International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 06 | June 2020 www.irjet.net p-ISSN: 2395-0072

Crop Prediction using Machine Learning

Mythresh A[1], Lavanya B[2], Meghana BS[3], Nisarga B[4]

[1,2,3,4] *Students of Dept. Computer Science and Engineering*
[1,23,4] *K S Institute of Technology, Bangalore, Karnataka, India*

Abstract - *Agriculture is one of the main sources of life in India. Without agriculture, the economy of the country will collapse. For decades, agriculture has been associated with the production of essential crops. Agriculture is the main source of national income for most developing countries and India being a developing country requires agriculture for its sustainability. Agriculture contributes around 20% to the national GDP. Our country is also progressing in the field of science and technology. The main objective that we are trying to achieve here is modernization of agriculture using machine learning concepts. There are many modern techniques used in agriculture but most of them require capital investment or IoT devices. We are trying to help farmers in making decision on which crop to grow so that optimum profit is obtained with minimum investment or by cutting down unnecessary investment. Naïve – Bayes algorithm is used to predict the crop. Considering factors such as soil, weather, temperature, rainfall etc., a machine learning model is designed. The developed model after training will predict crops according to the suitability of the land. The UI is made user friendly for easy interaction of the farmer with the android application.*

Key Words: *Agriculture, Crop prediction, Naïve-Bayes, Android application, Soil site suitability, Crop requirement*

1. INTRODUCTION

India is an agricultural country and its economy is predominantly dependent on agriculture yield and related industry products. Modernization of agriculture plays a key role in improving the economy of the country and thus will also lead the farmers of our country towards profit. Machine Learning techniques can be used in predicting crops. ML is still an emerging and challenging research field in agricultural data analysis.

1.1 Outline of the proposed system

Productivity of a particular crop greatly depends on land resources and the climate of the area along with other factors such as fertilizers. Identification of crop requirements and matching them with the resources available to optimize the productivity in sustainable manner assumes a greater importance. Crop management practices based on soil site suitability criteria and weather conditions will help to overcome this.

The proposed system will integrate the data obtained from repository, weather department and user inputs. A machine learning model is developed considering the various, different sets of data to obtain the output. The system takes input from various sources and repositories for weather, soil and crop requirement data and uses Naïve-Bayes algorithm to predict the best suitable crop for any given area. The developed user interface is flexible and highly interactive which will encourage the farmers to use this mobile application frequently

2. EXISTING METHODS

2.1 Machine Learning

Machine learning algorithm, Multiple Linear Regression is used by the system to predict the crop. The prediction is based on past production data of crops including identifying the suitable weather and soil parameters and comparing it with current conditions which will predict the crop more accurately and in a practical manner. Also, random forest method is used to predict the crops in some cases. [1]

2.2 Sensors

To modernize the agriculture, usage of IoT devices is a necessity. Many existing systems rely on sensors for data. [2] Our system requires no sensors or devices thereby reducing manual intervention as most of the processes are automated.

3. DATASETS

The main requirement for the proposed system is datasets. Along with inputs from user which includes mobile number and location access, two datasets are considered. The first dataset has details regarding soil site suitability criteria and the second dataset has details about the requirements required for the crop growth. The datasets are collected from NBSS and LUP. NBSS and LUP stands for National Bureau of Soil Survey and Land Use Planning.

3.1 Soil site suitability dataset

Soil site suitability dataset has details of the soil of the location. The dataset collected is from Hadonahalli Grama Panchayat, Doddaballapura Taluk, Bangalore rural District. Doddaballapur is a City municipal council in Bangalore Rural district in the state of Karnataka, India. Hadonahalli is an agriculture-based village. The prominent agricultural areas are divided and are named according to survey IDs. This dataset contains details about the soil texture, slope, erosion, gravel, rocky, depth, drainage and pH mapped to survey IDs.

3.2 Crop requirement dataset

Crop requirement dataset will have details about the conditions required for the growth of the crops. According to the climatic conditions of this location, there are 5 non-perishable crops that are prominently grown. The crops grown are Groundnut, maize, finger millet, castor and rice. For developing the prototype of the model, these are the crops considered

4. MACHINE LEARNING MODEL

Many machine learning models are compared for the obtained datasets. This is the box plot showing the best classifier for our dataset and the scores for it. As seen, Gaussian Naive Bayes Classifier is the algorithm that is best classifying our dataset with more than 60% accuracy.

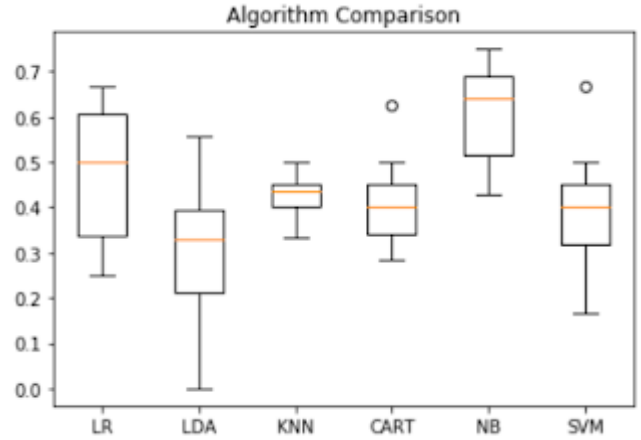

Fig 4(a) : Box Plot

```
LR: 0.476014 (0.145475)
LDA: 0.288214 (0.172097)
KNN: 0.430534 (0.047816)
CART: 0.419304 (0.105761)
NB: 0.608001 (0.100412)
SVM: 0.399304 (0.128136)
```

Fig 4(b) : Naïve-Bayes with 60% accuracy

A Naive Bayes Classifier is a supervised machine-learning algorithm that uses the Bayes' Theorem, which assumes that features are statistically independent. The theorem relies on the naive assumption that input variables are independent of each other, i.e. there is no way to know anything about other variables when given an additional variable. Regardless of this assumption, it has proven itself to be a classifier with good results.

