Crop yield prediction, forecasting and fertilizer recommendation using Data mining algorithm

K. Archana
ME Software Engineering, PSG College of Technology, Coimbatore

Dr.K.G.Saranya
Assistant Professor, Department of Computer Science,
PSG College of Technology, Coimbatore

ABSTRACT: Agriculture is the backbone of a developing economy such as India, for the income the majority of its population depends on agriculture. Data mining is a significant field of informatics that can be applied very efficiently to the agriculture sector. Agricultural stakeholders need to predict and forecast crop yields that can be acquired using data mining techniques. When the farmers are unaware of the soil nutrition and soil composition its results in minimal crop yield. The system developed which in effect focuses on macro nutrients (NPK) present in the soil to provide the most suitable crop suggestions. The system for crop yield prediction, forecasting and fertilizer recommendation are all separate and unique in the existing system. The proposed system constructs a collaborative system of crop yield prediction, forecasting and fertilizer recommendation. In this project a system is build that incorporates the agricultural dataset and the K Nearest Neighbor algorithm is applied to suggest the most appropriate crops. Crop yield prediction and forecasting will improve the agricultural production. Crop rotation will boosts the soil fertility and this system supports farmer friendly fertilization decision making.

Keywords: Nitrogen, Phosphorus, Potassium, soil nutrition, yield prediction, crop rotation, fertilizer recommendation.

1. INTRODUCTION

The choice of a crop for planting is one of the major challenges faced by farmers in the cultivation of crops. It is influenced by several factors. By recommending the most appropriate crops and suggesting the suitable fertilizer for the crop, the crop recommendation system can help farmers to select the appropriate crop for the crop yield. Thus there is a need for suggesting suitable crops and fertilizers using the data mining algorithm.

Data mining is a technique that uses complex algorithms and a series of predefined rules to function intelligently. It uses the past data to interpret patterns and then performs the intended task according to defined rules and algorithms based on an analysis it makes predictions. Data mining is a significant field of informatics that can be applied very efficiently to the agriculture sector. Together with big data technologies and high performance computing, new possibilities have arisen to unravel the quantification and comprehensive of the information intensive process in agricultural operating environments. Data mining is everywhere over the entire growing and harvesting process.

The crop recommendation system incorporates the agricultural datasets. The soil's Nitrogen (N), Phosphorus (P) and Potassium (K) levels are used as input to recommend crops. The proposed system constructs a collaborative system of crop yield prediction, forecasting and fertilizer recommendation. In this project a system is build that incorporates the agricultural dataset and the K Nearest Neighbor algorithm is applied to suggest the most appropriate crops.

The existing system focuses only on crop yield prediction and forecasting or crop recommendation or on crop rotation. the prediction was ineffective due to certain vital characteristics of the soil not considered. Existing works suggest fertilizer only based on the crop. The nutrients already present in the soil were not taken into account to make suggestions. Hence the proposed system considers the soil type, land type, soil texture and NPK values for suggesting the fertilizer.

2. RELATED WORK

Mansi Shinde et al [1], developed a system that provides farmers with recommendations for identifying suitable fertilizers and crops. Farmers can access the system via android based mobile devices, which helps to increase the crop yield. This application also permits user to buy recommended fertilizers from the purchase portal.

V.Sellam et al [2], Evaluated environmental parameters such as the annual rainfall, food price index of the region under cultivation that influences crop yield. The crop yield is a dependent variable that depends on the evaluation of all these environmental factors used in this paper to validate their effect on crop yield.

U.K.Diwan et al [3], developed a climate based model to provide a reliable forecast of crop yield in advance. The focus of this work was on the crop yield forecast model through the use of weather parameters and crop yield history. Temperature (maximum and minimum) and relative humidity were found to play a major role among all the weather factors in all the districts.

Rushika Ghadge et al [4], suggested a system to assist farmers by tracking soil quality on the basis of data mining techniques. The scheme focuses on inspecting soil quality to predict soil specific crops. This system also recommends an effective fertilizer to optimize the crop yield. This analyzes soil nutrition and soil based crop productivity to correctly recommend the crops.

P.Priya et al [5], developed a system that focuses on crop yield prediction based on the Kharif and Rabi season's dataset, which is used to build the model, the results obtained from this system will be useful for farmers to predict yield before they are grown on the farming land.

