International Journal For Technological Research In Engineering
Volume 6, Issue 7, March-2019 *ISSN (Online): 2347 - 4718*

IOT BASED CROP PREDICTION BY MONITORING TEMPERATURE AND RAINFALL USING DECISION TREE

Rishi Rathnavel.K[1], Sankaran.S[2], Koushik.T[3], Anil Kumar.P[4], Dr.Padmavathi[5]
[1,2,3,4]B.E, [5]Asst Prof, Department of Computer Science and Engineering
R.M.K College of Engineering and Technology, Puduvoyal, Chennai, TamilNadu, India.

ABSTRACT: Agriculture is the most important contributor to the Indian financial system. The agriculture crop production depends on the climatic, biological and economic factor. Today, agriculture, farmers are not only creating yield but also create the agriculture data. This guides to new methods and techniques such as data mining that can connect the knowledge of the data to the crop yield evaluation. To improve the crop yield, choose the best crop, thereby increases the quality and profitability of the agriculture sector using data mining techniques. An enormous amount of data is collected in agriculture and different data mining techniques are used to find efficient crop as per the climatic condition. Precision agriculture is already adopted in other countries, but we still need to involve IoT and cloud computing technologies for better production of crops. At present the climate differs in many areas around India due to various factors from human activities such as air pollution, deforestation, sewage and from natural changes such as distance of sea, wind direction, proximity to the equator. As per the changes in the climate, a farmer needs to predict which crop should be cultivated at which time. The dataset stores the details of the crop which should satisfy the requirements such as maximum and minimum temperature, maximum and minimum rainfall and location.

This research work presents a detailed analysis of improving crop yield by predicting the crop for the respective location as per the climatic changes. The current temperature and rainfall range data can be collected by using DHT11 Temperature Sensor and Soil Moisture Sensor connected to Raspberry Pi. The collected data's (location, temperature value and rainfall range) are stored in AWS IoT. Connections with remote locations can easily achieved by using messaging protocol such as MQTT (Message Queue Telemetry Transport). The Decision trees are versatile Machine Learning algorithm that can perform both classification and regression tasks in predicting the crop to be cultivated in a corresponding location as per the climatic changes. The caret package (Classification And REgression Training) is a set of functions that attempt to streamline the process for creating predictive models and to construct a decision tree which has a crop as a target field. Amazon QuickSight reads data from AWS storage services to provide ad-hoc exploration and analysis view by comparing the temperature and crop of corresponding location.

Keywords: Raspberry Pi, MQTT,DHT11, AWS IOT, Quick Sight, AWS, Machine Learning

I. INTRODUCTION

The process of classification is used to construct a model which helps to describe data classes. Classification is a process of the supervised machine learning algorithms. The objective of the classification model is to describe and distinguishes data classes. In classification, two sets are more important such as training set and testing set. The training set is used to build a model with the training data. Testing set is applied to the classification model and used to check the accuracy. The goal of the classification is to make a classifier based on some cases with some attributes to describe the objects or one attribute to describe the group of objects.

In India there are several agriculture crop productions and those crops depends on the various kinds of factors such as biology, economy and also the geographical factors. Applying such methodologies and techniques on historical yield of various crops, it is possible to obtain information or knowledge which can be helpful to farmers and government organizations for making improved decisions and for make better policy which help to improve production. In this research work, the main focus is on the application of data mining techniques which is used to extract knowledge from the agriculture data to estimate better yield.

In this dissertation, a current development in machine learning is discovered for classification problem. The objective of this dissertation is to classify the crops and solution to increase their yields using classification algorithms

II. PROBLEM DESCRIPTION

Nowadays, farmers are struggling to produce the yield because of unpredictable climatic changes, decrease / increase in rainfall and drastically decrease in water supply. So that an agricultural data has been collected, stored and analyzed for useful information. It is used to promote new advanced methods and approaches such as data mining that can give the information of the previous results to the crop yield estimation.

The analysis of a huge set of agriculture data is the major challenge in the agriculture and analyzed for useful information classify the agriculture crops based on temperature changes and rainfall received range. The relevant features are selected and the crops are classified based on location. In this research work, the crop yield is estimated, and the most excellent crop can be chosen by analyzing the climatic changes with the previous year data.

The climatic changes for the present year can't be known to overcome this problem, sensors are placed to find the temperature and rainfall received data's. The cloud AWS

interacted with Raspberry Pi in storing the captured values for analyzing purposes.

