IMPLEMENTATION PAPER ON AGRICULTURE ADVISORY SYSTEM

Pooja K C[1], Pooja K P[2], Pooja N G[3], Dr.A B Rajendra[4]

[1-3] *Students, Information Science and Engineering*
[2]*Head of Department, Dept. of Information Science and EngineeringVidyavardhaka College of Engineering, Mysuru, Karnataka, India*

Abstract

India is a country where the primary economic sectors are agriculture and industries that are closely related to it. It has become one of the nations that have had significant natural catastrophes like drought or flooding that have destroyed crops. Predicting acceptable crops to produce and suggesting proper fertilizers to improve crop output long before the harvest helps farmers and the government in making appropriate plans such as storing, selling, determining minimum price, importing/exporting, and so on. A thorough analysis of enormous volumes of data obtained from numerous variables, including soil quality, pH, EC, N, P, and K, is necessary for an advanced crop prediction. As it involves an enormous number of databases, this prediction method is an absolute option for data science applications. We take out insights from huge amounts of data using data science. This system provides a study on the various machine learning methods for predicting the best crops and recommending fertilizer. The performance of any crop prediction system depends critically on how accurately features are extracted and how well classifiers are employed. We have used two algorithms that are Naïve Bayes and KNN for predicting the crop and fertilizers respectively.

Key Words: **Crop Prediction, Data Science, Machine Learning, Naïve Bayes, KNN**

1. INTRODUCTION

India's main source of income is agriculture, which is an example of a sector that contributes only approximately 14% of GDP yet has a huge impact on the economy of the country. To raise the crop output and the farmers' living conditions, better methods are needed. Agriculture has evolved as a result of globalization, adopting the latest technologies and practices for a higher level of living. One of the more recent technologies and methods in agriculture is precision agriculture. Crop and fertilizer recommendations are one of the most important aspects of precision agriculture. Crop suggestion is based on several factors,[6] and precision agriculture technologies aid in detecting these factors, allowing for improved crop selection. Farmers may benefit from knowing the relative contributions of climate factors on agricultural output in order to plant and manage crops in the face of climate

Based on input and training data, machine learning can forecast suitability levels using a variety of techniques. In the suggested task, we forecast crops and provide the required fertilizers to increase agricultural output. For crop production estimation and fertilizer recommendations, we employ a variety of agricultural characteristics. The current agricultural field is beset by issues, one of which is inadequate for farmers. Farmers plant crops but do not receive adequate yields, resulting in lower profits. Crop recommendations and appropriate fertilizer recommendations are critical in the agriculture department, and appropriate crops and fertilizers are determined by a variety of factors like soil, rainfall, and temperature characteristics. It is critical to make early predictions based on these limits. ML can be utilized effectively to assist farmers in achieving the highest potential output. The list of partsthat make up this paper's organization is given below. Section 2 comprises the literature survey, which also contains a table that compares the paper of different authors. Section 3 comprises the methodology, whichhas 8 steps in it, and also contains a brief description of the algorithm used in the project. Section 4 contains the overview of the proposed system. Section 5 contains the conclusion and References are found in the Section 8 part.

2. LITERATURE SURVEY

1. Sally Jo Cunningham and Geoffrey Holmes,[10] proposed the Developing innovative applications in agriculture using data mining. The methods involved inthis are Weka classifiers: ZeroR, OneR, Naïve Byes, Decision Table, Ibk, J48, SMO, Linear Regression, M5Prime,LWR, Decision Stump and association rule include apriori algorithm and also it includes the EM clustering algorithm for the purpose of clustering. This produces a classifier, which is frequently in the decision tree form or a collection of guidelines that canbe used in predicting the categorization of newer data instances. This approach recognizes that machine learning technology is still expanding and improving, with learning algorithms that must be delivered to thepeople who deal with the data and are also familiar with the application domain from which it originates. Weka is a huge step forwardin bringing machine learning into the workplace.

2. D Ramesh, and B Vishnu Vardhan detailed Data Mining Techniques and Applications to Agricultural Yield Data [11]. The KNN and K-Means Algorithms are used in this. The system can anticipate the average yield production by examining the cluster to which the forecasted rainfall belongs in this procedure, given the rainfall in a specific year. It also mentions that the K- Means algorithm can split samples into clusters, but no consideration is given to the substances that cause this partition. This type of information can be obtained using bi-clustering.

