International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Unsupervised Fertigation and Machine Learning for Crop Vegetation Parameter Analysis

Mohd Izzat Mohd Rahman[1], Mohd Azraai Mohd Razman*[1], Anwar PP Abdul Majeed[2], Muhammad Nur Aiman Shapiee[1], Muhammad Amirul Abdullah[1], Rabiu Muazu Musa[3]

Submitted: 28/04/2023 **Revised**: 29/06/2023 **Accepted**: 09/07/2023

Abstract: This study proposes an IoT-based smart irrigation management system that can optimize water-resource utilization in a smart agricultural system. The system uses unsupervised learning-based clustering to predict the irrigation needs of a field based on the ground parameters sensed by automated monitoring devices. These parameters include soil moisture, light intensity, temperature, and humidity. The system extracts feature such as the maximum, minimum, mean, and standard deviation of four soil moisture sensors from the primary dataset of plants. Then, it applies lag features to enhance the accuracy of the classification model. The system uploads the dataset of 108 features to the Orange GUI and performs k-means clustering to assign cluster labels to the data as meta-attributes in a new dataset. The study evaluates the system using a month's worth of data and demonstrates its functionality and effectiveness. The system employs machine learning techniques such as Random Forest, Neural Network, and kNN, which achieve 100%, 99.9%, and 99.8% accuracy respectively.

Keywords: *Machine Learning; Feature Extraction; Classification; Fertigation System; Chili Plant*

1. Introduction

Plant care is a strategy and activity for maintaining the health and appearance of plants. Water and soil continue to be vital sources of life for all plants, assisting the photosynthetic process, particularly in diverse plants. Living systems of fertilization depend on nutrients from the water and soil. A fertigation system is a process of using fertilizer clarifications with irrigation water, commonly through a micro-sprinkler or a drip system [1]. Insufficient irrigation or excessive watering have been detrimental to production and nature, making automation of the rejuvenating greenery system important.

Jordan's Hashemite Kingdom is studying the Mediterranean region's climate via research. Snowfalls occur at short intervals on the majority of the kingdom's mountain highlands in the north, center, and south and are quite heavy and sometimes collected [2]. Whereas the Equatorial area surrounding Malaysia has a climate defined by year-round high average temperatures and high monthly precipitation. Malaysia's climate has a year-round high flat temperature line (above 25°C) and rainfall of more than 350mm in December, but less than 350mm for the remainder of the month as shown in Fig. 1. The irregularity of rainfall distribution in Malaysia throughout the year, the quality of the soil, the amount of water available in the region, and technological advancements in all spheres of our lives provide us with a sense of assisting farmers in watering the plants without time or effort by maintaining excellent plant production and providing irrigation water amounts [3].

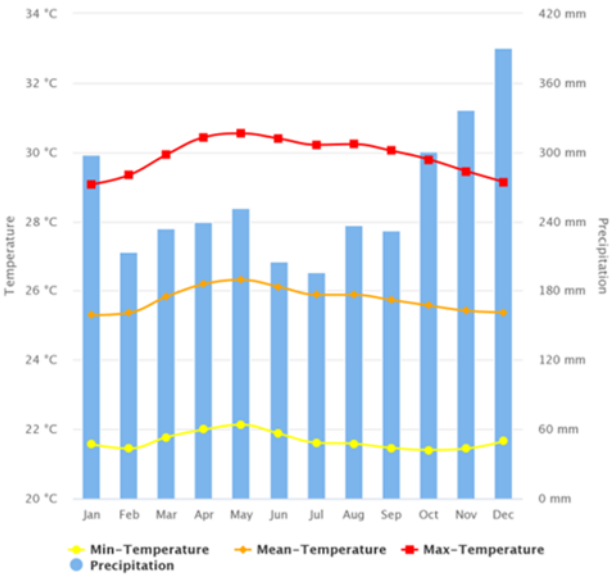


Fig. 1. Monthly Climatology of Min-Temperature, Mean-Temperature, Max-Temperature & Precipitation 1991-2020 Malaysia

As we recognize that the highest quality plants need, the first issue in this project is how to gather data from the set of plants that we want to apply. The sensors are adequate for them to interpret the data correctly, allowing us to go on to

1 Manufacturing and Mechatronic Engineering Technology (FTKPM), Universiti Malaysia Pahang, 26600, Pekan, Pahang, Malaysia
2 School of Robotics, XJTLU Entrepreneur College (Taicang), Xi'an Jiaotong-Liverpool University, 215127, Taicang, PR China
3 Center for Fundamental and Continuing Education, Department of Credited Co-curriculum, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Terengganu, Malaysia
** Corresponding Author Email: mohdazraai@ump.edu.my*

the next step of improving the plant's quality and healthiness [4]. The next phase will be data clustering to categorize the data as healthy or unhealthy plants. Another issue is classifying the machine learning model using the machine learning technique. Within the dataset received and extracted, we are not sure which model is the best to obtain the highest percentage of accuracy and prediction on getting the best model in this project.

The objective of this project is to develop a monitoring system to collect specific data from the plant based on the sensor applied. To identify classes of the plant growth rate through the feature extracted and the clustering of unsupervised learning and last is the formulation of machine learning models and classification performances to predict the growth based on the features extracted.

In this paper, an automatic fertigation system using unsupervised learning would be created to help water the plant efficiently and give a system that can reduce the number of workers needed for the watering plant. The remainder of this paper is organized as follows: Section 2 briefly presents a comprehensive assessment of the literature. Section 3 discusses the methods used in the study. Section 4 provides the experimental settings and discusses the experimental results. Section 5 summarises the article in the study.

