

Crop Prediction System using Machine Learning

Miss Vaishali Patil

Pro.Sachin Deshpande

Department of Computer Engineering
Vidyalankar Institute of Technology,
Mumbai, India

Department of Computer Engineering
Vidyalankar Institute of Technology,
Mumbai, India

Abstract—Looking at the current situation faced by farmers, we have observed that there is an increase in suicide rate over the years. The reasons behind this situation are weather conditions, debt, family issues. Sometimes farmers are not aware about the crop which suits their soil quality, soil nutrients and soil composition and rain fall value. The work proposes to help farmers check the soil quality depending on the analysis done based on data mining approach. Thus the system focuses on checking the soil quality and rainfall value to predict the suitable crop for cultivation and that can maximize the crop yield by using machine learning algorithm like decision tree algorithm and SMA.

Keywords— Soil, Crop, Machine learning, Predication, SMA

I. INTRODUCTION

The agriculture is one of the most important occupation in India. It is the big sector and plays an important role in overall development of the country. The type of soil and the type of crops becomes very significant from agriculture income point of view. The agriculture is very important and thus will lead the farmers of our country towards profit. Data analytic (DA) is the process of examining data sets in order to draw conclusions about the information they contain, increasingly with the aid of specialized systems and software. Earlier yield prediction was performed by considering the farmer's experience on a particular field and crop. As the conditions changes every day very rapidly, farmers are forced to cultivate more and more crops. Being this as the current situation, many of them farmer don't have enough knowledge about the new crops and are not completely aware of the benefits they get while farming them. Also, the farm productivity can be increased by understanding and forecasting crop performance in a variety of environmental conditions. Thus, the proposed system takes the location of the user and soil composition as an input. The processing part also takes into consideration two more datasets. The proposed system applies machine learning and prediction algorithm to identify the pattern among data and then process it as per input conditions. This in turn will propose the best alternative crops according to given conditions. Thus, this system will only require the location of the user and it will suggest number of profitable crops providing a choice directly to the farmer about which crop to cultivate and more information about crop also.

Machine Learning is a field of Computer Science, where new developments evolve at recent times, and also helps in automating the evaluation and processing done by the mankind, thus reducing the burden on the manual human power. Machine learning is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed.

Machine learning focuses on the development of computer programs that can change when exposed to new data. Finding out the suitable crops based on the soil's appearance becomes tedious for novice farmers. There also exists a need to prevent the agricultural decay.

Effective utilization of agricultural land is crucial for ensuring crop production of a country. Machine learning algorithms use for crop prediction system.

Machine learning tasks are typically classified into three broad categories, depending on the nature of the learning "signal" or "feedback" available to a learning system.

Broadly there are 3 types of Machine Learning algorithms:

Supervised Learning: The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule that maps inputs to output. Examples of Supervised Learning: Decision Tree, support vector machine, Random Forest, linear Regression etc.

Unsupervised Learning: In this algorithm, we do not have any target or outcome variable to predict/estimate. It is used for clustering population in different groups, which is widely used for segmenting customers in different groups for specific intervention. Examples of Unsupervised Learning: Apriori algorithm, K-means.

Reinforcement Learning: Using this algorithm, the machine is trained to make specific decisions. It works this way: the machine is exposed to an environment where it trains itself continually using trial and error. This machine learns from past experience and tries to capture the best possible knowledge to make accurate business decisions. Example of Reinforcement Learning: Markov Decision Process.

II. LITERATURE SURVEY:

[Mucherino, et. al. [9]] Agriculture sector plays a major role in Indian economy, as 70% households in India depends purely on this field. Agriculture in India contributes to about 17% of Gross Value Added as of 2015-16. But there is a continuous decline in agriculture's contribution to Gross Value Added. Food is essential for life and we depend on agricultural outputs, so farmers play a very important role.

The study in [P. Vinciya, et.al. [13]] used Multiple Linear Regression (MLR) technique for crop analysis. Decision tree algorithm and Classification is used to perform analysis of over 362 datasets and provide result. The training dataset here is classified into as organic, inorganic and real estate for predicting the type of soil. Results computed by this system are accurate as well as reliable.

