International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 11 | Nov 2018 www.irjet.net p-ISSN: 2395-0072

Survey of Crop Recommendation Systems

Deepti Dighe[1], Harshada Joshi[2], Aishwarya Katkar[3], Sneha Patil[4], Prof. Shrikant Kokate[5]

[1,2,3,4]Dept. Computer Engineering, Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India
[5]Assistant Professor, Dept. Computer Engineering, Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India

---

Abstract – Considering a nation where agriculture is still a dominant occupation and traditional farming methods are still in practice, thereby giving limited crop yields to the farmer which ultimately is less beneficial to the farmers compared to the inputs given by them. Hence, in order to maximize the yields of crop for given input, we are showcasing different methods which will be useful to develop a recommendation system for smart farming. The algorithms reviewed in this paper are CHAID, KNN, K-means, Decision Tree, Neural Network, Naïve Bayes, C4.5, LAD, IBK and SVM algorithms. We have also considered a paper which used Hadoop framework for the intensive calculations which are being performed. Use of techniques which use ensemble model for smart recommendation generation are also considered which help to get better accuracy for the system.

Key Words: Precision Agriculture, agriculture, maximum crop yield, minimum investment, environmental factors, economic factors, agricultural recommendation system.

1. Introduction

Diversity in India is unique which represents varieties of physical features and cultural features. Almost all families in India are dependent on agriculture and professions related to agriculture. IoT is playing major role in agriculture which is helping farmers from many problems and to focus on other related professions. Precision agriculture is the one of the best inventions. It is educating farmers in many ways like predicting disease in advanced so that farmer can take actions and get prevented from loss and recommending crop suitable for his field, weather information is also provided as well as it also provides information about marketing and he can export his products and helping to maintain field. Sensors and actuators automate his task such as irrigation and use of pesticides in proportionate. With all these techniques he is able to maximize profit and can continuously monitor his field.

In India, precision agriculture is not much evaluated. Now a day we found that every day the environment is changing continuously which is harmful to the crops and leading farmers towards debt and suicide. In many cases like this and with growing population to maximize yield farmers are using more pesticides and fertilizers which are leading to the soil infertility as well as decreasing the holding capacity of soil and increasing toxicity of soil. Farming land is used by growing industrialization, so again increasing rate of the soil pollution which affects the quality of plants. Various applications of precision agriculture are: prediction of diseases, prediction of weather forecasting, classification of soil, monitoring crop, yield prediction, automatic irrigation system, etc.

2. Motivation

Agriculture is the important factor of economy in India. In recent years due to industrialization; excessive use of pesticides the strength of soil is getting affected. Many of the methods followed by agriculture are not sufficient to increase the productivity. The common difficulty present among the Indian farmers is they don't have any information regarding the right crop based on their soil requirements so it affects the productivity.

Indian farmers face a lot of challenges in making decisions about which farming technique to opt for and which crop should be selected for which climate. The common problem existing among Indian farmers that they don't choose proper crops so as to obtain max yield which are based on topographic features and financial aspects. In agriculture sector, achieving max crop yield at min cost is goal of production.

3. Literature Survey

Agricultural crop recommendation systems are available in the market which consider various parameters like weather at the time the crop is to be planted, soil type, topography of the region, temperature and rainfall in the region, market prices of the

crop, crop duration, etc. Research has been carried out in this field and the following papers have been referred for the purpose of research and study.

Prof. Rakesh Shirsath and other co-author in paper [1] proposed a system which helps the users to make decisions regarding the crop to be planted. The system used is a subscription-based system which would have personalized information of every farmer registered. The system includes a module which maintains the information of the previous crops planted collected from various sources and shows a matching crop that can be planted. The whole process is done with the help of artificial neutral networks. At the end a feedback system is provided so that the developer can make changes required if the farmer finds some difficulty while using the system.

Big Data Analysis Technology Application in Agricultural Intelligence Decision System paper authors Ji-chun Zhao and Jian-xin Guo in paper [2] considers knowledge database as big data and inferences from the data is drawn. It considers various modules like users, knowledge engineer, domain expert, man-machine interface, inference engine and knowledge base. The knowledge acquisition system obtains knowledge for the decision system and establishes an effective knowledge base to solve the problem. The paper uses various Hadoop modules for the purpose of feature extraction. It uses the unstructured data and processes it using NoSQL, Hive, Mahout and uses HDFS to store the data. The data was just presented for wheat crop and other crops were not considered.