2. Related Work

Before starting with the experiment, it is important to understand the model and technique of machine learning that will be applied to its data. Prior to proceeding with the technique utilized to conduct the experiment, data collection is the initial step. Various studies show the utilization of machine learning has been applied such as agriculture, sports and transportation [5]–[9]. To determine the health and disease condition of a plant, machine learning methods were often used to enhance performance and reliability. Image processing is used in the majority of research because it can detect changes. Deep Convolutional Neural Networks (CNN) are specialized Feedforward Neural Networks that often use images as input and are used for categorization [10]. Convolution layers extract image features automatically rather than collecting them straight to send them into the network. CNN uses image processing filters to scale the input images to a preset size and convolve them. Fig. 2, depicts a deep CNN model that takes a 128x128 pixel picture and adds 32 convolution filters to it.

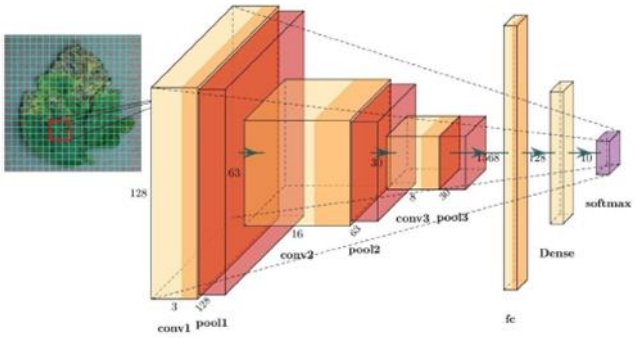


Fig. 2. Proposed CNN Model

The study [11] combined the Soil Moisture Differences (SMD) method with the SVR model. The methods used for feature extraction were k-mean, SVR, and an upgraded version of SVR, which resulted in SVR+k-mean. In the experiment, sensors for soil moisture, air temperature, humidity, UV, and soil temperature were used. In comparison to SVR, SVR+k-mean exhibited the greatest accuracy and lowest MSE during the whole investigation.

This experiment using rice yield was conducted by [12] utilizing Machine Learning (ML) and Multiple Linear Regression models (MLR). Numerous methodologies, including partial correlation analyses, meteorological and phenology factors, and the dependent variable, were utilized. In machine learning, numerous data types were utilized, including sowing, emergence, and three leaves. SVM was found to have the highest RMSE(R2) accuracy with a value of 737(0.33).

In the experiment conducted by [13], sensors such as the YL-60, the SEN13322 moisture sensor, the DS18S20, and the SHT10 (Temperature and Humidity) were utilized. In ANN and RF models, root means square error, mean absolute error, and Pearson's correlation coefficient is used to evaluate performance (R). Due to the low cost of the sensor, more dense deployment nodes were utilized to measure the field capacity profile with greater precision.

The paper [14] analyzes a report proposing an autonomous irrigation system for the Southern Jordan Valley using an IRIS mote and an MTS420 sensor board. The model for the intelligent autonomous irrigation system was developed using decision trees (DT) with an accuracy of 97.86%. In contrast, Anantnag was labelled as a high-risk zone for the apple scab forecast, whereas Srinagar was categorized as a low-risk zone.

For the study [15], a smart Farming System with three-layered architectures was developed (Agriculture Perception, Edge Computing, and Data Analytics). 50g samples of 5x1x1cm and 5x1x0.5cm were sliced for each drying procedure. For classifying outliers, the researcher suggests utilizing the Isolation Forest method (which has a 99% accuracy) and the Cubic Spline approach (RMSE lower than 0.085).

Collecting all the data from the literature study is a diverse sort of data, and machine learning was used to improve the accuracy of the forecasting of the plants and vegetables. This chapter covered determining the plant's health and development rate, as well as the optimal performance depending on the utilized models. This section concludes with a methodological technique for determining the growth rate of a plant utilizing an irrigation system and sensors, followed by a classification model based on machine learning. To find the significant value and features from the analysis of plants, feature extraction is required. Feature selection is described as selecting a distinct feature that will restrict the extraction feature and lower the complexity and computational time of machine learning. This strategy is always an excellent choice because it assists in determining the model's performance and saves time.

3. Methodology

3.1. Water Irrigation System with Sensor Applied

The current study was designed to extract the value data from the plant through the combination of an automated greenery system for rejuvenation and sensors attached to the plant using the Machine Learning Technique. Nevertheless, the system should collaborate with a workstation that reads and computes the machine learning algorithm based on the growth rate of the plants and consolidates data amongst the components. Fig. 3 depicts the modified automated rejuvenation greenery system, which comprises the applied sensors, the working station to compute the necessary factor of sensors data read, and the software used to determine the goals that must be attained.

