How to Cite:
Tambakhe, M. D., & Gulhane, V. S. (2022). An intelligent crop growth monitoring system using IoT and machine learning. *International Journal of Health Sciences*, *6*(S8), 230–241. https://doi.org/10.53730/ijhs.v6nS8.9686

An intelligent crop growth monitoring system using IoT and machine learning

Manoj D. Tambakhe
Ph.D. Scholar, Sipna College of Engineering & Technology, Amravati, Maharashra, India
Corresponding author email: manoj20t@gmail.com

Dr. Vijay S. Gulhane
Professor & Head, Sipna College of Engineering & Technology, Amravati, Maharash, India
Email: vijaygulhane27@gmail.com

***Abstract*---**Agriculture is the important part of any developing country. As population increases the food requirement also be increases so the traditional farming not sufficient to fulfill the requirement of food. The development of industries, decreases in land space, insufficient rainfall, chemicals used in farming and peoples are moving towards the urban cities are the challenges of the agricultural field. So we have to use another source of food which we called as hydroponics. In hydroponics we can developed our plant by supplying the nutrients through water without use of soil. In this study we had used the applicability of machine learning algorithm like support vector regression(SVR), Linear Regression, Lasso Regression, Decision Tree(DT), Ridge regression and Random forest to find the crop growth accuracy from the result we found that random forest has given the highest accuracy 95% and the performance are evaluated on various model like Decision Tree Regression having better performance on (R^2=0.86), Support Vector Regression having better performance on (MAE=12.65 and RMSE= 21.31) and Lasso Regression having better performance on (MSE=4.51). The whole dataset are stored on csv file and divided into two subpart 80% data is used for training and 20% is used for testing. In our system we had developed a intelligent hydroponics system for spinach plant and takes real time values continuously from the system such as Ph, EC, TDS, water temperature, temperature and humidity through various sensors. All parameters values are store on firebase.

***Keywords*---**hydroponics, IoT, machine learning.

Manuscript submitted: 27 March 2022, Manuscript revised: 9 May 2022, Accepted for publication: 18 June 2022

Introduction

Soil is the most important part of agriculture. Soil provides the necessary nutrients to plant which is required to grow [1]. With the growth of civilization, there are many challenges which are facing by the soil based agriculture. Day by day food requirement increases as the population Increases rapidly and the availability of land is decreases. The size of the earth remain change but the population increases variably. According to intelligent system it is observed that the amount of crop output available in today's world will become rare by 2050. The production of crop shortages produce due to climate change, water scarcity, chemical fertilizer used, urbanization and lack of working of human in agriculture. Increase in the pollution, change in environment, soil-erosion these are the factors that are degrading the land quality[2]. The major problem of climate changes like vidharbh region in Maharashtra where the summer temperature is more than 45^{o}c so that it is difficult to grow the vegetables like spinach and methi which requires a less temperature. So to fulfill the demand of food and to avoid the soil based traditional farming the hydroponics is the alternative. In hydroponics the plant are grown in the nutrient solution without the use of soil which is also called as soilless culture [3]. We can avoid the weather changes that impact on crop production by managing the temperature and humidity. The aim of modern agriculture is to increase the production and quality of the crop yield where less amount of fertilizers are used compare to soil based agriculture[4]. To address the current issues, to automate things and perform appropriate measure on time, technologies required. Artificial Intelligence(AI), Machine Learning(ML) and Deep Learning(DL) are the technologies that can be used to solve these issues. ML is enhancing the ability of computers to perform some task after trained for specific work. Before to think machine like human they should first learn like human because the mind of human takes a decision on the basis of past experience. There are many application of ML in hydroponics like controlling the crop growth, checking the nutrients values of solution and electrical conductivity values and so on [5].

Hydroponics

In hydroponics, crop can be grown in water solution without the use of soil. We will provide the necessary nutrients like Calcium(Ca), Nitrogen(N), Phosphorus (P), Potassium(K) and Micronutrients(Na) which dissolve in water and through channel we supply to the roots of plant.

IoT in Hydroponics

Internet of Things are use to connect the system to cloud over the internet. In our system internet of things manage the various parameters Ph, EC, TDS, Humidity, Temperature and water temperature with the used of IOT all these parameters are stored on cloud.

