Recommendation System for Crop Identification and Pest Control Technique in Agriculture

Conference Paper · April 2019

DOI: 10.1109/ICCSP.2019.8698099

3 authors, including:

Avinash Kumar
KIIT University
7 PUBLICATIONS **120** CITATIONS
SEE PROFILE

Sobhangi Sarkar
KIIT University
7 PUBLICATIONS **120** CITATIONS
SEE PROFILE

Recommendation System for Crop Identification and Pest Control Technique in Agriculture

Avinash Kumar, Sobhangi Sarkar and Chittaranjan Pradhan

***Abstract*—In Agriculture, mostly crops are loss due to the erroneous selection of the crop to be grown in a particular land holding. The farmers are generally not aware of the requirements of the crop i.e. the minerals, soil moisture and other soil requirements. This can cause a distress to farmer both mentally and financially. One more problem that a farmer generally encounters is the pest and diseases that can affect the crops they grow, which they are generally unaware of in an early stage. This problem of farmer is addressed in our paper and we have tried to solve it with the help of a Recommendation System. By the help of our model, we predict the best suitable crop to the farmer and detect the pest that may affect as well as suggest the pest control techniques. In this paper, we have applied SVM classification algorithm, Decision Tree algorithm and Logistic Regression algorithm and we have found that SVM classification model gives the better accuracy as compared to other algorithms.**

***Index Terms*—Crop Identification, Decision Tree, Logistic Regression, Pest Removal, Recommendation System, SVM Classification Model.**

I. Introduction

AGRICULTURE is an integral sector of Indian Economy. The contribution of Indian agriculture sector to India's Gross Domestic Product (GDP) is about 18% and which in average provides employment to 50% of the country's workforce, which clearly shows that agriculture plays a vital role in the gross economy.Poor result in agriculture can lead to intensification, anguish and discomfort to farmers, and leads to distress both at social and political level — all these reasons can lead to crisis and affect the economy . Recent years have seen the downfall of agriculture. Food and Agriculture Organization of the United Nations (FAO) states that globally one billion i.e. 33% of food or one-third of all food produced for our intake get wasted each year [1-3]. These losses are governed by many factors namely lack of land holding, exhaustion of soil due lack of manures, choice of unsuitable crops, climatic changes, pest, weed etc. Among all these factors, about 30-33% of total yield is lost in India due to pest as stated by P.K. Chakrabarty, Asst. Director General of plant protection and bio safety. So choosing a suitable crop and prevention of those crops hold an equal priority. It is vital as well as challenging as there are many factors that decide the crop that is suitable for growing and there are many insects which may affect the crops differently [4].

With advance in technology, many updated technology has been applied in agriculture sector to improve the health of crops named Precision agriculture. A better name for precision agriculture might be "site-specific agriculture" . Indian farmers tend to choose unsuitable crop for their soil and this problem can be solved by precision agriculture where the soil characteristics like soil type, texture, pH value etc are used for detecting that which crop is suitable for cultivation in that soil. This minimizes the risk of cultivating inexact crop which collectively results in better crop yield from a particular land holding . Once the seeds are sown it's important for farmers to prevent it from insects and weed. Generally Indian farmers prevent their crops by detecting the pest manually and removing them which requires intense labour [5-7]. So, using the technology of recommendation system we can recommend the crops suitable for growing and how to remove the pest which often attacks these crops. With our research work we can help farmers to be more technically sound and they can spread awareness among their fellow farmers about the correct crop to choose and also by preventing crop loss every year, they can be financially stable.

The rest of the paper is organized as follows. Section II describes the Related Work which had been done previously in this field. Section III describes the background technique that had been used in developing the predictive model. Section IV describes about the proposed work which involves various methodologies used. In Section V, we demonstrated our result analysis and finally in Section VI we have derived some conclusions.

