Smart Agriculture Crop Yield Maximization Technique

Chaithra M S[1*], Deepika N R[2], Bhoomika H A[3], Deeksha J[4], Anitha T G[5]
[1,2,3,4]*Student, Electronics and communication, EWIT, Bangalore, India*
[5] *Professor, Electronics and communication, EWIT, Bangalore, India*

Abstract - Agriculture and its allied sectors are undoubtedly the largest providers of livelihoods in rural India and is the backbone of the country. The agriculture sector is also a significant contributor factor to the country's Gross Domestic Product (GDP). This Project is an attempt to minimize the losses that occurs in Agricultural field due to climate changes, environmental changes and fluctuations in market price. Blessing to the country is the overwhelming size of the agricultural sector, but regrettable is the yield per hectare of crop is not up to international standards. Also, the poor knowledge of farmers about the crop is one of the possible causes for a higher suicide rate among marginal farmers in India. This paper proposes a viable and user-friendly yield prediction system for the farmers using Machine learning algorithms which are applied to get best predicted results through the web application.

***Keywords*:** Gross Domestic Product (GDP), yield prediction, Machine learning (ML)

1. INTRODUCTION

Agriculture is the most important sector and is the backbone of Indian economy. In India, agricultural yield primarily depends on weather conditions. Yield prediction is an important agricultural problem.

In the past farmers used to predict their yield from previous year yield experiences. Most of the farmers rely on traditional cultivation and on their long-terms experiences in the field for particular crop cultivation. Many times, it leads to lesser yield and not reasonable profit for the crops. Since the creation of new innovative technologies and techniques in the agriculture field is slowly degrading, abundant researchers are concentrating on cultivating artificial products that are hybrid products leading to an unhealthy life. Now a days, the seasonal climatic conditions change unexpectedly against the fundamental assets like soil, water and air that results in poor yield. The increase in yield in such climatic conditions can be done by appropriate selection of crop that suits the climate, temperature and several factors. Also, to gain the reasonable profit for the yield, it is necessary to predict the future crop price in the market before cultivation of a particular crop.

Indian agriculture sector accounts for 18 percent of India's GDP and provides employment to 50% of the country's workforce. In India, there are several ways to increase the economic growth in the field of agriculture. Also, there are multiple ways to improve the crop yield and the quality of the crop.

Data mining is one of the useful methods in predicting crop yield production. Agricultural researchers insist on the need for an efficient mechanism to predict and improve the crop growth. Majority of research works in agriculture are focusing on biological mechanisms to identify crop growth and improve its yield. The outcome of crop yield primarily depends on parameters such as variety of crop, seed type and environmental parameters such as Sunlight (Temperature), Soil (pH), Water (pH), Rainfall and Humidity.

Timely advice to predict the future crop productivity based on soil parameters and climatic conditions is required for farmers before cultivation of new crop. This prediction will help the farmers in choosing suitable crops for their farm according to the Soil type, Temperature,

Humidity, Water level, Spacing depth, Soil (pH), Season and Fertilizer. Thus, for crop prediction, based on the said parameters there are different techniques and algorithms are available. With the help of these algorithms crop yield can be predicted and Random Forest algorithm is one such algorithm.

In addition, information about surplus stock in the nearby districts is necessary in planning for particular crop cultivation and in predicting the crop price to gain profit. Furthermore, the selection of the same crop in every seasonal cycle makes the low soil fertility. Agricultural researchers insist on the need for an efficient mechanism to predict and improve the crop growth. The majority of research works in agricultural field focus on biological mechanisms to identify crop growth and improve its yield. Researchers have developed a number of algorithms to predict the most profitable crop and Machine learning algorithm is one such algorithm that allows users to choose the most profitable crop list.

Smart agriculture is the way of conveying information from traditional farmers to the educated farmers. To obtain estimates of aggregate physical production functions for the yields of various crops in specified states, considering various technological factors and a newly developed weather index as inputs. Regression and coefficient of determination analysis along with Average Error rate were carried out to make a decent comparison between our actual result which is called target and prediction model that is friendly interface for farmers, which gives the analysis of rice production based on available data. Different Data mining techniques were used to predict the crop yield for maximizing the crop productivity. Accurate and timely monitoring of agricultural crop conditions and estimating potential crop yields are essential processes for operational programs. The importance of crop yield prediction has initiated the study of application in several forecasting methods for evaluating crop yield estimation that aids in appropriate crop cultivation.

