© 2019 IJSRST | Volume 6 | Issue 1 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology
DOI : https://doi.org/10.32628/IJSRST196172

Crop Recommendation System to Maximize Crop Yield in Ramtek region using Machine Learning

D. Anantha Reddy, Bhagyashri Dadore, Aarti Watekar
Kavikulguru Institute of Technology and Science, Ramtek, Nagpur, Maharashtra, India

ABSTRACT

In Indian economy and employment agriculture plays major role. The most common problem faced by the Indian farmers is they do not opt crop based on the necessity of soil, as a result they face serious setback in productivity. This problem can be addressed through precision agriculture. This method takes three parameters into consideration, viz: soil characteristics, soil types and crop yield data collection based on these parameters suggesting the farmer suitable crop to be cultivated. Precision agriculture helps in reduction of non suitable crop which indeed increases productivity, apart from the following advantages like efficacy in input as well as output and better decision making for farming. This method gives solutions like proposing a recommendation system through an ensemble model with majority voting techniques using random tree, CHAID, K _ Nearest Neighbour and Naive Bayes as learner to recommend suitable crop based on soil parameters with high specific accuracy and efficiency. The classified image generated by these techniques consists of ground truth statistical data and parameters of it are weather, crop yield, state and district wise crops to predict the yield of a particular crop under particular weather condition.

Keywords : Precision agriculture, Recommendation system, Ensembling model, Majority voting techniques, Random tree, CHAID, K-Nearest Neighbor and Naive Bayes.

I. INTRODUCTION

India is one of the biggest producers of agricultural products and still has very less farm productivity. Productivity needs to be increased so that farmers can get more pay from the same piece of land with less labour. Precision agriculture provides a way to do it. Precision farming, as the name implies, refers to the applying of precise and proper total of comment like fertilizers, soil etc. but in recent times the trend in agriculture has drastically evolved due to globalization .various factors have affected the health of agriculture in India. many new technologies have been evolved to regain the health . One such technique is precision agriculture, at the proper time to the craw for increasing its productivity and increasing its yields Not all precision agriculture systems offer best results. precision agriculture is a technology of site-specific farming. But in agriculture it is important that the recommendations made are accurate and precise because in case of errors it may lead to heavy material and capital loss. recommendation of crops is one major domain in precision agriculture.

Ensembling is one such technique that is included in such research works. Among these various machine learning techniques that are being used in this field .ensembling is data mining model also known as the committee methods or model combiners, that combines the power of multiple models to acquire greater prediction, efficiency than any of its model

could achieve alone. Random forests square measure associate ensemble learning methodology for classification of algorithmic rules . All naive Bayes socio-economic classifiers adopts that the value of a particular feature is independent of the value of any other feature, given the class variable it's a classification technique supported Bayes' theorem with associate degree assumption of independence between predictors. . Naive Bayes is not single algorithm, but a clan, regression and different tasks, that operate by building a mess of call trees at coaching time and outputting the category that's the mode of the categories or mean prediction of the individual trees. 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. It is a management strategy that employs detailed, site specific information to precisely manage production inputs. This concept is sometimes called precision agriculture, prescription farming, or site-specific management. The idea is to know the soil and crop characteristics unique to each part of the field, and to optimize the production inputs within small portions of the field. The philosophy behind precision agriculture is that production inputs (seed, fertilizer, chemicals, etc.) should be applied only as needed and where needed for the most economic production. Why should producers be interested in precision agriculture? Precision farming techniques can improve the economic and environmental sustainability of crop production. In today's agriculture, producers tend to farm each field as a single unit. Although they often recognize in-field variability, they have had few tools with which to manage that variability.

