International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 9, Issue.7, July 2021 E-ISSN: 2347-2693

A Machine Learning Based Crop and Fertilizer Recommendation System

Supriya M.S.[1*], Nagarathna[2]

[1,2]Dept. of Computer Science, PES College of Engineering, Mandya, Karnataka, India

[*]*Corresponding Author: Supriygowdams@gmail.com, Tel.: +91-89704 51808*

DOI: https://doi.org/10.26438/ijcse/v9i7.6468 | Available online at: www.ijcseonline.org

Received: 19/Jul/2021, Accepted: 20/Jul/2021, Published: 31/Jul/2021

***Abstract*—** India is a country where agricultural and agriculture-related sectors provide the majority of the country's income. Agriculture is the country's main source of revenue. It is also one of the countries that has major natural disasters such as drought or flooding, which have caused crop devastation and repeated crop cultivation leads to soil degradations due to this farmers suffer significant financial losses as a result of this, leading to suicide. The goal is to build a machine learning model for crop and fertilizer recommendations system based on soil features which includes different types of parameters value such as PH, Organic Carbon, Nitrogen, phosphorus, potassium, sulphur, zinc, iron, temperature, rainfall. Naïve Bayes and LVQ algorithms are used for crop recommendations and KNN classifier are used for fertilizer recommendations. This system displays the results of a study on the machine learning approaches and compare with the neural networks to forecast the best crops recommendations. The Machine Learning algorithm gives more accurate results than CNN.

***Keywords*—** Crop and Fertilizer Recommendation, Naïve Bayes (NB), Machine Learning, Agriculture, Learning vector quantization (LVQ)

I. INTRODUCTION

Agriculture is having a great impact on the country's economy .Agriculture and associated areas like horticulture, sericulture, fishing, forestry and logging are responsible for 17% of the GDP and has provide around 49% of employment the total labor force in 2014. Agriculture is having a great impact on the country's economy. In the last decade India has seen serious natural calamities like drought or flood. Due to such disasters there is a huge loss to crop production and ultimately to the farmers. Due to such financial loss many farmers are committing suicide. If natural calamities are not present then there may be sudden pest attack destroying the crop. In any case farmer and the crop are always at the edge of risk. Government policies are there but that is not sufficient. Use of modern machineries and technologies has greatly improved the yield and quality of crops. The management of recurrent fruit crops in each crop cycle is also important for the yield in farms. There is a need of proper attention predominantly to monitor diseases which affect the production severely. Increase in use of pesticides has resulted into induced diseases which reduce the yield and quality significantly.

More than 60% of the land in the country is used for agriculture in order to suffice the needs of 1.3 billion people Thus adopting new agriculture technologies is very important. This will be leads the farmers of our country towards profit [1]. Prior crop prediction and yield prediction was performed on the basis of farmers experience on a particular location. They will prefer the prior or neighborhood or more trend crop in the surrounding region only for their land and they don't have enough of knowledge about soil nutrients content such as nitrogen, phosphorus, potassium in the land. Being this as the current situation without the rotation of the crop and apply an inadequate amount of nutrients to soil it leads to reduce in the yield and soil pollution (soil acidification) and damages the top layer. Considering all these problems takes into the account we designed the system using machine learning for betterment of the farmer. Machine learning (ML) is a game changer for agriculture sector. Machine learning is the part of artificial intelligence, has emerged together with big data technologies and high-performance computing to create new opportunities for data intensive science in the multi-disciplinary agri-technology domain. In the Agriculture field machine learning for instance is not a mysterious trick or magic, it is a set of well define model that collect specific data and apply specific algorithms to achieve expected results [7].

II. LITERATURE SURVEY

Fatin et al. [1] two different Machine Learning (ML) algorithms are proposed to analyze the crops' yield. These two algorithms, Support Vector Regression (SVR) and Linear Regression (LR), are quite suitable for validating the variable parameters in the predicting the continuous variable estimation with 140 data points that were acquired. The parameters mentioned above are key factors affecting the yield of crops. The error rate was measured with the help of Mean Square Error (MSE) and Coefficient of Determination (R2), where MSE gave out

approximately 0.005 and R2 gave around 0.85. The same dataset has been used for quick comparison between the algorithms' performances.