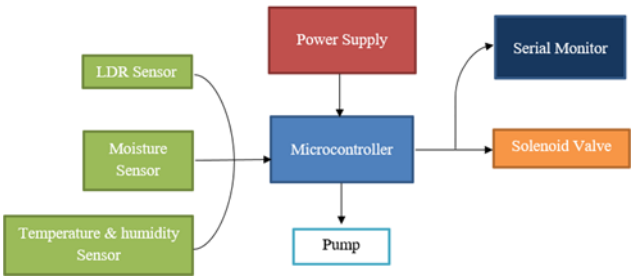


Fig. 3. Schematic Diagram of Automated Rejuvenation Greenery System

3.2. Experimental Setup

The plants' data were collected in the IMAM's laboratory at Universiti Malaysia Pahang. A total of 4 plants were carried out and split into 2 plants with watering and the other 2 plants without watering. A 9.5-litre water tank is filled with water to serve as a standby for the water irrigation system until it has to be refilled. An LDR sensor, a temperature and humidity sensor, and four soil moisture sensors were installed. Two plants get watered in the morning and evening, while the others two are not. Each soil moisture sensor is installed on the soil of each plant to assess the plant's moisture level. Each plant's data is gathered. Fig. 4(a) shows both chilli plants are watered, while Fig. 4(b) shows both plants are non-watered. All four chilli plants are placed in a semi-outdoor environment.

(a) Water plant *(b) Non-water plant*

Fig. 4. Water (a) and Non-water (b) of the Chili Plants in a Semi-Outdoor Environment

3.3. Machine Learning Technique

To get into a deeper depth of the flow in this section, Fig. 5 illustrates the process of the ML techniques that had executed in this study. The processes that will be intentionally are solely emphasized on the ML methodology, analysis, and classification of the plant growth rate. The progression of the Machine Learning technique is split into two junctures, the first interval is necessarily focused on the data pre-processing which later to be extracted to feature extraction.

For the Rodríguez [13] study, a smart Farming System with three-layered architectures was developed (Agriculture Perception, Edge Computing, and Data Analytics). 50g samples of 5x1x1cm and 5x1x0.5cm were sliced for each drying procedure. For classifying outliers, the researcher suggests utilizing the Isolation Forest method (which has a 99% accuracy) and the Cubic Spline approach (RMSE lower than 0.085).

The data pre-processing part is to organize the clustered results of the signal according to the type of data taken from each sensor. The organized data will then proceed with the feature extraction to gain more extraction. Feature extraction is decent enough to proceed with clustering and classification, but increasing more features helps to improve the accuracy to obtain the best features. The lag feature is a name for a variable that contains data from preceding time steps. Applying the lag feature will help increase double the features that have been obtained from extraction.

The dataset taken from the experiment contains a huge number of data extracted from the sensors. The window sample is made into this dataset and resized dataset into 5 minutes each. The result of window size sampling is from 5

seconds to 5 minutes, there are not a lot of differences. Thus, we proceed by using the 5 minutes data sampling. For each temporal feature, the mean, standard derivation, maximum, and minimum values are applied. When the 54 retrieved features are combined with the lag features, a total of 108 features were created. The greater the number of characteristics, the more accurate the result would be.

Random Forest, kNN, and Neural Network will be used and tested to determine whether the proposed methods have met the hypothesis stated. Next, the second phase of ML methodology will broaden the analysis sector including event identification, feature analysis and selection, model comparison, and lastly, the important element of hyperparameter tuning. The expectation from both phases is that it can produce the advancement of the proposed methodology. This subsection will systematically express each of the methods from pre-processing, classification, the model used, and the performance evaluation.

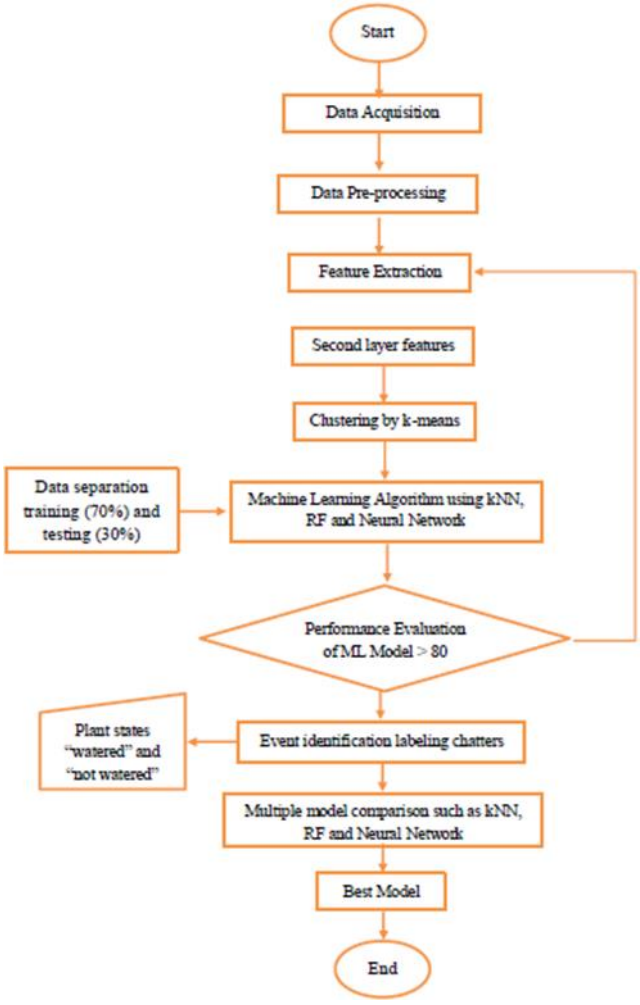


Fig. 5. Machine Learning Techniques Flowchart

3.4. Classification Performance Evaluation

The classification model technique was derived from the theoretical concept of Machine Learning, which entails automated learning through trials based on a specific example. The benefit of the ML model above other traditional methods that need complexity for modelling is the improved results it may provide. The acquired data's history may be used to build a categorized impression and to cluster the chilli plant dataset we possess. In this example, many distinct kinds of classifier models will be used, and the theoretical ideas underlying each will be explored. According to the norm separation of the datasets was permitted by separating the data into 70% training data and 30% test data for assessment and evaluation. The performance of the classifiers is assessed based on their classification accuracy, precision, F1 score, recall, and confusion matrix. The formula of each performance is as shown in Fig. 6, Equation (1), Equation (2), Equation (3), and Equation (4).