II. Related Work

Our S. Babu suggested the requirement of precision agriculture in India through a software model with an aim to reach every small farmer in Kerala region of India [8]. It laid emphasis on the idea of precision algorithm and moved forward to develop software to support it. The cardinal

Avinash Kumar is with School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India (e-mail: avinash1605@gmail.com).

Sobhangi Sarkar is with School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India (e-mail: shobhangisarkar@gmail.com).

Chittaranjan Pradhan is with School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India (e-mail: chitaprakash@gmail.com).

978-1-5386-7595-3/19/$31.00 ©2019 IEEE

objective of this model is to provide advice to the all farmers even the ones having smaller land holdings. This model is primarily developed for Kerala because the average size of land holdings here is smaller than most of India but however the same model can be implemented anywhere with minimal modifications.

S. Pudumalar et. al. discussed the problem farmers face due to inappropriate selection of crop to be cultivated in kind of the soil available to them reason being the lack of appropriate knowledge [9]. They designed a recommendation system that can act as a guide for farmers to select the proper fit crop for the soil. After a series of experiment on the algorithms they have used algorithms like Random Forest, Navie Bayes, CHAID model, K-nearest neighbour and gave a conclusion that the system designed can predict appropriate results to a fair and decent level with an accuracy of 88%. Although they examined the model for a particular state, the model with hardly any modification in dataset and can give a more accurate recommendation for a lager geometric area.

A. Savla et. al. described how pest affect the crops the most and leads to a considerable amount of loss for the farmers and the importance of the proper removal techniques required [10]. The work proposed here is a PCT-O ontology model to describe the occurrence of the pest and how it is removed whether using chemicals or physical means. The information is retrieved using Information Retrieval (IR) system and the pest that would affect the crop is predicted using Recommendation System.

R. Kumar et. al. described a model that uses farmer's location to predict the crop suitable in that area by detecting his or her location then works with different climatic statistics and economical statistic in sub-district level [11]. It aims at collecting the information of that region such as temperature, seasonal crops, crop growing period and on one hand selects similar upazila on basis of the above stated parameters and the seasonal crop to be grown and among them selects the top-k crops. The front end of the model has been developed using XML and SQLlite Manager is used for storing data the model has been developed using JAVA and Android Studio. The developed model predicts the crop to a satisfactory level.

III. Background Techniques

A. Dataset Collection

The dataset file consists of the integral attributes of soil required to choose the crop to be cultivated [12].Various online sources providing information about the general crops were used. The crops considered in this model includes bajra, cashew nut, chick pea, coffee (Arabica), coffee (white robusta), cotton, jowar, jute, pulses, ragi, rice, sesame seed, sugarcane, tea, wheat [13-14].

Fig. 1 gives an outline of the dataset. The total number of instances of each crop available in the datatset used for training is illustrated. In a dataset of 52611 data values, 1485 data values are of bajra, 7623 data values are of cashew nut, 784 data values are of chick pea, 7392 data values are of coffee (Arabica), 5082 data values are of coffee (robusta), 576 data values are of cotton, 3840 data values are of jowar, 400 data values are of jute, 756 data values are of pulses, 7128 data values are of ragi, 5632 data values are of rice, 1170 data values are of seasme seed, 5859 data values are of sugarcane, 2420 data values are of tea and 2464 data values are of wheat.

