

IoT-Integrated Machine Learning for Precision Watering in Bamboo Mushroom Farming

Pinit Nuangpirom¹, Siwasit Pitjarnit², Parida Jewpanya^{2*}, Nopadon Maneetien³

¹ Department of Industrial Education and Technology, Faculty of Engineering, Rajamangala University of Technology Lanna, Chiang Mai, 50300, THAILAND

² Department of Industrial Engineering, Faculty of Engineering, Rajamangala University of Technology Lanna Tak, Tak 63000, THAILAND

³ Department of Electrical Engineering, Faculty of Engineering, Rajamangala University of Technology Lanna, Chiang Mai, 50300, THAILAND

*Corresponding Author: parida.jewpanya@rmutl.ac.th

DOI: <https://doi.org/10.30880/ijie.2025.17.01.003>

Article Info

Received: 5 August 2024

Accepted: 27 October 2024

Available online: 30 April 2025

Keywords

Bamboo mushroom, machine learning, regression algorithm, Internet of Thing.

Abstract

This study offers a method for enhancing bamboo mushroom farming. We aim to increase productivity by combining machine learning methods for device-level computing with the Internet of Things (IoT). The first step is to record the ideal environmental conditions in the bamboo mushroom greenhouse. The IoT devices collect data on temperature, humidity, soil moisture, and water usage, storing it in the cloud. The regression model is then formulated for irrigation control and predicting water consumption in the bamboo mushroom farm. Later, the microcontroller is programmed to control the water pump in a systematic manner to release water. The study found that temperature, soil moisture, and relative humidity are the primary factors affecting water content. The proposed method increased mushroom volume by 46.67% and saved 22% of water usage, demonstrating the successful integration of machine learning into smart farming at the device level.

1. Introduction

Bamboo mushrooms, with their delicious flavor and flower-like shape, are popular in cooking and vegetable gardens. They have a soft, sweet taste and are rich in protein, vitamins, and minerals like iron and salt [1]. Their stems can be brown or green, and they are suitable for growing in small houses or vegetable gardens [2]. Bamboo mushroom production requires expertise and adaptability to local climates. It involves setting up equipment, assembling leaves, adding cultures, fibers, maintaining humidity, waiting for mushrooms to grow, and harvesting them when they're ready for cooking or processing. Bamboo mushroom cultivation presents several challenges primarily due to their specific environmental requirements and complex life cycle. These mushrooms need a highly controlled and humid environment, mimicking their natural tropical and subtropical habitats, which can be difficult and costly to replicate on a large scale [3]. Additionally, bamboo mushrooms have a unique and intricate development process, including a dependency on specific soil microbes and organic material for optimal growth. This complexity makes cultivation labor-intensive and requires precise monitoring of conditions such as temperature, humidity, and substrate composition. Furthermore, their delicate structure makes them susceptible to damage and contamination, posing additional hurdles for successful and sustainable cultivation [4].

Smart greenhouses are a technological advancement that can significantly address and solve various agricultural problems by utilizing modern technologies to improve efficiency through the incorporation of sensor technologies and automated control systems. This reduces resource waste and develops durable solutions like

temperature and humidity management, thereby improving livestock and crop cultivation efficiency [5]. Smart greenhouses, for instance, use Internet of Things (IoT) technology to aggregate data from various sensors for analysis and well-informed decision-making when measuring water content, food content, or air quality [6]. People are adopting greenhouse lighting technologies, such as LED systems that can adjust light intensity to suit plant needs [7]. This makes it possible to regulate the perfect amount of color and lighting duration. Using energy-saving technologies, like solar energy and smart cooling systems, to reduce energy use that affects the environment [8]. By utilizing innovation and technology, smart greenhouses can operate more efficiently. Smart greenhouses can reduce resource usage and implement a more comprehensive management system to meet the demands of farmers and advanced industries [9].

Nowadays, modern smart technology is increasingly incorporating machine learning methods into IoT systems. Machine learning is the ability of computers to learn from experience and improve performance. It uses data as a starting point for learning and creating models, helping anticipate future events and make judgments. There are various types of computer learning, including supervised, unsupervised, and reinforcement learning [10]. In machine learning, there are several algorithms, each with different functions and uses depending on the nature of the data and the problems encountered. Linear regression algorithm is used for predicting continuous values using linear generation to estimate the value of the dependent variable [11]. In another form, logistic regression algorithm is performed for classification, it is often used for classification tasks such as classifying data into two groups, 0 or 1 [12]. Decision trees (DT) are used for classification and prediction. Slicing data into smaller sets according to a given condition creates a tree structure [13]. Random forest algorithm (RFR) is the application of several decision trees simultaneously and combining results to reduce errors and increase stability [14]. Support Vector Machines (SVM) is used for linear and non-linear classification. It tries to find the highest dividing line between data groups [15]. K-Nearest Neighbors (K-NN) method locates information near the data for classification or prediction [16]. K-Means Clustering algorithm is used for data grouping by grouping data into groups based on similarity [17]. Finally, neural networks frequently tackle complex and large-scale problems, including deep learning [18].

Machine learning (ML) is utilized in various IoT applications, including device-level applications for real-time data processing, data layers for integrity verification, connectivity-level applications for communication automation, system-level applications for system analysis, and service-level applications for IoT-enabled apps and services. [19]. Embedded IoT devices use sensors to communicate with the environment, ranging from smart to connected devices. They have hierarchical architecture with multiple processing layers, and linked devices rely on users or cloud processing for decision-making and remote control [20]. The increasing number of IoT devices is causing challenges for central clouds in managing data collection due to the network's burden and unpredictable latency [21]. The edge computing paradigm suggests moving data processing to the collection point, allowing dispersion across layers. [22-25]. After examining relevant literature, it was determined that the majority of ML model processing takes place at the cloud level, which presents a significant challenge for integrating IoT with ML. This method is expensive and necessitates intricate connectivity. Creating machine learning models that can be processed by edge computing systems or by IoT devices themselves should be the main goal of solving this problem. This allows for real-time processing and decision-making in addition to cost reduction.

However, despite the advancement of smart greenhouses, there remains a gap in precision water management solutions specifically tailored for bamboo mushroom cultivation. Existing IoT systems in agriculture primarily focus on generalization, which may not cater to the delicate and highly specific environmental requirements of bamboo mushrooms. This research introduces a system that combines IoT technology with machine learning (ML) to develop a customized water control system. What distinguishes this method is its focus on edge computing, which allows for real-time, on-site data processing directly within the IoT devices themselves rather than relying heavily on cloud-based processing. This approach not only reduces operational costs but also improves the immediacy of decision-making, making it particularly effective for managing the sensitive growing conditions of bamboo mushrooms.

This paper presents the use of IoT systems integrated with the ML algorithm for automated water supply control for bamboo mushroom cultivation. The objective of this system is to offer several benefits to the bamboo mushroom farm, including optimized water usage, real-time monitoring, and enhanced control over environmental conditions [26]. The challenge with this system lies in making informed decisions based on accurate data from sensors, as inaccurate decisions can lead to overwatering or underwatering. A significant research gap lies in developing precise mathematical models or algorithms to predict water quantities for bamboo mushrooms under varying environmental conditions while addressing the operational constraints of IoT systems. This highlights the need for methods that integrate machine learning models with IoT capabilities to improve efficiency and sustainability. The proposed system uses IoT devices to gather data on the temperature, humidity, soil moisture, and water usage at each stage of the first crop in the bamboo mushroom farm, which uses the farmer's experience to control water supply. The ML algorithm then uses those data to perform the fit regression model [27], which is statistically suitable for determining the relationships among multiple independent variables and the dependent variable—the optimal water quantity for bamboo mushrooms under varying conditions [28–

29]. After that, microcontroller boards directly execute the ML algorithm and automate the water supply based on environmental factors for the next crop of the bamboo mushroom farm. The following sections provide the details.

