Nikhara Krushi –A Crop Recommendation System for Precision Agriculture

Nasreen Taj M B, Kavya H C, Nayana R R, Bindu H S, Meghana D P
Department of information science,
GMIT, Davangere

***Abstract*— Data mining is the practice of examining large pre-existing databases in order to generate new information. Data mining has its applications in many fields like agriculture, finance, medicine etc. Data mining is used for analyzing biotic and abiotic factors. Agriculture is the most important sector of Indian Economy. Indian agriculture sector accounts for 18 percent if Indian's gross domestic product and provides employement to 50% of the countries workforce. The common problem existing among the Indian farmers are they don't choose the right crop based on their soil requirements. Due to this they face a serious difficulty in productivity. This problem of farmers has been addressed through precision agriculture. "Precision agriculture or site specific crop management (SSCM) is a farming management concept based on observing ,measuring and responding to inter and intra-field variability in crops". This reduces the wrong choice on a crop and increase in productivity. In this paper, this problem of farmers can be solved by proposing a recommendation system by using K-Nearest Neighbor algorithm to recommend a crop for the site specific parameters with high accuracy and efficiency.**

Keywords* - *K-Nearest Neighbor, Recommendation System.

I. INTRODUCTION

India is one among the oldest countries which is still practicing agriculture. But in recent times the trends in agriculture has drastically evolved due to globalization. Various factors have been evolved to regain the health of agriculture in India. Many new technologies have been evolved to regain the health. One such technique is precision agriculture. Precision agriculture is budding in India. Precision agriculture is the technology of "site-specific" farming. It has provided us with the advantage of efficient input, output and better decisions regarding farming.

Although precision agriculture has delivered better improvements it is still facing certain issues. There exists many systems which propose the inputs for a particular farming land. Systems propose crops, fertilizers and even farming techniques. Recommendation of crops is one major domain in precision agriculture.

Recommendation of crops is dependent on various parameters. Precision agriculture aims is in identifying these parameters in a site-specific manner in order to resolve issues regarding crop selection. The "site-specific" technique has improved the results yet there is a need to supervise the results of such systems. Not all precision agriculture systems provide accurate results. But in agriculture it is important that the recommendations made are accurate and precise because incase of errors it may lead to heavy material and capital loss.

Many research work is being carried out, in order to attain an accurate and efficient model for crop prediction. This paper proposes a system that uses k-nearest neighbor algorithm to build an efficient and accurate model.

II. LITERATURE SURVEY

The paper [1] states the requirements and planning needed for developing a software model for precision farming is discussed. It deeply analyses the basics of precision farming. The author's start from the basics of precision farming and move towards developing a model that would support it. This paper describes a model that applies Precision Agriculture (PA) principles to small, open farms at the individual farmer and crop level, to affect a degree of control over variability. The comprehensive objective of the model is to deliver direct advisory services to even the smallest farmer at the level of his/her smallest plot of crop, using the most accessible technologies such as SMS and email. This model has been designed for the scenario in Kerala State where the average holding size is much lower than most of India. Hence this model can be deployed elsewhere in India only with minor modifications. The paper [2] makes a comparative study of classification algorithms and their performance in yield prediction in precision agriculture. These algorithms are implemented in a data set collected for several years in yield prediction on soya bean crop. The paper [3] states the usage of agricultural data with data mining and visual data mining techniques are depicted. This paper reduces the high dimensional agricultural data to smaller size to acquire useful knowledge related to yield, input application(like fertilizers).The techniques used is Self-organizing maps and multi-dimensional scaling techniques (Sammon's mapping) to reduce the data. The conclusion derived is that Self-organizing maps is suitable when dataset is large and Sammon's maping is suitable when data set is small. The paper [4] depicts the importance of crop selection and the factors deciding the crop selection like production rate, market price and government policies are discussed. This paper proposes a Crop Selection Method (CSM) which solves the crop selection problem and improves net yield rate of the crop. It suggests a series of crop to be selected over a season considering factors like weather, soil type, water density, crop type. The predicted value of influential parameters determines the accuracy of CSM. Hence there exists a need to include a prediction method with improved accuracy and preformence. Data mining techniques in paper [5] are used to estimate the crop yield for the cereal crops in one of the major districts of Bangladesh.

The methodology comproses of two parts namely Clustering (for creating district clusters) and Classification using k-NN (k- nearest neighbor), Linear Regression , (ANN) artificial neural network in rapid miner tool. The accuracy of prediction lies in the range of 90-95. The data set included 5 environmental variables, 3 biotic variables and 2 area related variables to determine the crop yield in different districts. The paper proposed a future work of geospatial analysis to improve accuracy.

The paper [6] aims to solve the crucial problem of selecting the classifiers for the ensemble learning. A method to select a best classifier set from a pool of classifiers has been proposed. The proposal aims to achieve higher accuracy and performance. A method called SAD (Selection by accuracy and Diversity) was proposed based on accuracy and classification performance. The paper [7] tries to solve the problem of food insecurity in Egypt. It proposes a framework which would predict the production, and import for that particular year. It uses Artificial Neural Networks to build the prediction. At the end of the process we would be able to visualize the amount of production import, need and availability. Therefore it would help to make decisions on whether food has to be further imported or not. The soil datasets in paper [8] are analyzed and a category is predicted. From the predicted soil category the crop yield is identified as a Classification rule. Naïve Bayes and k-Nearest Neighbor algorithms are used for crop yield prediction.