RSF as mentioned in paper [3] is a recommendation system for famers which considers a location detection module, data analysis and storage module, crop growing database, physiographic database. The similar location detection module identifies the locations which are similar to the user's locations and checks the similar crops that are planted in those locations. Accordingly, using similarity matrix, the recommendations for the user is generated. Location detection module uses the Google API services to get the current location of the user to identify the similar locations. But the system does not get user feedback to improve the process.

The system in paper [4] suggested by authors S.Pudumalar and associated co-authors uses an ensemble technique called Majority Voting Technique which combines the power of multiple models to achieve greater prediction accuracy. The methods used are Random Trees, KNN, CHAID and Naïve Bayes for ensemble so that even if one method predicts incorrectly, the other models are likely to make correct predictions and since the majority voting technique is used, the final prediction is correct one. If-then rules are the main components which are used in the prediction process. The accuracy obtained is 88% using the ensemble model.

Paper [5] is a review paper for studying various algorithms and their accuracy in the agricultural field proposed by Yogesh Gandge and Sandhya. It was observed that Multiple Linear Regression gave an accuracy of 90-95% for rice yield. Decision tree using ID3 algorithm was considered for soybean crop and the recommendations were generated. The third algorithm was SVM which was used on all the crops and the accuracy was good with computationally less requirements. Neural network was used on corn data to achieve 95% of accuracy. Other algorithms were also used which are KNN, C4.5, K-means, J48, LAD Tree and Naïve Bayes. The conclusion was that still improvement is needed for the algorithms to achieve better accuracy.

In Use of Data Mining in Crop Yield Prediction [6], paper [6], the dataset used was collected from Kaggle.com The author has analysed the data using WEKA tool for algorithms which are LWL, J48, LAD Tree and IBK. The accuracy was measured using specificity, sensitivity, accuracy, RMSE and mean absolute error. For each classifier, confusion matrix was used to get the correctly identified instances. The observation was that better accuracy can be obtained if pruning is used.

Paper [7] presented by Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh proposed use of seven machine learning techniques i.e. ANN, SVM, KNN, Decision Tree, Random Forest, GBDT and Regularized Gradient Forest for crop selection. The system is designed to retrieve all the crops sowed and time of growing at a particular time of the year. Yield rate of each crop is obtained and the crops giving higher yields are selected. The system also proposes a sequence of crops to be planted to get the higher yields.

4. Comparative Study:

Table 1. Comparative Study

Sr.no	Name of researcher, Year of Publication	Paper title	Methodology Adopted / Modules Used	Observations Noted
1	Prof. Rakesh Shirsath, 2017	Agriculture decision support system using data mining	1.Subscription based system 2. ANN	1. Android app with a login module 2. Previously planted crops

			3. Android application 4. Personalized content	known to system 3. User feedback mechanism 4. Maintenance of crop.
2	Ji-chun Zhao, Jian-xin Guo, 2018	Big Data Analysis Technology Application in Agricultural Intelligence Decision System	1.Inference engine 2.Domain expertise 3.Knowledge engineering 4.Knowledge acquisition module 5.Knowledge base for recommendation system	1. Large database of crops 2. Processed using Hadoop 3. Professional knowledge 4. Past experiences 5. Feature selection using HDFS 6. Future Scope: Using Hadoop with Artificial Neural Networks.
3	Miftahul Jannat Mokarrama, 2017	RSF: A Recommendation System for Farmers	1.Location Detection 2.Data analysis and storage 3.Similar location detection 4. Recommendation generation module.	1. Physiographic, thermal, crop growing period, crop production rate 2. Seasonal crop database 2. Similar location detection 3. Generating the set of crops 4. Similarity between the crops planted in a region
4	S.Pudumalar, E.Ramanujam, 2016	Crop Recommendation System for Precision Agriculture	1. Random tree 2. CHAID 3. KNN 4. Naïve Bayes 5. WEKA tool	1. Pre-processing of data 2. Handling missing and out-of-range values 3. Feature extraction 4. Ensemble model to get higher accuracy 5. Rule generation
5	Yogesh Gandge, Sandhya, 2017	A Study on Various Data Mining Techniques for Crop Yield Prediction	1. Attribute selection 2. Multiple Linear Regression 3. Decision Tree using ID3 4. SVM 5. Neural Networks 6. C4.5 8. K-means and KNN	1. Selection of agricultural field 2. Selection of crop previously planted 3. Input from user 4. Preprocess 5. Attribute Selection 6. Classification algorithm on data 7. Crop is recommended
6	Shruti Mishra Priyanka Paygude, 2018	Use of Data Mining in Crop Yield Prediction	1. J48 2. LAD tree 3. LWL 4. IBK algorithm	1. WEKA tool 2. LAD tree showed the lowest accuracy 3. Errors can be minimized by pruning the tree 4. IBK was observed to achieve higher accuracy