Machine Learning Algorithm

ML is the branch of Artificial Intelligence(AI). Artificial Intelligence helps computer to take decision from the train data which is use for certain task. Like human, work and take decision from past experience same ML take a decision from past experience. ML algorithms are use for controlling the growth of plant and nutrient intake solution of plant in hydroponics system [15]. There are Machine Learning algorithm which are use for finding the accuracy of growth of plant.

Literature Review

Ahmed F subahi et al. [1] developed intelligent greenhouse system in which internal temperature can be maintain in the proper range through IoT technologies. A petri nets model was developed for the monitoring and generating suitable temperature. Construct Energy Efficient(EE) system to manage the huge amount of big data generated by sensors and predict crop growth rate, future analysis, energy consumption, production by graph data model.

In paper[3] develop hydroponics system and predict the crop growth by applying the machine learning algorithm like KNN and Lasso regression. The working of the system to take parameters values from nutrient solution through IoT, plant image analysis using OpenCV algorithm and display on user mobile. To apply ML data is collected from sensors, store on flask server, make a csv file, train the data and predict the growth of plant.

In paper[4] author works on the prediction of Crop Growth Rate (CGR), using different machine learning algorithms like random forest and deep neural network. For case study tomato plant used and take a input from system for total 167 days. To find the CGR they collect the parameters from the system like Electrical conductivity (EC), nutrient solution(NS) which contents the mixture of Ca, Mg, Na, K and N, ion concentration, dry wet of fruits.

In this paper[6] authors apply the four algorithm such as support vector machine, deep learning, artificial neural network and extreme gradient boosting to find the accuracy of daily maize plant transpiration(T). To find result they had used combination of four input parameters like (Tmax, Tmin, RH, U and Rs) , soil and crop variables and result shows that DNN gives the more effective result among the four in daily maize.

Yuda prasetia et al. [7] developed an indoor and outdoor hydroponics system to find the growth of plant for bok choy. Monitor the plant growth in sunlight as well in LED light results shows that the growth of the plant in LED light is superior to the sunlight in terms of number of leaves, weight and plant height In [8] authors developed a plant emotional expresser and collect various values in real time mode and predict the crop growth.

In [9] crop growth majorly depends on the environment changes, so developed a system in which suitable moisture is maintain using timer based water spraying system. According to external changes of weather the proper moisture level maintain also the system automatically monitors the Ph, light intensity and proper ratio of nutrients in water tank. Evaluate two hydroponics techniques Nutrient Film Techniques (NFT) and Deep Water Culture (DWC) and compared with the conventional soil based method. Result shows the deep water culture gives more suitable in terms duration of crop growth by 15 days, higher values for the photosynthesis and translated into good range of plant/yield for the system. Whereas in NFT techniques resulted in more water saving than any other techniques [10]. Automatic hydroponics system was developed to captured real time data through IoT and stored on cloud apply deep neural network to find highest accuracy 88% the system work for many week to capture data and trained 10000 times to achieve highest accuracy[16].

Proposed System

We planted spinach in our hydroponics system and monitored its growth for 45 days, collecting data from the system via IoT devices connected to the internet and upload data to the cloud (firebase). After storing the parameters, create a CSV file to use with the machine learning algorithm to calculate the accuracy of the plant's growth. For a better result, we use deep learning figure 1 shows the working flow of our system

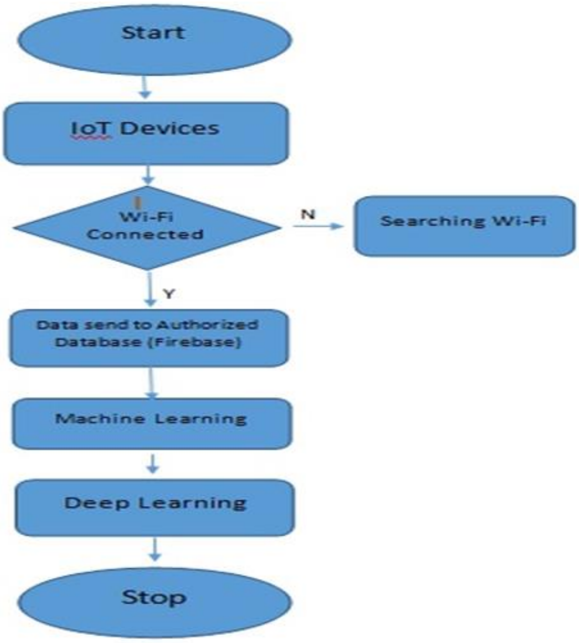

Figure 1 Working Flow of System

Working with Hydroponics

The actual implementation of our system is depicted in Figure 2. The process of a plant's development from seedling to harvest. A three-phase hydroponics system that includes plant seedling, IOT setup, and harvesting. IoT devices connect to gathered parameter values from sensors which are then stored on Firebase.

