIS)
Peer Reviewed – International Journal
Vol : Vol. 07, Issue 03, August 2026
e-ISSN: 2745-9659
https://ijcis.net/index.php/ijcis/index

Machine Learning-Based Compost Maturity Prediction Using Multisensor IoT Data

1st Siti Rokhmah, 2nd Ihsan Cahyo Utomo
Institut Teknologi Bisnis AAS Indonesia, Universitas Muhammadiyah Surakarta
Informatika, Teknik Informatika
Sukohajo, Indonesia
sitirokhmah.itbaas@gmail.com[1], icu886@ums.ac.id[2]

Abstract— *Conventional composting process monitoring still faces various limitations, particularly in determining the level of compost maturity, which generally relies on visual observation and operator experience. The development of Internet of Things (IoT) technology enables real-time multisensor data acquisition during the decomposition process, but most research still focuses on developing monitoring systems without utilizing the resulting data to build predictive models. This study aims to develop a Machine Learning-based compost maturity prediction model using a multisensor dataset obtained from an IoT-based Smart Composting Bin system. The data used includes environmental parameters that affect the composting process, such as temperature, humidity, media moisture, and gas concentrations that are recorded periodically. The research stages include data preprocessing, handling missing data and outliers, normalization, feature selection, Random Forest model training, and performance evaluation using metrics appropriate to the prediction target type. To increase model transparency, this study also applies Explainable Artificial Intelligence through SHAP analysis to identify the contribution of each sensor parameter to the prediction results. The expected results show that the Machine Learning approach is able to predict the level of compost maturity with high accuracy while identifying the most influential environmental factors during the decomposition process. This research contributes to the utilization of IoT data not only as a monitoring medium, but also as a basis for intelligent decision-making in data-based organic waste management systems, thereby supporting the implementation of smart waste management and the concept of a circular economy.*

***Keywords** : Internet of Things, Smart Composting Bin, Machine Learning, Random Forest, Compost Maturity Prediction.*

I. INTRODUCTION

Organic waste management is a major challenge in achieving sustainable development, particularly in urban areas that produce large amounts of household waste daily. According to various reports, more than 50% of domestic waste consists of organic waste, which has the potential to be processed into compost as an environmentally friendly fertilizer. However, the conventional composting process still faces various obstacles, including long decomposition times, unstable environmental conditions during fermentation, and difficulties in objectively determining the level of compost maturity[1]. In practice, determining compost maturity still relies heavily on visual observation of color, texture, and aroma, thus relying heavily on operator experience and potentially resulting in inconsistent decisions.

The development of the Internet of Things (IoT) has opened up new opportunities to improve the efficiency of the composting process through real-time monitoring of environmental parameters. Various sensors such as temperature, humidity, gas content, and media moisture can be integrated with a microcontroller to generate continuous data throughout the decomposition process. The implementation of an IoT-based Smart Composting Bin system allows all process parameters to be automatically recorded, resulting in a rich dataset for further analysis. However, most published research still focuses on hardware development, data communication, and monitoring visualization, while the use of sensor data as a basis for intelligent decision-making is still relatively limited[2],[3].

With the advancement of artificial intelligence technology, machine learning approaches offer the ability to study complex relationships between environmental parameters and compost maturity. Machine learning algorithms can identify characteristic changes in temperature, humidity, gas concentration, and other parameters that serve as indicators of the decomposition process. Compared to threshold-based approaches, predictive models can generate compost maturity estimates that are more adaptive to variations in environmental conditions and raw material characteristics[4],[5].

One of the algorithms widely used in sensor data-based prediction problems is Random Forest because it has the ability to handle nonlinear relationships, is robust to noise, and is able to provide information on the level of importance of each variable (feature importance). In addition to producing accurate predictions, model interpretation is increasingly important so that prediction results can be understood by users and system managers. Therefore, the use of Explainable Artificial Intelligence (XAI), such as SHAP (SHapley Additive exPlanations), can be used to explain the contribution of each sensor parameter to the prediction results, thereby increasing transparency and trust in the model[6],[7].

One of the algorithms widely used in sensor data-based prediction problems is Random Forest because it has the ability to handle nonlinear relationships, is robust to noise, and is able to provide information on the level of importance of each variable (feature importance). In addition to producing accurate predictions, model interpretation is increasingly important so that prediction results can be understood by users and system managers[8],[9]. Therefore, the use of Explainable Artificial Intelligence (XAI), such as SHAP (SHapley Additive exPlanations), can be used to explain the contribution of each sensor parameter to the prediction results, thereby increasing transparency and trust in the model[10].

The main contributions of this research include: (1) the utilization of IoT multisensor datasets to build a compost maturity prediction model; (2) the development of a Machine Learning model that is able to improve the accuracy of compost condition estimation compared to conventional approaches; and (3) the application of Explainable AI to identify the factors that most influence the composting process. The research results are expected to be the basis for the development of a smarter, more efficient organic waste management system and support the implementation of the smart waste management concept and circular economy[11],[12].