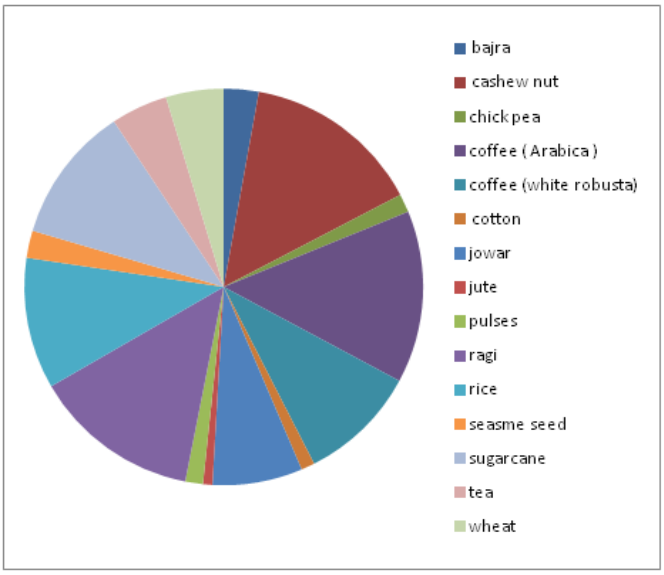

Fig. 1. Dataset

The attributes considered are soil color, pH, average rainfall and temperature. These soil characteristics play a major role in predicting a crop.

Soil color does not influence the behavior and functionality of soil; however, it can point out to the composition of the soil and give clues to the conditions that the soil is subjected to.

The ***pH*** is important because it influences the availability of essential nutrients.

Rainfall is also an important factor for predicting any crop. The water requirement of each crop is different so we have considered this parameter too for predicting the crop.

Temperature is a primary factor affecting the rate of plant development.

Hence, for the above mentioned reasons the soil characteristics considered are essential for predicting a crop seed to be cultivated.

B. Crop Prediction using Classification

1) SVM Algorithm: Classification can be accomplished on structured or unstructured facts [15]. Classification is an approach where we designate data into a given number of classes. The main intention of a classification problem is to recognize the group/class to which a updated data will fall under. Support Vector Machine (SVM) is a supervised machine learning algorithm or model which can be used for classification and as well as for regression challenges. However, we mainly use it in classification challenges. SVM is generally represented as training data points in space which is

divided into groups by intelligible gap which is as far as possible. SVMs can accurately carry out a non-linear classification using a technique called the kernel trick, which is indirectly mapping the inputs into high-dimensional feature spaces.

We have applied SVM approach in our model as:

(i) Importing library SVC from sklearn.svm Class
(ii) Now we create SVM classification object
(iii) At last we fit our data

```
from sklearn.svm import SVC
classifier=SVC (kernel='linear',
random_state=0)
classifier.fit(X_train, Y_train)
```

2) Decision Tree: Decision Tree algorithm is a member of supervised learning algorithms [16] and like other algorithms belonging to that same group we use decision tree algorithm for finding an answer to many regression and classification problems. The main purpose of using Decision Tree is to form a training prototype which we can use to foresee class or value of target variables by learning decision rules deduced from previous data (training data).

The Decision tree can be described by two distinct types, namely decision nodes and leaves. The leaves are the results or the final end results. Each node in the tree acts as a test case for some attribute, and each edge descending from that node corresponds to one of the possible answers to the test case. This process is recursive in nature and is repeated for every sub-tree rooted at the new nodes.

We have applied Decision tree approach in our model as:

(i) Importing library DecisionTreeClassifier from sklearn.tree Class
(ii) Now we create DecisionTreeClassifier object
(iii) In the last we fit our data

```
from sklearn.tree import
DecisionTreeClassifier
classifier=DecisionTreeClassifier
(criterion='entropy', random_state=0)
classifier.fit (X_train, Y_train)
```

3) Logistic Regression: The Logistic Regression model (or Logit model) is a broadly used statistical model that, in its basic form, uses a logistic function to model a binary dependent variable; many more complex extensions exist [17]. In Regression Examination, Logistic regression (or Logit regression) is predicting the parameters of a logistic model; it is a form of Binomial regression.

We have applied Logistic Regression in our model as:

(i) Importing library LogisticRegression from sklearn.linear Class
(ii) Now we create LogisticRegression object
(iii) In the last we fit our data

```
from sklearn.linear_model import
LogisticRegression
classfier=LogisticRegression
(random_state=0)
classifier.fit (X_train, Y_train)
```

IV. Proposed Work

In this model, we have used classification models like linear SVM algorithm and decision tree to predict the crop. At first, we have trained our model with a dataset and then applied linear SVM and decision tree algorithm. After applying both the algorithms, we found linear SVM was predicting more accurate results.