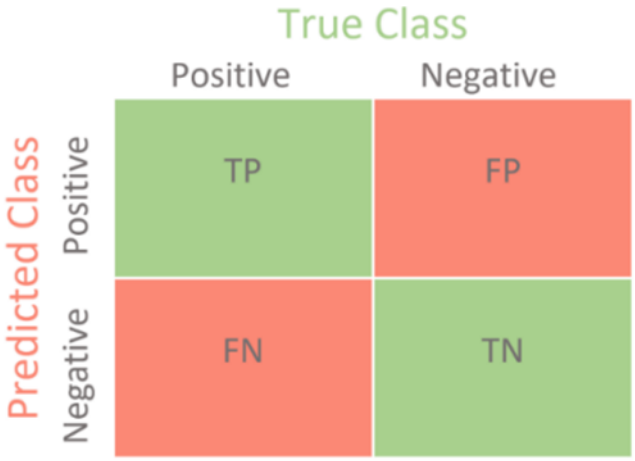


Fig. 6. Confusion Matrix for Binary Classification

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1\ Score = \frac{2(Precision * Recall)}{Precision\ +\ Recall} \quad (4)$$

4. Experimental Results

This section will show each step starting with the original dataset that was collected from chill plants using the various types of sensors. It will also specify the number of features that will be obtained after applying feature extraction and lag features to the original dataset using the data table in Orange GUI. Additionally, all the classification models that were used in the GUI will appear with their accuracy results after training and testing the dataset, with the great features in Orange indicating how we can obtain the correct number for misclassified data using the scatter plot for this dataset

and the confusion matrix. Additionally, there are other characteristics that will be used to aid in our understanding of chill plants, including silhouette plots, rank, hierarchical clustering, data sampling, and box plots. All of these elements will be included in the Orange GUI, as seen in Fig. 7.

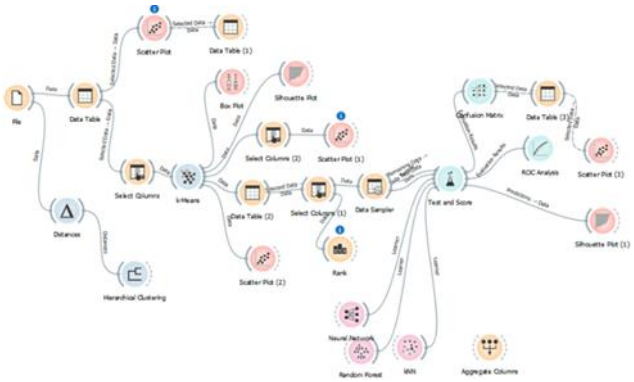
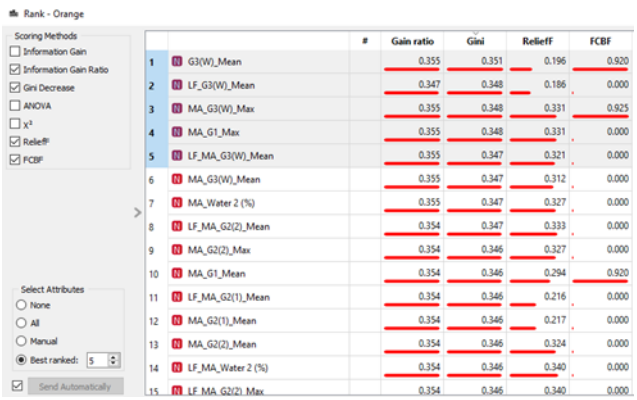


Fig. 7. The Workflow of the Project of GUI in Orange Software

After uploading the dataset file in the GUI, it is necessary to select the features that should be clustered and ignore columns that cannot be included in the clustering process, such as the time and date column. Additionally, it may allow for the selection of the number of columns to cluster and the exclusion of features that are not yet ready to be included in the clustering stage. Thus, there are a total of 1147 rows and 108 features were applied in the clustering process.

4.1. Clustering by k-means

K-means clustering is a widely used unsupervised machine learning approach. The k-means algorithm determines the location of k centroids and then assigns each data point to the cluster with the fewest centroids. The k-means method is used to cluster the data and store the cluster label as a meta-attribute in a new dataset. Additionally, the silhouette scores for each k's clustering results were displayed. Two clusters out of a total of 14 clusters are optimum for this project's data because 2 have the greatest silhouette score of 0.479 in comparison to the other clusters.