Figure 2 Hydroponics Farming system with IoT

In figure 3 shows the real time database of all the parameters like Ph, EC, temperature, TDS, Humidity, Water temperature stored on firebase.

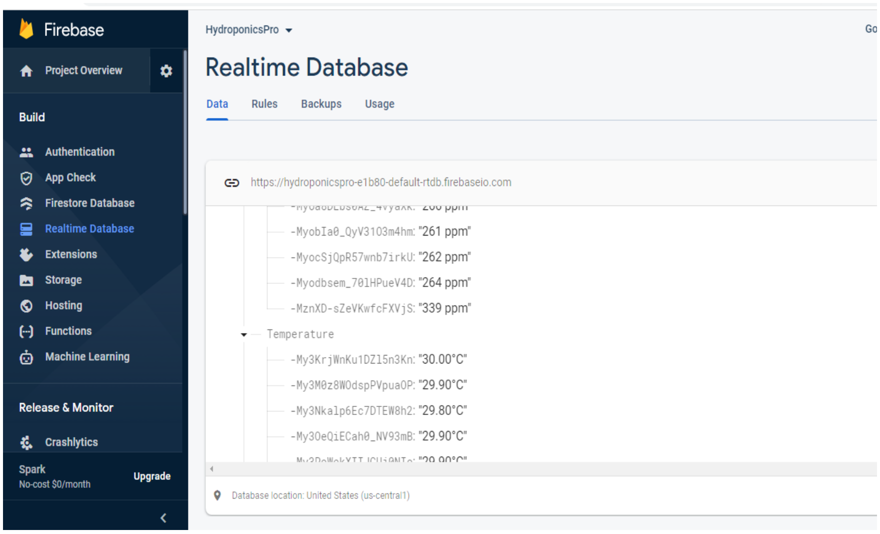

Figure 3 real time database

Methodology

Dataset Description

Data collection plays a vital role in this project as, without data, the whole project loses its meaning, we had collected real time data from the hydroponics system continuously, system update data on firebase after each 5 minute. To automate plant growth, a smart IoT-based plant monitoring hydroponics system is deployed as shown fig 2. There are six types of hydroponics farming system used for plant grow such as Deep water culture (DWC), Wick system (WC), Ebb and flow, Drip system, Nutrient film technique(NFT) and Aeroponics [11]. In our system we use the NFT technique in which the water with nutrients was flow with thin layer in channel. In NFT water are re-circulated and pass through the plant root with a nutrient contain that required for the growth of plant. This technique save a lot of water due to recirculation [8]. To transfer the readings from the sensor to the cloud, we used the Nodemcu(ESP8266) wi-fi module. The pH value, Electrical Conductivity (EC), water temperature, Total dissolved solids (TDS), Temperature, and Humidity were all continuously monitored in real time and send to cloud firebase, followings sensors are used.

pH Sensor

pH sensor are used to monitor the acidity and alkalinity in water solution. It determines the hydrogen power of water solution. The glass pH probe consists of voltmeter and two electrodes one is sensor electrode and other is reference electrode which are fitted in glass tubes . When we insert the pH sensor inside the water sensor which finds the total ions present . The most of the plant required a Ph value in between the range of 6 to 7.

EC and TDS Sensor

Electrical conductivity(EC) and Total Dissolved Solids(TDS) is an important factor in water to check water quality. Both parameters are used to indicate the salinity of water and correlated to each other. EC is used to measure the electric charge of liquid and its capability is determined by dissolved ion ionic strength, concentrations, and measuring temperature. Concentration of dissolved ion is measured by tds sensor.

DHT11 Sensor

DHT11 sensor with digital signal calibration gives information of temperature and humidity. It provides the high quality product, accurate calibration and fast reading with a low price. The uses of DHT11 are in many application it is in small size with signal transmission upto 20 meters. The DHT11 is a digital sensor that can detect temperature and humidity in the air. With Raspberries, this sensor is incredibly simple to use. It has a high level of stability and a highly accurate calibration [14].