3. Monali Paul, at el.[12] Described the Analysis of soil Behavior and Prediction of Crop Yield using the Data Mining Approach in the year 2015. For its system to work, it employs the KNN and the Naïve Bayes Algorithm. This demonstrates that the category with the highest confidence value is projected to be the category of that specific soil.It also suggests that this research can assist soil analysts and farmers in determining which land to seed on in order to maximize agricultural yield.

4. Ami Mistry and Vinita Shah[13] described the Brief Survey of data mining Techniques Applied to applications of Agriculture in the year 2016. This system employs a variety of approaches, including Linear Regression, KNN, Regression Tree, and SVM for classification. K-means clustering, Self-organized maps, Density-based clustering, and Weight-based clustering are examples of clustering techniques. The results indicate that rice production variability is influenced by sunshine hours and daily temperature variation in the current research area. According to this, farmers might plant different crops in different regions based on basic forecasts produced by this research, and if that happens, every farmer would have a chance to boost the country's overall productivity.

5. Yogesh Gandge, and Sandhya[14] described A Study on Various Data Mining Techniques for Crop Yield Prediction in the year 2017. This system uses a Classification Algorithm. The agricultural yield prediction per acre is the outcome, along with some recommendations. It is noted that the method used by most experts does not employ a unified approach in which all elements impacting crop production can be used simultaneously for crop yield prediction.

6. Md. Tahmid Shakoor, et al.[15] described Agricultural Production Output Prediction Using Supervised Machine Learning Techniques in the year 2017. This system uses the KNN algorithm, Decision Tree algorithm, and ID3 (Iterative Dichotomis) algorithm. Without eliminating the dataset's outliers, the Decision Tree Learning- ID3 approach produces alower percentage error value than the KNN technique. Though the study is confined to a single set of data, more data will be added in the future, which will be analyzed using more machine learning algorithms.

7. Umid Kumar Dey et al.[16] described the Rice Yield Prediction Model Using Data Mining in the year 2017.This system uses the k-means algorithm, Linear Regression, SVM Regression, and Modified Non-linear Regression. This system has been found to be quite good at predicting yields, with SVM regression providing the best performance. The modified Non- linear regression equation is found to perform better than the other three predefined models. It also establishes that the MNR equation is the most appropriate.

8. Kuljit Kaur and Kanwalpreet Singh Attwal[17] proposed "Effect of Temperature and rainfall on Paddyyield using Data Mining in the year 2017. They used the Apriori Algorithm for their study, which gave the resultthat predicts the growth of the paddy yield. This depended on various parameters like Rainfall and Temperature. In the end, they concluded that with the increase in rainfall the paddy yields also increased. During the reproductive phase, the rainfall and temperature did not influence, and also during the maturation phase, the yield was better expected at thelower temperature and worse at a higher temperature.

Table -1: Literature Review

S.No.	Title	Author	Method used	Results
1.	Developing innovative applications in agriculture using data mining	Sally Jo Cunninghamand Geoffrey Holmes	Weka classifier: ZeroR, OneR, Naive Bayes, Decision- Table, Ibk, J48, SMO, Linear Regression. Association Rules: Apriori Algorithm Clustering: EM Clustering Algorithm.	The end result is a classifier, which is typically in the shape of a decision tree or guidelines thatcan be used to anticipate how fresh data examples will be classified.
2	Data Mining Techniquesand Applications to Agricultural Yield Data.	D Ramesh, B Vishnu Vardhan	KNN, K-Means algorithm.	By examining the cluster in whichthe projected rainfalls, the system can predict the average yield production based on the rainfall in a certain year.
3	Analysis of Soil Behavior and Prediction of Crop Yield using Data MiningApproach.	Monali Paul, Santosh K. Vishwakar- ma, Ashok Verma.	KNN and Naive Bayes.	We can see from the results that the category with the highest confidence value is expected to be the soil category for that specific soil.
4	Brief Survey of data mining Techniques Applied to applications of Agriculture	Ami Mistry, Vinita Shah	Classification technique: Linear Regression, K - nearest neighbor Clustering Technique: K- means clustering, Self-organized maps, Density-based clustering.	The results indicate that in the current study area, sunlight hoursand the range of daily temperature have important roles in rice yield variations.
5	A Study on Various Data Mining Techniques for Crop Yield Prediction	Yogesh Gande and Sandhya.	Classification Algorithm.	The output is a projection of agricultural yield per acre along with some suggestions.
6	Agricultural Production Output Prediction Using Supervised Machine Learning Techniques	Md. Tahmid Shakoorand others.	KNN, Decision Tree algorithm, ID3(Iterative Dichotomis) algorithm	The result reveals that without eliminating the dataset's outliers,the Decision Tree Learning- ID3 approach delivers a lower percentage error value than the KNN technique.
7	Rice Yield Prediction Model Using Data Mining	Umid Kumar Dey and others.	K-Means, Multiple Linear Regression, Modified Nonlinear regression.	It can be observed that it predicts yields well, with SVM as best result.