II. METHODS AND MATERIAL

Dataset Collection: The dataset comprising the soil specific attributes which are collected for Ramtek town tested at soil testing lab Nagpur, Maharashtra, India. The crops considered in our model groundnut, pulses, cotton, vegetables, paddy, sugarcane, coriander. The number of instances of each crop available in the training dataset is depicted. The attributes considered where Depth, Texture, PH, Soil Color, Permeability, Drainage, Water holding and Erosion. The above stated parameters of soil play a major role in the crop's ability to extract water and nutrients from the soil. For crop growth to their fullest potential, the soil must provide a satisfactory environment for it. Soil is the anchor of the roots. The water holding capacity determines the crop's ability to absorb nutrients and other nutrients that are changed into ions, which is the form that the plant can use. Texture determines how porous the soil is and the comfort of air and water movement which is essential to prevent the plants from becoming waterlogged. Soil texture which affects the soil's ability to hold onto nutrients. The level of acidity or alkalinity (Ph) is a master variable which affects the availability of soil nutrients. The activity of microorganisms present in the soil and also the level of exchangeable aluminium can be affected by PH. The water holding and drainage determine the penetration of roots. Hence for the following reasons the above stated parameters are considered for choosing a crop.

Crop Prediction using ensemble technique: Ensemble is a data mining model also known as the Comminttee Methods or Model Combiners, That combiners, that combine the power of multiple models to acquire greater prediction, efficiency that any of its models could achieve alone. In our system, we use one of the most familiarensembling technique called Majority voting technique. In the voting technique any number of base learners can be used. There has to be at least two base learners. The learners are chosen in a way that they are competent

to each other yet being complimentary also. Higher the competition higher is the chance of better prediction. But it is necessary for the learners to be complimentary because when one or few members correcting this error would be high. Each learner builds itself into a model. The model gets trained using the training data set provided. When a new sample has to be classified, each model predicts the class on its own. Finally, the class which is predicted by majority of the learners is voted to be the class label of the new sample. This method is implemented in Rapid miner tool depicts the process implemented in rapid miner.

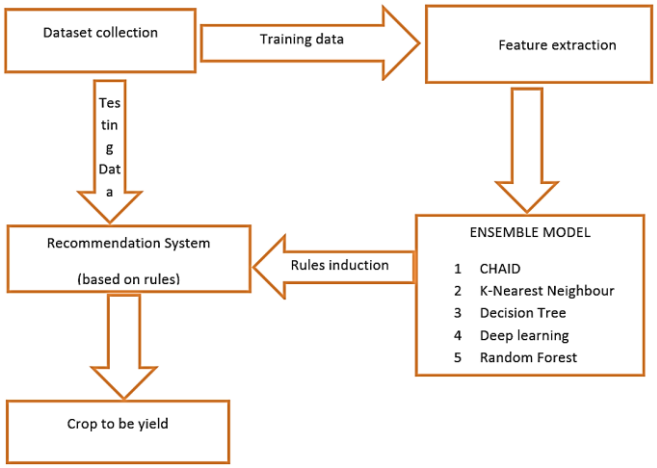

Fig 1. Proposed System Architecture.

It consist of four machine learners namely Naïve bayes,K-Nearest neighbour and CHAID and Random tree. The operator corresponding to each learner is positioned. The operator performancethe classification correspondingly. The tree to rules operator is used to induce rules directly from the CHAID and random tree.

Rules include from the Model:

The rule below demonstrates an example of the proposed recommendation system.

IF ph is mid alkaline
AND depth is above 100
AND water holding capacity is LOW
AND drainage is moderately well
AND erosion is moderate
THEN PADDY

The IF part of the rule states the soil specification needed for the cultivation of the recommended crop which is specified in the THEN PART of the rule.

III. RESULTS AND DISCUSSION

Learners Used in the Model:

1. RANDOM TREE:

Random tree similar to that of a decision tree.But it differs from random tree in a way that for each split only a random subset of attribute are available. Random tree can be built for both nominal and numerical data. The Random tree is similar to C4.5 or CART butt it selects only a random subset of attributes. At each node it considers K randomly chosen attributes. The subset ratio parameter specifies the size of the subset.