III. ID3 ALGORITHM

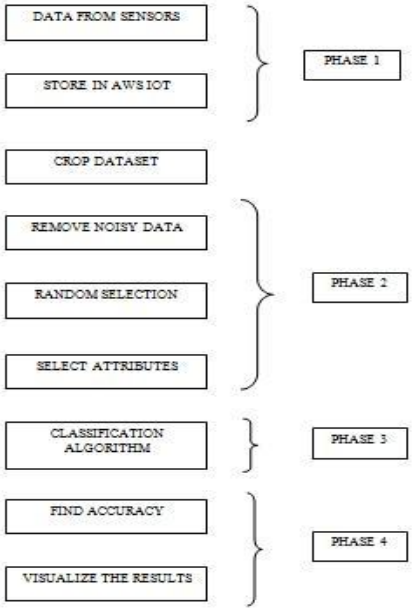

Figure 1: Classification Block Diagram

In decision tree learning, ID3 (Iterative Dichotomiser 3) is an algorithm used to generate a decision tree from a dataset. ID3 is the precursor to the C4.5 algorithm, and is typically used in the machine learning and natural language processing domains. The ID3 algorithm it takes rainfall receives for the respective season as the root node. The entropy (or information gain) can be calculated on each iteration of the algorithm, during this process every unused attribute are used for the construction of tree. Next the process will select the attribute which has the smallest entropy (or largest information gain) value. The selected attributes are grouped and form a set to produce subsets of the data. The decision tree is constructed with each non-terminal node (internal node - temperature, season) representing the selected attribute on which the data was split, and terminal nodes (leaf nodes - crop) representing the class label of the final subset of this branch.

3.1. Information Gain

The information gain IG can be measured as
[IG(T,a)=H(T)-sum_{vin vals(a)}frac{|{textbf{x}in T|x_a=v}|}{|T|} cdot H({textbf{x}in T|x_a=v})]

The information gain measures the *difference* between the entropy *before the split*, and the weighted sum of the entropies *after the split.*

3.2. IG in R

```
InformationGain <- function( tble ) {
tble <- as.data.frame.matrix(tble)
entropyBefore <- Entropy(colSums(tble))
s <- rowSums(tble)
entropyAfter <- sum (s / sum(s) * apply(tble, MARGIN = 1, FUN = Entropy ))
informationGain <- entropyBefore - entropyAfter
return (informationGain)
}
```

IV. DATA SET

The agriculture database contains details about the crop growth of previous years with climate and rainfall details. This data set contains number of attributes such as state of India, district of respective state, crop year, season, annual rainfall received on the corresponding season, minimum and maximum temperature during that season and crop get cultivated. The crop production data set was referred from data world website. The crop production for the duration between 1997 and 2011 was maintained in the dataset. As per the climatic changes, the crop cultivation changes from season to season. By analyzing all those previous year crop production data, choosing a crop as per the climatic changes can be predicted.

V. EXPERIMENTAL RESULTS

5.1 Predict Temperature

The DHT11 sensor is capable of measuring both temperature and relative humidity and provide fully calibrated digital outputs. Its temperature measuring range is from -40 to +125 degrees Celsius with +-0.5 degrees accuracy. The humidity measuring range, from 0 to 100% with 2-5% accuracy.

Figure 2: DHT11 Sensor with Raspberry Pi

5.2. Predict Rainfall Range

The Soil Moisture Sensor uses capacitance to measure dielectric permittivity of the surrounding medium. In soil, dielectric permittivity is a function of the water content. The sensor creates a voltage proportional to the dielectric permittivity, and therefore the water content of the soil. The sensor averages the water content over the entire length of the sensor. There is a 2 cm zone of influence with respect to the flat surface of the sensor, but it has little or no sensitivity at the extreme edges. It helps to predict three level of soil content (dry soil, humid soil and wet soil)

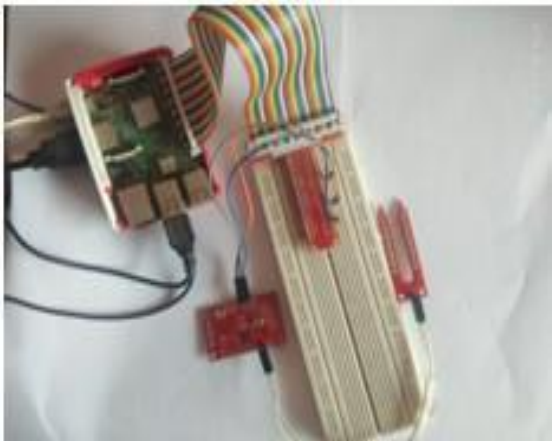

Figure 3: Soil Moisture Sensor with Raspberry Pi

5.3. Discovering Knowledge

In decision tree learning, ID3 (Iterative Dichotomiser 3) is an algorithm used to generate a decision tree from a dataset. ID3 is the precursor to the C4.5 algorithm, and is typically used in the machine learning and natural language processing domains. The ID3 algorithm begins with the original set as the root node (rainfall). The decision tree is constructed with each non-terminal node (internal node) representing the selected attribute on which the data was split, and terminal nodes (leaf nodes - crop) representing the class label of the final subset of this branch.