3. METHODOLOGY

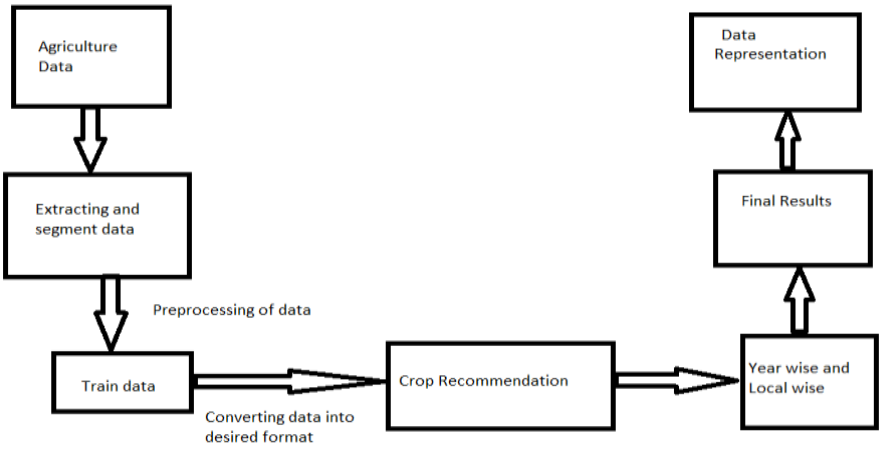

Fig 1 : Methodology

Step 1: Raw data and Weather Statistics

This is the first step in the crop recommendation process where we collect agriculture data. Agriculture data, including agriculture characteristics, crop details, farmer details, and yield details, was obtained from the region "Mysore." Rainfall, temperature, and soil [6] characteristics such as PH, nitrogen, potassium, iron, and so on are all considered agriculture parameters.

Step 2: Extract and Segment Data (Data Pre-processing)

Agriculture data is analyzed here, and only relevant information is taken. The information needed for processing is extracted and split into various regions. Because the complete agriculture data is not necessary for processing, and because if we input all data, processing would take too long, data processing is performed

Step 3: Train Data

After the relevant data has been extracted
and segmented, the data must be trained, which involves translating the data into the required format, such as numerical values, binary, string, and so on. The type of algorithm determines the conversion.

Step 4: Supervised Learning

Designing and analyzing data-learning systems is a major focus in the field of machine learning.

KNN algorithm

It is one of the effective algorithms that just operate on numbers. For numerical data, it functions and moves through data more quickly than other algorithms. It operates based on distance calculation.

Flow of Algorithm.

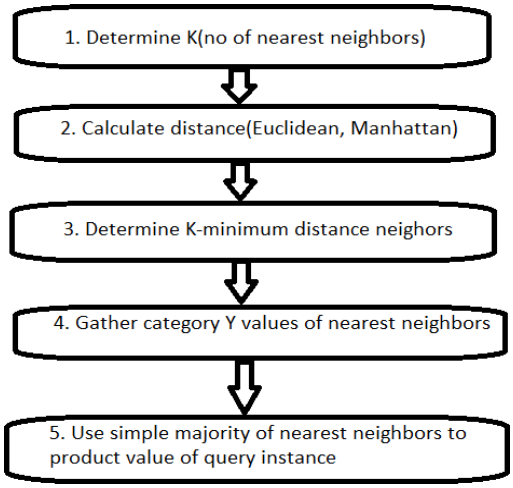

Fig 2: Flow of KNN algorithm

Supervised learning is a type of machine learning that employs training data that includes predicted responses.