2. Materials and Methods

This study focuses on an IoT-integrated machine learning algorithm for precision watering in bamboo mushroom farming. The proposed system is divided into four phases (Fig. 1). In the initial phase, the expertise and traditional practices of farmers are leveraged to establish a baseline for water management in bamboo mushroom cultivation. Farmers use their knowledge and experience to water the farm as they normally would. Concurrently, IoT devices are deployed to monitor and record various environmental parameters, including: temperature, humidity, soil moisture and water usage. By collecting and analysing this data, the system aims to capture the traditional farming practices and use them as a reference for developing more efficient and automated watering solutions in subsequent phases. This approach ensures that the automated system aligns with the proven techniques of experienced farmers while also identifying opportunities for optimization and innovation.

In the second phase, the data collected from the IoT devices is analysed using a machine learning algorithm to generate a regression model. This model is designed to predict optimal water usage based on the key environmental factors of temperature, humidity, and soil moisture. The multiple regression models are compared and evaluated to identify the best-performing model for predicting water usage in the third phase. This step involves rigorous testing and evaluation using key performance metrics including, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 score.

In the final phase, the selected regression model is integrated into the IoT systems to automate and optimize the watering process for bamboo mushroom cultivation. This implementation ensures that the real-time data collected by IoT devices is utilized effectively to maintain optimal growing conditions. By the end of this phase, the regression model will be fully implemented into the IoT system, automating the watering process to ensure optimal water usage based on real-time environmental data. This integration will lead to more efficient water management, improved crop yields, and reduced labor costs.

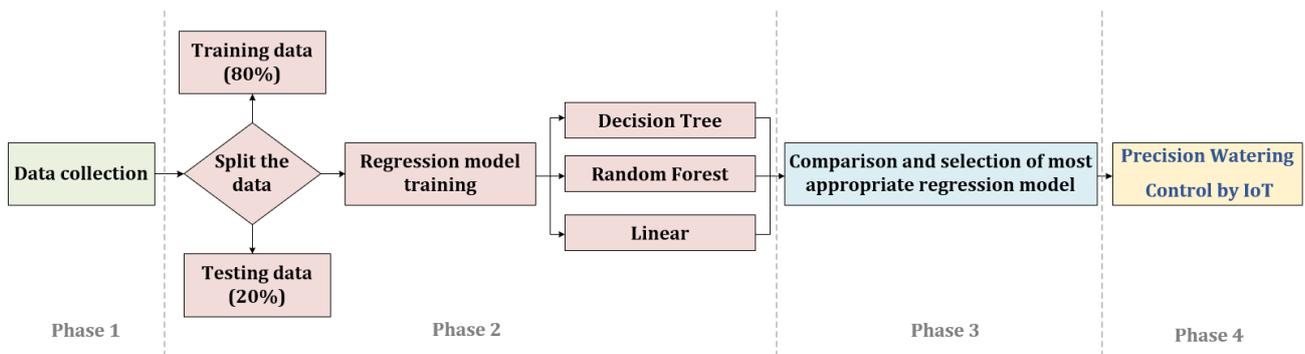


Fig. 1 System architecture of the IoT-Integrated ML for smart bamboo mushroom farming

2.1 IoT Design for Data Collection

2.1.1 Hardware Design

Fig.2 illustrates the IoT system architecture for recording environmental data in this study. The ESP32 controller is utilized. It has a smaller size, large input and output pins, built-in WiFi function, affordability, and support for a micro-USB cable for power supply and communication between the computer and microcontroller. The specifications of the ESP32 development board are shown in Table 1. The process of collecting data begins with measuring temperature and humidity using a DHT22 sensor, which transmits digital data to the D32 pin of the ESP32 microcontroller board. Soil moisture is measured using a soil moisture sensor, which sends analog data to the D34 pin of the ESP32 board. Water flow is measured with a flow sensor that sends digital data to the D12 pin of the ESP32 board, and the water flow data is used to calculate the water volume applied to the bamboo mushrooms. On the output side, the ESP32 board connects to the cloud system via the internet, storing data in a Google Sheets database and displaying it through the Blynk application [30].

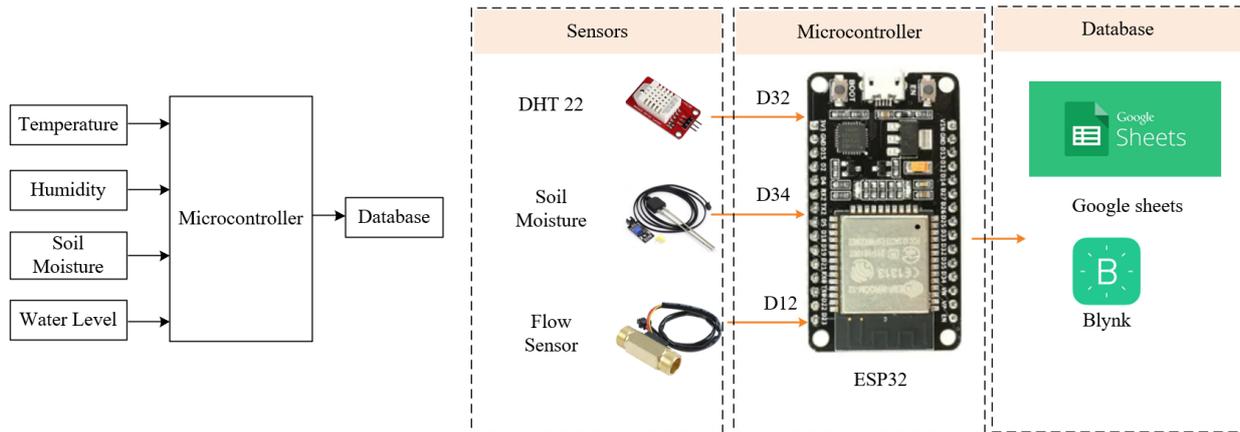


Fig. 2 IoT system for data collection

Table 1 Technical specification of ESP32

Feature	Description
Processor	Dual-core Xtensa LX6, 160-240 MHz
RAM	520 KB SRAM
Flash Memory	4 MB (external)
Wi-Fi	802.11 b/g/n, Wi-Fi Direct (P2P), soft-AP mode
Bluetooth	Bluetooth v4.2 BR/EDR and BLE (dual mode)
GPIO Pins	36 GPIO pins
Analog Inputs	16 channels of 12-bit SAR ADC
Analog Outputs	2 channels of 8-bit DAC
Digital I/O	PWM, I2C, I2S, SPI, UART
PWM Channels	Up to 16 PWM output channels
Operating Voltage	2.3 to 3.6 V
Current Consumption	Light sleep: 0.8 mA, Deep sleep: 10 µA, Hibernation: 5 µA, Active: 160-240 mA
Peripheral Interfaces	SPI, I2C, I2S, UART, CAN, Ethernet MAC (requires external PHY), SD/SDIO/MMC, RMT, PWM, Motor PWM
Temperature Range	-40°C to +125°C (industrial grade)