Model Implemented

Linear Regression

Linear regression is simple Machine learning algorithm used for predictive analysis. This algorithm shows the linear relationship between the dependant variable (Y) and independent variable(X). The system having single variable like x this system is called as simple linear regression and if there are more than one input variable such regression is called as multiple linear regression. The linear regression is the straight line show the relationship between the variables.

Traditional slope-intercept form is used to calculate best-fit line linear regression.

$$y = mx + b \quad (1)$$
$$y = a_0 + a_1 x \quad (2)$$

y= Dependent Variable.
x= Independent Variable.
a0= intercept of the line.
a1 = Linear regression coefficient.

Decision Tree

It is supervised machine learning algorithm. It is graphical representation of all possible solutions. Decision tree is a data structure that used supervised learning. Decision is depends on some condition which start from root node and go down to each node(children) with input feature and split to possible solution just like a simple tree [15].

Lasso Regression (Least Absolute Shrinkage and Selection Operator)

The regression procedure assumes a linear relationship between the inputs and the target variable, which is known as linear regression. Adding penalties to the loss function during training encourages simpler models with smaller coefficient values, which is an extension of linear regression. Regularized linear regression and penalized linear regression are terms used to describe these expansions. Lasso Regression is a sort of regularized linear regression with an L1 penalty that is widely used. This allows the coefficients for input variables that do not even contribute much to the prediction task to shrink. This penalty allows some coefficient values to be set to zero, thereby removing input variables from the model and allowing for automatic feature selection [12].

$$\sum_{i=1}^{n}(y_i - \sum_j x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \quad (3)$$

λ is the amount of shrinkage, strength of L1 penalty control by **λ**. When **λ**=0 parameters are not eliminated, if **λ** increases bias increases and if **λ** decreases variance increases.

Ridge Regression

Ridge regression is used to analyse the data. This regression was important to reduce the complexity of model and multi-collinearity, it also used to shrink the coefficients. It is a popular linear regression that has L2 penalty. The cost function is calculated by adding penalty equivalent to square of the magnitude of the coefficients. Following equation shows the relationship.

$$\sum_{i=1}^{M}(y_i - \hat{y_i})^2 = \sum_{i=1}^{M}(y_i - \sum_{j=0}^{p} w_j \times x_{ij})^2 + \lambda \sum_{j=0}^{p} w_j^2 \quad (4)$$

The constraints on the coefficients are (w). The penalty term (lambda) regularizes the coefficients so that the optimization function is penalized if the coefficients take high values [13].

Random Forest

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It creates decision trees from various samples, using the majority vote for classification and the average for regression. Important essential characteristics of Random Forest is that it can handle data sets of continues variables in regression as well categorical variables in classification For classification problems, it produces superior results [13].

In addition, the random forest contributes to the model's overall randomness while the trees are growing. Nodes are divided by seeking the best trait among a random bunch of characteristics, rather than aiming for the most fundamental property. There is a great deal of variety as a consequence, which often leads to a superior model. For classification, the Random Forest algorithm uses Gini impurity by default but offer Entropy as an alternative. The mathematical function for Gini Impurity is as follows:

$$\sum_{i=1}^{C} f_i(1 - f_i) \tag{5}$$

Where fi is the frequency of label i at node and C is the number of unique labels

Result and Discussions

Model Evaluation

The performance of Linear Regression, Decision Tree Regression, Support Vector Regression, Lasso Regression, Ridge Regression, Random Forest Regression and Keras Regression(DNN) of hydroponics system were evaluated by more widely used mean absolute error(MAE), mean squared error(MSE), root mean square error(RMSE) and coefficient of determination (R^2) [4] . The performance statistics is given as below.

$$\mathrm{MAE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} |\mathrm{y_i} - \hat{\mathrm{y}}| \tag{6}$$

$$\mathrm{MSE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{y_i} - \hat{\mathrm{y}})^2 \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{y_i} - \hat{\mathrm{y}})^2} \tag{8}$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \tag{9}$$

Where

$\mathrm{y_i} = \text{Actual valuesof y}$

$\hat{y}$ = predicted values of y

$\bar{y}$ = mean *values of y*

Influences of various model on input parameters

We had collected a real-time dataset for hydroponic farming. The data was stored in the cloud using the Firebase database. We had collected attributes values from sensors like Ph, Electrical Conductivity (EC), Humidity, TDS, Temperature and Water temperature. The dataset from the Firebase database was downloaded from the cloud to a local drive using the python library. The downloaded data was stored in a CSV file.