#		Gain ratio	Gini	ReliefF	FCBF
1	G3(W)_Mean	0.355	0.351	0.196	0.920
2	LF_G3(W)_Mean	0.347	0.348	0.186	0.000
3	MA_G3(W)_Max	0.355	0.348	0.331	0.925
4	MA_G1_Max	0.355	0.348	0.331	0.000
5	LF_MA_G3(W)_Mean	0.355	0.347	0.321	0.000
6	MA_G3(W)_Mean	0.355	0.347	0.312	0.000
7	MA_Water 2 (%)	0.355	0.347	0.327	0.000
8	LF_MA_G2(2)_Mean	0.354	0.347	0.333	0.000
9	MA_G2(2)_Max	0.354	0.346	0.327	0.000
10	MA_G1_Mean	0.354	0.346	0.294	0.920
11	LF_MA_G2(1)_Mean	0.354	0.346	0.216	0.000
12	MA_G2(1)_Mean	0.354	0.346	0.217	0.000
13	MA_G2(2)_Mean	0.354	0.346	0.324	0.000
14	LF_MA_Water 2 (%)	0.354	0.346	0.340	0.000
15	LF_MA_G2(2)_Max	0.354	0.346	0.340	0.000

Fig. 8. Ranking of Best Features of the Dataset

After the clustering dataset, it's a good idea to compare the best 5 features as shown in Fig. 8. There are four scoring methods chosen to determine the best 5 features which are information gain ratio, gini decrease, reliefF, and FCBF. The greater the mean decrease accuracy or Gini score, the more significant the variable is in the model. The greater the gain ratio, the more significant the variable is in the model. G3(W)_Mean had the best rank from the total of 108 features compared to LF_G3(W)_Mean. This is because the gain ratio of G3(W)_Mean is higher and the decrease gini is also higher compared to the rest of the features. G3(W)_Mean, LF_G3(W)_Mean, MA_G3(W)_Max, MA_G1_Max, and LF_MA_G3(W)_Mean scores the top 5 ranks out of 108 features, so these 5 features are the importance of the features in the model.

4.2. Events Identification

Fig. 9 shows a graph created with data from watered plant 1 and one of its feature extractions, which is the third mean for this sensor. The mean features of the soil moisture sensor of the watered plant are used to produce the graph, which begins with the first cluster, then moves to the second cluster, and so on, until the end. This indicates that for C2, the soil is dry due to low soil moisture, but for C1, the soil is wet. Thus, we state that C2 Required Water (RW), but C1 is Enough Water (EW).

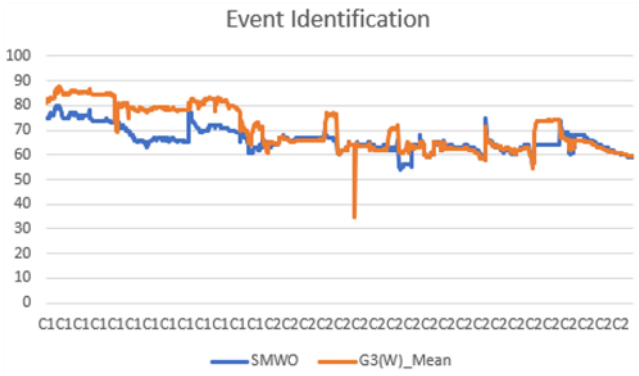


Fig. 9. Event Identification for Soil Monitoring Sensors Water Outside with Mean Feature

4.3. Good versus Bad Separation of Box Plot

MA_G1_Max box plots indicate that the first cluster's lowest value is 89.1538, the middle value is 91.2308, and the highest value is 94.1538. However, the lowest value for the second cluster is 63.0769, the median value is 64.3077, and the highest value is 67. The fact that the centroid of the first cluster is quite far away from the centroid of the second cluster indicates that this dataset's clustering is more organized into two different groups. Fig. 10 shows the good separation of the box plot and demonstrates the significance of the features chosen since the two boxes do not overlap.

While for Fig. 11, we created a box plot that shows bad separation using the Orange GUI to provide more information about the features in this dataset. This box plot for MA_Temperature (Degree Celcius) provides further information about the two clusters for MA_Temperature

(Degree Celcius). The box plots for MA_Temperature (Degree Celcius) indicate that the minimum value is 27.7462, the medium value is 29.1462, and the maximum value is 31.1385. However, the minimum value for the second cluster is 29.2885, the median value is 31.2038, and the maximum value is 32.9385. For this two-cluster box plot, the region between 27 and 33 values is shared by both clusters. The box plot for these two clusters has a lot of similarities because the destines between the two clusters are so close to each other, that the clustering for this feature will be inadequate. The graphic demonstrates that the feature selected is irrelevant due to the overlapping of two boxes.

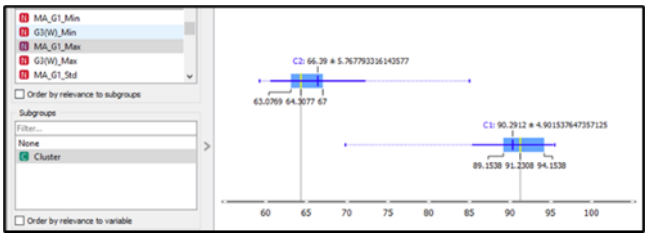


Fig. 10. Example of Good Separation for Box Plot (MA_G1_Max)

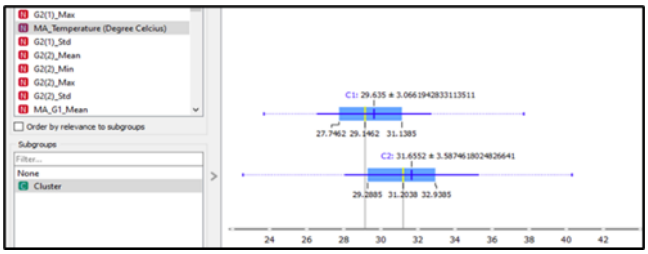


Fig. 11. Example of Good Separation for Box Plot (MA_G1_Temperature)

4.4. Good versus Bad Separation of Scatter Plot

The figure below illustrates the difference in clustering results when 108 features are used against the five best features from the dataset. The best way to get information about how to read or analyse a scatter plot is to examine the scatters of data plots. A good separation occurs when data scatter directly proportionally, are highly organised, and cluster between two groups, as seen in Fig. 12. Thus, these two features selected are the important variables in the model. Noted that there is an outlier in the plot shows that there is a reading error that occurred due to the malfunction of sensors during that time. A poor separation occurs when the data scatter apart between two distinct groups, as seen in Fig. 13. Hence, these two features selected are not important variables in the model.