The study in [Shivnath Ghosh, et. al [14]] fed data to a Back Propagation Network to evaluate the test data set. Back Propagation Network uses a hidden layer which helps in better performance in predicting soil properties. Back Propagation Network here, is employed to develop a self-trained function to predict soil properties with parameters. This gives more accuracy and performs better than the traditionally used methods, however, sometimes the system becomes slow and inconsistency is seen in the output.

In [Zhihao Hong, et.al [15]] two regression supervised machine learning methods are used: Support Vector Machine (SVM) and Relevance Vector Machine (RVM) to show effectiveness in soil quality prediction. a smart wireless device for sensing soil moisture and meteorological data. The wireless device gives an error rate of 15% and 95% accuracy. However, it has not been tested for real time data.

The paper [Sabri Arik, et. al [16]] involves a check for Soil Fertility and Plant Nutrient by using back propagation algorithm. The results are accurate and enable improvement in soil properties. It performs better as compared to traditional methods. However, system is slowing inefficient and not stable.

According to paper [Vaneesbeer Singh, et.al [17]], three methods are used which includes Decision tree, Naive Bayes Classifier, and KNN Classifier which analyses soil and predicts crop yield, However rule based induction and SVM can be used for more accuracy as results are not accurate.

This [Jeetendra Shenoy, et.al [2]] paper posits a solution to reduce the transportation cost and uses an IOT based approach in order to reduce the number of middle hops and agents between the farmers and end user which in turn will help the farmer. This paper turns out to be the motivation for our project.

In [M.R. Bendre, et.al [10]] order to predict the yielding of the crops, the crops are analysed and based on analysis they are categorized. This categorization is done based on data mining algorithms. This paper gives insight into various classification rules like Naive Bayes, K-Nearest Neighbour. Using this paper, we analysed the classification rules and identified which will be appropriate for data set which we will be using in our project.

This [Monali Paul, et.al. [11]] paper provides a smart phone based application which will measure the PH value of the soil, temperature and humidity in real time. The system uses a micro-controller block, sensing block and communication block. Sensors are employed in farm which can communicate with smartphones using Bluetooth in real time. This paper provides means of remote analysis of soil through various techniques. This paper encouraged us to look for various techniques through which we can transfer the data we will obtain from sensors for processing and eventually generating the output

This [Dr.N.Suma, et.al. [12]] paper theorizes an Explorative data analysis and discusses about designing of various predictive model. A sample data set is taken and various regression techniques are applied in order to identify and analyse the properties of each. Various regression techniques which are discussed in this paper are Linear, Multiple Linear, non-Linear, Logistic, Polynomial and Ridge regression. Using this paper, comparative study of various data analytics algorithm is obtained. This helps us to judge which algorithm best suits our proposed system.

III. EXITING SYSTEM

Machine learning in Agriculture is a Novel field still a lot of work has been done in field of Agriculture using Machine learning.

[R.Sujatha et.al.[1]] proposed a yield prediction model which used Data Mining techniques for classification and Prediction. This model worked on input parameters crop name, land area, soil type, soil pH, pest details, weather, water level, seed type and this model predicted the plant growth and plant diseases and thus enabled to select the best crop based on weather information and required parameters.

[Shweta Taneja et.al.[7]] proposed an approach which used unsupervised learning technique K Means Clustering technique to classify the soils into clusters based on the salinity factors. This work classified the soils as SodicSaline -Sodic and Acidic. This model enabled the analysts to select the best soil for crop productivity.

[D Ramesh et.al. [8]] proposed a crop yield prediction model that implanted two Data mining techniques namely Multiple Linear Regression and Density Based Clustering techniques. The predict ants were Year', 'Rainfall', 'Area of Sowing', 'Yield', 'Fertilizers' (Nitrogen, Phosphorous and Potassium) and Response variable was 'Production'. In Kg/Hectares.

[A. Mucherino et.al.[9]] conducted a study on the different data mining techniques used in Agriculture. The techniques like K Means.KNN,ANN,SVM were studied related to Agriculture field and concluded that these techniques in combination with GPS and Remote sensing techniques can be used to study the characteristics' of soil, classify soils , classify crops and for prediction too.