Vaneesbeer Singh et al [6], presented an approach using various machine learning techniques to predict the yield category based on macro and micro nutrients contained in the soil. After evaluation the predicted yield category of the product determines the yield of the crops.

Vrushal Milan Dolas et al [7], presented a modified decision tree algorithm and the classifier used in this application incorporated soil dataset. The soil is classified into groups or classes that have similar behavior, so that the farmers are familiar with the soil type and they are able to plant the crops accordingly.

R.Sujatha et al [8], presented the basic principles for understanding phenology and following acceptable planting dates over time and space separations for distinct genotypes, this paper presents the estimation of crop yield and the choice of the most suitable crop to increase the value and benefit of the agricultural area.

Supriya.D.M et al [9], established a model where data mining techniques have been used to predict the category of analyzed soil dataset. The predicted category would indicate the crop yield. Classification algorithm programmed to classify unknown samples using the information provided to purchase a collection of classified samples.

S.Veenadhari et al [10], created a user friendly web application to predict the effect of climate parameters on crop yield and to find the most powerful weather parameter on crop yield for the selected plant in particular regions of Madhya Pradesh. The application presented an overview of the relative impact of various weather parameters that is responsible for crop yield.

3. METHODOLOGY

K Nearest Neighbor (KNN) Algorithm is used for crop recommendation. Initial step is to get the input parameters like soil type, land type, soil texture, Nitrogen, Phosphorus and Potassium from the user. The next step is filtering the dataset based on the soil type, land type and soil texture provided by the user. Then Euclidean distance is computed between the Nitrogen, Phosphorus and Potassium values provided by the user and the NPK values present in the filtered dataset. Finally the computed Euclidean distance is sorted in the ascending order and top five crops is recommended with its appropriate fertilizer.

4. PROPOSED SYSTEM

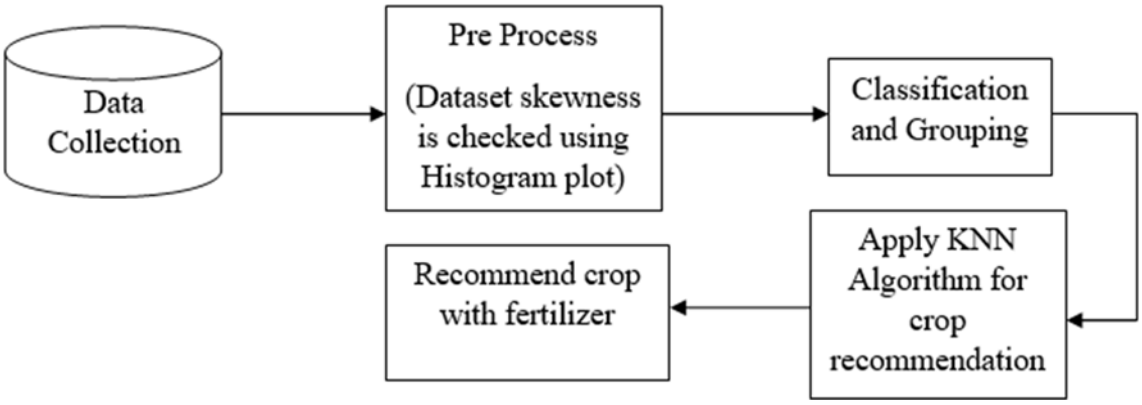

Figure 1 Proposed System Architecture

DATA COLLECTION:

Agriculture data set is collected and it is checked for the presence of any missing values. The dataset used is consistent and complete without any missing values. The data set is further analysed for thorough understanding and data clarity and the dataset is loaded to the crop recommendation system for further proceedings.

PRE PROCESS:

The histogram chart is plotted for the dataset to check if there is any skewness. The Agriculture dataset contains two types of soil Black soil and Red Sandy soil. In the pre processing task the user can analyse the dataset by viewing all instance for each kind of soil type (Black soil and Red sandy soil) separately for better understanding of the crop with its appropriate soil type. This leads in choosing the right crop at the right time and thus maximizes the yield.

CLASSIFICATION AND GROUPING

In classification and grouping step the classification process is done based on three attributes namely land classification, month classification and soil texture classification. Finally the grouping process is done for the crop attribute.

KNN ALGORITHM:

After detailed analysis of the dataset K-Nearest Neighbor Algorithm is applied to recommend the crops.