2.1.2 Software Design

Designing software for environmental data logging involves sending data to Google Sheets and displaying it via a mobile application using the Blynk application. Here are the steps for both designs:

1) Setting Up Google Sheets Integration:

- Create a Google Sheets Spreadsheet: Begin by creating a new spreadsheet in Google Sheets specifically designed to store the environmental data collected from this project.
- Enable Google Sheets API: Navigate to the Google Cloud Console, where you can either create a new project or select an existing one. To activate the Google Sheets API, navigate to "APIs & Services" > "Library" in the console, then search for "Google Sheets API." Proceed to make it possible for a new project.
- Establish Credentials: Create a dedicated service account for the project's use of the Google Sheets API. Download the JSON file containing the service account credentials. These credentials are essential for authenticating software and granting it programmatic access to Google Sheets.
- Write Software Code: Use a suitable programming language, such as C or Python, to create software that interacts with the Google Sheets API. Implement authentication mechanisms using service account credentials, and leverage the appropriate APIs to facilitate data reading and writing to the designated spreadsheet.

2) Integrating with the Blynk application:

- Create a Blynk Project: To display or control ESP32 data, download the Blynk app, create a new project, and customize it with widgets like buttons and graphs.
- Obtain Blynk Auth Token: After creating a project, receive an authentication token via email, which is crucial for ESP32 to connect to the Blynk server.
- Use Arduino IDE or Platform IO with Blynk libraries to write ESP32 code. To establish a connection, include Wi-Fi credentials and an Auth Token. Implement functions to send sensor data to Blynk widgets and receive commands.

An example of a software design is shown in Fig. 3.

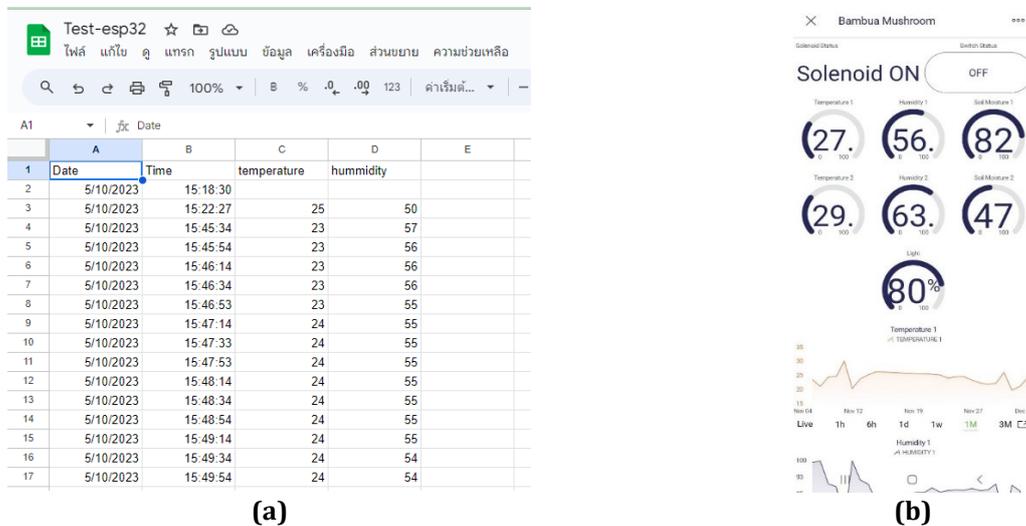


Fig. 3 Software design (a) Google sheets database; (b) Blynk application

2.2 Machine Learning Model Development

The development of the machine learning model is a critical phase that involves transforming the collected data into actionable insights for predicting optimal water usage. This phase includes several key steps: data preprocessing, model selection, training, and validation.

- 1) Data preprocessing: compile the data collected from the IoT devices during the first phase. Remove any outliers or anomalies to ensure the data is accurate and reliable. Standardize the data to facilitate effective analysis. This study performs descriptive statistics on the dataset by examining various statistical measures, such as the mean, standard deviation, minimum, and maximum values of all factors. By computing these statistical values, it is able to identify probable abnormalities that would impact the analysis, giving a thorough picture of the dataset.
- 2) Correlation analysis: identify relationships between temperature, humidity, soil moisture, and water usage. The analysis evaluated each feature's predictive potential, creating a correlation matrix to highlight strong positive or negative correlations.
- 3) ML algorithm selection: Select an appropriate machine learning algorithm for regression analysis. In this work, we selected three models of regression, including decision tree regression (DTR), random forest regression (RFR), and linear regression (LR).
- 4) Training and Testing the Model: Use the historical data to train and test the regression model, enabling it to predict future water usage based on current environmental conditions. At this point, the data was divided into training and testing sets with a ratio of 80:20 using the train-test split approach. This indicates that 80% of the entire dataset was selected for training, and the remaining 20% was designated for testing. This splitting ratio allows for training the model on a sizable enough subset of the data and testing its performance on test data.

2.2.1 Decision Tree Regression (DTR)

The decision tree algorithm uses a tree structure to compute the value of a dependent variable from independent variables. Each node in the tree divides the data into subgroups using Boolean conditions to predict the desired dependent variable [13]. The equation for DTR can be represented, as shown in Equation (1).

$$\bar{y}_i = \frac{1}{n} \sum_{i=1}^n y_i \quad (1)$$

where:

\bar{y}_i is the mean predicted value at terminal nodes

y_i is the mean value of the observed output

n is the number of samples at the current node

2.2.2 Random Forest Regression (RFR)

The random forest algorithm is a method used to address classification and regression problems. It combines multiple decision trees to improve prediction accuracy. The forest builds each tree using a random subset of the training data and features [14]. The equation for RFR can be represented, as shown in Equation (2).

$$\bar{Y} = \frac{1}{q} \sum_{l=1}^q \bar{Y}_l \quad (2)$$

where:

\bar{Y} is the final predicted value

\bar{Y}_l is the output of l -th tree, and $l = 1, 2, \dots, q$

q is the expected number of trees

2.2.3 Linear Regression (LR)

Linear regression aims to find the linear relationship between the independent variables X and a dependent variable Y . The model assumes a linear relationship of the form, as shown in the equation (3) [11]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon \quad (3)$$

when

Y is a dependent variable (response or target)

X_k is independent variable (predictors or features), and $k = 1, 2, 3, \dots, m$

β_0 is the intercept (constant term)

β_k are the coefficients (slope parameters) associated with each X_k

ε is the error term representing the random noise or residuals

2.3 Model Comparison and Selection

The evaluation stage aims to provide a thorough analysis of regression approaches and performance metrics, enabling users to select the classifier with the most relevant accuracy specifications. The explanation is as follows [20-21]. Mean Absolute Error (MAE) is a metric that calculates the average of model errors, indicating the difference between actual (y) and predicted values (y'), as shown in the equation (4):

$$MAE = \frac{1}{n} \sum_{i=1}^n (y_i - y'_i) \quad (4)$$

Root Mean Square Error (RMSE) is a statistical tool used to evaluate regression models by calculating the average deviation between predicted and actual values, with lower values indicating better performance. It can be calculated using the equation (5):

$$RMSE = \sqrt{\frac{1}{n} \left(\sum_{i=1}^n (y_i - y'_i)^2 \right)} \quad (5)$$

R^2 score, which ranges from 0 to 1, measures the variance of target variables in a model, with higher values indicating better model fit. It is calculated using the following the equation (6):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (6)$$

2.4 Implementing the Regression Model into IoT Systems

In this phase, the selected regression model is integrated into the IoT systems to automate and optimize the watering process for bamboo mushroom cultivation. As illustrated in Fig. 4, the workflow begins with collecting environmental data, such as temperature and humidity from the DHT22 sensor and soil moisture from the Soil Moisture sensor. These collected values serve as the independent variable in the regression equation. The calculated dependent variable represents the appropriate water quantity for the bamboo mushrooms based on the current environmental conditions.