2. PROPOSED SYSTEM

The proposed system is based on real-time data of the soil parameters by sensors such as humidity sensors, soil moisture sensor, pH sensor and nutrients sensor. The real-time sensor data is saved in database on ThingSpeak cloud. These data are used in ML algorithms such as classification algorithm (K-nearest neighbor, Support vector machine, Random Forest, Decision tree) to recommend the appropriate crop for cultivation based on soil parameters and climate conditions. Regression algorithms (Linear, Logistic, Random Forest and Decision tree regression) are applied for further data analysis for prediction of crop yield and market price to gain profit, on the basis of external dataset collected from the various agricultural organizations. Python programming language is used for algorithm implementation.

3. METHODOLOGY

Dataset: It must have the following attributes:

• Soil Parameters: Soil Type, Soil pH value.

•Climatic Parameters: Humidity, Temperature, Rainfall.

•Production cost of cultivation previous year yield details for that region.

The objective of this project is crop prediction in our surroundings, hence it is required to find the dataset, details of climatic conditions and soil parameters like rainfall, temperature, moisture, soil contents etc.

These factors will help in the prediction of the crops by using various classifiers on the given dataset. Thus, various factors are assessed and the factors strongly leading to accurate prediction of the crops.

Crop Yield Prediction:

The outcome of crop yield primarily depends on parameters such as variety of crop, seed type and environmental parameters such as sunlight (Temperature), soil (pH), water (pH), rainfall and

humidity. By analyzing the soil and atmosphere at particular region in order to have more crop yield and the net crop yield can be predicted.

This prediction will help the farmers to choose appropriate crops for their farm according to the soil type, temperature, humidity, water level, soil pH, season, fertilizer.

Fertilizer Prediction:

India is a highly populated country and randomly change in the climatic conditions need to secure the world food resources. Farmers face serious problems in drought conditions. Type of soil plays a major role in the crop yield. Suggesting the use of fertilizers may help the farmers to make the best decision for their cropping situation.

Pre-processing: The dataset that is used needs to be pre-processed because of the presence of redundant attributes, noisy data in it. Initially, data cleaning operation is performed where the redundant factors are determined and removed.

As part of the exploratory data analysis, the categorical factors are split and are assigned values as 0 and 1 based on whether the factor is present or not. These assigned values assist in further classification based on that particular factor.

It is a method that is used to convert the raw data into a clean data set. The data are gathered from different sources, it is collected in raw format which is not feasible for the analysis. By replacing missing values and null values, we can transform data into an understandable format. The final step is splitting of training and testing data (Here 80% of dataset is taken as training dataset).

Feature engineering refers to manipulation- addition, deletion, combination, mutation of dataset to improve ML model training, leading to better performance and greater accuracy.

Effective feature engineering is based on sound knowledge of the business problem and the available data sources.

Splitting the dataset: By using similar data for training and testing, you can minimize the effects of data discrepancies and better understand the characteristics of the model. After a model has been processed by using the training set, you test the model by making predictions against the test set.

Machine Learning: An “algorithm” in machine learning is a procedure that is run on data to create a machine learning model. Machine learning algorithms perform “pattern recognition”. Algorithms “learn” from data, or are “fit” on a dataset. There are many machine learning algorithms the two algorithms use in this technique are:

Classification algorithms: K-nearest neighbor, Support vector, Decision tree, Random Forest.

Regression algorithms: Linear, Logistic, Random Forest, Decision tree.

●Comparison and Selection of Machine Learning Algorithm.

●Before deciding on algorithm to use, first we need to evaluate and compare, then choose the best one that fits this specific data.

a) Random Forest Algorithm:

It creates decision trees on different data samples and then predict the data from each subset and then by voting gives better solution for the system. RF uses the bagging method to train the data which increases the accuracy of the result. Random Forests are the aggregation of tree predictors in such a way that each tree depends on the values of a random subset sampled independently and with the same distribution for all tree in the forest.

b) Decision Tree Algorithm:

A decision tree is a flowchart like structure in which each internal node represents a test on a feature (e.g. whether a coin flip comes up heads or tails), each leaf node represents a class label and branches represent

conjunctions of features that lead to those class. Decision Trees are a type of Supervised Machine Learning where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves.