D Ramesh, et al [2], paper they focus on the applications of Data Mining techniques in agricultural field. Different Data Mining techniques are in use, such as K-Means, K-Nearest Neighbor (KNN), Artificial Neural Networks (ANN) and Support Vector Machines (SVM) for very recent applications of Data Mining techniques in agriculture field. In this paper consider the problem of predicting yield production and that remains to be solved based on the available data. The problem of yield prediction can be solved by employing Data Mining techniques. This work aims at finding suitable data models that achieve a high accuracy and a high generality in terms of yield prediction capabilities. For this purpose, different types of Data Mining techniques were evaluated on different data sets.

Monali Paul, et al [3], work presents a system, which uses data mining techniques in order to predict the category of the analyzed soil datasets. The category, thus predicted will indicate the yielding of crops. The problem of predicting the crop yield is formalized as a classification rule, where Naive Bayes and K-Nearest Neighbor methods are used. In this work, classification of soil into low, medium and high categories are done by adopting data mining techniques in order to predict the crop yield using available dataset. This study can help the soil analysts and farmers to decide sowing in which land may result in better crop production.

Ami Mistry, et al [4], in this research, they have considered the effects of environmental (weather), biotic (pH, soil salinity) and area of production as factors towards crop production in any place. Taking these factors into consideration as datasets for various districts and applied suitable data mining techniques to obtain crop yield predictions.

Sandhya et al [5], According to this study, there is still scope for the improvement in result. During the study which they have carried out it is observed that the algorithm which is used by most of the authors does not uses a unified approach where in all the factors affecting the crop yield can be utilized simultaneously for predicting the crop yield. There is still further scope of improvement as the dataset which is considered is small in some cases. Therefore the result can be improved by using a large dataset.

Chakrabarty et al [6], in the research suggests area based beneficial crop rank before the cultivation process. It indicates the crops that are cost effective for cultivation for a particular area of land. To achieve these results, they are considered six major crops which are Aus rice, Aman rice, Boro rice, Potato, Jute and Wheat. The prediction is based on analyzing a static set of data using Supervised Machine Learning techniques. This static dataset contains previous years' data according to the area. The research has an intent to use Decision Tree Learning- ID3 (Iterative Dichotomiser 3) and K-Nearest Neighbors Regression algorithms.

III. PROPOSED METHOD

Proposed system contains 2 major objectives, one recommending suitable Crop and secondly recommending the suitable Fertilizer based on the soil features and environment conditions. This system presents the study about the various Machine Learning techniques and compared with CNN technique to recommend the suitable crop.

We use suitable technology to work with real time application, that is "visual studio" as front end technology and "SQL server" as back end technology. These technologies are preferred because it supports more suitable libraries, tools and concepts required to work with real time application compared to other technologies. Proposed system helps farmers to cultivate right type of crops in right time and also helps farmers to increase crop yield by suggesting suitable fertilizers. Major population in India will benefit from this application. Supervised learning algorithms used for the recommendations such as "Naïve Bayes" and "K nearest neighbour" algorithm. These algorithms are preferred as they work efficiently, generated faster results and also work for all formats of data. and also few survey papers suggests these algorithms are efficient and good for agriculture data-sets.

A. Crop Recommendation

Learning vector quantization and Naïve bayes algorithms are used to recommend suitable crop. Figure 1 show the block diagram for crop recommendation where the real time data are the input for the system. Which is pre-processed to remove nosie from the original data and converted into desired format. After converted the data into desired format we train the data sets in order to recommend the suitable crop the machine learning and CNN algorithms are used to recommend the suitable crop.

LVQ is a prototype-based supervised classification algorithm. LVQ is the supervised counterpart of vector quantization system. There are "n" number of input units and "m" number of output units. The layers are fully interconnected with having weights on them. here we are using 10 parameters (i.e. PH value, organic carbon, nitrogen, potassium, zinc, phosphorous, iron , sulfur, temperature, Rainfall, crop) and 6 different types of crops(paddy, wheat, maize, barley, groundnut, mustard).this parameters are used to train the system and which follows the below steps :