To ensure the effective use of the IoT system and regression model, it is essential to provide users with a clear and intuitive interface for visualizing real-time data, model predictions, and watering schedules. There are two ways to display the results: 1) LCD Display: Directly displays the water amount on the local display; 2) Blynk Application: Transmits data over the internet for display on the Blynk app. This dual-display approach guarantees real-time monitoring and remote access to the water dispensing recommendations. The implementing system setup is illustrated in Fig. 5.

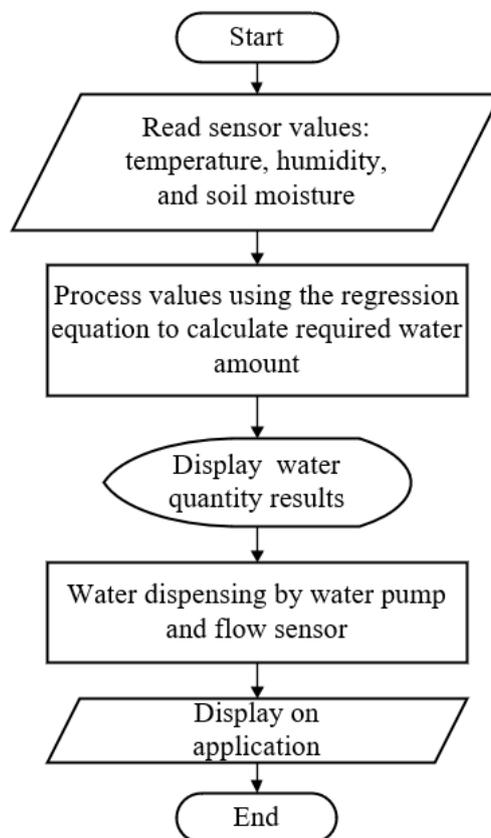


Fig. 4 Experimental setup of bamboo mushroom farm

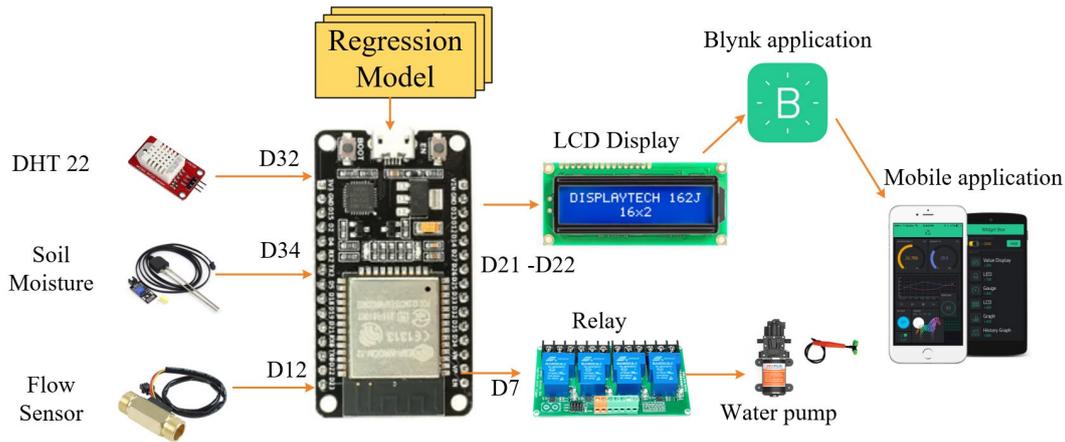


Fig. 5 Experimental setup of bamboo mushroom farm

3. Results

This section outlines the findings from a comprehensive study on the integration of IoT and ML for precise water management in bamboo mushroom cultivation. The results of environmental data collection using the IoT system are first discussed. The next section provides an evaluation of various machine learning models, leading to the selection of the most effective regression mode. Finally, the practical implementation of the integrated IoT-ML system for automated water control in the bamboo mushroom farm is investigated.

3.1 Environmental Data Collection Using IoT

The temperature, humidity, soil moisture, and daily water usage, were collected using IoT devices. This research utilized supervised learning in the experiment, which took place in a closed greenhouse with dimensions of 3 m x 6 m x 2 m (W x L x H), in Nong Luang Subdistrict, Umphang District, Tak Province, Thailand (at coordinates 16.07922805119336, 98.7534778151535). The cultivation period spanned from September 1, 2023, to September 26, 2023. Fig. 6 shows the results of temperature and humidity measurements using the DHT22 sensor. The sensor transmits digital data to the ESP32, which is then stored in a Google Sheets database. The recorded data reveal an inverse relationship between temperature and humidity: as the temperature increases, humidity tends to decrease. In greenhouse conditions, the highest recorded temperature reached 63 °C, while the maximum humidity observed was 95%. Fig. 7 illustrates the soil moisture levels measured using a soil moisture sensor integrated with the IoT system. The measurements show that soil moisture ranged from a minimum of 50% to a maximum of 77%. Fig. 8 displays the farmer's manual watering results, with water usage ranging from 108 to 144 liters. It is important to note that this experiment's constraints require successful bamboo mushroom cultivation. In other words, the data collected from the model farmer must come from an expert who has successfully cultivated bamboo mushrooms during the recording period for the training model to be effective.