CROP RECOMMENDATION:

The KNN algorithm recommends the most suitable crop along with its appropriate fertilizer by calculating the Euclidean distance.

YIELD CALCULATION:

As the final the step the crop yield can be calculated for the required crop by just proving the area of the cultivation land in acres.

5. EXPERIMENTAL RESULTS

RESPONSE TIME METRIC

Response time is used to calculate the processing time of the K Nearest Neighbor Algorithm and providing the crop suggestion to the user. Stopwatch class is used to measure the elapsed time. It is used to calculate the execution time of the Euclidean distance function and it is discussed in the table 1.

ACCURACY METRIC

Accuracy is used to measure the degree to which the result of our prediction is correct. Accuracy is computed by the formula,

$$\textbf{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Where, TP = True Positives

TN =True Negatives

FP = False Positives

FN = False Negatives

PERFORMANCE EVALUATION

Table 1 Performance Evaluation

N	P	K	Response Time	Accuracy
80	3	280	664 ms	83.83%
60	7	180	507 ms	83%
78	3	205	512 ms	82.21%
92	9	350	715 ms	83%
86	5	380	489 ms	84%

PERFORMANCE EVALUATION RESULTS

From the performance evaluation results the overall accuracy of the system was around 80% when tested with different NPK values. The response time of the system was below 800 ms. Thus the system developed is efficient and effective in recommending crop for agricultural stakeholders accurately and precisely.

6. CONCLUSION AND FUTURE WORK

By categorizing the soil samples according to the soil type, land type and macro nutrients Nitrogen (N), Phosphorus (P) and Potassium (K) present in the soil the suitable crop along with its appropriate fertilizer is suggested to the agricultural stakeholder. The month in which the yield will be high is also suggested to the user. The yield calculation is also provided for the crop selected by the farmer. The proposed crop recommendation system provides 82% of accuracy.

The future work is to implement Machine Learning Algorithms like Ensemble Classifiers to predict the crop yield and recommend the crop with appropriate fertilizer. In the existing system only soil characteristics were considered to provide crop recommendations. In the future work the climatic parameters will also be taken into account to provide crop recommendations.

7. REFERENCES

[1] Mansi Shinde, Kimaya Ekbote, Sonali Ghorpade, "Crop Recommendation and Fertilizer Purchase System", IJCSIT International Journal of Computer Science and Information Technologies, Volume 7, Issue 2, 2016.

[2] V. Sellam, E. Poovammal, "Prediction of Crop Yield using Regression Analysis", IJST Indian Journal of Science and Technology, Volume 9, Issue 38, October 2016.

[3] U.K. Diwan, H.V. Puranik, G.K. Das, J.L. Chaudhary, "Yield Prediction of Wheat at Pre-Harvest Stage Using Regression Based Statistical Model for 8 District of Chhattisgarh, India", IJCMAS International Journal of Current Microbiology and Applied Sciences, Volume 7, Issue 1,2018.

[4] Rushika Ghadge, Juilee Kulkarni, Pooja More, Sachee Nene, Priya.RL, "Prediction of Crop Yield using Machine Learning", IRJET International Research Journal of Engineering and Technology, Volume 5, Issue 2, February 2018 .

[5] P.Priya, U.Muthaiah, M.Balamurugan, "Predicting yield of the crop using machine learning algorithm", IJESRT International Journal of Engineering Sciences & Research Technology, Volume 7, Issue 4, April 2018.

[6] Vaneesbeer Singh, Vinod Sharma, Abid Sarwar, "Analysis of soil and prediction of crop yield (Rice) using Machine Learning approach", IJARCS International Journal of Advanced Research in Computer Science, Vol. 8, June 2017.

[7] Vrushal Milan Dolas, Prof. Uday Joshi, "A Novel Approach for Classification of Soil and Crop Prediction", IJCSMC International Journal of Computer Science and Mobile Computing, Volume 7, Issue 3, March 2018.

[8] R.Sujatha , Dr.P.Isakki, "A Study on Crop Yield Forecasting Using Classification Techniques", 2016 IEEE .

[9] Supriya D M, "Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining Approach", IJIRCCE International Journal of Innovative Research in Computer and Communication Engineering, Volume 5, Issue 5, May 2017.

[10] S.Veenadhari, Dr. Bharat Misra, Dr. CD Singh , "Machine learning approach for forecasting crop yield based on climatic parameters", International Conference on Computer Communication and Informatics (ICCCI -2014).