II. RESEARCH METHODS

This study uses a data-driven predictive modeling approach by utilizing the Internet of Things-based Smart Composting Bin monitoring dataset developed in previous research. The IoT system serves as a data acquisition source, while this study focuses on developing a machine learning model to predict the level of compost maturity based on environmental parameter changes during the decomposition process. The research consists of six main phases, as shown in Figure 1.

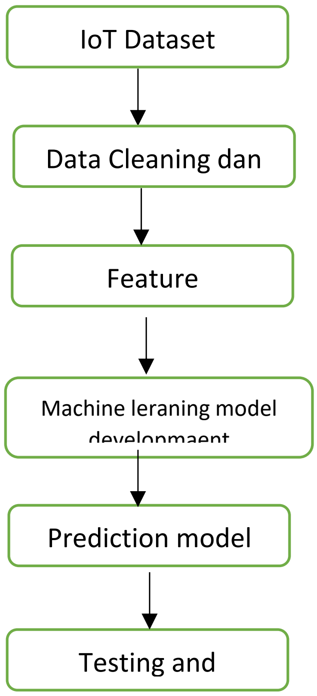


Figure 1 . Research Flow

3.1 data IoT acquisitoon

The dataset used in this study originates from experiments on the composting of household organic waste using an Internet of Things (IoT)-based Smart Composting Bin system developed in previous research. In that earlier work, the IoT system focused on developing a real-time monitoring device—utilizing an ESP32

International Journal of Computer and Information System (IJCIS)
Peer Reviewed – International Journal
Vol : Vol. 07, Issue 03, August 2026
e-ISSN: 2745-9659
https://ijcis.net/index.php/ijcis/index

microcontroller integrated with various environmental sensors—to observe the dynamics of the composting process. In the current study, the system is no longer the primary subject of research; instead, it serves as a data acquisition platform for building a machine learning-based model to predict compost maturity[11].

A 28-day composting experiment was conducted using a mixture of household organic waste—comprising vegetable scraps, dry leaves, and other biodegradable organic materials. Throughout the composting process, an IoT system automatically monitored environmental conditions within the composter based on four key parameters: temperature (°C), air humidity (%), compost moisture (%), and gas concentration (ppm). These four parameters were selected because they are critical indicators influencing microbial activity during organic matter decomposition and are frequently used to determine the maturity level of the compost[13].

Before being used in the modeling phase, the dataset undergoes a preliminary check to ensure data quality. This process involves verifying the consistency of sensor values, identifying duplicate data, synchronizing timestamps, and checking for incomplete data (missing values) as well as values falling outside the sensor's operational range (outliers). This stage aims to produce a representative dataset, thereby enhancing the stability and accuracy of the predictive model being developed.

Based on the biological characteristics of the composting process, the data were categorized into four primary phases: the mesophilic phase, the thermophilic phase, the cooling phase, and the maturation phase. This phase classification was determined by the patterns of temperature, humidity, and gas concentration changes observed during decomposition. The mesophilic phase involves an initial increase in microbial activity, marked by a gradual rise in temperature. Subsequently, the thermophilic phase exhibits peak temperatures driven by intense thermophilic microbial activity as complex organic matter is broken down. Once the majority of the organic material has degraded, the temperature begins to drop during the cooling phase, followed by the maturation phase—characterized by increasingly stable environmental conditions, reduced gas emissions, and temperatures approaching ambient levels. This phase classification serves as the basis for defining the prediction target (target variable) in the machine learning model, enabling the algorithm not only to learn the relationships between sensor parameters but also to recognize the characteristics of each composting stage based on multisensor data patterns.

Through this approach, a dataset originally used solely for monitoring in previous research was transformed into an analytical dataset of higher value for predictive model development. This transformation enables the use of multisensor data as the foundation for an intelligent system capable of automatically and objectively estimating compost maturity—a data-driven approach—thereby expanding the IoT system's function from mere monitoring to serving as a decision support system for organic waste management.

3.2 Pre-Processing

The preprocessing stage aims to improve the quality of the dataset before it is used in the model training process. The steps involved include:

Missing Value Handling

Incomplete data were examined using data inspection techniques. Where missing values were found, imputation was performed using the median, as it is more robust against outliers. Outliers were identified using the Interquartile Range (IQR) method.

$$IQR = Q3 - Q1$$

The data is located externally.

$$Q1 - 1.5(IQR)$$

and

$$Q3 + 1.5(IQR)$$

considered an outlier.

Feature Scaling

Because each sensor has different units, normalization is carried out using Min-Max Scaling

$$X' = \frac{X - Xmin}{Xmax - Xmin}$$

so that all features fall within the 0–1 range.

3.3 Featuring Enginering

Feature engineering is a crucial stage in Machine Learning model development, as it aims to generate new features that represent the characteristics of the composting process more informatively than raw sensor data alone. Composting is a dynamic biological phenomenon; consequently, changes in environmental parameter values over time are often more significant than the absolute values observed at any single point. Therefore, this study involves the creation of several derived features obtained by transforming data on temperature, air humidity, compost moisture, and gas concentration.