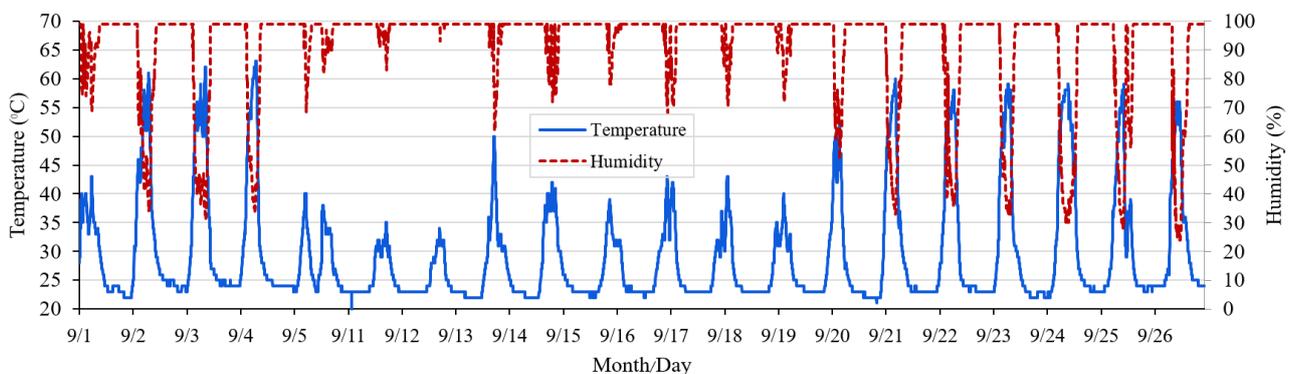


Fig. 6 Measurements of humidity and temperature in the greenhouse

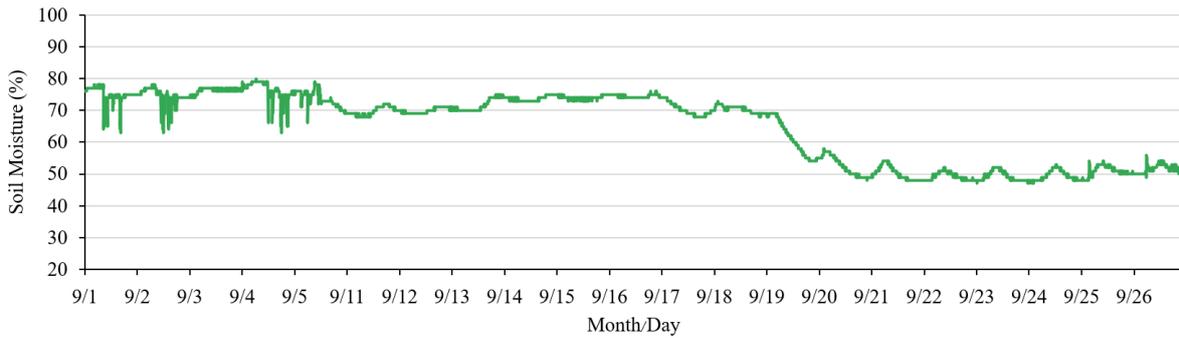


Fig. 7 Soil moisture

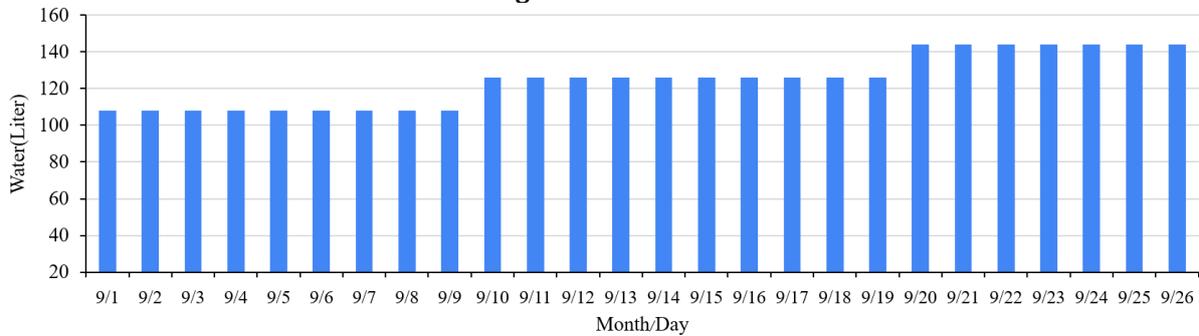


Fig. 8 The volume of water consumed each day

3.2 Development and Evaluation of Machine Learning Models

3.2.1 Data Pre-processing

This process performs descriptive statistics on the dataset by analyzing various statistical measures, including the mean, standard deviation, minimum, and maximum values for the bamboo mushroom farm. By computing these statistical values, it is able to identify probable abnormalities that would have an impact on the analysis, giving a thorough picture of the dataset. The environmental (temperature, humidity, soil moisture, and water) information is recorded for 26 days, yielding 37,440 records per variable. The data necessitates basic data management, which involves averaging daily data and reducing it to 52 records using Google's Colab software in Python and the Sklearn library, as illustrated in Table 2.

Table 2 Fundamental feature of the dataset

Feature	Mean	Std	Min	25%	50%	75%	Max
Temperature (°C)	28.34	3.14	23	26	27.5	31	34
Humidity (%)	92.8	5.89	81	87	96	97	99
Soil moisture (%)	67.4	10.7	49	58	72.5	75	78
Day	13.5	7.6	1	7.25	13.5	19.75	26
Water (liter)	129.7	14.8	108	115	130	143	150

In this research, feature selection becomes important after the steps of exploratory data analysis and data pretreatment. The method entails identifying and selecting the irrigation-related aspects that are most relevant to the overall quality. Temperature, humidity, and soil moisture levels are only a few of the data included in the characteristics of this study, which is based on a preprocessed dataset. Correlation analysis was used to determine the relationship between features and water management in bamboo mushroom farms. The analysis evaluated each feature's predictive potential, creating a correlation matrix to highlight strong positive or negative correlations. The study's findings showed that the majority of day values exhibited a positive correlation with water values, indicating that the number of days has an impact on the amount of water required for mushroom cultivation. There is also an inverse relationship between the day and soil moisture, humidity, and temperature. Additional validation is required, along with comparisons to other models, as shown in Fig. 9

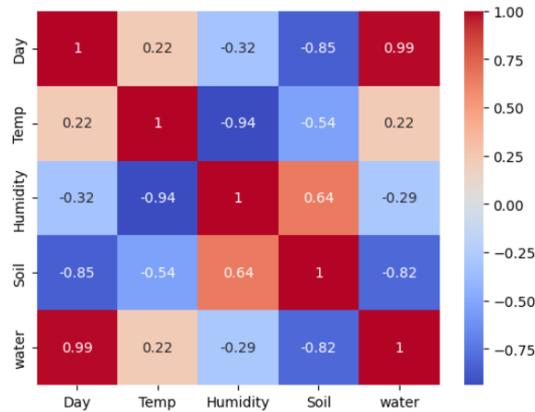


Fig. 9 Correlation of bamboo mushroom farm environment

3.2.2 Regression Model Construction

The process of developing regression models in this study utilized Google's Colab software. It began with importing the necessary libraries, specifically the Sklearn library for machine learning tasks. Next, a CSV file containing historical mushroom cultivation data from farmers' manual records was loaded. The independent and dependent variables were defined, followed by splitting the data into 80/20 for training and testing purposes. Subsequently, regression models—Decision Tree Regression (DTR), Random Forest Regression (RFR), and Linear Regression (LR)—were created and trained.

The models were then evaluated to determine the best-performing one by testing their prediction errors. This evaluation was crucial in selecting the model to integrate into the IoT system for accurate water supply control in mushroom cultivation. The comparison between actual and predicted values across all three models is illustrated in Figure 9, and their performance metrics are summarized in Table 3.