Naive Bayes Classifiers rely on the Bayes' Theorem, which is based on conditional probability or in simple terms, the likelihood that an event (A) will happen given that another event (B) has already happened. Essentially, the theorem allows a hypothesis to be updated each time new evidence is introduced. The equation below expresses Bayes' Theorem in the language of probability:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Where,

- "P" is the symbol to denote probability.
- P(A | B) = The probability of event A (hypothesis) occurring given that B (evidence) has occurred.
- P(B | A) = The probability of the event B (evidence) occurring given that A (hypothesis) has occurred.
- P(A) = The probability of event B (hypothesis) occurring.
- P(B) = The probability of event A (evidence) occurring.

5. WORKFLOW OF THE SYSTEM

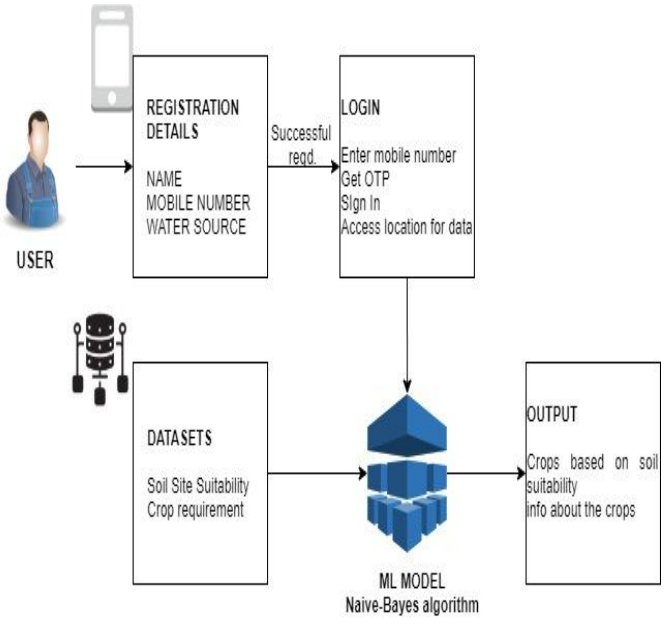

Fig 5 : Workflow

In the front end, the user will register initially providing all the necessary details. For successful registration, the user has to provide his name, mobile number and the water source which he is dependent on. Once the registration is done, the farmer can login into the system through OTP. OTP stands for one-time password which will become invalid after some time. This allows hassle free login as compared to password login.

In the back end, considering all the data from datasets, location and weather details, a machine learning model is designed. Naïve-Bayes algorithm is used to predict the output. Crops are predicted based on soil suitability. The user can choose any one among the list of crops predicted. Also, the information regarding each predicted crop is displayed. The UI is very interactive and will work in the users' native language.

6. IMPLEMENTATION AND RESULTS

6.1 IMPLEMENTATION

The developed system currently works only in Hadonahalli, as the dataset is confined to this location. The only requirement from user end is a smart phone which supports android application and can access location through GPS.

6.2 RESULT

Fig 6.2 : Output

This is the snapshot of the android mobile application. Here, for the survey ID 37, the possible crops that can be grown are groundnut, castor and sugarcane.

MODULE	GIVEN INPUT	EXPECTED OUTPUT	ACTUAL OUTPUT
REGISTRATION	User details including name, mobile number and access to GPS	On giving details, the data should be stored in the back end	Registration successful
LOGIN	Once registration is done, Login is done by entering the mobile number in the app for which an OTP will be sent	OTP is sent to the registered mobile number. The OTP will be automatically fetched from SMS	Login successful
TRAINING DATA	Two sets of datasets, soil site suitability and crop requirement datasets, and user input are considered	ML algorithm is applied, machine is trained to predict the expected output	Training successful
PREDICTION	Real time data which includes location and weather details	Machine will predict the output	Prediction successful with accepted accuracy
RESULT	User info	Crops that can be grown along with its suitability for the location	Output displayed

Table 6.2: Testing results

7. FUTURE ENHANCEMENT

Collect sufficient amount of data from all the districts and states of India, so that the application can be used by farmers all over the country. Also improve model efficiency. There should be no language barrier, the app should be flexible to everyone irrespective of the diverse nature of this country

REFERENCES

1. Crop prediction system using machine learning Prof. D.S. Zingade, Omkar Buchade, Nilesh Mehta, Shubham Ghodekar, Chandan Mehta, Department of computer science and engineering, All India Shri Shivaji Memorial Society's Institute of Information Technology
2. Machine learning approach for Crop selection based on agro-climatic conditions
3. Prof K.D Yesugade, Aditi Kharde, Ketki Mesati, Dept of computer science engineering, BVCOEW, Pune, India
4. Rice crop yield prediction in India using SVM Niketa GandhiUniversity of Mumbai
5. Soil Classification using machine learning methods and crop suggestion based on soil series.
6. Sk Al Zaminur Rahman, Kaushik Chandra Mitra, S.M. Mohidul Islam, Computer science and engineering, Kuhana University