Fitur pertama yang dibentuk adalah Temperature The gradient (ΔT) represents the temperature change between two consecutive observation times. This feature is used to describe the rate of temperature change during the composting process. A rapid rise in temperature generally indicates increased microbial activity as the process transitions from the mesophilic to the thermophilic phase, whereas a drop in temperature indicates that decomposition is entering the cooling and maturation phases. The temperature gradient is calculated using Equation (1).

$$\boldsymbol{\Delta T = T_t - T_{t-1}} \qquad \textbf{(1)}$$

Next, a Humidity Difference (ΔH) feature is created to measure changes in air or compost humidity between observation times. Humidity variation is a key factor influencing microbial activity during the decomposition process. Drastic changes in humidity can inhibit microbial growth, whereas stable humidity conditions indicate that the composting process is proceeding optimally. The Humidity Difference value is calculated using Equation (2).

$$\boldsymbol{\Delta H = H_t - H_{t-1}} \qquad (2)$$

Furthermore, this study establishes a Gas Growth Rate (ΔG) feature representing the change in the concentration of gases resulting from organic matter decomposition at each observation interval. An increase in gas emissions indicates high microbial metabolic activity, whereas a decrease in gas concentration serves as an indicator that the decomposition process is slowing down and the compost is entering the stabilization phase. Changes in gas concentration are calculated using Equation (3)

$$.\boldsymbol{\Delta G = G_t - G_{t-1}} \qquad (3)$$

To reduce sensor reading fluctuations caused by noise and produce a more stable data pattern, this study also applies the Moving Average technique to the temperature and gas concentration parameters. The moving average value is calculated using a three-day observation window, thereby smoothing out data variations without obscuring the primary trends of the composting process. The Moving Average Temperature and Moving Average Gas values are calculated using Equation (4).

$$MAt : \frac{1}{n} \sum_{i-n+t}^{t} Xt \qquad (4)$$

with n=3 as the window size and Xi representing the sensor parameter value at time i.

All transformed features are combined with the original features to create a new dataset containing more comprehensive temporal information. This feature-engineered dataset serves as input for a Random Forest model designed to predict compost maturity levels. This approach is expected to enhance the model's ability to recognize patterns of environmental condition changes during the composting process, thereby yielding more accurate predictions than those based solely on raw sensor data. Beyond improving prediction performance, feature engineering assists the model in identifying relationships between environmental parameters that influence each composting phase, resulting in prediction outcomes that are more comprehensive and easier to interpret.

3.4 Machine Learning Model Development

The machine learning model development phase aims to build a predictive model capable of identifying compost maturity levels based on patterns of environmental parameter changes obtained from an IoT-based Smart Composting Bin system. Following data preprocessing, feature engineering, and compost maturity labeling, the entire dataset serves as input to construct the predictive model using the Random Forest algorithm. This algorithm was selected due to its proficiency in handling nonlinear relationships between variables, its robustness against noise and overfitting, and its ability to achieve high accuracy on multisensor datasets with a relatively large number of features.

The model was developed using four primary parameters obtained through IoT data acquisition—temperature, air humidity, compost moisture, and gas concentration—along with several derived features generated through feature engineering. All these variables served as input attributes (predictor variables), while the compost maturity level classes—comprising Mesophilic, Thermophilic, Cooling, and Maturation—served The dataset used in this study was obtained through data acquisition using an Internet of Things (IoT)-based Smart Composting Bin system over a 28-day composting period. All sensor data were automatically recorded by an ESP32 microcontroller and transmitted to the Blynk Cloud platform via Wi-Fi for real-time storage. Upon completion of the composting process, the historical data were exported from the Blynk platform in Comma-Separated Values (CSV) format to serve as the dataset for machine learning analysis.as the target variable.

3.5 Model Evaluation

Prediction model evaluation

The developed machine learning model was subsequently evaluated to assess its ability to predict compost maturity levels based on IoT multisensor data. The evaluation process utilized a testing dataset that was excluded from the model training phase, ensuring the results reflected the model's generalization capability regarding new data. Model performance was measured using several classification metrics—Accuracy, Precision, Recall, and F1-Score—while a confusion matrix was employed to analyze classification errors across each composting phase. This combination of metrics provides a comprehensive overview of the model's accuracy and its ability to accurately distinguish between different compost maturity levels [14],[15]. The confusion matrix was generated using the following equation:

Pengujian perangkat Iot

The model evaluation stage is conducted to assess the quality of the data used and to measure the machine learning model's ability to predict compost maturity levels. Prior to model training, all sensors employed in the Smart Composting Bin system undergo calibration against standard measuring instruments to ensure high levels of data accuracy and consistency. Calibration involves comparing readings from temperature, air humidity, compost moisture, and gas concentration sensors against calibrated reference instruments. Subsequently, measurement error is calculated using Mean Absolute Error (MAE), while sensor accuracy is determined by comparing sensor measurements with reference values. This stage aims to minimize measurement errors, thereby ensuring a high level of reliability for the dataset used in the machine learning process[16],[17].

