WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Crop Recommendation and Yield Prediction using ML Techniques

Amulya K. S[1], Radhika K. P[2]

[1,2] B.E., M. Tech., Assistant Professor, CSE Department, Cauvery Institute of Technology, Mandya

ABSTRACT

Agriculture has the largest contribution in the GDP of our country. But still the farmer's don't get worth price of the crops. It is mostly happens due to improper irrigation or inappropriate crops selection or also sometimes the crop yield is less than that of expected. By analyzing the soil and atmosphere at particular region best crop in order to have more crop yield and the net crop yield can be predict. This prediction will help the farmers to choose appropriate crops for their farm according to the soil type, temperature, humidity, water level, spacing depth, soil PH, season, fertilizer and months. This prediction can be carried out using Random Forest classification machine learning algorithm. The prediction of crop yield has direct impact on national and international economies and play important role in the food management and food security. Deep learning gains importance on crop monitoring, crop type classification and crop yield estimation applications with the recent advances in image classification using Deep Convolutional Neural Networks. Traditional crop yield prediction approaches based on remote sensing consist of classical Machine Learning methods such as Support Vector Machines and Decision Trees. India being an agriculture country, its economy predominantly depends on agriculture yield growth and agro industry products. Data Mining is an emerging research field in crop yield analysis. Yield prediction is a very important issue in agricultural.

Keywords-- Deep Learning, CNN, SVM and Decision Tree

INTRODUCTION

The Data Analysis is process of inspecting cleansing, modelling data with the goal of discovering useful information and conclusions. It is a process of analysing, extracting and predicting the meaningful information from huge data to extract some pattern. This process is used by companies to turn the raw data of their customer to useful information. This analysis can also be used in the field of Agriculture.

It is mostly happens due to improper irrigation or inappropriate crops selection or also sometimes the crop yield is less than that of expected. Agricultural researchers insist on the need for an efficient mechanism to predict and improve the crop growth and Majority of research works in agriculture focus 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.

By analysing the soil and atmosphere at particular region best crop in order to have more crop yield and the net crop yield can be predict. This prediction will help the farmers. To choose appropriate crops for their farm according to the soil type, temperature, humidity, water level, spacing depth, soil PH, season, fertilizer and months.Crop yield estimation is a difficult task since it is affected by various factors such as genetic potential of crop cultivar, soil, weather, cultivation practices (date of sowing, amount of irrigation and fertilizer, etc.) and biotic stress . Several methods of crop yield. estimation have been developed such as statistical, agro-meteorological, empirical, biophysical, mechanistic.

Problem Statement

In India farmers still follow the traditional technology which they adopted from their ancestor. But the problem is that in earliest time climate was very healthy everything was happened on time. But now most of the things have been changed due to global warming and many other factors. The main problem with agriculture in India is lack of rainfall in seasonal time. Humidity is also necessary for crops but it has been excessive, it also converts as drawback. Winter season is been affected so Rabi crops are widely affected. Since few years the rainfall in winter season was high as expected.

To overcome these above issues we need to develop a system which will able to find the hidden facts or results, patterns and insights. The farmer can predict which crop should sow so that he/she can get more benefit. In proposed system we are applying data analytics techniques on agriculture production baseddatasets and find the insights so that it can help to the farmers and their decision making.

Existing System

This section provides the proposed methodology used for crop yield prediction. The purpose of crop yield prediction is to estimate production in agriculture sector for better crop management and make strategic decisions for improvingcrop yield in future. The Existing model can be incorporated with a decision support system (DSS) that canbe used in precision agriculture which aims at complete farm management.

Proposed System

A program, software component, or other computer system is created by programming and deploying a technical specification or algorithm. In the case of a given interface, there are likely to be a number of observations. The class that implements an interface is essentially its implementation of the interface, so if the interface does not specify any methods, then the class also includes methods that implement the methods specified in the interface In order to implement a system successfully; interrelated tasks must be performed in an appropriate order.

System Analysis Design

System design is the process of the defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements .Systems design could be seen as the application of systems theory to product development.

