CROP HEALTH MONITORING USING IOT SENSORS BASED REGRESSION MODEL

Tharageswari K [1] M. Jamunarani* [2]*

ABSTRACT

Farmers of today face the problem of decreased crop yield. One of the main things a farmer can do to increase crop yield is to select the correct crop. Not performing analyses like metrological factors checking and soil analysis and having less knowledge about soil fertility leads to less crop yield. Knowing the above details will boost crop production. Farmers face considerable challenges in selecting the right crop since the climate does not follow a pattern, and the farmers need to gain basic knowledge regarding crop selection and modern farming methods. When farmers select the same crop every season, the soil will lose fertility, leading to low crop yield. Machine learning (ML) algorithms and IoT devices are used in the proposed study to make a system that can perform accurate, practical and valuable decision-making. The farmer can achieve maximal yield with the help of the system. This system will suggest to the farmer how to select the right crop. Compared to the laboratory testing performed in the olden days, the proposed regression model is reliable, where manual errors can be avoided. In the area of agriculture, the main priority is the selection of the correct crop. As a contribution to agriculture, Smart Crop Selection model based on machine learning and IoT was developed. Metrological information and soil factors are the data used for our system. Potassium, CO2, EC, temperature, soil's humidity, rainfall, nitrogen, phosphorus and pH value are the factors used by the system for crop selection. Only some of these factors are employed by the existing system, making it inefficient compared to our proposed model. Our proposed system sends real-time sensory data to analyze the various factors.

Computer Science and Engineering[1],
Karpagam Academy of Higher Education, Coimbatore, India[1]
tharageswari.k@kahedu.edu.in[1]
Kalaingar Karunanidhi Institute of Technology, Coimbatore, India[2]
* Corresponding Author

Keywords: Yield, Crop Monitor, IoT, Machine learning, prediction

I. INTRODUCTION

The agricultural field plays a vital role in the world's economy. It is humanity's lifeline for survival. Farmers select crops with fundamental knowledge of conventional farming methods [1]. Typically, the crops planted in neighbouring areas or the currently popular crop will influence the crop selection of the farmers. Land fertility is affected adversely by selecting crops without scientific knowledge regarding crop rotation and land fertility [2]. The quality of the crop is determined by factors like fertilizer used, nutrients in soil and groundwater level; significant factors contributing to the crop quality are groundwater level and type of use. Farmers recurrently face the same challenges. Owing to the fewer soil nutrients and wrong crop selection, the soil's acidity might go up [3]. Soil fertility plays a vital role in choosing crops and their healthiness. Another major factor that affects yield and quality is the volatile climate.

Supporting farmers and encouraging them to choose the right crop is the motivation of the proposed study; for this, we need to find the problems farmers face [4]. We introduced a system based on a machine learning algorithm that uses the Internet of Things called Smart Crop Seleto (SCS) to overcome farmers' problems regarding crop selection. Farmers face most issues due to a lack of proper systems or approaches [5]. As illustrated in Figure 1, the Smart Crop Selection model considers metrological parameters like temperature, air's carbon dioxide (CO_2) level, soil pH level, soil type, humidity, EC, and rainfall. The production of the crop and its growth is directly affected by metrological factors. Soil analysis is carried out to check the fertility of the soil [6].

Values of macronutrients present in the soil are identified in the process of soil analysis. A few of the macronutrients in the soil are potassium, nitrogen and phosphorus. Nutrients influence the crop's well-being; they are helpful in their prevention. The pH value of the soil presents its alkalinity, and it also controls the availability of nutrients in the soil by regulating the chemical reactions that convert them into different minerals [7]. Crop growth is affected by lesser or higher EC values in the soil. EC value indicates data such as salinity, quality of water and its fertility. Crop health can be maintained well with the optimum levels of CO_2 in the air [8]. In the plant, the process of photosynthesis is made possible by CO_2. The method introduced can be applied to two soil varieties: 1. clay and 2. loamy.

A high moisture and humidity level is maintained in these two soil types, which is needed for most crops. Healthy crops always need good rainfall. Various plants require various water levels for their growth. Before the crop sowing, information like the season's average rainfall can significantly help. Predicting rainfall can be challenging, but ML algorithms obtain promising outcomes. When techniques like precision agriculture are used, a 50% to 90% growth rate for crop yield can be achieved. It paves the way for better resource utilization and correct decisions [9]. Utilizing these techniques will maintain the fertility of the soil.

Using the IoT can be valuable for precision agriculture technology. Unwelcome circumstances can be avoided by making practical decisions based on an IoT-based farming system. Smart agriculture using an automation system will be precise and relatively inexpensive compared to conventional farming. The structure of IoT systems has the following three layers: the perception layer, the network layer, and the application layer. Physical devices such as cameras, RFID tags, and sensors are in the perception layer to collect data. Forwarding of data and its communication are performed in the network layer. The domain of use and the Internet of Things are combined using the application layer [10].