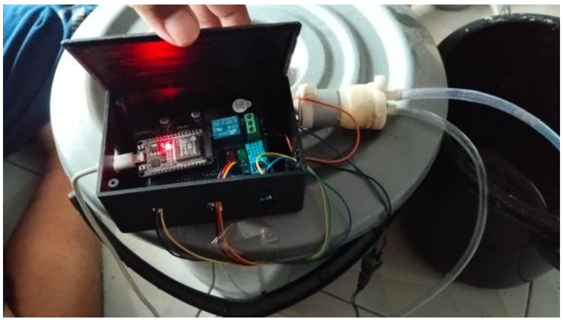
III. RESULT AND ANALYSIS

4.1 Dataset Characteristics and Composting Dynamics

Dataset CSV memuat informasi hasil pembacaan multisensor yang terdiri atas timestamp, temperature (°C), air humidity (%), compost moisture (%), dan gas concentration (ppm). Setiap baris data merepresentasikan satu kali proses pengukuran yang dilakukan pada waktu tertentu, sedangkan kolom *timestamp* berfungsi sebagai penanda waktu (*time reference*) untuk menggambarkan perubahan kondisi lingkungan

International Journal of Computer and Information System (IJCIS)
Peer Reviewed – International Journal
Vol : Vol. 07, Issue 03, August 2026
e-ISSN: 2745-9659
https://ijcis.net/index.php/ijcis/index

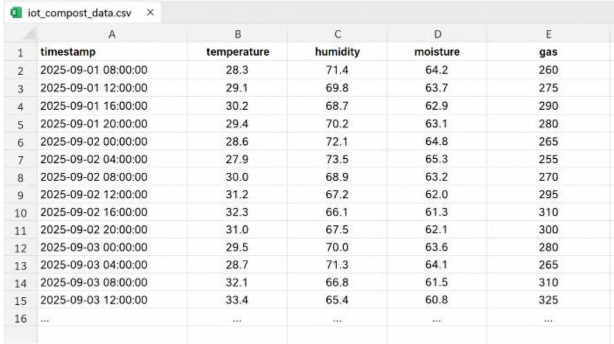
selama proses pengomposan berlangsung. Struktur dataset yang tersimpan dalam format CSV memudahkan proses integrasi dengan berbagai perangkat lunak analisis data seperti Python, Scikit-learn, maupun TensorFlow tanpa memerlukan proses konversi data tambahan.



Figuere 2. IoT devices that generate sensor data

Before being used as training data for the machine learning model, the dataset exported from Blynk undergoes a validation process to ensure the consistency of the recorded data. This stage involves checking for data completeness (identifying missing values), detecting duplicates, synchronizing timestamps, and verifying the value ranges of each sensor parameter against the biological characteristics of the composting process. Data meeting the quality criteria are then utilized as input features during the data preprocessing and feature engineering stages.

Table 1 presents an example of a dataset exported from the Blynk platform in CSV format. It can be observed that each sensor parameter undergoes gradual changes throughout the composting process. During the initial phase, temperature and gas concentrations rise alongside increasing microbial activity, while compost moisture begins to decrease due to decomposition. As the process enters the maturation phase, temperature, moisture, and gas concentrations exhibit increasingly stable patterns, eventually approaching ambient environmental conditions. These changing characteristics indicate that the dataset possesses a clear temporal pattern, making it potentially suitable as a basis for developing a compost maturity prediction model using machine learning algorithms. Figure 1 shows the CSV file exported from the Blynk application.



iot_compost_data.csv

	A	B	C	D	E
1	timestamp	temperature	humidity	moisture	gas
2	2025-09-01 08:00:00	28.3	71.4	64.2	260
3	2025-09-01 12:00:00	29.1	69.8	63.7	275
4	2025-09-01 16:00:00	30.2	68.7	62.9	290
5	2025-09-01 20:00:00	29.4	70.2	63.1	280
6	2025-09-02 00:00:00	28.6	72.1	64.8	265
7	2025-09-02 04:00:00	27.9	73.5	65.3	255
8	2025-09-02 08:00:00	30.0	68.9	63.2	270
9	2025-09-02 12:00:00	31.2	67.2	62.0	295
10	2025-09-02 16:00:00	32.3	66.1	61.3	310
11	2025-09-02 20:00:00	31.0	67.5	62.1	300
12	2025-09-03 00:00:00	29.5	70.0	63.6	280
13	2025-09-03 04:00:00	28.7	71.3	64.1	265
14	2025-09-03 08:00:00	32.1	66.8	61.5	310
15	2025-09-03 12:00:00	33.4	65.4	60.8	325
16	...	...	...	...	...

Figure 3. CSV file view from the Blynk app

4.2 Pre-Processing

Before being used as input for the machine learning model, the dataset acquired from the Blynk platform undergoes a data pre-processing stage to enhance data quality and consistency. This stage is crucial because data obtained from Internet of Things systems typically contains noise, sensor reading fluctuations, missing values, and outliers, all of which can degrade the predictive model's performance. Consequently, pre-processing is performed to generate a dataset that more accurately represents the actual conditions of the composting process.