Our model predicts the crop to be selected by the farmer for growing, by entering some soil parameter. The input is taken in a form a .csv file which is shown in Fig. 2.

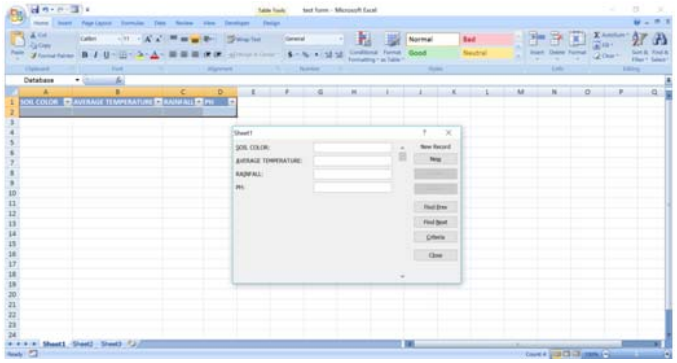

Fig. 2. .csv file format

The entered values are stored in a file named test.csv. The file created is used as the test dataset in our model to predict the crop on the basis of the soil parameters entered by the user.

```
y_pred_new=classifier.predict(X_test_new)
```

The output of our model gives the crop that is best suitable to grow in a soil that has similar soil parameters as that of the parameters entered by the user. Our model also tells the user about the pests that generally affect the crop predicted and its removal measurement i.e. the commonly used pesticides used for removal of such kind of pests.

The overall block diagram of our work is shown in Fig. 3.

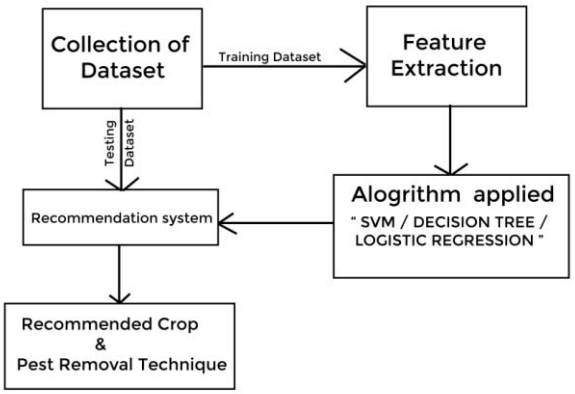

Fig. 3. Block Diagram of Overall Methodology of Proposed System

V. Result Analysis

Our model is designed by using Spyder IDE (ANACONDA 3) and the accuracy of our accounts to 89.66%. After training our model with an ample amount of data values we have successfully predicted the crop and the pest control technique to a quiet a satisfactory level and the rules are induced from algorithm like SVM and Decision Tree. The rules induced from these models helps in building RECOMMENDATION SYSTEM. The input is taken in a form and the training set formed is precisely classified. This model can predict for any situation and for any crop and if it fails to predict it will show a message that couldn't predict.

Table I
Sample Of Dataset

Soil Type	Average Temperature	Average Temperature	Ph	Predicted Crop
Coastal Sands	30	150	7.5	Cashew Nut
Black Soil	27	155	4.5	Coffee (White Robusta)
Sandy Soil	26	75	5	Sugarcane

The above mentioned Table I is a sample of the dataset used to predict the crop and suggest the pest which may affect the predicted crop.