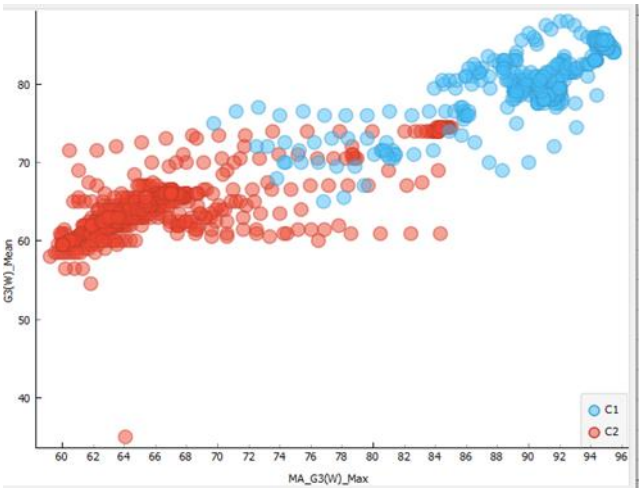


Fig. 12. Scatter Plot for MA_G3(W)_Max vs G3(W)_Mean from the Best 5 Features

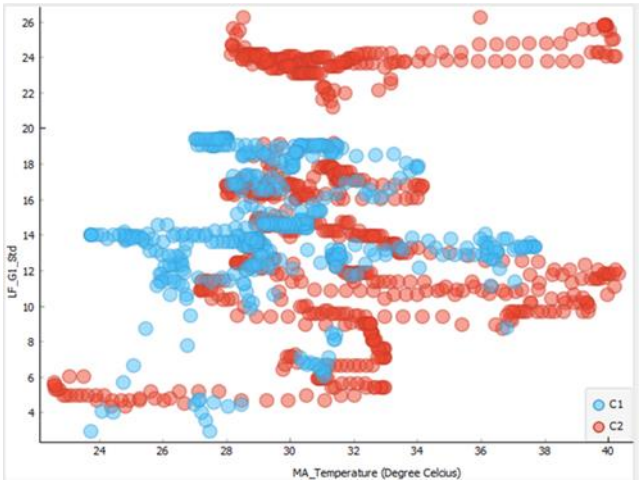


Fig. 13. Scatter Plot for MA_Temperature (Degree Celcius) vs LF_G1_Std Out from the Best 5 Features

4.5. Data Classification

Fig. 14 illustrates how the classification accuracy of Random Forest (RF), kNN, and Neural Networks varies. The most accurate is RF, as the percentage on the test is 99.7% and the percentage on the train is 100%. The second is a Neural Network with a testing set of 99.7% and a training set of 99.9%. The least accurate is kNN, with a testing set of 99.1% and a training set of 99.8%.

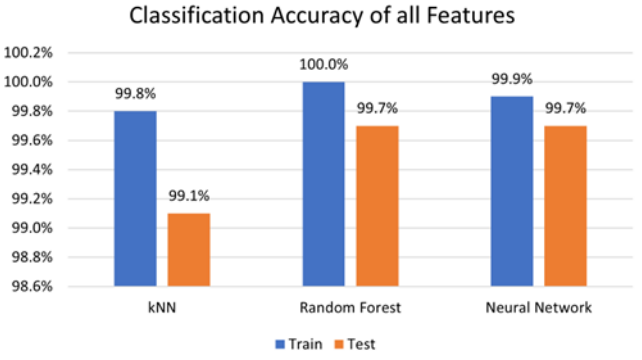


Fig. 14. Classification Accuracy for All Features

From Table 1, we can see that the classifier works best for Random Forest with the highest F1 score in the train set and test set, 1.000 in the training set, and 1.000 in the testing set. The equation of the F1 score as shown in Table 1. had proved that higher precision and recall would produce a higher F1 Score. Thus, Random Forest has the highest precision and recall compared to Neural Networks and kNN. kNN classifier had the lowest precision, recall, and F1 score with 0.998 in the training set, and 0.991 in the testing set.

Model	F1	Precision	Recall
Random Forest	1.000	1.000	1.000
	0.997	0.997	0.997
Neural Network	0.999	0.999	0.999
	0.997	0.997	0.997
kNN	0.998	0.998	0.998
	0.991	0.991	0.991

Train
Test

Table. 1. Performance Evaluation of All Features

For the best 5 features that were chosen, kNN, Random Forest, and Neural Network will be used again. Since this feature is the best feature among the whole features, the algorithm results have a high accuracy level, as shown in the bar chart, but the accuracy for other features still has a high accuracy performance result. The high accuracy percentage in the train and test set was Random Forest 99.8% in the training set and 98.3% in the testing set, as shown in Fig. 15. The second-best performer was kNN, which scored 98.9% for the training set and 96.2% for the testing set. Finally, the Neural Network model had the lowest accuracy results in the training set at 94.8%, but 92.2% in the testing set.