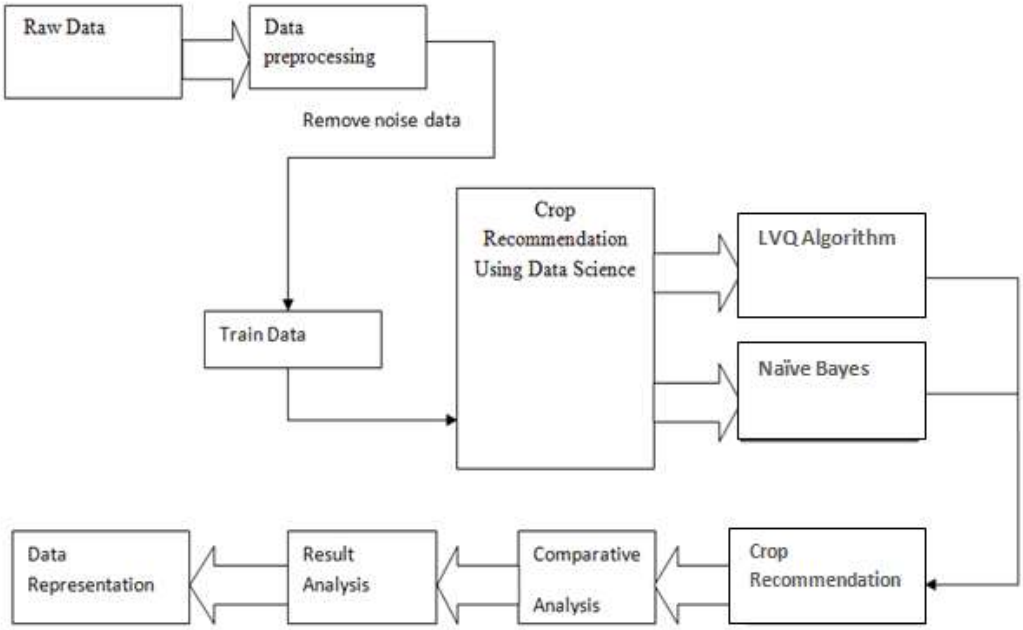

Figure 1: Block Diagram for Crop Recommendation

- For each training input vector x, do the following steps.
- Calculating distance using Euclidean distance.
- Compute the wining unit J where D(j) is minimum.
- Update the new weights of the J neuron as follows:
  - i. If target class T and class Cj is equal to J then wj(new)=wj(old)+α(X-wj(old))
  - ii. If target class T and class Cj is not equal to J then wj(new)=wj(old)-α(X-wj(old))
- Compare the values to classify the results.

After training the input vector calculate the distance between input vector with training vector for all six different crops using Euclidean distance where minimum distance is J then update the weights of the J neuron as shown in the above steps and compare the values to classify the results .To improve the accuracy of the system, one more classifier is used i.e, Naïve Bayes.

Naïve Bayes classifiers are simple "Probabilistic Classifiers" based on applying Bayes theorem and here we are using naïve bayes classifier for crop recommendation which consists different types of soil parameters and this are inputed to the system then it retrieves the required data to calculate the each attribute value i.e attribute value vj , input value v, outcome s where , n is the number of training examples for which v=vj, n_c is the number of examples is classified to s for which v=vj and p is a priori estimate for p(aj), number of attributes is m. After calculating the attributes value then apply the formulae i.e p=(n_c+mp)/(n+m) for all six types of crops and Multiplies the probabilities by p then classifies the results based on this probabilities value the crop which contains maximum probabilities values that crop is classified for new input data. Comparing the result of naïve bayes classifier with LVQ ,the naïve bayes gives more accurate result than LVQ.

B. Fertilizer Recommendation.

The K-nearest neighbor algorithm (KNN) are used to recommend suitable fertilizer and figure 2 shows the block diagram for Fertilizer Recommendation. KNN is a non-parametric classification approach that is employed. The dataset contains different types of parameters (i.e. PH, organic, nitrogen, phosphorus, potassium, sulphur, zinc, iron, temperature, rainfall, crop and fertilizer) and recommending fertilizer for 6 different crops (paddy, wheat, maize, barley, groundnut, mustard). This set of parameters are input to the system to train the dataset then determine the k values and calculate the distance. The Euclidean formula is used to calculate the distance, for the crop paddy the parameters values are 6.1, 7.98, 136.4 for PH, Organic Carbon, Nitrogen respectively and the training datasets parameters PH, Organic Carbon, Nitrogen contains following values are 5.7, 7, 188.43 respectively. Calculate the Euclidean Distance is as follows: d (pi, rj)=√(5.7- 6.1)2+(7.7.98)2+(188.43-136.4)2 where , pi is a training dataset and rj is a new input values.