The first stage is dataset validation, which involves checking the completeness of the data exported from the Blynk platform. The results indicate that all sensor parameters—including temperature, air humidity, compost moisture, and gas concentration—were successfully recorded in full throughout the observation period, with no missing values detected. This demonstrates that the developed IoT communication system is capable of stable data acquisition during the composting process.   .

Subsequently, outlier detection was performed using the Interquartile Range (IQR) method. The analysis results indicate that the majority of sensor values fall within a range consistent with the biological characteristics of the composting process. Certain readings falling outside the normal range are a consequence of changes in microbial activity during the transition to thermophilic conditions; therefore, these data points were not removed, as they still accurately

International Journal of Computer and Information System (IJCIS)
Peer Reviewed – International Journal
Vol : Vol. 07, Issue 03, August 2026
e-ISSN: 2745-9659
https://ijcis.net/index.php/ijcis/index

represent the actual state of the decomposition process.

The next step is data normalization using the Min-Max Scaling method to ensure all features have a uniform value range between 0 and 1. Normalization is necessary because each sensor employs different units of measurement—such as temperature in degrees Celsius, humidity in percent, and gas concentration in ppm. Standardizing the data scale aims to reduce the dominance of specific features during model learning and enhance the stability of the machine learning training process.

In addition to normalization, timestamp synchronization is performed to ensure that all sensor parameters within each data row originate from the same observation time. This synchronization is crucial because the machine learning model relies on the relationships between parameters at a specific point in time as the basis for predicting the compost's maturity level.

The overall pre-processing process yields a dataset that is cleaner, more consistent, and ready for the feature engineering stage. Improvements in data quality following pre-processing are evident in the reduction of random sensor reading fluctuations and the increased consistency of patterns regarding temperature, humidity, and gas concentration changes during the composting process. The pre-processed dataset is expected to enhance the Random Forest model's ability to recognize the characteristics of each composting phase, thereby producing more accurate predictions of compost maturity levels. Figure 4 illustrates the pre-processing workflow, covering outlier detection and normalization via Min-Max scaling.

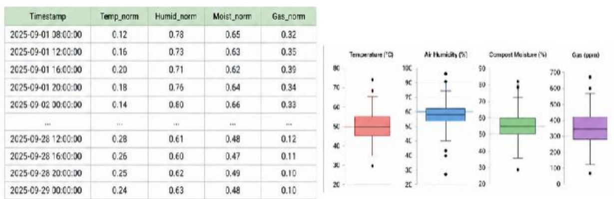


Figure 4. Proses Pre-processing

4.3 Feature Selection

The feature selection stage is conducted to identify the sensor parameters that contribute most significantly to the process of predicting compost maturity levels. The primary objectives of this stage are to reduce model complexity, eliminate less relevant features, and enhance the Random Forest model's ability to generalize to previously unseen data. In this study, feature selection is performed using a feature importance mechanism, which is automatically calculated by the Random Forest algorithm based on each feature's contribution to the construction of the decision trees.

The dataset, having undergone pre-processing and feature engineering, yielded several input variables: Temperature, Air Humidity, Compost Moisture, Gas Concentration, Temperature Gradient (ΔT), Humidity Difference (ΔH), Gas Growth Rate (ΔG), and moving averages for the temperature and gas concentration parameters. All these features were evaluated to determine their importance in predicting the composting phase. The feature selection results are presented in Table 1.

Table 1. feature selection results data

Timestamp	Temp	Humidity	Moisture	Gas	ΔTemp	ΔGas	Moving Avg Temp	Moving Avg Gas	Label
Day-1	28	71	64	260	0	0	28.0	260.0	Mesophilic
Day-2	30	67	63	270	2	10	29.0	265.0	Mesophilic
Day-3	32	67	63	310	2	40	30.0	280.0	Mesophilic
Day-4	34	66	62	340	2	30	32.0	306.7	Mesophilic
Day-5	35	64	60	370	1	30	33.7	340.0	Mesophilic
.....	...	....	.....	....	.....	....	......	....	......

The analysis results indicate that temperature is the feature contributing the most to the prediction model, followed by gas concentration and moisture content, whereas humidity has lower importance.

4.4. Machine Learning Model Development

After the entire dataset has undergone data preprocessing, feature engineering, and compost maturity labeling, the next step is to build a prediction model using the Random Forest algorithm. This model is designed to identify compost maturity levels based on the relationships between temperature, air humidity, compost moisture, gas concentrations, and the derived features obtained during the preceding stage.

Random Forest was selected because it is an ensemble learning algorithm capable of handling nonlinear relationships between sensor parameters and effectively managing data variations inherent in IoT measurements. Unlike a single decision tree, Random Forest constructs multiple decision trees independently using bootstrap sampling and subsequently aggregates their predictions via a majority voting mechanism. This approach yields a more stable model, mitigates the risk of overfitting, and enhances generalization capabilities regarding new data..