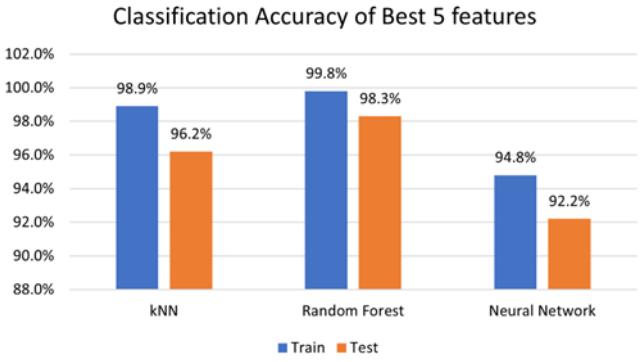


Fig. 15. Classification Accuracy for the Best 5 Features

Table 2 shows the performance evaluation of the best 5 features. From the table, we can see that the classifier works best for Random Forest with the highest F1 score, precision, and recall in the train set and test set, 0.998 in the training set, and 0.983 in the testing set. The Neural Network classifier had the lowest precision, recall, and F1 score with 0.948 in the training set, and 0.921 in the testing set.

Model	F1	Precision	Recall
Random Forest	0.998	0.998	0.998
	0.983	0.983	0.983
Neural Network	0.948	0.948	0.948
	0.921	0.921	0.922
kNN	0.989	0.989	0.989
	0.962	0.962	0.962

Train
Test

Table. 2. Performance Evaluation of the Best 5 Features

4.6. Comparison Between Classification Accuracy of All Features and Classification Accuracy of the Best 5 Features.

Fig. 16 depicts a comparison of the classification accuracy of all features vs the classification accuracy of the best five features in this section. As a result, the classification accuracy of all features was higher than that of the best five features. This demonstrates that the more features included in the data, the more accurate the classification model obtains. Thus, it is important to use all the features to determine the classification accuracy as compared to the best 5 features.

Besides, Random Forest achieves the highest classification accuracy of 99.7% on the testing set, and 100% on the training set for all features. While for the best 5 features, Random Forest also achieves the highest classification accuracy of 98.3% on the testing set and 99.8% on the training set. Thus, we can conclude that Random Forest performs better accuracy compared to kNN and Neural Networks.

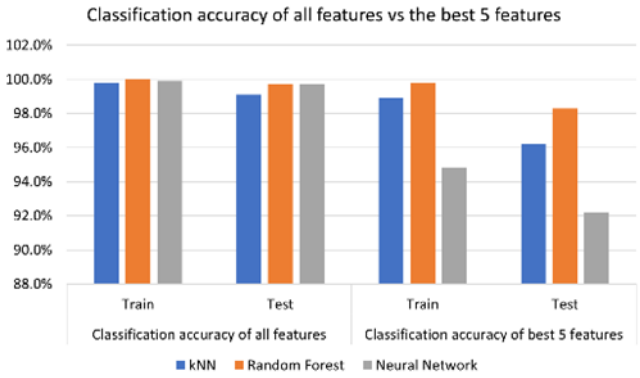


Fig. 16. Classification Accuracy for All Features VS Classification Accuracy of the Best 5 Features

4.7. The Growth Change for the Plants

After examining the results in the various sub-chapters of this chapter, it is clear that for optimum plant development, plants must be well cared for with enough watering, as determined by data on parameters for watered and unwatered plants placed in a semi-outdoor environment. Here is a comparison of the chilli plants from the beginning to the finish of this project's data collecting. The number of leaves on each plant is counted to assess how much it has grown in a semi-outdoor environment. On each plant, the letters A and B stand for both non-watered plants placed in

a semi-outdoor environment. The remaining two plants (C and D) are watered and planted in a semi-outdoor environment. Table 3 summarises the data collected during the course of the trial.

Table. 3. Numbers of Leaves for Chili Plants

Plant	Number of Leaves		
	4/12/2021	*13/12/2021*	*10/1/2021*
A	14	25	42
B	13	21	30
C	15	33	96
D	7	13	82

5. Conclusion

As shown by this project's goals, the overall project objectives have been accomplished. The first objective is to construct and develop a monitoring system and to gather data from the chilli plants using a variety of sensors that can read the data from the plants. The Controllino was used to read and store sensor data on the computer. After constructing the prototype system required to reach the aim, data is collected for the plant to be used for feature extraction and lag features to obtain further features and columns. The last objective was to formulate the machine learning models for the clustering part using K-means clustering, followed by selecting the best features that can be used to get a more accurate classification result. The best model is determined by the one with the greatest classification percentage accuracy among the three classification models. The best model is selected based on its ability to accurately predict the development of plants.

The processes of window sampling, feature extraction, and lag features help in obtaining the best number of features for machine learning. The greater the number of features in a dataset, the more precise the outcome would be. When the plots determined the condition of the plants by separating them into clusters containing enough water and required water for the plant, the outcome was clearly visible. Utilizing the dataset for window sampling prior to feature extraction, the growth rate of each plant was forecasted. At the conclusion of this operation, the goals had been accomplished 100%.

Acknowledgements

The authors would like to thank Universiti Malaysia Pahang for supporting this work via the Research Grant programme (RDU200332).