5. Existing Systems:

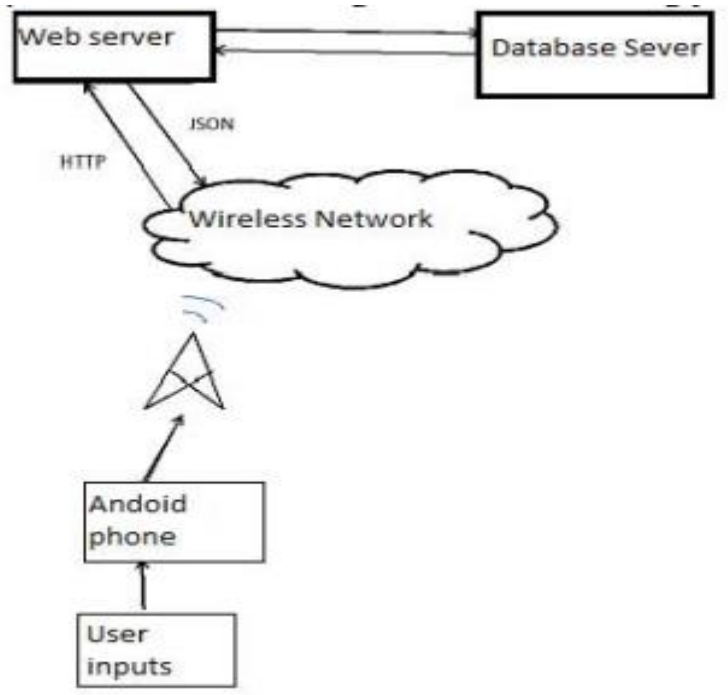

Fig. 1. A recommendation system [1]

Fig. 1. Shows the interaction of user with the system. The user inputs are given through an app on an android phone which is then sent to a web server through a wireless network on the user's phone using HTTP and JSON formats. The required data is then fetched from the database server and recommendations are generated. The android application has two modules i.e. user and admin who have permission for accessing the content accordingly. Considering various factors like ph level of soil, month of cultivation, weather in the region, temperature, type of soil, etc. factors were considered to select maximum likely crops for plantation. After the generation of the set of crops, the farmers are given information about the location to buy the seeds, how to cultivate and maintenance of the crop. But the conclusion was that best matching crops needs to be identified out of the final set of crops generated.

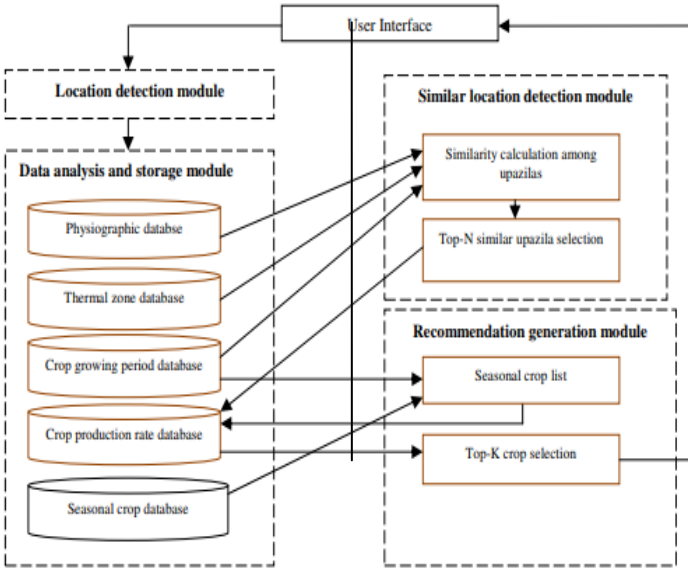

Fig. 2. Architecture of a recommendation system [3]

As shown in fig. 2. Various modules have been included which help to recommend the best crops to be planted. Module physiographic database has the information about various regions and its topology to consider the area and crop similarity, thermal zone database which has information of temperature, crop growing period to consider how much time the crop would last in the field to manage the future crop recommendations and seasonal crop database to identify which crop grows in which season. Then the crops matching all the criteria is identified and top-K crops are selected using similar crop detection module. The authors have used a similarity matrix to calculate the accuracy of the system. The system requires to be upscaled for a larger region.