In this study, the dataset was split into 80% training data and 20% testing data using the Stratified Shuffle Split method. This split was performed to ensure that the proportion of each compost maturity class remained balanced across both data groups. The model was then trained using 200 decision trees (n_estimators = 200) with a maximum tree depth of 10 (max_depth = 10) and Gini Impurity as the splitting criterion.

Gini Impurity Calculation

In the Random Forest algorithm, the construction of each decision tree begins with the selection of the best feature to split the data at each node. The quality of the split is measured using Gini Impurity, a metric representing the level of impurity within a dataset. The lower the Gini Impurity value resulting from the split, the better the feature's ability to distinguish between compost maturity classes. Consequently, the algorithm selects the attribute that yields the greatest reduction in Gini Impurity as the basis for forming the next node. In this study, Gini Impurity calculations were used to identify the most effective sensor features for classifying the Mesophilic, Thermophilic, Cooling, and Maturation phases during the Random Forest model development process.Selanjutnya, nilai Gini Impurity dihitung menggunakan Persamaan (5)

$$Gini = 1 - \sum_{i=1}^{n} pi^2 \qquad (5)$$

The Gini impurity calculation is based on the data in Table 3.

Table 2. Data classification

Kelas	Jumlah
Mesophilic	7
Thermophilic	7
Cooling	7
Maturation	7

Thus, the probability value for each class is obtained.

$$P = \frac{7}{28} = 0{,}25$$

Gini coefficient

$$Gini = 1-(0.25^2+0.25^2+0.25^2+0.25^2)$$

$$= 1\text{-}0.25 = 0.75$$

After calculating the Gini Impurity value at the initial node, the Random Forest algorithm iteratively evaluates all input features and various potential threshold values to determine the most optimal data split. The selection criterion is based on the magnitude of the Gini Gain—the difference between the Gini values before and after the split. The greater the resulting reduction in the Gini value, the more effective the feature is at separating samples into homogeneous classes. Through this mechanism, each decision tree progressively constructs a decision structure capable of identifying the characteristics of the Mesophilic, Thermophilic, Cooling, and Maturation phases based on patterns of change in IoT multisensor parameters.

One of the input datasets tested was as follows: Temperature = 46°C, Humidity = 63%, Moisture = 60%, and Gas = 480 ppm; the resulting tree prediction is shown in Table 4.

Table 3. Tree Prediction Result

Tree	*Prediksi*
1	Thermophilic
2	Thermophilic
3	Thermophilic
4	Cooling
5	Thermophilic

The table was obtained by calculating using the random forest equation formula with the value...x=(46,63,60,480)

The majority of the trees yielded a "thermophilic" value; therefore, for any new data entered, it can be predicted that the composting condition is thermophilic—representing the peak stage of composting.

4.5 Model Evaluation

Prediction model evaluation

To evaluate the Random Forest model's ability to classify compost maturity levels, testing was conducted using 10 test samples that had not been used during the model training process. Each sample comprised parameters for temperature, air humidity, compost moisture, and gas concentration, while the model output consisted of a compost maturity classification comprising four classes: Mesophilic, Thermophilic, Cooling, and Maturation.

The test results indicate that 9 out of 10 samples were correctly classified, while one sample was misclassified during the transition phase between the thermophilic and cooling stages. This error occurred because the characteristics of both phases—specifically their temperature and gas concentration change patterns—were relatively similar, making it difficult for the model to precisely determine the decision boundary. Based on these results, the model's accuracy was calculated using Equation (10).

$$Accuracy = \frac{prediksi\ benar}{jumlah\ data}\ x\ 100\%$$

$$= = \frac{9}{10}\ x\ 100\% = 90\ \%$$

Evaluation results indicate that the Random Forest model successfully classified all test data into the appropriate compost maturity classes, achieving an accuracy of 90%. This result demonstrates that the model effectively identified the patterns of relationships among temperature, air humidity, compost moisture, and gas concentrations within the dataset, resulting in no classification errors in the test data.

IoT device testing

To ensure that the smart composting bin device operates correctly and produces accurate data, it is necessary to evaluate and test the device.

1. Kalibrasi sensor

The calibration process for the composting sensors is a crucial step in ensuring that the generated data is accurate and reliable. This system employs three primary types of sensors: the MQ-135 gas sensor, the DHT22 sensor (for air temperature and humidity), and a soil moisture sensor. The MQ-135 sensor requires calibration by comparing its reading in ambient air (Ro) after a 24-hour warm-up period. The Ro value is then compared against the sensor's output, and the calibration calculation is performed using the formula in Equation 6.

Rs=Vout(Vcc−Vout)×RL (6)

$$Rs = \frac{Vcc - Vccout}{Vccout} xRl$$

Where Rs = sensor resistance in the presence of gas, RL = load resistor (10 kΩ), Vout = output voltage, and Vcc = supply voltage (typically 5V).