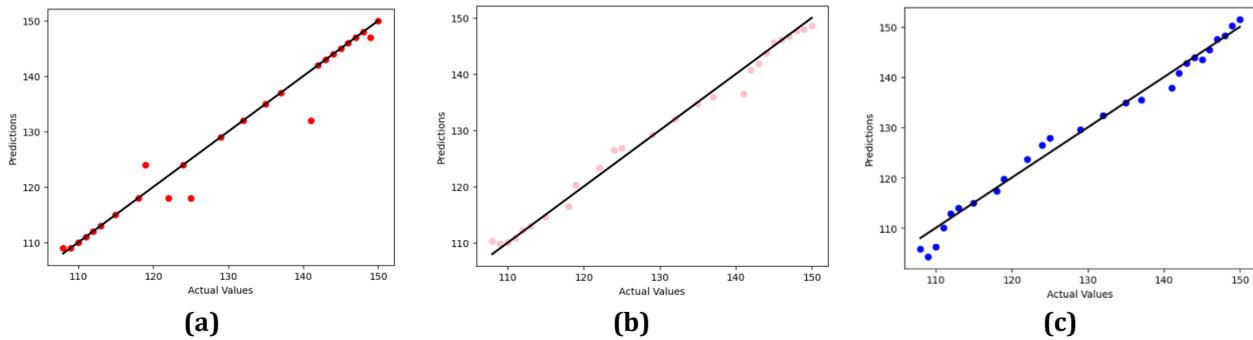


Fig. 10 Actual vs predicted for regression models (a) DTR; (b) RFR; (c) LR

Figure 10 depicts the comparison between the actual values and the predicted values of the three machine learning models used to forecast water usage for bamboo mushroom cultivation. The predictions of all three models showed that the Decision Tree (DT) method had higher variability in predicting water usage compared to actual recorded values, especially in moderate water usage ranges. Conversely, random forest regression (RFR) and linear regression (LR) showed closer alignment between predicted and actual values. The reason linear regression performed better than other models is that, as seen in Figure 9, the correlation values of the relevant parameters show that most of the data exhibit a strong linear relationship. In particular, the independent variables, day and soil moisture, have high correlations with the dependent variable (water) at 0.99 and -0.82, respectively. However, a comparison with more advanced models might reveal even higher accuracy. However, the advantage of the linear regression model for a low-cost IoT system is its direct coding into the microcontroller, eliminating the need for cloud-based processing.

To ensure model accuracy, the study evaluated their performance using RMSE, MAE, and R^2 metrics. Linear regression emerged as the most accurate model, achieving an MAE of 1.65 and an RMSE of 1.84 when implemented on Google's Colab software. In contrast, decision tree regression showed poorer performance, with an MAE of 3.16 and an RMSE of 3.81. These findings underscore linear regression as the optimal choice for precise water supply prediction in bamboo mushroom farming. Detailed comparison results are presented in Table 3. By examining past data and identifying trends and patterns, linear regression can produce accurate forecasts of the

water supply. The following is the derivation of the regression model in the Bamboo Mushroom Farm Case. the dependent variable y' is the liters of water used during each planting period. There are four independent variables, x_1 is the order of the dates, x_2 is the greenhouse's temperature (degrees Celsius), x_3 is the relative humidity in greenhouses (%), and x_4 is the soil moisture (%). The regression equation obtained by analyzing with machine learning techniques via using Google's Colab software, shown as Equation (7). It represents the model equation for predicting the water usage in bamboo mushroom cultivation. This equation was derived through training and testing various environmental variables until achieving optimal performance. In the next section, researchers will implement this equation into a program for the IoT system to make precise decisions regarding water distribution. Subsequently, they will conduct another round of bamboo mushroom cultivation to test and assess the system's operational efficiency.

Table 3 Evaluation results of training dataset

Model	MAE	RMSE	R^2
Decision Tree Regression	3.16	3.81	0.92
Random Forest Regression	2.01	2.37	0.97
Linear Regression	1.65	1.84	0.98

$$y' = -33.63 + 1.02x_1 + 2.23x_2 + 1.41x_3 - 0.74x_4 \quad (7)$$

3.3 The Real-world Implementation

This section contains the findings from the tests conducted to assess how well the bamboo mushroom's irrigation system is working. Fig. 11 illustrates the real-world deployment of the suggested system and presents the farm prototype's elements. Fig. 12 shows the variation in environmental data over 26 days. The temperature during the day averaged 29 °C, and while the average humidity was 92%, the soil moisture levels remained within the range of 50%–78%.

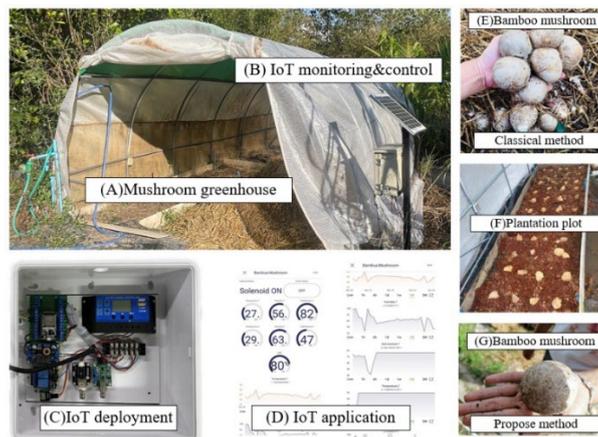


Fig. 11 Real-world deployment

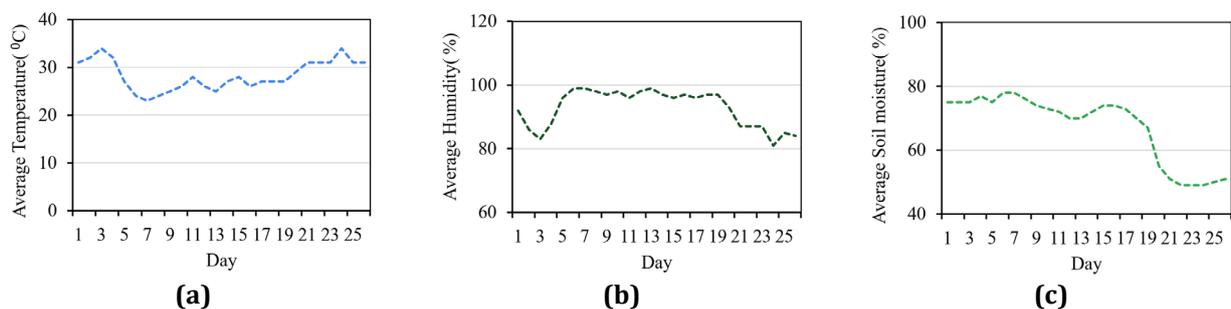


Fig. 12 Daily average environment in bamboo mushroom farm (a) Real-time temperature; (b) Real-time humidity; (c) Real-time soil moisture

In order to see the performance of the proposed technique in the bamboo mushroom farm, the comparison of mushroom production between the traditional process (arbitrary control of the watering by farmers) and the proposed process (automatically controlled by the IoT system) is investigated. The results in Fig. 13 indicate that the proposed method produced mushrooms with an average diameter of 70 mm, which is larger than the traditional method that yielded a size of 61 mm. The primary reason for this improvement is the use of linear regression in deciding the water supply for bamboo mushrooms. This ensures that the moisture levels and environmental conditions closely match the actual needs of the mushrooms, leading to better growth compared to the traditional method. Additionally, watering based on the mushrooms' specific requirements helps conserve water more effectively than using a preset automatic watering schedule.