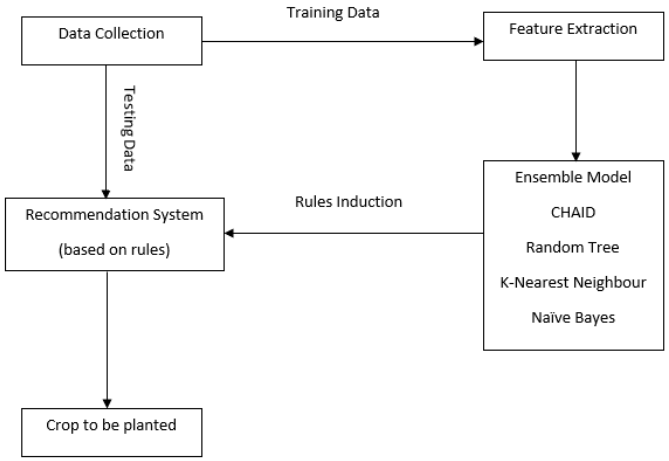

Fig. 3. Use of ensemble techniques for crop recommendation [4]

The system shown in fig. 3. uses an ensemble model which gathers the prediction from every model mentioned i.e. CHAID, Random Tree, KNN and Naive Bayes and uses a majority voting technique which will help to make accurate predictions even if one algorithm fails to make the correct prediction. Higher the competition, higher is the chance of getting a correct output [4]. Hence, choosing the models to be used for ensemble has to be done considering the accuracy and results obtained from the models. If-then rules are generated by the CHAID and Random Forest models. They are fed to the recommendation engine and the best set of rules are given as recommendation.

6. Conclusion

From the study in this paper, we concluded that there is still a need of research in the agricultural field to get better accuracy. Using ensemble methods is a good way to ensure better accuracy of the system. Also, if we want to consider only one algorithm for the recommendation system then we can use SVM due to its simple computational requirements.
Future work in this field can be geospatial analysis combining and using all the seasonal, soil, weather, temperature, topographical, crop production and economic condition of the farmer into one single model to provide a robust and centralized system for the user to access. Particularly, the study showed that for the previously proposed recommendation systems, factors such as user's investment in farming, number of people for the maintenance and the stretch of land available for cultivation were not considered. These parameters also play a major role in the profits of the farmer which is the basic reason why a farmer would use a crop recommendation system.

7. References

[1] 2017 International Conference on I2C2, "Agriculture decision support system using data mining", Prof. Rakesh Shirsath; Neha Khadke; Divya More.

[2] 2018 the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis, "Big Data Analysis Technology Application in Agricultural Intelligence Decision System", Ji-chun Zhao; Jian-xin Guo.

[3] 2017 IEEE Region 10 Humanitarian Technology Conference, "RSF: A Recommendation System for Farmers", Miftahul Jannat Mokarrama; Mohammad Shamsul Arefin.

[4] 2016 IEEE Eighth International Conference on Advanced Computing (ICoAC), "Crop Recommendation System for Precision Agriculture", S.Pudumalar*, E.Ramanujam*, R.Harine Rajashreeń, C.Kavyań, T.Kiruthikań, J.Nishań.

[5] 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), "A Study on Various Data Mining Techniques for Crop Yield Prediction", Yogesh Gandge, Sandhya.

[6] Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018), "Use of Data Mining in Crop Yield Prediction", Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, Sonali Idate.

[7] 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, T.N., India. 6 - 8 May 2015. pp.138-145, "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique", Rakesh Kumar , M.P. Singh , Prabhat Kumar and J.P. Singh

[8] IEEE Sponsored 9th International Conference on Intelligent Systems and Control (ISCO) 2015, "XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction", Aakunuri Manjula, Dr.G .Narsimha