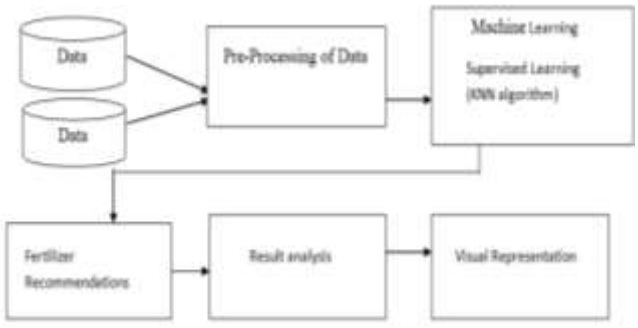

Figure 2: Block Diagram for Fertilizer Recommendation.

After calculating the distance between input data and training data of each row sort the result in ascending order from the sorted array chooses the top k rows and classify the results.

IV. EXPERIMENTAL STUDY

The dataset considered for usage in the given proposed work is a soil dataset primarily comprising of soil physical and chemical properties, along with the environment conditions details. The Data-sets collected from agriculture department of two region of Karnataka i.e. Mandya, Mysore. Which contains 4 sets of datasets crop and fertilizer datasets of Mandya region and crop and fertilizer datasets of Mysore region.

The datasets which contains following parameters PH value, Organic Carbon, Nitrogen, Potassium, Zinc, Phosphorous, Iron , Sulfur, Temperature, Rainfall, Crop, yield, fertilizer. The crops considered are paddy, wheat, maize, barley, groundnut, mustard and all these parameters are used for crop and fertilizer recommendation.

V. RESULTS

The results in Table 1 shows the soil features such as PH, OC, Nitrogen ,Phosphorus ,Potassium , Sulphur , Zinc ,Iron , Temperature, Rainfall and Crop which their want to grow. Our model suggest the suitable fertilizer based on this soil features using KNN Classifier for the following six different crops are paddy, Maize, Wheat, Groundnut, Mustard, Barley.

TABLE 1 - Fertilizer Recommendation Results

Crops	PH	Organic carbon	Nitrogen	phosphorus	Potassium	sulphur	Zinc	Iron	Temperature	Rainfall	Recommended Fertilizer
Paddy	6.7	7.15	125	7.15	6	92.67	125	6.7	26.28	105.43	Urea super granulated
Maize	7.4	5.9	126.3	5.9	6.4	177.8	126.3	7.4	25.28	105.43	Nitrate nitrogen
Wheat	7	5.7	123	5.7	6.2	127	123	7	25.75	105.14	Amide nitrogen
Groundnut	6.4	7	158	7	7.15	191.2	158	6.4	26.12	68.21	Ammonium chloride
Mustard	7.2	6	127.4	6	6.3	110	158	7.2	26.12	68.21	Urea_45N
Barley	5.7	6.2	161.2	6.2	5.7	120.6	127.6	5.76	25.95	140.6	Urea super granulated

TABLE 2 - Crop Recommendation Results

H	Organic carbon	Nitrogen	phosphorus	Potassium	sulphur	Zinc	iron	Temperature	Rainfall	Recommended Crop
6.7	7.15	125	7.15	6	92.67	125	6.7	26.28	105.43	Paddy
7.4	5.9	126.3	5.9	6.4	177.8	126.3	7.4	25.28	105.43	Maize
7	5.7	123	5.7	6.2	127	123	7	25.75	105.14	Wheat
6.4	7	158	7	7.15	191.2	158	6.4	26.12	68.21	Groundnut
7.2	6	127.4	6	6.3	110	158	7.2	26.12	68.21	Mustard
5.7	6.2	161.2	6.2	5.7	120.6	127.6	5.76	25.95	140.6	Barley

TABLE 3 - Naïve Bayes and LVQ Comparison table for crop recommendation

Constraint	Naïve Bayes Algorithm		LVQ Algorithm	
Regions	Mysore	Mandya	Mysore	Mandya
Accuracy	96%	95%	81%	88.25%
Time (milli secs)	2609	16148	2164	10721
Precision	96%	95%	81%	88.25%
Recall	4%	5%	19%	11.75%

The results in Table 2 shows the soil features such as PH, OC, Nitrogen, Phosphorus, Potassium, Sulphur, Zinc, Iron, Temperature and Rainfall. Our model suggest the suitable crop based on this soil features using Naïve Bayes.