After simulation of the dataset these are some of the test cases presented below in Fig. 4, Fig. 5 and Fig. 6:

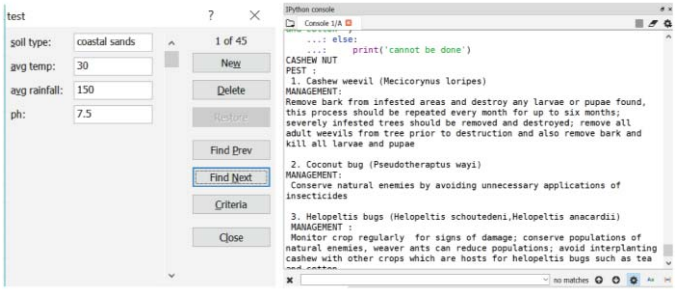

Fig. 4. Testcase 1

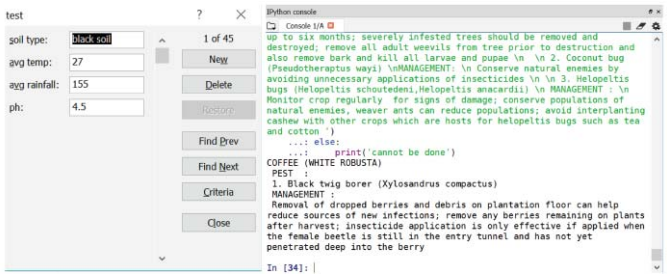

Fig. 5. Testcase 2

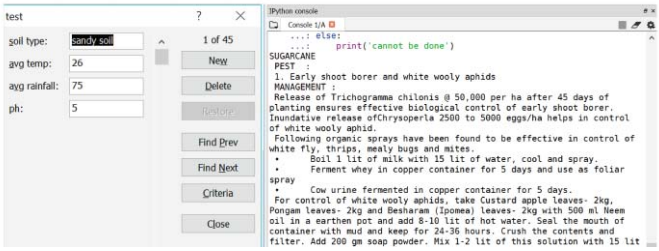

Fig. 6. Testcase 3

The accuracy percentage of different algorithms is plotted and shown in Fig. 7. The accuracy percentage shown here are obtained by applying the corresponding algorithms on respective data sets. The graph shown in red color depicts the accuracy of algorithms achieved in our model and the blue color depicts the accuracy obtained in the model of other algorithms as cited by 2.

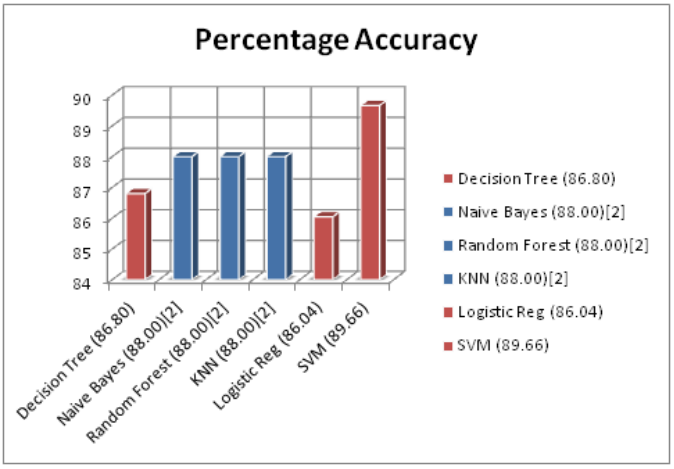

Fig. 7. Percentage Accuracy

Hence from the above graph of comparison, we can conclude that when data set similar to the data set we have used SVM algorithm suits the best as depicted by the graph.

VI. Conclusion

Agriculture being an important part of our economy, it is essential to ensure that the even the smallest investment done in the agriculture sector should be taken care of and when it comes to investment, crop seeds are one of them. So it is essential to check if the correct crop has been chosen for a land holding with matches its requirements to benefit the nation in general and farmer in particular. Our future work aims at developing this model with more soil attributes and with larger data set.

References

[1] L. Madhusudhan, "Agriculture Role on Indian Economy", Business and Economics Journal, 2015.

[2] V. Kumar, V. Dave, R. Bhadauriya, S. Chaudhary, "KrishiMantra: Agricultural Recommendation System", ACM Symposium on Computing for Development, Bangalore, India, 2013, pp. 1-2.