2. K-NEAREST NEIGHBOR:

K-Nearest Neighbour can be used for both classification and regression. K-Nearest Neighbours is a non-complex algorithm which stores all the available cases and classifires new cased based on some similarity measure. The sampled set is classified based upon the ”closeness” sssthat is the distance measure such as Euclidean distance or Manhattan distance.

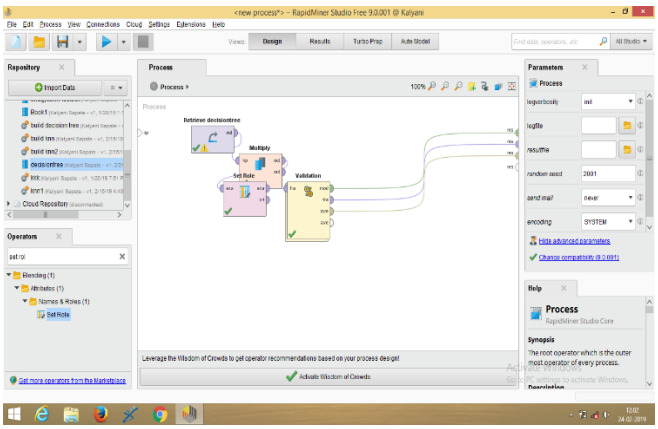

Fig 2: Illustrates the entire process work flow.

It shows three operators namely retrieve, set role, validation. the retrieve operator retrieves the dataset that is being uploaded in the tool. the set role operator used to set the target attribute or special attributes. In order to estimate the statistical performance of learning operator a cross-validation is performed by the validation operator.

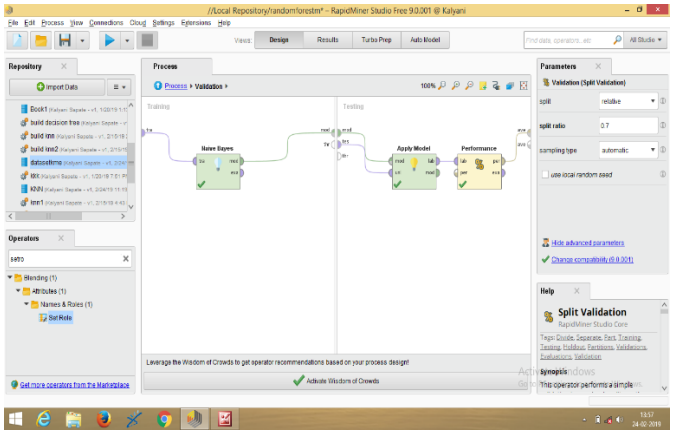

Fig 3: Illustrates the sub-process of validation operator

The training process consist of the voting operator which is the technique that we process for the better results. On the testing sub-process lies the apply model and performance operator which evaluate the correctness of the model.

3. RANDOM FOREST:

Random forest square measure associate ensemble learning methodology for classification, regression and different tasks, that operate by building a mess of call trees at coaching time and outputting the category that's the mode of categories or mean prediction of the individual trees. Random call forests correct for call tree custom of over fitting to their coaching set. The primary rule for random call forest was created by Tin KamHo victimization the random mathematical space methodology, which, in Ho's formulation, could be a thanks to implement the "stochastic discrimination" approach to classification.

4. DECISION TREE:

Classifies data using the attributes. Tree consists of decision nodes and decision leafs. Nodes can have two or more branches which represents the value for the attributes tested. Leafs nodes produces a homogeneous result.