Data Flow Diagram

DFDs can also be used for the visualization of data processing. A DFD shows what kind of information will be input to and output from the system, how the data will advance through the system, and where the data will be stored

DFD Level One

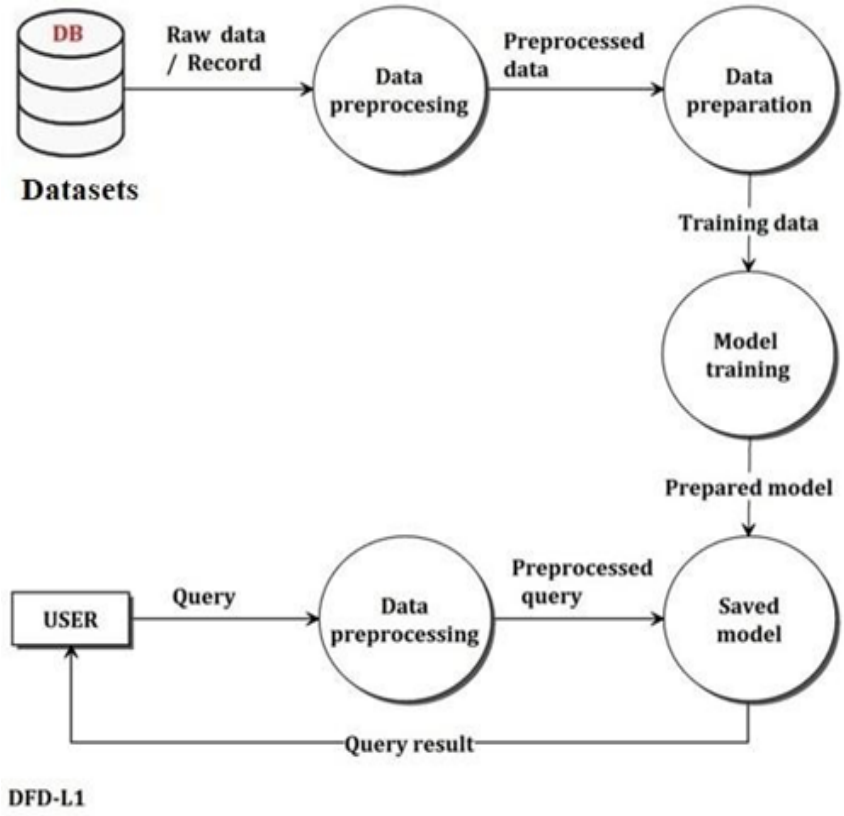

Fig 1 : DFD Level One

DFD Level Two

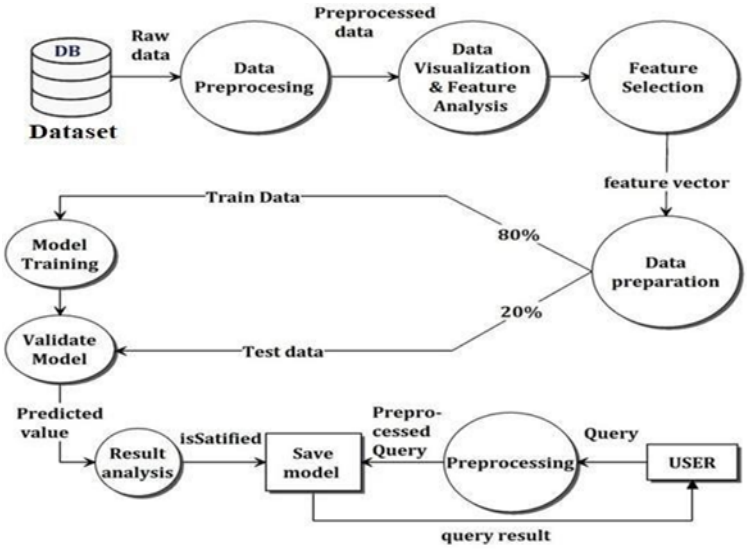

Fig 2 : DFD Level Two

Architecture Design

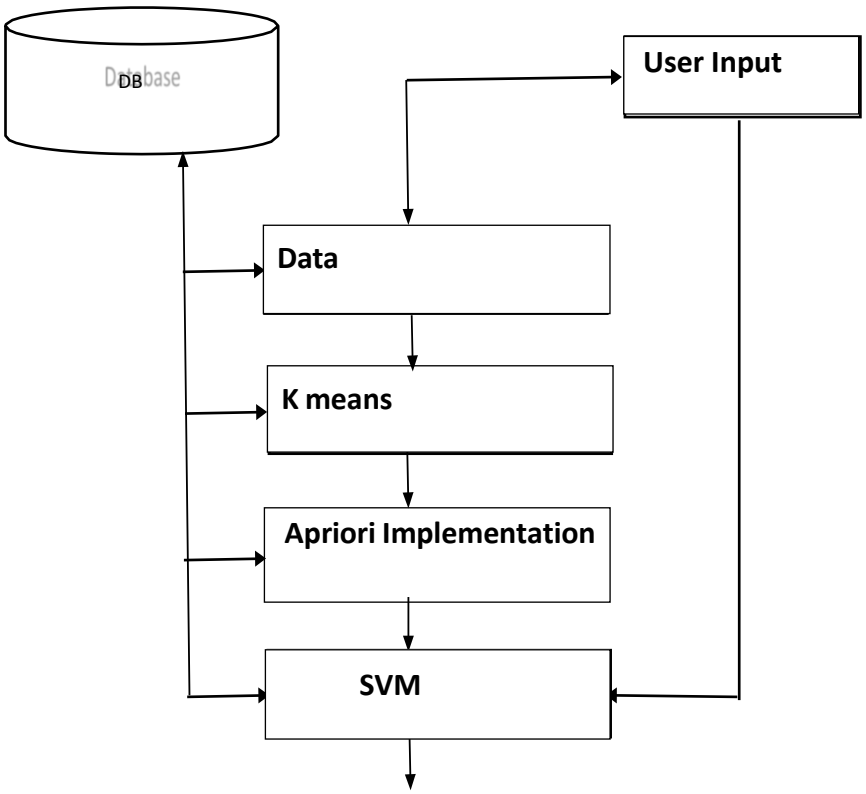

Fig. 3: Architecture Design of Project

Implementation:

This project is used to predict the crop yield and suitable crop by considering the information such as soil type, temperature, humidity, season, fertilizer and months. The system provides easier and faster access to All the basic information regarding the District, rainfall, area under irrigation, crop, season yield, fertilizers used through which user can analysis the crop and also select the option of Prediction where he can select the crop production parameters to get the suitable crop for his farm. This system provides simple visualization so that user can understand and analysis things in easy way.

Algorithm used

Machine Learning Models:

Classification Machine learning technique for dataset classification is useful, according to the examination of datasets. The Classification algorithm is a Supervised Learning technique that determines the category of new observations using training data.