[3] SAVE FOOD: Global Initiative on Food Loss and Waste reduction. http://www.fao.org/save-food/resources/keyfindings/en/.

[4] P. S. Vijayabaskar, R. Sreemathi, E. Keertanaa, "Crop Prediction using Predictive Analytics", IEEE International Conference on Computation of Power, Energy Information and Communication, Melmaruvathur, India, 2017, pp. 370-373.

[5] J. Lacasta, F. J. Lopez-Pellicer, B. Espejo-Garcia, J. Nogueras-Iso, F. J. Zarazaga-Soria, "Agriculture Recommendation System for Crop Protection", Computers and Electronic in Agriculture, Elsevier, vol. 152, 2018, pp. 82-89.

[6] S. Kanaga Subh Raja, R. Rishi, E. Sundaresan, V. Srijit, "Demand based Crop Recommender System for Farmers", IEEE Technological Innovations in ICT for Agriculture and Rural Development, Chennai, India, 2017, pp. 194-199.

[7] M. J. Mokarrama, M. S. Arefin, "RSF: A Recommendation System for Farmers", IEEE Region 10 Humanitarian Technology Conference, Dhaka, Bangladesh, 2017, pp. 843-850.

[8] S. Babu, "A Software Model for Precision Agriculture for Small and Marginal Farmers", IEEE Global Humanitarian Technology Conference: South Asia Satellite, Trivandrum, India, 2013, pp. 352-355.

[9] S. Pudumalar, E. Ramanujam, R. Harine Rajashree, C. Kavya, T. Kiruthika, J. Nisha, "Crop Recommendation System for Precision Agriculture", IEEE International Conference on Advanced Computing, Chennai, India, 2017, pp. 32-36.

[10] A. Savla, P. Dhawan, H. Bhadada, N. Israni, A. Mandholia, S. Bhardwaj, "Survey of Classfication Algorithms for Formulating Yield Prediction Accuracy in Precision Agriculture", IEEE International Conference on Innovations in Information, Embedded and Communication Systems, Coimbatore, India, 2015, pp. 1-5.

[11] R. Kumar, M. P. Singh, P. Kumar, J. P. Singh, "Crop Selection Method to Maximize Crop Yield rate using Machine Learning Technique", IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials, Chennai, India, 2015, pp. 138-145.

[12] M. Paul, S. K. Vishwakarma, A. Verma, "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining Approach", IEEE International Conference on Computational Intelligence and Communication Networks, Jabalpur, India, 2015, pp. 766-771.

[13] List of Major Crops of India | Temperature, Rainfall & Soil, https://www.jagranjosh.com/general-knowledge/list-of-major-crops-of-india-temperature-rainfall-soil-1473918924-1.

[14] Table of Main Crops, Their Climatic Conditions, https://www.topperlearning.com/doubts-solutions/can-you-give-me-a-table-of-the-main-%20%20%0dcrops-rice-wheat-millets-maize-jowar-bajra-ragi-cotton-jute-pulses-sugarcane-coffee-tea-%20%0drubber-their-climatic-cond-7999/.

[15] N. Gandhi, L. J. Armstrong, O. Petkar, A. K. Tripathy, "Rice Crop Yield Prediction in India using Support Vector Machines", IEEE International Joint Conference on Computer Science and Software Engineering, Khon Kaen, Thailand, 2016, pp. 1-5.

[16] P. Singh, B. Jagyasi, N. Rai, S. Gharge, "Decision Tree based Mobile Crowdsourcing for Agriculture Advisory System", IEEE India Conference, Pune, India, 2014, pp. 1-6.

[17] T. Huang, B. Li, D. Shen, J. Cao, B. Mao, "Analysis of the Grain Loss in Harvest based on Logistic Regression", Elsevier Information Technology and Quantitative Management, vol. 122, 2017, pp. 698-705.