Soil moisture, Humidity, pH and N, P, K (Nitrogen, Phosphorous, Potassium) values are measured using sensors and the data is sent to cloud. Signal conditioning can include amplification, filtering, converting range matching, isolation and any other processes required to make sensor output suitable for processing after conditioning.

Graphical User Interface (GUI) is a form of user interface that allows users to interact with electronic devices through graphical icons and audio indicator such as primary notation. GUI autofill's the cloud values and provides it as an input to the trained model and the data is stored in the bytes format. It compares output of different algorithms based on accuracy and determines which approach is more suitable. The predicted result is analyzed and stored in the string datatype.

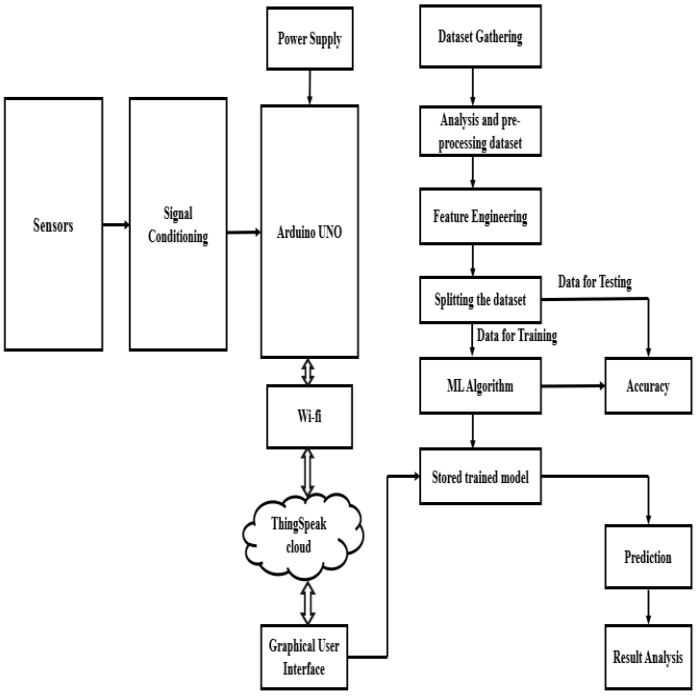

Fig-1: Block Diagram

4. RESULT

In this paper, yield manufacturing inspection is done and is operated by applying both the Compression and Regression Algorithms. These models are investigated with various kind of crops in different regions across India to anticipate the harvest. Indeed, even manure information was developed utilizing the back-propagation calculation and estimated to acquire the output of how much nitrogen, phosphorus is necessary for the sector of land. The models for the crop are examined in expecting the yield and by different parameter concerning the fault rate.

The result of harvest anticipation. The process of gathering requirements is usually more than simply asking the users. Depending on the complexity of the application, the process for gathering requirements has a clearly defined process of its own. This process consists of a group of repeatable processes that utilize certain techniques to capture, document, communicate, and manage requirements.

The below Fig-2, Fig-3, Fig-4, Fig-5, Fig-6 represents Soil parameters Crop Recommendation, Yield Prediction, Price Prediction and Fertilizer Recommendation respectively.

The charts shown below Fig-7, Fig-8 represents Algorithm accuracy comparison and Seasonal crop estimation respectively.