Naive Bayes Algorithm is used for crop recommendation because of the following reasons:

- Efficient classifier
- Works fine for a smaller number of parameters as well as a greater number of parameters.
- Works fine for small and bigdata-set.
- More accurate results

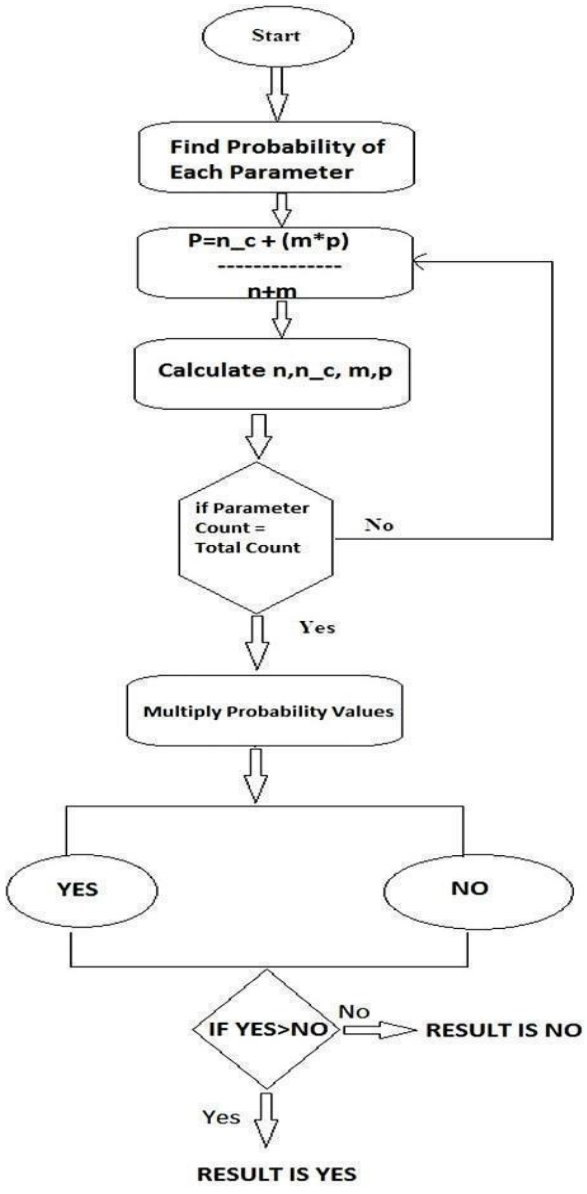

Fig 3: Flow of Naïve Bayes algorithm

Step 5: Crop Recommendation (Priority- wise)

Here suitable crops are recommended for the farmers which may yield high profits. The naïve Bayes algorithm generates outputs (crop recommendations) based on priority wise.

Step 6: Location and Year-based

The crop recommendation is based on region-wise and year- wise.

Step 7: Results

Advising suitable and high-profit crops to farmers is done in priority order. Here high probability crops are extracted and sorted and the top 3 crops are recommended for the farmers.

Step 8: Visual Representation

Crops are recommended for the farmers on GUI. When users get to log in, the application system recommends suitable and high-profit crops for the farmers on a GUI.

4. PROPOSED WORK

The The proposed system contains 2 major objectives, one recommending suitable fertilizer based on the soil type and features, and secondly suggesting the suitable crop dependent on environmental conditions. Machine learning algorithms are applied to analyze data and to recommend suitable fertilizers and suitable crops. Data sets were collected from agriculture departments. Rainfall, temperature, PH value, nitrogen, potassium, zinc, phosphorous, iron, etc. all these parameters are used for recommendation. The system developed as a real-time application that is useful for agriculture departments and farmers.

We use suitable technology to work with real-time applications, in the front-end we use visual studio, and in the back-end SQL server is used. These technologies are preferred because it supports more suitable libraries, tools, and concepts required to work with real-time application compared to other technologies. The recommended system helps farmers to grow the suitable type of crops at the correct time and also by suggesting suitable fertilizer for increased crop yield. This application will be helpful to the majority of Indians. The Bayesian classifier, K nearest neighbor, or Random Forest algorithms are examples of supervised learning algorithms that are utilized for the recommendations. These algorithms are chosen because they produce quicker results, operate effectively, and support all data formats. A few survey publications have also suggested that these algorithms are effective and suitable for data sets related to agriculture.