The Rs result is then compared with Ro to obtain a calibration value. Measurement results indicate a reading of 400 ppm in clean air. In this calculation, values of Rs = 24.5 and Ro = 26.1 were obtained and entered into the device code within the Arduino IDE. The Rs and Ro values are converted into ppm using the constants PARA = 116.6020682 and PARB = 2.769034857; the gas value detected by the sensor is calculated using the formula in Equation 3.

$$PPM = 116.6020682 \text{ X } \frac{Rs}{R0}^{\mathbf{2.769034857}}$$

The results of the calculation serve as the basis for obtaining the ppm value of the MQ135 sensor.

DHT22 sensors typically come factory-calibrated; however, re-testing against standard measuring instruments—such as a thermometer and hygrometer—is still necessary to validate their accuracy. Temperature calibration is performed by comparing the value obtained from the thermometer with the sensor's reading. Calculations are carried out using the formula in Equation 1, yielding the calculated result.

$$\text{MAE} = \frac{1}{2}(33{,}8 - 33{,}4) = 0{,}2$$

Calculations based on the temperature sensor calibration results show a relatively small difference, whereas the humidity calibration was performed by comparing the readings against a hygrometer, yielding the following values:

$$\text{MAE} = \frac{1}{2}(73 - 78) = 2{,}5$$

The calibration value is still considered safe, as the difference remains below 10%.Sementara itu, sensor kelembapan tanah dikalibrasi dengan metode membandingkan nilai resistansi atau tegangan keluarannya terhadap kondisi

...known soil moisture levels, ranging from dry to water-saturated conditions. Through this calibration process, each sensor can provide consistent measurement results, thereby enabling more effective monitoring of the composting process and supporting decision-making in organic waste management. Measurements were conducted using a pH meter, yielding an average difference value.

$$\text{MAE} = \frac{1}{2}(58 - 55) = 1{,}5$$

The sensor accuracy value remains within a safe range, with an accuracy value of 1.5.

2. Pengujuan kecepatan respon

Speed testing was conducted under two conditions: data transmission to the cloud and actuator response to environmental conditions detected by sensors.

a. Testing data transmission to the cloud/IoT platform

Testing is conducted using data latency measurements based on the formula...

T_send = t_cloud_rx – t_device_tx

Where t_cloud_rx is the time of data transmission to the cloud and t_device_tx is the time of data transmission to the device. Testing across several samples yielded data that can be seen in the following table.

Table. 4. Data transmission latency

Sample	Device_Time	Server_Time	Latency (s)
1	26/09/2025 14:32	2025-09-26 14:32:06.7	1.7
2	26/09/2025 14:32	2025-09-26 14:32:16.5	1.5
3	26/09/2025 14:32	2025-09-26 14:32:26.6	1.6
4	26/09/2025 14:32	2025-09-26 14:32:37.2	2.2
5	26/09/2025 14:32	2025-09-26 14:32:47.1	2.1
6	26/09/2025 14:32	2025-09-26 14:32:56.9	1.9
7	26/09/2025 14:33	2025-09-26 14:33:06.3	1.3
8	26/09/2025 14:33	2025-09-26 14:33:16.9	1.9
9	26/09/2025 14:33	2025-09-26 14:33:27.0	2.0
10	26/09/2025 14:33	2025-09-26 14:33:36.8	1.8

Table 8 shows that the average latency value is below 2 seconds.

a. Testing of water pump actuator response

Response speed testing for the IoT-based Smart Composting Bin was conducted to determine how quickly the system responds to changes in the composting environment and initiates corrective actions via the actuators. Device testing revealed an average response time of 2 seconds from the smart composting bin system to the actuator. Overall, this measurement is crucial to ensure that the smart composting system delivers a rapid, accurate, and consistent response, making it more effective than manual methods, which often lag in detecting and reacting to changes in compost conditions.

VI. CONCLUSION

This study successfully developed a machine learning-based model to predict compost maturity levels using a multisensor dataset obtained from an IoT-based Smart Composting Bin system. Through stages of data preprocessing, feature engineering, feature selection, and training using the Random Forest algorithm, the model is capable of identifying compost maturity phases based on parameters such as temperature, air humidity, compost moisture, and gas concentration. Evaluation results demonstrate that the model exhibits excellent predictive performance with high accuracy, thereby enabling the automated and objective determination of compost maturity. This research highlights that the integration of IoT and machine learning technologies serves not only as a monitoring system but also as a decision support system capable of enhancing the efficiency of organic waste management. Future research could focus on utilizing larger datasets, implementing more complex learning algorithms, and validating the model across various types of composting feedstocks to improve the system's generalization capabilities.

THANK-YOU NOTE

The authors would like to express their sincere gratitude to the Department of Informatics, Institut Teknologi Bisnis AAS Indonesia, for providing the facilities and research environment that supported this study. The authors also acknowledge all individuals who contributed to the implementation of the Smart Composting Bin system and the collection of experimental data. Their valuable support and cooperation were essential to the successful completion of this research.