a) **Decision Tree Classifier:** The Decision Tree algorithm is one of the supervised learning algorithms family. The decision tree approach, unlike other supervised learning algorithms, can also be used to solve regression and classification issues By learning simple decision rules inferred from past data, the purpose of employing a Decision Tree is to develop a training model that can be used to predict the class or value of the target variable (training data).

b) **Support Vector Machine:** The SVM algorithm's purpose is to find the optimum line or decision boundary for categorising n-dimensional space into classes so that additional data points can be readily placed in the correct category in the future

c) **Random Forest Classifier**: " According to the name, "Random Forest" is a classifier that "contains a number of decision trees on various subsets of a given dataset and takes the average to improve the dataset's projected accuracy. Instead, then relying on a single decision tree, the random forest collects the forecasts from each tree and predicts the final output based on the majority votes of predictions.

Hardware Requirements

1) **Arduino UNO:** The Arduino Uno is a microcontroller board that uses the ATmega328 microcontroller (datasheet). It contains 14 digital input/output pins, 6 analogue inputs, a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button.

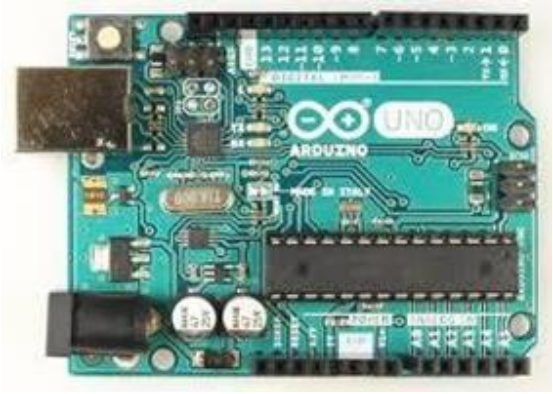

Fig 4: Arduino UNO

2) **DHT11 Sensor**: This sensor is a digital humidity and temperature sensor with a low cost. Despite the use of an ADC, this sensor produces digital output and may thus be directly attached to the data pins of a microcontroller. It also has an eight-bit microprocessor for serial data transmission of temperature and humidity information.