The results in Table 3 shows the comparison of Naïve Bayes and LVQ Algorithm for Crop Recommendations and displays their accuracy which shows that Naïve Bayes gives more accurate result than LVQ for both the regions.

VI. CONCLUSION

Proposed system recommend the crop and fertilizer using data science techniques based on the soil tested results which shows the Naïve Bayes gives more accurate result than LVQ and its accuracies are 96% and 81% respectively for Mysore region and for Mandya region its accuracies are 95% and 88.25% but Naïve Bayes taken more time compare to LVQ for crop recommendation and for fertilizer recommendation KNN gives 92.22% accuracy for Mysore region and 79.67% accuracy for Mandya region. This system also useful to agriculture departments to recommend the right crop in right time. The goals that have been achieved by the developed system are, simplified and reduced the manual work, Large volumes of data can be stored and It provides Smooth work flow. Here, in this work, six different crops of two regions are chosen for classification, it can be enhanced further by choosing large number of datasets and crops.

REFERENCES

[1] F. F. Haque, A. Abdelgawad, V. P. Yanambaka and K. Yelamarthi, "Crop Yield AnalysisUsing Machine Learning Algorithms," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), **pp. 1-2, 2020,** doi: 10.1109/WF-IoT48130.2020.9221459.

[2] D. Ramesh , B. Vishnu Vardhan, "Data Mining Techniques and Applications to Agricultural Yield Data", International Journal of Advanced Research in Computer and Communication Engineering **Vol. 2, Issue 9, September 2013.**

[3] Monali Paul, Santosh K. Vishwakarma, Ashok Verma , "Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach", International Conference on Computational Intelligence and Communication Networks. **2015.**

[4] Ami Mistry1, Vinita Shah , "Brief Survey of data mining Techniques Applied to applications of Agriculture" , International Journal of Advanced Research in Computer and Communication Engineering **Vol. 5, Issue 2, February 2016** Copyright to IJARCCE DOI 10.17148/IJARCCE.2016.5263 301 .

[5] Yogesh Gandge, Sandhya , "A Study on Various Data Mining Techniques for Crop Yield Prediction", International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques **2017** (ICEECCOT) 978-1-5386-2361-9/17/$31.00 ©2017 IEEE.

[6] Md. Tahmid Shakoor, Karishma Rahman, Sumaiya Nasrin Rayta, Amitabha Chakrabarty, "Agricultural Production Output Prediction Using Supervised Machine Learning Techniques", 978-1-5386-3831-6/17/$31.00 2017 IEEE.

[7] R. N. Bhimanpallewar and M. R. Narasingarao, "Alternative approaches of Machine Learning for Agriculture Advisory System," 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), **pp. 27-31, 2020,** doi: 10.1109/Confluence47617.2020.9058152.

[8] Nischitha K, DhanushVishwakarma, Mahendra N, Ashwini, Manjuraju M.R, "Crop Prediction using Machine Learning Approaches", International Journal of Engineering Research & Technology (IJERT), **Vol. 9 Issue 08, August-2020.**

[9] Fahim Jawad, Tawsif Ur Rahman Choudhury, S M Asif Sazed, Shamima Yasmin , Kanaz Iffat Rishva, Fouzia Tamanna, Rashedur M Rahman ,"Analysis of Optimum Crop Cultivation Using Fuzzy System", 978-1-5090-0806-3/16/$31.00 copyright 2016 IEEE ICIS **2016, June 26-29, 2016,** Okayama, Japan.

[10] R. B. Guruprasad, K. Saurav and S. Randhawa, "Machine Learning Methodologies for Paddy Yield Estimation in India: a Case Study," IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, **pp. 7254-7257, 2019,** doi: 10.1109/IGARSS.2019.8900339.

AUTHORS PROFILE

Miss. *Supriya M.S.* Pursuing Master of Technology in Computer Science and Engineering from PES College of Engineering, Mandya. She has received her Bachelor's degree in Computer Science and Engineering from Academy for Technical and Management Excellence (ATME), Mysuru.

Dr. Nagarathna currently working as Professor in Department of Computer Science and Engineering, PES College of Engineering, Mandya