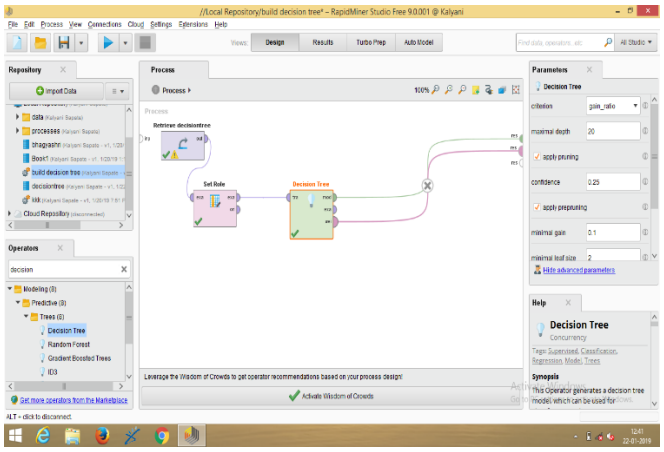

Fig 4: Illustrates the entire process work flow.

IV. CONCLUSION

India is a nation in which agriculture plays a prime role. In prosperity of the farmers, prospers the nation. Thus our work would help farmers in sowing the right seed based on soil requirements to increase productivity of the nation. Our future work is aimed at an improved data set with large number of attributes and also implements yield prediction.

V. REFERENCES

[1]. SatishBabu (2013),'A Software Model for Precision Agriculture for small and Marginal Farmers'. At the International Centre for and Open Source Software(ICFOSS) Trivandrum, India.

[2]. AnshalSavla, ParulDhawan, HimtanayaBhadada, NiveditaIsrani, Alisha Mandholia, SanyaBhardwaj (2015),'Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture', Innovation in Information, Embeddedand Communication system(ICIIECS).

[3]. AakunuriManjula, Dr. G. Narsimha (2015),'XCYPF:AFlexibleand Extensible Framework for Agriculture Crop Yield Prediction', Conference on Intelligent Systems and Control(ISCO).

[4]. YashSanghvi , Harsh Gupta, HarmishDoshi, DivyaKoli, AmogghAnsh , DivyaKoli, Umang Gupta(2015), Comparison of Self Organizing Maps and Sammon's Mapping on agriculture datasets for precision agriculture', International Confereence on Innovations in information, Embedded and Communication Systems(ICIIECS).

[5]. Rakeshkumar, M.P. Singh, Prabhat Kumar and J.P. Singh(2015),'Crop Selection Methode to Maximize Crop Yeild Rate using Machine Learning Technique',International conference on smart Technologies and Management for Computing,Communication,Controls,Energy and Materials(ICSTM).

[6]. A.T.M. ShakilAhamed, NavidTanzeemMahmood, NazmulHossain, Mohammad TanzirKabir, Kallal Das, FaridurRahman, Rashedur M Rahman(2015),'Applying Data Mining Techniques to Predict Annual Yeild of Major Crop and Recommend Planing Different Crop in Different District in Bangladesh',(SNPD) IEEE/ACIS International Conference.

[7]. LiyingYang(2011),'Classifiers selection of enssebling learning based on accuracy and diversity'Published by Elsevier Ltd. Selection and/or peer-review under responsibility of[CEIS].

[8]. Tapas RanajnBaitharua, Subhendu Kumar Panib(2016),'Analysis of Data Mining Techniques for Healthcare Decision Support System Using Liver Discover Dataset' International conference on Computational Modeling and Security(CMS).

[9]. Monali Paul, Santosh K. Vishwakarma, Ashok Verma(2015),'Analysis of soil Behaviour and Prediction of Crop Yield using Data Mining Approach',International Conference on Computational Intelligence and Communication Networks.

[10]. Bhuvana, Dr. C. Yamini(2015),'Survey on Classification Algorithms in Data mining'International conference on resent Advance in Engineering Science and Management.

Cite this article as :

D. Anantha Reddy, Bhagyashri Dadore, Aarti Watekar, "Crop Recommendation System to Maximize Crop Yield in Ramtek region using Machine Learning", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 6 Issue 1, pp. 485-489, January-February 2019. Available at doi : https://doi.org/10.32628/IJSRST196172

Journal URL : http://ijsrst.com/IJSRST196172