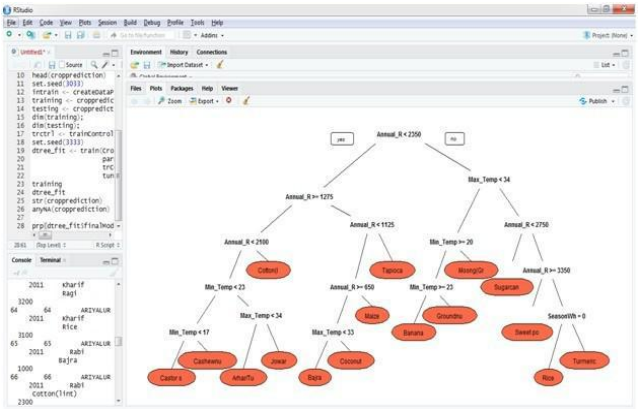

Figure 4 Decision Tree with Crop as a Target Node

VI. COMPARATIVE RESULTS

Amazon QuickSight is a fast, cloud-powered BI service that makes it easy to build visualizations, perform ad-hoc analysis, and quickly get business insights from huge data. Using cloud-based service user can easily connect to the data, perform advanced analysis, and create stunning visualizations and rich dashboards that can be accessed from any browser or mobile device. Amazon QuickSight also referred as Super-fast, Parallel, In-memory, Calculation Engine (SPICE) uses a combination of columnar storage, in-memory technologies enabled through hardware, machine code generation, and data compression to allow users to run interactive queries on large datasets and get rapid responses.

The crop get compared with temperature changes and season using Amazon QuickSight.

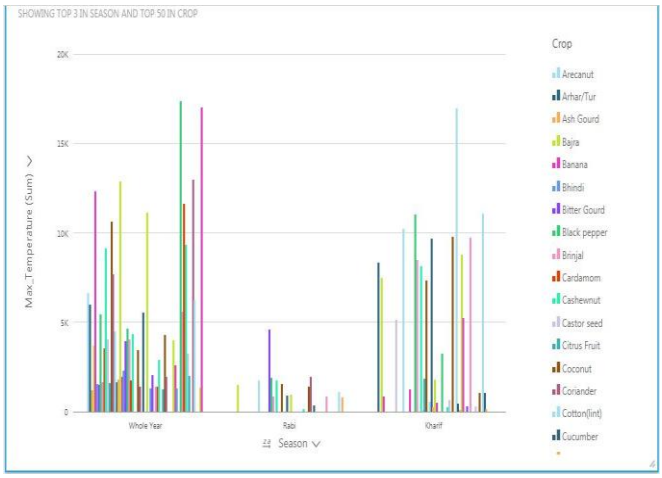

Figure 5 Comparison of season and temperature

VII. CONCLUSION

Data mining is the latest research area of agriculture. This is reasonably a fresh research field and it is projected to grow in the future. Nowadays, farmers are struggling to produce the yield because of unpredictable climatic changes and range of rainfall. The experiments conducted to analyze a small number of traits contained within the dataset to decide their effectiveness when compared with standard statistical techniques. Feature selection is very important to classify the crop data. In this research work, decision tree concept was used to build a model in finding a solution of which crop is better for cultivation. ID3 algorithm is used to classify the crops and a solution to increase their yield by analyzing the climatic changes. This work helps to predict the best crop for analyzing the temperature and rainfall. This will help farmers to increase the yield and income level. The obtained accuracy of ID3 algorithm is 95.83 % and it is better than other algorithms.

REFERENCES

[1] Aurobindo Sarkar and Amit Shah, "Learning AWS: Design, build, and deploy responsive applications using AWS Cloud components", pp.10–21, February 2018.

[2] Jyotshna Solanki and Yusuf Mulge, "Different Techniques Used in Data Mining in Agriculture" International Journal of Advanced Research in Computer Science and Software Engineering, pp. 1223-1227, May 2015.

[3] R. Kalpana, N. Shanthi and S. Arumugam, "A Survey on Data Mining Techniques in Agriculture" International Journal of Advances in Computer Science and Technology, Vol. 3, No. 8, August 2014, pp. 426-431.

[4] A.K. Yegna Narayan Aiyer, "Field Crops of India", pp. 10-117, 1958.

[5] Deepak B. Andore, "AWS IOT Platform based Remote Monitoring by using Raspberry Pi" International Journal of Latest Technology in Engineering, Management & Applied Science, Vol. IV, Issue X, October 2017, pp. 38-42.