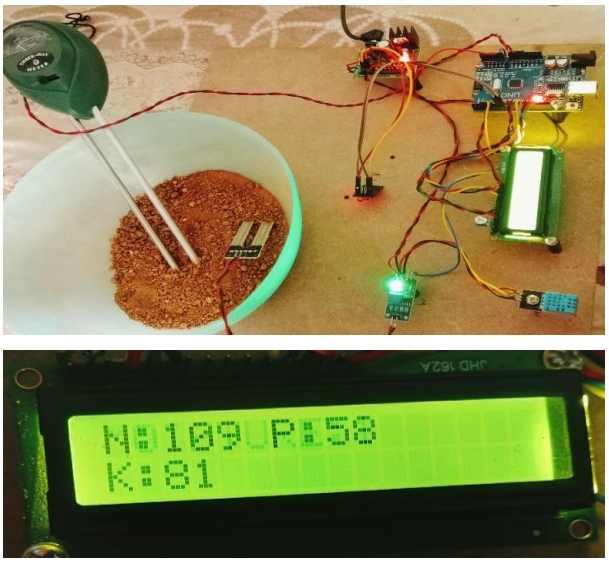

Fig-2: Soil parameters

Fig-3: Recommended Crop

Fig-6: Recommended Fertilizer

Charts

Fig-4: Predicted Yield

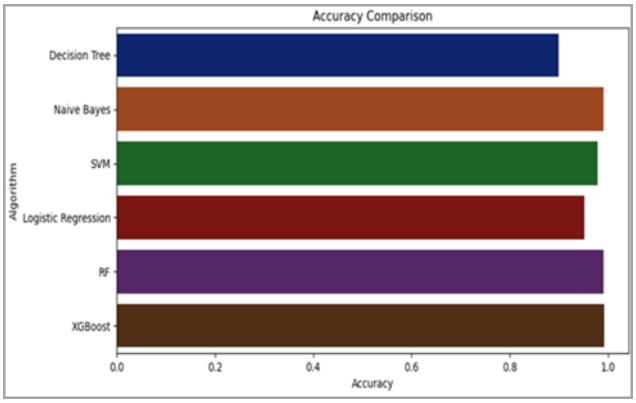

Fig-7: Algorithm Accuracy Comparison

Fig-5: Price Prediction

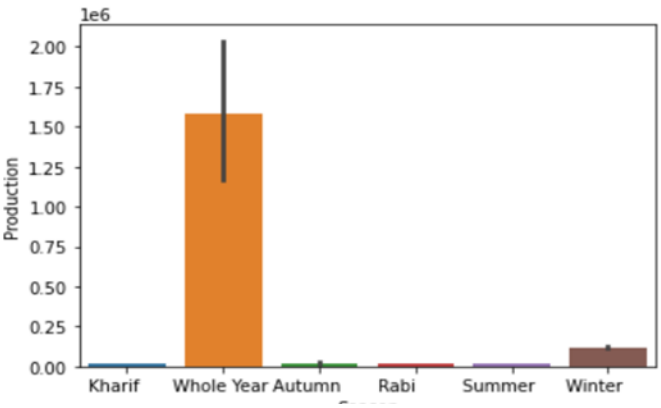

Fig-8: Seasonal Crop Estimation

5. CONCLUSION

The sensor value is updated every few seconds in the ThingSpeak cloud and is tested against the ML trained model. The acquired result will be beneficial to know the Yield of the harvest so that we can go for the better one. Also, farmers can be guided for the necessary quantity of manures for that farm.

To improve Agribusiness, this approach reduces farmers' losses and improves yield. In this phase of the project, we reviewed the literature on the price prediction, crop-yield prediction. This literature helps us understand the challenges that we face in the price dataset to identify the price and yield. According to the literature, experimental results it was concluded that random forest and back propagation helps in crop-price identification and helps to eradicate the challenges mentioned in the dataset. The accuracy rate of this algorithm is far better than another algorithm. By integrating this with other departments like sericulture and other development initiatives on the village level, farmers are able to gain a deeper understanding of agricultural ecosystems.

6. FUTURE ENHANCEMENT

- By integrating IOT sensors into the system architecture and not relying solely on user input, instead bringing in data directly from the cloud to the input parameters using node MCU.
- It would be beneficial to further develop the prediction of crop price which would include a wide variety of datasets, methods for forecasting, and implementation.
- Implementation of GPS sensors for obtaining the location and inputting it directly from the API.

REFERENCES

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

[2] A. 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

[3] G. 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.

[4] W. Mupangwa, R. Makanza, L. Chipindu et al., "Temporal rainfall trend analysis in different agro-ecological regions of southern Africa," Water SA, vol. 47, no. 4, pp. 466–479, 2021.

[5] A. 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.

[6] E. Manjula, S. Djodiltachoumy, "A Model for Prediction of Crop Yield" International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017.

[7] Garg, A., Garg, B. (2017). A robust and novel regression based fuzzy time series algorithm for prediction of rice yield. 2017 International Conference on Intelligent Communication and Computational Techniques (ICCT).

[8] S. Imran, "Effective crop selection and conservative irrigation using IoT," International Journal of Science and Research (IJSR), 2016.