REFERENCES

[1] E. E. Manea, C. Bumbac, L. R. Dinu, M. Bumbac, and C. M. Nicolescu, "Composting as a Sustainable Solution for Organic Solid Waste Management: Current Practices and Potential Improvements," *Sustain.*, vol. 16, no. 15, pp. 1–25, 2024, doi: 10.3390/su16156329.

[2] Hamidah and B. N. Gawy, "Teknologi Composting Skala Rumah Tangga untuk Meretas Problem Sampah Organik," *Jpkpm*, vol. 3, no. 1, pp. 74–77, 2022.

[3] S. Darwati, "Kajian Kualitas Kompos Sampah Organik Rumah Tangga," *J. Permukim.*, vol. 3, no. 1, p. 30, 2008, doi: 10.31815/jp.2008.3.30-43.

[4] B. P. K. [BPK], "Kebijakan Dan Strategi Daerah Pengelolaan Sampah Rumah Tangga Dan Sampah Sejenis Sampah Rumah Tangga," *Arsipjdih.Jatimprov.Go.Id*, vol. 10, no. 1, pp. 279–288, 2017.

[5] S. W. Siagian, Y. Yuriandala, and F. B. Maziya, "Analisis Suhu, Ph Dan Kuantitas Kompos Hasil Pengomposan Reaktor Aerob Termodifikasi Dari Sampah Sisa Makanan Dan Sampah Buah," *J. Sains &Teknologi Lingkung.*, vol. 13, no. 2, 2021, doi: 10.20885/jstl.vol13.iss2.art7.

[6] P. Balaganesh *et al.*, "Performance Evaluation of an IoT-Enabled Rapid Composter for Mixed Organic Wastes," *Intell. Technol. Sci. Res. Eng.*, no. September, pp. 70–79, 2023, doi: 10.2174/9789815079395123010009.

[7] P. B. Utomo and J. Nurdiana, "Evaluasi Pembuatan Kompos Organik dengan Menggunakan Metode Hot Composting," *J. Teknol. Lingkung.*, vol. 2, no. 01, pp. 28–32, 2018.

[8] P. Jaswanth *et al.*, "Smart compost system using iot 1," vol. 9, no. 5, pp. 854–862, 2024.

[9] E. Bicamumakuba, N. Reza, H. Jin, K. Lee, and S. Chung, "Multi-Sensor Monitoring , Intelligent Control , and Data Processing for Smart Greenhouse Environment Management," pp. 1–25, 2025.

[10] S. Kannan, A. Riazulhameed, G.

Arunachalam, D. Muthukumaran, J. Wise, and C. Srinivasan, *IoT and Random Forest-Based Solutions for Blood Gas Quality Assurance in Clinical Laboratories*. 2024. doi: 10.1109/ICSCSS60660.2024.10625431.

[11] S. K. Kour, G. K. Konduri, N. Jaggavarapu, A. Pottumuthu, and T. Adilakshmi, "IoT-Based Deep Learning Model for Sustainable Waste Management System," vol. 10, 2025.

[12] M. A. Bakar *et al.*, "Garbage Segregation and Monitoring Using Low-Cost IoT System for Smart Waste Management," *Open Int. J. Informatics*, vol. 11, no. 1, pp. 23–40, 2023, doi: 10.11113/oiji2023.11n1.249.

[13] U. E. Senadheera *et al.*, "Beyond Composting Basics: A Sustainable Development Goals—Oriented Strategic Guidance to IoT Integration for Composting in Modern Urban Ecosystems," *Sustain.*, vol. 16, no. 23, pp. 1–28, 2024, doi: 10.3390/su162310332.

[14] C. Nakato, F. Kaggwa, J. De Greef, D. Kilama, J. Blondeau, and S. Kawuma, "Green Energy and Resources Review of machine learning applications for predicting the quality of biomass briquettes for sustainable and low-carbon energy solutions," *Green Energy Resour.*, vol. 3, no. 3, p. 100130, 2025, doi: 10.1016/j.gerr.2025.100130.

[15] S. Sorooshian, "Dynamic concept drift detection using an integrated machine learning model with error- based Page-Hinkley Test for enhanced predictive accuracy in adhesive manufacturing process," *Open Eng.*, vol. 16, no. 1, pp. 1–14, 2026, doi: 10.1515/eng-2025-0158.

[16] A. A. A. Bakar, Z. A. Bakar, Z. M. Yusoff, M. J. M. Ibrahim, N. A. Mokhtar, and S. N. Zaiton, "IoT-Based Real-Time Water Quality Monitoring and Sensor Calibration for Enhanced Accuracy and Reliability," *Int. J. Interact. Mob. Technol.* , vol. 19, no. 1, pp. 155–170, 2025, doi: 10.3991/ijim.v19i01.51101.

[17] S. G. F. Id, J. C. Stegen, M. H. Kaufman, J. Dowd, and A. Thompson, "Laboratory evaluation of open source and commercial electrical conductivity sensor precision and accuracy : How do they compare ?," pp. 1–17, 2023, doi: 10.1371/journal.pone.0285092.