IV. PROBLEM DEFINATION

A. Problem Statement

More than half of population is dependent on agriculture. However farmers still lack in basic knowledge about their soils, which crop to sow in what type of soil, the efficient use of fertilizers and in addition to it they also lack in advanced techniques of prediction of crop yield at the time of sowing. Prior prediction of crop yield can be helpful in early determination of factors leading to decline in production. It can also be helpful in proper use of fertilizer and pesticides, proper selection of different varieties of crop and in addition to it early prediction gives farmers an option for prior arrangements of storage and marketing and avoids losses. Many researchers have been conducted to develop an efficient method for yield prediction but focus has been always on statistical techniques and not much has been done in machine learning approach.

B. Objective

1) Broad Objective:

In this project, we aim to design expert prediction system which will give best crop and alternate crop to the farmer. Also get the direction for use of fertilizer and pesticide. This system can be used by farmer as well as by agriculture department for better productivity.

2) Specific Objective:

- a) To provide a user-friendly GUI to the farmer for handling.
- b) To identify best crop for soil and rain condition.
- c) To identify the possible alternate crop and increase productivity.
- d) To enable the farmer a platform for getting more information about crop like fertilizer and pesticide.
- e) To develop crop prediction system for farmer.

V. PROPOSE SYSTEM

The system aims to help farmers to cultivate proper crop for better yield production. To be precise and accurate in predicting crops, the project analyze the nutrients present in the soil composition and rainfall value for the crop productivity based on location. It can be achieved using unsupervised and supervised learning algorithms, like

Decision Tree and Simple moving average. Some datasets are use like crop’s composition, rainfall value, and crop fertility.

A. Features of propose system—

- 1) One or more alternate crop recommendation.
- 2) Recommendation of fertilizer and pesticide for crop.
- 3) Basic information about fertilizer and pesticide.

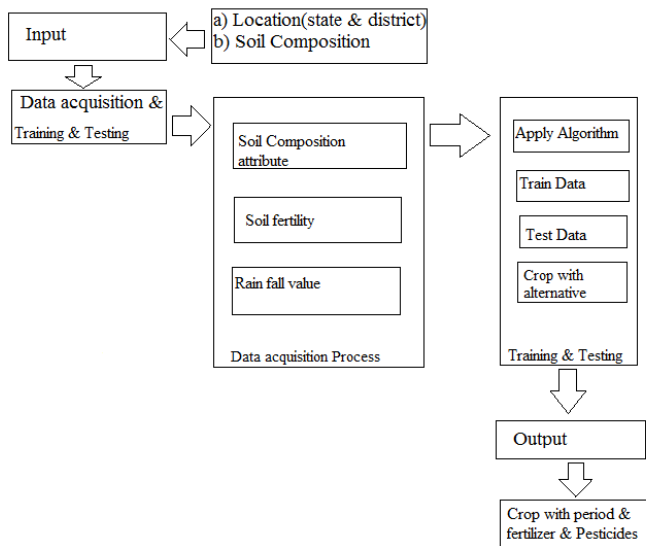


Fig. 1. System Architecture

B. Module Diagram

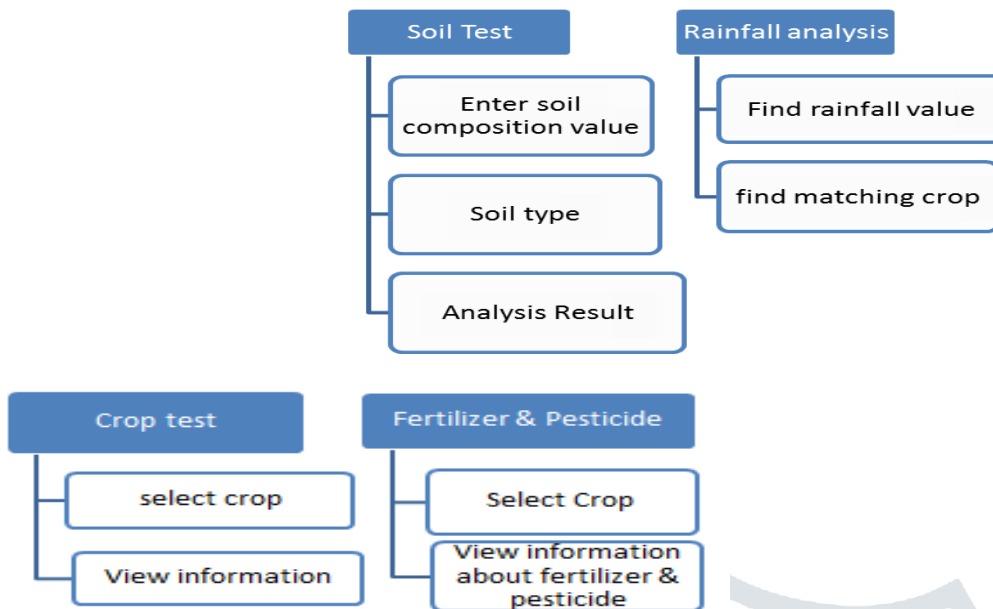


Fig.2. Module Diagram

The system has four modules as shown in above figure. In soil test analysis user enters pH and location. Output of this module is analysis result of the percentage of nutrients in that soil. Soil crop matching module finds the matching crop that could be grown in that soil by comparison with the crop database. In fertilizer and pesticide recommendation module, user is recommended with fertilizer & pesticide that will give the highest crop yield. In the crop test module, user can select a crop and view information about it.

C. Algorithms

Decision Tree Algorithm-

- Need-

To create a training model which can use to predict class of target variables by learning decision rules inferred from training data.

- Advantages-
 - 1) Can handle both categorical and numerical data
 - 2) Simple to understand, interpret, visualize.
 - 3) Decision tree implicitly perform variable screening.

Simple Moving average-

- Need

- 1) To work on historical data
- 2) To work on time series of data

- Advantages-

- 1) This method is applicable to linear and non-linear values
- 2) It shows true average over the time.
- 3) A simple moving average is customizable in that it can be calculated for a different number of time periods.

VI. SUMMARY

The system uses supervised and unsupervised Machine learning algorithms and gives best result based on accuracy. The results of the machine learning algorithms will be one or more alternative crops and their basic requirement. Thus the system will help reduce the difficulties faced by the farmers. It will act as a medium to provide the farmers efficient information required to get high yield and thus maximize profits which in turn will reduce rates of difficulties.

REFERENCES

- [1] R Sujatha, Dr P. Issaki, "A Study of Crop Yield Prediction Using Data Mining Techniques" IEEE 2016.
- [2] Jeetendra Shenoy, Yogesh Pingle, "IOT in agriculture", IEEE 2016
- [3] N. Heemageetha, "A survey on Application of Data Mining Techniques to Analyze the soil for agricultural purpose", IEEE 2016
- [4] https://en.wikipedia.org/wiki/Data_analysis.
- [5] <https://en.wikipedia.org/wiki/Agriculture>
- [6] Essentials of Machine Learning Algorithms. [Online] Available at <https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/> (Accessed on Dec. 9, 2017).

- [7] Shweta Taneja, Rashmi Arora, Savneet Kaur, "Mining of Soil Data Using Unsupervised Learning Technique", International Journal of Applied Engineering Research, ISSN 0973-4562 Vol. 7 No.11, 2012.
- [8] D Ramesh , B Vishnu Vardhan, "Data mining technique and applications to agriculture yield data", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 9, September 2013
- [9] Mucherino, P. Papajorgji, P.M. Pardalos, "Data Mining in Agriculture", Springer, 2009
- [10] M.R. Bendre, R.C. Thool, V.R.Thool, "Big Data in Precision agriculture", Sept,2015 NGCT
- [11] Monali Paul, Santosh K. Vishwakarma, Ashok Verma, "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining approach", 2015 International Conference on Computational Intelligence and Communication Networks.
- [12] Dr.N.Suma, Sandra Rhea Samson, S.Saranya, G.Shanmugapriya, R.Subhashri, "IOT Based Smart Agriculture Monitoring System", Feb 2017 IJRITCC.