3. CONCLUSIONS

Nowadays farmers face ample issues and they don't recognize the correct info concerning crops to grow and cultivate. This planned system helps farmers to understand the correct crop to grow. The planned system predicts the crop's victimization information science techniques supported the soil tested results. This method is additionally helpful to agriculture departments to predict the correct crop right time. If we've got such reasonable automation, is going to be helpful to farmers and agricultural fields. The created system achieves the objectives by streamlining and reducing manual labor, preserving enormous amounts of knowledge, and facilitating an efficient workflow. We've seen how techniques like data mining and machine learning algorithms can enable us to analyze data and recommend the most appropriate fertilizer and crops for a given piece of land. The recommendations are made using a supervised learning algorithm, like a Bayesian classifier, K nearest neighbor, or Random Forest algorithm. These algorithms have been chosen because they work reliably, provide quick results, and are compatible with every data format. The system makes use of data gathered by the agriculture department. However, the proposed system is a real- world application aimed at agriculture departments, assisting farmers in cultivating suitable crops and high profit.

REFERENCES

[1] Maria Rossana C.de Leon,Eugene Rex L.Jalao, "A prediction model framework for crop yield prediction," Asia Pacific Industrial Engineering and Management System.

[2] Ramesh A. Medar,Vijay S. Rajpurohit, "A survey on Data Mining Techniques for crop yield prediction", International Journal of Advance Research in Computer Science and Management Studies. Volume 2, Issue 9, Sept 2014.

[3] D Ramesh, B Vishnu Vardhan, "Region specific crop yield Analysis: A data Mining approach ", UACEE International Journal of Advancees in computer Science and its Applications-IJCSIA volume 3: issue 2.

[4] George RuB, "Data Mining of Agricultural Yield Data:A comparison of Regression models.

[5] Sudhanshu Sekhar Panda, Daniel P. Ames, and Suranjan Panigrahi, "Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques", Remote Sensing 2010, 2, 673-696; doi:10.3390/rs2030673

[6] S.Veenadhari,Dr. Bharat Misra ,Dr. CD Singh," Machine Learning Approach for forecasting crop yield based on climatic parameters", 978-1- 4799-2352- 6/14/$31.00 ©2014 IEEE

[7]https://www.ifad.org/en/web/latest/event/asset/41949435 June, 2020

[8] Digital Green (2019). Farm Stack. A Digital Green Web Publication. Available online at: https://farmstack.co/.

[9] Das, A., Basu, D., and Goswami, R. (2016). Accessing agricultural information through mobile phone: lessons of IKSL services in West Bengal. Indian Res. J. Extension Educ. 12, 102–107. https://api.semanticscholar.org/CorpusID:168547286

[10] https://www.researchgate.net/publication/2606155_Developing_Innovative_Applications_in_Agriculture_Using_Data_Mining

[11] Data Mining Techniques and Applications to Agricultural Yield Data D Ramesh, B Vishnu Vardhan Associate Professor of CSE, JNTUH College of Engineering, Karimnagar Dist., Andhra Pradesh, India1Professor of CSE, JNTUH College of Engineering, Karimnagar Dist., Andhra Pradesh, India2

[12] Monali Paul, at el. Described the Analysis of soil Behaviour and Prediction of Crop Yield using the Data Mining Approach in the year 2015 https://www.researchgate.net/publication/306302068_Analysis_of_Soil_Behaviour_and_Prediction_of_Crop_Yield_Using_Data_Mining_Approach

[13] Umid Kumar Dey et al. described the Rice Yield Prediction Model Using Data Mining in theyear 2017. https://www.researchgate.net/publication/316906792_Rice_yield_prediction_model_using_data_mining

[14] Ami Mistry and Vinita Shah described the Brief Survey of data mining Techniques Applied to applications of Agriculture in the year 2016 https://www.researchgate.net/publication/301214822_IJARCCE_Brief_Survey_of_data_mining_Techniques_Applied_to_applications_of_Agri culture

[15] Kuljit Kaur and Kanwalpreet Singh Attwal proposed “Effect of Temperature and rainfall on Paddy yield using Data Mining in the year 2017. https://www.researchgate.net/publication/317424300_Effect_of_temperature_and_rainfall_on_paddy_yield_using_data_mining.