We had apply the following step on CSV file

a. The data in CSV format need to be preprocessed
b. Being the data got stored in object format, we had converted the data from the object to float values using pandas library and regular expression library.
c. All missing values and other unwanted data had been removed to get clean data into CSV format which will be useful for the further machine learning process.
d. As the data is not on a proper scale, we applied a standard scaling function on the given data using the Standard-Scaler function. The data is then split into training and testing purposes in an 80:20 ratio.

Figure 4 shows the Machine learning algorithms with accuracy of the plant growth. From the graph it shows that the Random Forest gave accuracy with 95 %.

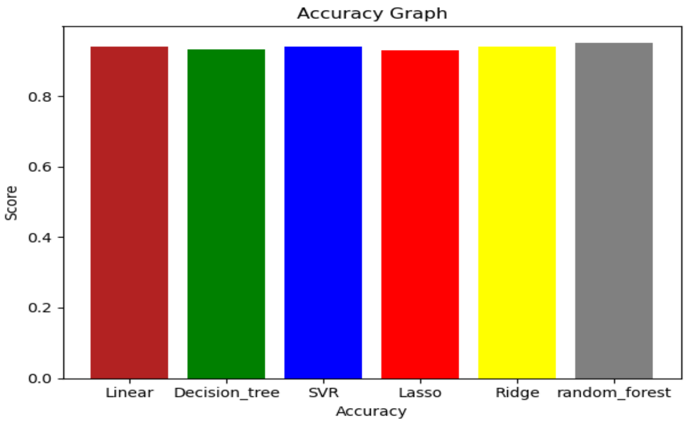

Fig. ML algorithm with Accuracy

We had check our machine learning model with mean absolute error(MAE), mean squared error(MSE), root mean square error(RMSE) and coefficient of determination (R^2). Table 1 gives the performance matrix evaluation of various ML models. Support Vector Machine and Lasso has better performance on (R^2=0.93), Support Vector Regression having better performance on (MAE=12.65 and RMSE= 21.31) and Lasso Regression having better performance on (MSE=4.51).

Sr.No	Machine Learning Models	MAE	MSE	RMSE	R2
1	Linear Regression	12.74	4.64	21.55	0.9288
2	Decision Tree Regression	16.9	8.69	29.48	0.8668
3	Support Vector Regression	12.65	4.54	21.31	0.9304
4	Lasso Regression	14.38	4.51	21.23	0.9304
5	Ridge Regression	12.74	4.64	21.55	0.9288
6	Random Forest Regression	13.44	4.99	22.35	0.9234

Table 1: Performance Matrix Evaluation of ML model

Conclusion and Future Scope

Agriculture is the backbone of every developing country like India. Hydroponics provides the alternative solution to the traditional farming to fulfill the requirement of the food demand. IoT and Machine learning provides the better solution in terms of accuracy to find the plant growth. In our system we had developed a real time hydroponics system by taking the spinach plant as case study. In our model we had takes the various input parameters values in real time mode and store on cloud. Apply the different a Machine learning algorithms to find the accuracy so we had found that Random forest gave us 95% accuracy and check the performance by various evaluation models and found that SVR and Lasso has better in evaluation. In future hydroponics farming system could be extended in real time mode for another plant growth to increase the higher accuracy.

References

1. Surbhi A. F. and Bouanza K. E. , "An Intelligent IoT-Based System Design for Controlling and Monitoring Greenhouse Temperature," in *IEEE Access*, vol. 8, pp. 125488-125500, 2020, doi: 10.1109/ACCESS.2020.3007955.
2. Lakshmi, J.V.N and HemanthK.S. “Hydroponic Farming using Leafy Green Machines for Agriculture by Applying Deep Learning Methods”, IJRECE, 7(1): 2062-2065,ISSN: 2348-2281
3. H. K. Srinidhi, H. S. Shreenidhi and G. S. Vishnu, "Smart Hydroponics system integrating with IoT and Machine learning algorithm," 2020 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), 2020, pp. 261-264, doi: 10.1109/RTEICT49044.2020.9315549.
4. Ms Swapnil Verma & Dr.Sushopti D.Gawade, “A machine learning approach for prediction system and analysis of nutrients uptake for better crop growth in the Hydroponics system”, 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 150-156, doi: 10.1109/ICAIS50930.2021.9395956.
5. Mokhtar A, El-Ssawy et. al.,”Using Machine Learning Models to Predict Hydroponically Grown Lettuce Yield.”,Front. Plant Sci. 13:706042., 2022, vol 13, pp. 1-10, doi: 10.3389/fpls.2022.706042.

6. Junliang F., Jing Z., Lifeng W., Fucang Z.,"Estimation of daily maize transpiration using support vector machines, extreme gradient boosting, artificial and deep neural networks models", Agricultural Water Management, Volume 245, 2021,106547, ISSN 0378-3774, https://doi.org/10.1016/j.agwat.2020.106547.
7. Yuda P., Aji G. P., Andrian R., "Evaluation of IoT-Based Grow Light Automation on Hydroponic Plant Growth", JITEKI, Vol. 7, No. 2, August 2021, pp. 314-325 ISSN: 2338-3070, DOI: 10.26555/jiteki.v7i2.21424
8. C.H. Vanipriya, Maruyi, S. Malladi et al.," Artificial intelligence enabled plant emotion xpresser in the development ydroponics system", Materials Today: Proceedings 45 (2021) 5034–5040, https://doi.org/10.1016/j.matpr.2021.01.512
9. Muralimohan G, Arjun S V, Sakthivel G "Design and Development of IoT based Hydroponic Farming Setup for Production of Green Fodder" Nat.Volatiles & Essent.Oils,2021;8(4):4325-4340
10. Maliqa Majid, Junaid N et al." Evaluation of hydroponic systems for the cultivation of Lettuce (Lactucasativa L., var. Longifolia) and comparison with protectedsoil-based cultivation" , Agricultural Water Management, Volume 245, 2021,106547, ISSN 0378-3774, https://doi.org/10.1016/j.agwat.2020.106572
11. Ghorbel, Roukaya & Chakchak, Jamel et al. (2021). Hydroponics "Soilless Farming": The Future of Food and Agriculture – A Review. 10.52460/issc.2021.007.
12. Yaping Cai, Kaiyu Guan, et al. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches", Agricultural and Forest Meteorology, Volume 274, 2019, Pages 144-159, ISSN 0168-1923, https://doi.org/10.1016/j.agrformet.2019.03.010.
13. Juan Cao, Zhao Zhang et. al."Integrating Multi-Source Data for Rice Yield Prediction across China using Machine Learning and Deep Learning Approaches", Agricultural and Forest Meteorology, Volume 297, 2021, 108275, ISSN 0168-1923, https://doi.org/10.1016/j.agrformet.2020.108275.
14. Shaiz Akhtar Mohammad, Daggumalli Sai NikhilaChowdary, Dr. R. Jebakumar, "An Smart Aquaponic System Using IoT" Journal of Positive School Psychology, 2022, Vol. 6, No. 4, 226-235, http://journalppw.com.
15. Yemeserach Mekonnen, Srikanth Namuduri et. al. "Review—Machine Learning Techniques inWireless Sensor Network Based Precision Agriculture", Journal of The Electrochemical Society, 2020, vol 167, 037522, DOI: 10.1149/2.0222003JES.
16. M. Mehra, S. Saxena, S. Sankaranarayanan, R. J. Tom, and M. Veeramanikandan, "IoT based hydroponics system using Deep Neural Networks," Comput. Electron. Agric., vol. 155, no. October, pp. 473–486, 2018, doi: 10.1016/j.compag.2018.10.015.
17. Rinartha, K., & Suryasa, W. (2017). Comparative study for better result on query suggestion of article searching with MySQL pattern matching and Jaccard similarity. In *2017 5th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-4). IEEE.
18. Rinartha, K., Suryasa, W., & Kartika, L. G. S. (2018). Comparative Analysis of String Similarity on Dynamic Query Suggestions. In *2018 Electrical Power,*

Electronics, Communications, Controls and Informatics Seminar (EECCIS) (pp. 399-404). IEEE.

19. Susilo, C. B., Jayanto, I., & Kusumawaty, I. (2021). Understanding digital technology trends in healthcare and preventive strategy. *International Journal of Health & Medical Sciences*, *4*(3), 347-354. https://doi.org/10.31295/ijhms.v4n3.1769