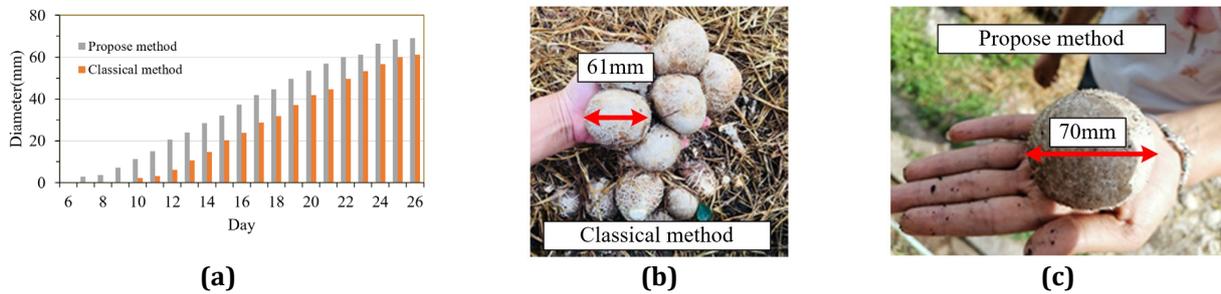


Fig. 13 Daily average environment in bamboo mushroom farm (a) Day vs diameter of mushroom; (b) Traditional method; (c) Propose method

4. Discussion

This study employs a regression model with IoT systems to forecast the appropriate irrigation of bamboo mushrooms. The regression model is selected because it generates results in the form of a linear equation, where the initial variable (x) directly influences the dependent variable (y). Moreover, its low cost and reduced cloud computing further contribute to its appeal. However, other methods, such as DT and K-NN, also yield effective results, but both are suitable for use with Python-enabled control devices such as Raspberry Pi boards, which have complex and expensive implementation procedures. Compared to ESP32 boards, most of these methods require cloud-level computing. Therefore, future applications should be considered according to the suitability and efficiency of that work.

The study discusses the potential of innovative smart water management to significantly enhance farming practices. The study focuses on recording the environmental conditions that influence bamboo mushrooms' growth. The operation principle of IoT technology starts with gathering data from sensors and recording the results in the database. The result implies that farmers increased their daily water intake from day 1 (108 liters) to day 26 (150 liters), with a correlation of 0.99 between the number of days and watering. This is because mushrooms have a high-water self-retention rate. This result concurred with the work of Chukwu S.C. et al. [26], which states that the water content of mushrooms is about 90–95% of body weight. Considering Figure 4, the initial evaluation dataset shows the effectiveness of the method in forecasting the water supply for bamboo mushrooms. The linear regression model's accuracy depends on the degree of alignment between the data points and the ideal regression line. The data points align well with the regression line, indicating linear regression's suitability for predicting bamboo mushroom water supply. This analysis suggests that the linear regression model can effectively predict the appropriate amount of water needed for irrigating bamboo mushrooms. This is consistent with the theory that linear regression can handle data with linear relationships effectively [11],[27]. Therefore, this study supports the application of the linear regression model for forecasting water supply in the bamboo mushroom cultivation process.

Table 2 demonstrates that, when executed on Google's Colab software, linear regression outperforms other regression models in error evaluation metrics, specifically RMSE and MAE. It achieves minimal average error and variability, whereas decision tree regression has poor performance. This is because decision tree regression is good for modeling non-linear relationships between output variables and input features [27], while the data exhibit linear relationships. For random forest regression, it is effective in the middle range because it is a supervised learning technique that combines multiple decision trees and can be used for regression problems. This result concurred with the work of Kaushlendra [14]. However, each of the three models displays a different error value.

This study aims to integrate the linear regression model into the ESP32 board, as shown in Fig. 1. This control technique is consistent with the research of Firdaus et al. [28], which introduced the IoT and fuzzy logic integration for improved substrate environment management in mushroom cultivation. As a result, the control of watering the mushrooms are more precise. Moreover, the results of recording environmental data with the IoT system

show that the average humidity is 92% and the temperature suitable for bamboo mushrooms averages 29 degrees Celsius. The results are consistent with those of Md. Taiqul Islam [29], who studied the characteristics of indoor mushroom cultivation of grey oysters at different stages of humidifying treatment. This humidity will result in an increase in the weight of the mushrooms. Finally, automatic watering based on weather conditions and the decision of the regression model results in mushrooms that are 9 mm larger than those obtained from the classical method, which is equivalent to a weight increase of 46.67%. Consequently, the successful cultivation of bamboo mushrooms and the efficient increase in production for farmers are the results of the method of determining the appropriate watering of bamboo mushrooms, which is based on a regression algorithm and an IoT system.

5. Conclusions

The study explores an IoT-integrated machine learning algorithm for precision watering in bamboo mushroom farming. The system establishes a baseline for water management using the farmers' expertise and traditional practices. IoT devices monitor environmental parameters like temperature, humidity, soil moisture, and water usage. This data is used to develop efficient and automated watering solutions. Next, a machine learning algorithm is used to generate a regression model that predicts optimal water usage based on temperature, humidity, and soil moisture. The best-performing model is then integrated into the IoT systems to automate and optimize the watering process. This ensures that real-time data collected by IoT devices is effectively utilized to maintain optimal growing conditions. The integration of the regression model will lead to more efficient water management, improved crop yields, and reduced labor costs. The following is a summary of the findings: (1) Design the Internet of Things. For recording the resulting environmental values. We found that the average temperature was 29 degrees Celsius. In terms of humidity value, the average moisture value is 92%, and in terms of soil moisture value, it has an average moisture value of 70%. (2) Google's Colab software for regression analysis was used to determine the optimal watering schedule for bamboo mushroom growth, revealing temperature, soil moisture, and relative humidity as the primary factors affecting water content. (3) The study on bamboo mushroom cultivation revealed that the average diameter of the mushrooms exceeded the yield from the classical method, and their volume increased by 46.67%. Moreover, the proposed method saved 22% of the water usage. Therefore, this research's methodology successfully integrates machine learning approaches into smart farming at the device level.

Acknowledgement

The authors would like to thank the Faculty of Engineering, Rajamangala University of Technology Lanna for supporting this project.

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: **study conception and design:** Pinit Nuangpirom, Parida Jewpanya; **data collection:** Pinit Nuangpirom, Parida Jewpanya; **analysis and interpretation of results:** Pinit Nuangpirom, Parida Jewpanya; **draft manuscript preparation:** Pinit Nuangpirom, Siwasit Pitjamit, Parida Jewpanya, Nopadon Maneetien. All authors reviewed the results and approved the final version of the manuscript.

References

- [1] Ruksiriwanich, W., Khantham, C., Linsaenkart, P., Chaitep, T., Rachtanapun, P., Jantanasakulwong, K., ... & Sringarm, K. (2022) Anti-inflammation of bioactive compounds from ethanolic extracts of edible bamboo mushroom (*Dictyophora indusiata*) as functional health promoting food ingredients, *International Journal of Food Science & Technology*, 57(1), 110-122, <https://doi.org/10.1111/ijfs.14878>
- [2] Mazumder, M., Roy, S., & Sarkar, A. K. (2022) Pharmacological and therapeutic value of bamboo mushroom *Phallus indusiatus* (Agaricomycetes), *Italian Journal of Mycology*, 51, 47-57, <https://doi.org/10.6092/issn.2531-7342/12971>
- [3] Cabral, T. S., Silva, B. D., Martín, M. P., Clement, C. R., Hosaka, K., & Baseia, I. G. (2019) Behind the veil—exploring the diversity in *Phallus indusiatus* sl (Phallomycetidae, Basidiomycota), *MycoKeys*, 58, 103-127, <https://doi.org/10.3897/mycokeys.58.36539>
- [4] Chen, C., Wang, J., Li, Q., Fu, R., Jin, X., Huang, W., & Lu, D. (2020) Mitogenomes of two phallus mushroom species reveal gene rearrangement, intron dynamics, and basidiomycete phylogeny, *Frontiers in Microbiology*, 11, 1-14, <https://doi.org/10.3389/fmicb.2020.01346>