III. METHODOLOGY

3.1 Dataset Collection

In this study, the first step is collection of data,thedataset is collected for different locations of the soil. In addition, similar online sources of general crop data were also used. The attributes considered for crop prediction are Depth, Texture, Ph, Soil Color, Permeability, Drainage, Water holding and Erosion.

These 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 aluminum 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.

3.2 Algorithm used in the model

K-NEAREST NEIGHBOR: K-Nearest Neighbor can be used for both classification and regression. K-Nearest Neighbors is a non-complex algorithm which stores all the available cases and classifies new cases based on some similarity measure .The sample set is classified based upon the “closeness” that is the distance measure such as Euclidean distance or Manhattan distance.

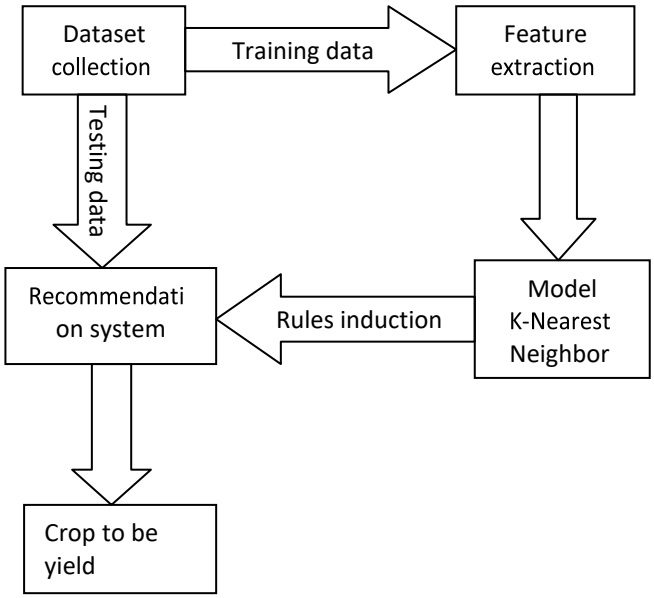

Fig 1: This figure depicts the overall methodology of proposed system.

The rule below demonstrates an example of the proposed recommendation system.
IF ph is mild 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 specifications needed for the cultivation of the recommended crop which is specified in the THEN PART of the rule.

IV. RESULT

The rules generated from the above model is used to develop a RECOMMENDATION SYSTEM. This is achieved by creating a GUI. The GUI is deployed as a web portal. The model which is trained with the training data set is tested with inputs from the user from the portal. The scripting done will respond to any test case suggesting a crop. If the test case doesn’t match any of the predictions then a no match output is produced.

V.CONCLUSION

India is a nation in which agriculture plays a prominent role. Thus our work would help farmers in sowing the right seed based on soil requirements to increase productivity of the crops and acquire profit out of such a technique. Thus the farmer’s can plant the right crop increasing his yield and also increasing the overall productivity of the nation.

VI. REFERENCES

[1]. Satish Babu (2013), 'A Software Model for Precision Agriculture for Small and Marginal Farmers', at the International Centre for Free and Open Source Software (ICFOSS) Trivandrum, India.

[2]. Anshal Savla, Parul Dhawan, Himtanaya Bhadada, Nivedita Israni, Alisha Mandholia , Sanya Bhardwaj (2015), 'Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture', Innovations in Information,Embedded and Communication systems (ICIIECS).

[3]. Yash Sanghvi, Harsh Gupta, Harmish Doshi, Divya Koli, Amogh Ansh Divya Koli, Umang Gupta (2015), 'Comparison of Self Organizing Maps and Sammon's Mapping on agricultural datasets for precision agriculture', International Conference on Innovations in Information, Embedded and Communication systems (ICIIECS).

[4]. Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh (2015), 'Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique', International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM).

[5]. A.T.M Shakil Ahamed, Navid Tanzeem Mahmood, Nazmul Hossain, Mohammad Tanzir Kabir, Kallal Das, Faridur Rahman, Rashedur M Rahman (2015) , 'Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh' , (SNPD) IEEE/ACIS International Conference.

[6]. Liying Yang (2011), 'Classifiers selection for ensemble learning based on accuracy and diversity' Published by Elsevier Ltd. Selection and/or peer- review under responsibility of [CEIS].

[7]. Aymen E Khedr, Mona Kadry, Ghada Walid (2015), 'Proposed Framework for Implementing Data Mining Techniques to Enhance Decisions in Agriculture Sector Applied Case on Food Security Information Center Ministry of Agriculture, Egypt', International Conference on Communications, management, and Information technology (ICCMIT').

[8]. 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.

[9]. Bhuvana, Dr.C.Yamini (2015), 'Survey on Classification Algorithms in Data mining.' International Conference on Recent Advances in Engineering Science and Management.

[10]. Roshani Ade, P.R.Deshmukh (2014), 'Efficient Knowledge Transformation System Using Pair of Classifiers for Prediction of Students Career Choice', International Conference on Information and Communication Technologies (ICICT).