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] D. Mattos, D. M. Kadyampakeni, A. Q. Oliver, R. M. Boaretto, K. T. Morgan, and J. A. Quaggio, *Soil and nutrition interactions*. Elsevier Inc., 2020. doi: 10.1016/B978-0-12-812163-4.00015-2.

A. H. Blasi, M. A. Abbadi, and R. Al-Huweimel, "Machine Learning Approach for an Automatic Irrigation System in Southern Jordan Valley," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6609–6613, 2021, doi: 10.48084/etasr.3944.

[2] S. Sayari, A. Mahdavi-Meymand, and M. Zounemat-Kermani, "Irrigation water infiltration modeling using machine learning," *Comput Electron Agric*, vol. 180, p. 105921, Jan. 2021, doi: 10.1016/j.compag.2020.105921.

[3] H. Navarro-Hellín, J. Martínez-del-Rincon, R. Domingo-Miguel, F. Soto-Valles, and R. Torres-Sánchez, "A decision support system for managing irrigation in agriculture," *Comput Electron Agric*, vol. 124, pp. 121–131, Jun. 2016, doi: 10.1016/j.compag.2016.04.003.

[4] M. A. Abdullah, M. A. R. Ibrahim, M. N. A. bin Shapiee, M. A. Mohd Razman, R. M. Musa, and A. P. P. Abdul Majeed, "The classification of skateboarding trick manoeuvres through the integration of IMU and machine learning," *Lecture Notes in Mechanical Engineering*, pp. 67–74, 2020, doi: 10.1007/978-981-13-9539-0_7/FIGURES/7.

[5] S. Puteh, N. F. M. Rodzali, M. A. M. Razman, Z. Z. Ibrahim, M. N. A. Shapiee, and M. A. M. Razman, "Features Extraction of Capsicum Frutescens (C.F) NDVI Values using Image Processing," *MEKATRONIKA*, vol. 2, no. 1, pp. 38–46, Jun. 2020, doi: 10.15282/MEKATRONIKA.V2I1.6727.

[6] M. N. A. Shapiee, M. A. R. Ibrahim, M. A. M. Razman, M. A. Abdullah, R. M. Musa, and A. P. P. Abdul Majeed, "The Classification of Skateboarding Tricks by Means of the Integration of Transfer Learning and Machine Learning Models," *Lecture Notes in Electrical Engineering*, vol. 678, pp. 219–226, 2020, doi: 10.1007/978-981-15-6025-5_20/COVER.

[7] N. F. Mohd. Ali, A. F. Mohd. Sadullah, A. P. A. Majeed, M. A. Mohd. Razman, M. A. Zakaria, and A. F. Ab. Nasir, "Travel Mode Choice Modeling: Predictive Efficacy between Machine Learning Models and Discrete Choice Model," *The Open Transportation Journal*, vol. 15, no. 1, pp. 241–255, Dec. 2021, doi: 10.2174/1874447802115010241.

[8] M. N. A. Shapiee, A. A. Abdul Manan, M. A. Mohd Razman, I. Mohd Khairuddin, and A. P. P. Abdul Majeed, "Chili Plant Classification Using Transfer Learning Models Through Object Detection," *Lecture Notes in Electrical Engineering*, vol. 900, pp. 541–551, 2022, doi: 10.1007/978-981-19-2095-0_46/COVER.

[9] M. Agarwal, S. K. Gupta, and K. K. Biswas, "Development of Efficient CNN model for Tomato crop disease identification," *Sustainable Computing:*

Informatics and Systems, vol. 28, p. 100407, 2020, doi: 10.1016/j.suscom.2020.100407.

[10] L. M. S. Campoverde, M. Tropea, and F. De Rango, "An IoT based Smart Irrigation Management System using Reinforcement Learning modeled through a Markov Decision Process," *Proceedings of the 2021 IEEE/ACM 25th International Symposium on Distributed Simulation and Real Time Applications, DS-RT 2021*, 2021, doi: 10.1109/DS-RT52167.2021.9576130.

[11] Y. Guo *et al.*, "Integrated phenology and climate in rice yields prediction using machine learning methods," *Ecol Indic*, vol. 120, no. August 2020, p. 106935, 2021, doi: 10.1016/j.ecolind.2020.106935.

[12] Zaman, N. Jain, and A. Forster, "Artificial neural network based soil VWC and field capacity estimation using low cost sensors," *2018 IFIP/IEEE International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks, PEMWN 2018*, no. September, 2018, doi: 10.23919/PEMWN.2018.8548808.

[13] R. Akhter and S. A. Sofi, "Precision agriculture using IoT data analytics and machine learning," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 5602–5618, 2022, doi: 10.1016/j.jksuci.2021.05.013.

[14] P. Rodríguez, A. I. Montoya-Munoz, C. Rodriguez-Pabon, J. Hoyos, and J. C. Corrales, "IoT-Agro: A smart farming system to Colombian coffee farms," *Comput Electron Agric*, vol. 190, no. September, p. 106442, 2021, doi: 10.1016/j.compag.2021.106442.

[15] Manikandan, J. ., & Uppalapati, S. L. . (2023). Critical Analysis on Detection and Mitigation of Security Vulnerabilities in Virtualization Data Centers. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 238–246. https://doi.org/10.17762/ijritcc.v11i3s.6187

[16] Jackson, B., Lewis, M., Herrera, J., Fernández, M., & González, A. Machine Learning Applications for Performance Evaluation in Engineering Management. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/126