- [5] Bersani, C., Ruggiero, C., Sacile, R., Soussi, A., & Zero, E. (2022) Internet of things approaches for monitoring and control of smart greenhouses in industry 4.0, *Energies*, 15(10), 1-30, <https://doi.org/10.3390/en15103847>
- [6] Rayhana, R., Xiao, G., & Liu, Z. (2020) Internet of things empowered smart greenhouse farming, *IEEE Journal of Radio Frequency Identification*, 4(3), 195-211, <https://doi.org/10.1109/JRFID.2020.3013673>
- [7] Paradiso, R., & Proietti, S. (2022) Light-quality manipulation to control plant growth and photomorphogenesis in greenhouse horticulture: The state of the art and the opportunities of modern LED systems, *Journal of Plant Growth Regulation*, 41(2), 742-780, <https://doi.org/10.1007/s00344-021-10337-4>
- [8] Ostergaard, P. A., Duic, N., Noorollahi, Y., Mikulcic, H., & Kalogirou, S. (2020) Sustainable development using renewable energy technology, *Renewable energy*, 146, 2430-2437. <https://doi.org/10.1016/j.renene.2019.08.094>
- [9] Maraveas, C. (2022) Incorporating artificial intelligence technology in smart greenhouses: Current State of the Art, *Applied Sciences*, 13(1), 1-35, <https://doi.org/10.3390/app13010001>
- [10] Wickramasinghe, I. (2022) Applications of Machine Learning in cricket: A systematic review, *Machine Learning with Applications*, 10, 100435, <https://doi.org/10.1016/j.mlwa.2022.100435>
- [11] Maulud, D., & Abdulazeez, A. M. (2020) A review on linear regression comprehensive in machine learning, *Journal of Applied Science and Technology Trends*, 1(2), 140-147, <https://doi.org/10.38094/jastt1224>
- [12] Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019) A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models, *Journal of clinical epidemiology*, 110, 12-22, <https://doi.org/10.1016/j.jclinepi.2019.02.004>
- [13] Charbuty, B., & Abdulazeez, A. (2021) Classification based on decision tree algorithm for machine learning, *Journal of Applied Science and Technology Trends*, 2(01), 20-28, <https://doi.org/10.38094/jastt20165>
- [14] Kaushlendra, S., & Vedant, S. (2023) Symptoms Diagnosis Using Machine Learning Model Random Forest, *International Journal for Science Technology and Engineering*, 11(5), 2004-2007, <https://doi.org/10.9790/1816-110520042007>
- [15] Yahyaoui, A., & Yumusak, N. (2018) Decision support system based on the support vector machines and the adaptive support vector machines algorithm for solving chest disease diagnosis problems, *Biomedical Research*, 29(7), 1-6, <https://doi.org/10.4066/biomedicalresearch.29-18-719>
- [16] Cunningham, P., & Delany, S. J. (2021) K-nearest neighbour classifiers-a tutorial, *ACM computing surveys (CSUR)*, 54(6), 1-25, <https://doi.org/10.1145/3459665>
- [17] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002) An efficient k-means clustering algorithm: Analysis and implementation, *IEEE transactions on pattern analysis and machine intelligence*, 24(7), 881-892. <https://doi.org/10.1109/TPAMI.2002.1017616>
- [18] Liu, Y., Liu, S., Wang, Y., Lombardi, F., & Han, J. (2020) A survey of stochastic computing neural networks for machine learning applications, *IEEE Transactions on Neural Networks and Learning Systems*, 32(7), 2809-2824, <https://doi.org/10.1109/TNNLS.2020.2980083>
- [19] Samie, F., Bauer, L., & Henkel, J. (2019) From cloud down to things: An overview of machine learning in internet of things, *IEEE Internet of Things Journal*, 6(3), 4921-4934. <https://doi.org/10.1109/IIOT.2019.2901849>
- [20] Mahdavinnejad, M. S., Rezvan, M., Barekatin, M., Adibi, P., Barnaghi, P., & Sheth, A. P. (2018) Machine learning for Internet of Things data analysis: A survey, *Digital Communications and Networks*, 4(3), 161-175, <https://doi.org/10.1016/j.dcan.2017.10.002>
- [21] Mohammad, A. S., & Pradhan, M. R. (2021) Machine learning with big data analytics for cloud security, *Computers & Electrical Engineering*, 96, 107527, <https://doi.org/10.1016/j.compeleceng.2021.107527>
- [22] Tahsien, S. M., Karimipour, H., & Spachos, P. (2020) Machine learning based solutions for security of Internet of Things (IoT): A survey, *Journal of Network and Computer Applications*, 161, 102630, <https://doi.org/10.1016/j.jnca.2020.102630>
- [23] Tran, M. Q., Doan, H. P., Vu, V. Q., & Vu, L. T. (2023) Machine learning and IoT-based approach for tool condition monitoring: A review and future prospects, *Measurement*, 207, 112351, <https://doi.org/10.1016/j.measurement.2022.112351>
- [24] Messaoud, S., Bradai, A., Bukhari, S. H. R., Quang, P. T. A., Ahmed, O. B., & Atri, M. (2020) A survey on machine learning in Internet of Things: Algorithms, strategies, and applications, *Internet of Things*, 12, 100314, <https://doi.org/10.1016/j.iot.2020.100314>
- [25] Azman, N. N. ., Mohd Addi, M., & Al-Abed, A. (2023). Development of an IoT Embedded Wearable Device with Non-Contact Temperature Detector for Early Detection of Fever. *International Journal of Integrated Engineering*, 15(3), 15-27. <https://publisher.uthm.edu.my/ojs/index.php/ijie/article/view/12810>
- [26] Chukwu, S. C., Ekesiobi, S. I., Nwankwo, I. A., & Arah, E. C. (2022) Effects of watering frequency and watering method on the growth and yield of oyster mushroom (*pleurotus ostreatus*), *Nigerian Agricultural Journal*, 53(1), 297-302, <https://doi.org/10.52331/naj.v53i1.289>

- [27] Shorouq, A. E., Salloum, S. A., Shaalan, K., & Abdallah, S. (2023) Comparative Analysis Study for Air Quality Prediction in Smart Cities Using Regression Techniques, *IEEE Access*, 11, 115140-115149, <https://doi.org/10.1109/ACCESS.2023.3292004>
- [28] Firdaus, I., Balakrishnan, V., Mohd, Z. S., Abd-Rahman, S. S., Manaf, A. A., Samingan, M. I., & Ibrahim, M. S. (2024) IoT and fuzzy logic integration for improved substrate environment management in mushroom cultivation, *Smart Agricultural Technology*, 7, 100427, <https://doi.org/10.1016/j.atech.2023.100427>
- [29] Md. Tariqul, I., Zarina, Z., Nasrul, H., & Mohd Azlan, B. M. I. (2016) Characteristics of Indoor Mushroom Cultivation of Grey Oyster (*Pleurotus pulmonarius*) by Different Stages of Humidifying Treatment, *World Applied Sciences Journal*, 34 (8), 1066-1075, <https://doi.org/10.5829/idosi.wasj.2016.1066.1075>
- [30] Nila, F. S. ., Tan, W.-H., Ooi, C. P., & Tan, Y.-F. . (2024). IoT-Based Embedded System for Streamlined Thermal Comfort Data Collection in Buildings. *International Journal of Integrated Engineering*, 16(3), 78-91. <https://publisher.uthm.edu.my/ojs/index.php/ijie/article/view/15892>