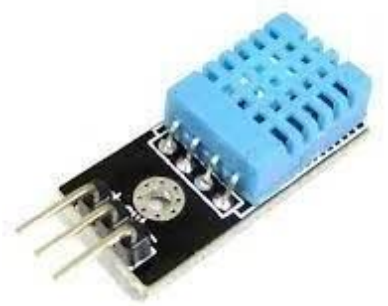

Fig 5: DHT 11 Sensor

3) **Soil Moisture Sensor**: The moisture sensor has three pins: one for voltage, one for ground, and one for analogue input. This sensor measures the moisture content of the soil (in volume percent). Because moisture content is measured in percentages, the analogue value must be mapped to a range of 0-100. The electrical resistance of the soil is the attribute exploited by this sensor. This sensor has two probes that allow the current to travel through the soil.

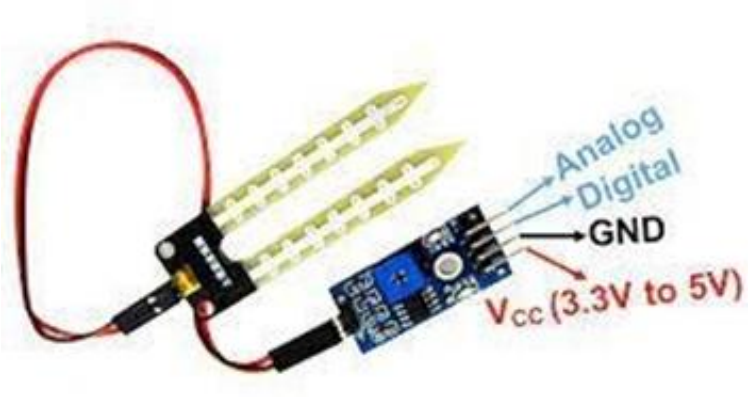

Fig 6: Soil Moisture Sensor

4) **Wi-Fi Module**: The ESP8266 Wi-Fi Module is a self-contained SOC with an integrated TCP/IP protocol stack that can provide access to your Wi-Fi network to any microcontroller. The ESP8266 may either host an application or offload all Wi-Fi networking functionality to a separate application processor.

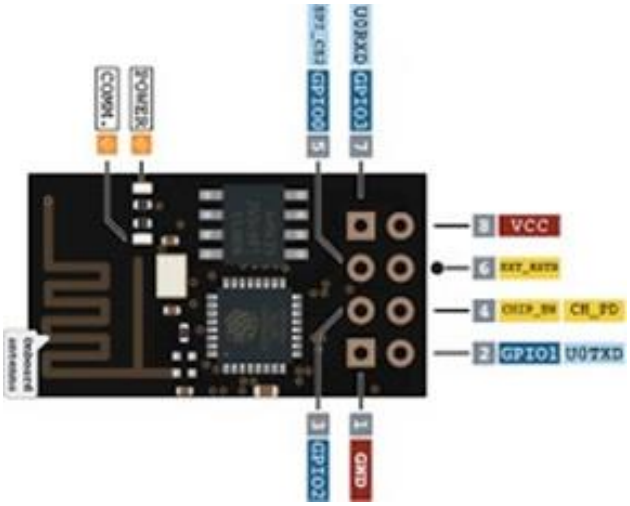

Fig 7: ESP8266 Wi-Fi Module

Software Requirements

1) **Arduino IDE:** Arduino IDE is an open-source programming environment that is used to write and compile code for the Arduino module. It primarily consists of a text editor for writing code, a text console, a message area, and a toolbar with buttons for common functions. The programmes produced using this software are referred to as sketches. This software mostly uses C/C++ functions for coding.

2) **Thingspeak Cloud Server**: It's a free and open-source programme. This platform offers services for visualising, analysing, and aggregating live data flowing on a cloud server. It provides real-time visualisations of data submitted to this cloud server by various devices. It has the ability to run MATLAB code, so we can evaluate and process statistics as they occur in real time. Prototyping is a common application for it.