In Artificial Intelligence, machine learning is the most exciting application. With no explicit program, it can provide experiences using learning. A less complicated and low-cost hardware device is proposed in this paper, which can be of great use to farmers and agriculture officers to obtain maximum yield. SCS system is tested subsequently and trained using dataset classification. The size and kind of dataset determine the ML classifier's performance and accuracy. The dataset used for preparing the SCS system. Our dataset for training the model has 2,200, spanning over 11 crop kinds. RF, NB, DT, KNN and SVM are the five supervised machine-learning algorithms for dataset classification [10]. ML algorithms are ensembled to achieve better accuracy and overcome drawbacks. In real-time testing, the outcome of our system shows an accuracy of 97% to 98%. SCS model uses innovative aspects, which are given below:

(i) The proposed model linear regression considers extra factors of metrology, whereas only a few factors were taken into consideration by the studies conducted previously,

(ii) An effective but inexpensive sensors are employed in the proposed system, but the existing system uses costly and complex sensors for innovative agriculture systems.

(iii) Analyzing soil in the lab manually is not cost-effective and consumes too much time. The existing system procures soil for analysis manually, so it is prone to human errors. So, we introduced linear analysis model, which is an automated system that provides accurate outcomes.

The work is drafted as follows: section 2 gives detailed analysis of prevailing works. The methodology is provided in section 3. The outcomes are given in section 4. The conclusion is given in section 5.

II. RELATED WORKS

Majumdar et al. adopted agricultural methods based on IoT to monitor weather conditions. This system is tested for scientific and commercial aspects such as weather condition-based crop irrigation, security issues and IoT component cost [11]. A smart irrigation system was suggested by Imran that considers factors such as soil humidity, the intensity of light, moisture level and temperature. The experiments were conducted on five soil types: late, rite, alluvial, loamy, silt, and black. The experiment results show that we can select crops depending on various lands and their soil characteristics. An application named Think Speak is employed to analyze the data. Through an android application, the water requirement for crop irrigation is intimated to the farmers [12].

An Internet of Things framework was proposed. It states that profit maximization and crop production can be improved by finding the best farming methods and land selection. To monitor the field correctly, a wireless sensor network was placed for data sensing to identify various meteorological factors. Fertilizer requirements can be found using a pH sensor. It helps to find the nutrients in the soil nutrients. Irrigating the field concerning current weather conditions is suggested to farmers using an explicitly developed Android app [13]. A method for crop prediction was proposed by Mulge et al. to maximize the yield and achieve good crop quality by taking real-time metrological factors like solar light, humidity, precipitation, and temperature into account using ML algorithms [15].

Gupta and Nahar proposed a novel two-tier ML methodology for predicting the crops. Adaptive K-Nearest Centroid Neighbor (aKNCN) is present in the first tier; it is a classifier that will analyze the soil quality and categorize the soil samples into various classes depending on the soil properties. An ML algorithm is used in the second tier to predict crop yield. The performance of the three algorithms (Decision tree, SVM and KNN) were compared. For proper crop selection, processes like weather analysis, disease detection, and monitoring conditions are done by training another model. CNN is used for the detection of disease. Information leakages and security issues were identified in the IoT system's network layer. The yield prediction model was proposed [15] as a model for suitable crop selection. A dataset from an external source is obtained to analyze the soil. Soil is analyzed for various macronutrients and micronutrients.

III. METHODOLOGY

A. Data Set

There are many areas for improvement in detecting diseases in plants using traditional technologies. To solve this issue, our detection technique for disease employs the Kaggle dataset. Three classifications, healthy, bacterial spot and early blight, are done in the image collection of 895 images.

B. Prediction model

Once the WSN was put to work, the data from the sensors was logged into a Postgres database. The captured data is analyzed to find the best possible prediction model. Extracted data from the DB table is compiled in. CSV file format. Data pre-processing: Unwanted columns not required for us are dropped from the. CSV file. Then, abnormal values like impractical temperature peak needs to be eliminated from the table. Our outcomes will be negatively affected by these values. So, they are destroyed.

Analysis using linear regression: The same data is fed into the machine learning algorithm for disease prediction and advanced treatment plans using a simple linear regression model. The captured data was analyzed to find a suitable prediction model.

In the linear regression-based prediction model, three variables are used: T- temperature, W- wetness of leaf and its duration, and I- incubation period of pathogen causing the infection. So, collected data is analyzed to find a suitable prediction model. Three variables used are temperature (T), wetness (W) duration and incubation period for pathogen (I). The objective is to find the relationship of I as T and W's function. This Equation helps identify the susceptibility of a definite blend of temperature and wetness duration regarding an infection attack. A single response measurement is linked to a single predictor like covariate or regression X in every observation in a simple linear regression model. The complex model with conditional mean function is linear is in Eq. (1):

$$E(Y \mid X) = \beta_0 + \bar{\beta_1} * X \quad 1$$

More than one predictor variables are present in most problems. This leads to the "multiple-regression" mean function as provided in Eq. (2):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_q x_q + \varepsilon \quad 2$$

Here, *X* variable's intercept is called q, , and the coefficients or slope is *x*. Other predictors are kept constant; the difference in mean response per unit increases in *X* with each coefficient estimate. The respective Eq. is given in Eq. (3):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_q x_q + \quad 3$$

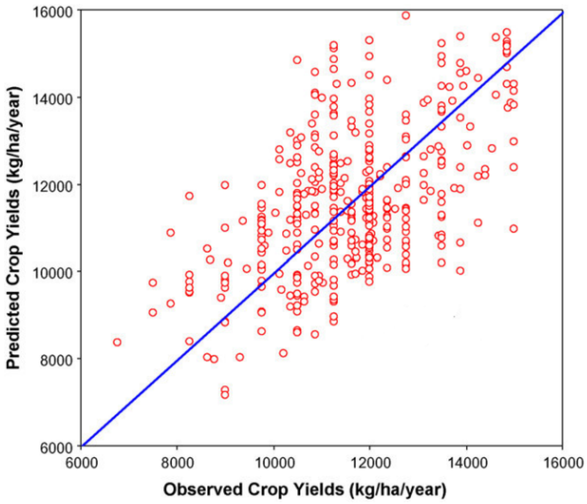

Figure 1. Linear regression for crop yield prediction

The linear regression model is used to get Eq. (3), which states that the wetting period value and temperature value predict the incubation period as in Fig 1. The inverse relationship between the incubation period with wetting duration and temperature indicates the negative coefficients 0.3329 and 0.887. The *I* value increases with the wetting period and the temperature decreases.

IV. RESULTS

Our proposed system shows improved accuracy and performance by applying linear analysis compared to individual models. Accuracy is calculated by dividing accurate predictions (calculated through an algorithm) by the quantity of dataset values.

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \quad 4$$

Here, the true positive is TP, the actual negative is TN, the false positive is FP, and the false negative is FN. The following formula is used to compute classification error rate:

$$Error\ rate = (1 - accuracy) * 100 \quad 5$$

Our model's reliability is checked by computing the following parameters: Recall, F1-score and precision. Accurate predictions in positive classes are calculated to produce precision results.

$$Precision = \frac{TP}{TP + FP} \quad 6$$

Number of positive cases predicted by our model is given in Recall.

$$Recall = \frac{TP}{TP + FN} \quad 7$$

Combination of Recall and is used for computing the F1-score:

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \quad 8$$

ML algorithm's performance was compared. Compared to DT, SVM, and KNN algorithms, NB and RF algorithms show better classification accuracy, as illustrated in Fig []. For binary classification, SVM is highly suitable. An accurate decision tree and a stable boundary are achieved because of Multiple DTs.

For every class, instances that are wrongly predicted and predicted correctly are shown in the confusion matrix. Accuracy for the proposed model is computed using the Eq. (9):

$$Accuracy = \frac{sum\ of\ total\ predicted\ value}{total\ predicted\ value} \quad 9$$

Table 1. Comparison with other approaches

Approaches	Accuracy %
DT	96%
SVM	96.2%
RF	93%
k-NN	90%
Proposed model	**98.5%**

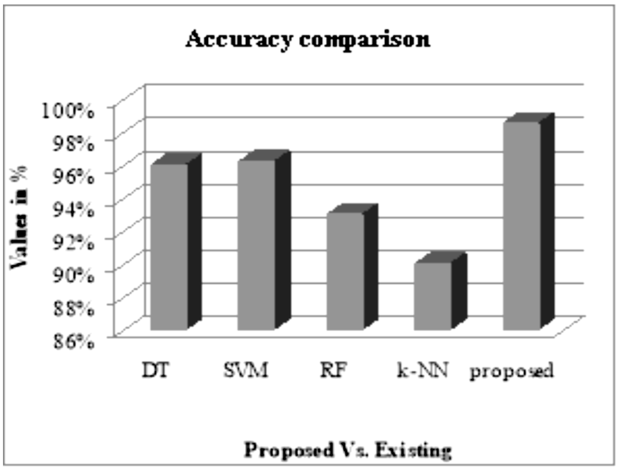

Figure 2. Accuracy comparison

Figure 3. Yielded crop sample

The values from the confusion matrix are used to calculate the performance metrics of "black gram", which is one of the output label classes as in Fig 3. By setting values and applying formulas where the following values of performance metrics are determined by percentage.

V. CONCLUSION

Due to climatic change, wrong crop selection, unpredicted weather, and wrong amount of nutrients and water, farmers using conventional agricultural methods face many issues like poor crop quality and minimum crop yield. Prior research studies used factors to predict crops to sow less, resulting in inaccurate prediction and poor crop quality. The proposed automation model intends to select a suitable crop for maximum yield by choosing the right crop. Evidence-based analysis and novel regression model is used in our model to overcome the issues faced by traditional methodologies. For example, proper nutrition in the right amount will provide maximum crop yield. Influential parameters are chosen correctly in our proposed model. Compared to previous research ML, algorithms-based proposed research provides less computational cost and improved accuracy. AS future enhancement, more crops and parameters can be included in this system. ML algorithms are capable of being more efficient and accurate, can be included. Security can be integrated with the regression model to protect crop data. Drone cameras can be deployed to monitor the crop. With soil nutrients' real-time sensory data, a recommendation system for fertilizer can also be built.

REFERENCES

[1] Kanchan and N. Shardoor, "Krashignyan: a farmer support system," Asian Journal For Convergence In Technology (AJCT) ISSN-2350-1146, vol. 7, no. 3, pp. 1–7, 2021.

[2] Issad, R. Aoudjit, and J. J. Rodrigues, "A comprehensive review of data mining techniques in smart agriculture," Engineering in Agriculture, Environment and Food, vol. 12, no. 4, pp. 511–525, 2019.

[3] Kavita, "Satellite-based crop yield prediction using

machine learning algorithm," in 2021 Asian Conference on Innovation in Technology (ASIANCON), pp. 1466–1470, PUNE, India, August 2021

[4] Jain, A. Kumar, S. Garud, V. Pradhan, and P. Kulkarni, "Crop selection method based on various environmental factors using machine learning," International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 2, pp. 1530–1533, 2017.

[5] Bhojwani, R. Singh, R. Reddy, and B. Perumal, "Crop selection and IoT based monitoring system for precision agriculture," International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 2, 2017.

[6] Majumdar, S. Mitra, and D. Bhattacharya, "IoT for promoting agriculture 4.0: a review from the perspective of weather monitoring, yield prediction, security of WSN protocols, and hardware cost analysis," Journal of Biosystems Engineering, vol. 46, no. 4, pp. 440–461, 2021.

[7] Rekha, V. P. Rangan, M. V. Ramesh, and K. V. Nibi, "High yield groundnut agronomy: an IoT based precision farming framework," in 2017 IEEE Global Humanitarian Technology Conference (GHTC), vol. 2017, pp. 1–5, San Jose, CA, USA, Dec. 2017

[8] Nagasubramanian, R. K. Sakthivel, R. Patan, M. Sankayya, M. Daneshmand, and A. H. Gandomi, "Ensemble classification and IoT-based pattern recognition for crop disease monitoring system," IEEE Internet of Things Journal, vol. 8, no. 16, pp. 12847–12854, 2021

[9] Ram and R. Kumar, Prediction of the crop cultivating using resembling and IoT techniques in agricultural fields for increasing productivity 1, European Journal of Molecular & Clinical Medicine, 2020.

[10] Bakthavatchalam, B. Karthik, V. Thiruvengadam et al., "IoT framework for measurement and precision agriculture: predicting the crop using machine learning algorithms," Technologies, vol. 10, no. 1, p. 13, 2022.

[11] Gupta, D. Nagda, P. Nikhare, and A. Sandbhor, "Smart crop prediction using IoT and machine learning," International Journal of Engineering Research & Technology (IJERT), pp. 18–21, 2021.

[12] Gupta and P. Nahar, "Classification and yield prediction in smart agriculture system using IoT," Journal of Ambient Intelligence and Humanized Computing, 2022

13] Kishan Das, D. Mishra, and D. Deepa, "Automation and integration of growth monitoring in plants (with disease prediction) and crop prediction," Materials Today: Proceedings, vol. 43, no. 1, pp. 3922–3927, 2021.

[14] Khan, Z. Ali, M. Waqas, and G. A. S. Sidhu, "Efficient power allocation with individual QoS guarantees in future small-cell networks," AEU-International Journal of Electronics and Communications, vol. 105, pp. 36–41, 2019.

[15] Parween, A. Pal, I. Snigdh, and V. Kumar, "An IoT and machine learning-based crop prediction system for precision agriculture," in In Emerging Technologies for Smart Cities, pp. 9–16, Springer, Singapore, 2021.