System Testing

- **Functional Testing**

This is a type of black-box testing that is based on the specifications of the software that is to be tested. The application is tested by providing input and then the results are examined that need to conform to the functionality it was intended for. Functional testing of software is conducted on a complete, integrated system to evaluate the system's compliance

- **Non- Functional Testing**

This section is based upon testing an application from its non-functional attributes. Non- functional testing involves testing software from the requirements which are non- functional in nature but important such as performance, security, user interface, etc.

Loading the image Module Testing

The below Table 6.1 shows the successful test case for loading the Crop Yield detection image that is selected by the user to do the processing.

Test Name	Load the Dataset.
Test Feature	Checks whether the data selected are loaded or not.
Output Expected	The crop yield data to be loaded from the dataset.
Output Obtained	Loaded crop yield dataset which is used for the Processing.
Result	Successful

Table 1: Test case for loading the image

Pre-Processing Module Testing

The original image is given as input. In the preprocessing step noise reduction method is used.

Test Name	Noise Reduction
Test Feature	Data Labelling
Output Expected	Unwanted attributes Removal
Output Obtained	Unwanted data files removed
Result	Successful.

Table 2: Test case for Pre-Processing

Training Module Testing

The below table shows the successful test case for training the Crop Yield dataset. Here in training the image is loaded and then checks whether it is already loaded and trained or not, if not trained then it goes to the process of training. The data selected for training undergoes preprocessing and then feature extraction and then finally the images are trained. These processes are done in Training GUI.

Test Name	Training data.
Test Feature	Training the Crop Yield dataset.
Output Expected	The crop data to be trained.
Output Obtained	Trained dataset
Result	Successful.

Table 3: Training Module Testing

Conclusion

Features, Agriculture is the backbone of counties like India. However, the usage of technology towards agriculture is to be given paramount importance towards agriculture. This paper proposes a system which will help farmers to have an idea of yield estimates based on weather parameters and area under cultivation Using this farmer can make decisions on whether to grow that particular crop or go for alternate crop in case yield predictions are unfavorable. This research work can be enhancing to the next level. We can build a recommender system of agriculture production and distribution for farmer. By which farmers can make decision in which season which crop should sow so that they can get more benefit. This system is work for structured dataset. In future we can implement data independent system also. It means format of data whatever, our system should work with same efficiency.

REFERENCES

[1] Ponce-Guevara, K. L., Palacios-Echeverria, J. A., Maya-Olalla, E.,Dominguez Limaico, H. M., Suarez-Zambrano, L. E., Rosero-Montalvo, P.D., Alvarado-Perez, J. C. (2017). GreenFarm-DM: A tool for analyzing vegetable crops data from a greenhouse using data mining techniques (First trial). 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM).

[2] Jheng, T.-Z., Li, T.-H., Lee, C.-P. (2018). Using hybrid support vector regression to predict agricultural output. 2018 27th Wireless and Optical Communication Conference (WOCC).

[3] Manjunatha, M., Parkavi, A. (2018). Estimation of Arecanut Yield in Various Climatic Zones of Karnataka using Data Mining Technique: A Survey. 2018 International Conference on Current Trends Towards Con- verging Technologies (ICCTCT).

[4] Shakoor, M. T., Rahman, K., Rayta, S. N., Chakrabarty, A. (2017). Agricultural production output prediction using Supervised Machine Learning techniques.2017 1st International Conference on Next Generation Computing Applications (NextComp).

[5] Grajales, D. F. P., Mejia, F., Mosquera, G. J. A., Piedrahita, L. C.,Basurto,C. (2015). Crop- planning, making smarter agriculture with climate data. 2015 Fourth International Conference on Agro-Geoinformatics (Agro-Geoinformatics).

[6] Shah, P., Hiremath, D., Chaudhary, S. (2017). Towards development of spark based agricultural information system including geo-spatial data. 2017 IEEE International Conference on Big Data (Big Data).

[7] Sekhar, C. C., Sekhar, C. (2017). Productivity improvement in agriculture sector using big data tools. 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC).

[8] Afrin, S., Khan, A. T., Mahia, M., Ahsan, R., Mishal, M. R., Ahmed, W.,Rahman, R. M. (2018). Analysis of Soil Properties and Climatic Data to Predict Crop Yields and Cluster Di_erent Agricultural Regions of Bangladesh.2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS).