FARMER'S ANALYTICAL ASSISTANT:A SURVEY

Aishwarya Jadhav, Ankita Mande, Dhanashree Gurav, Meghana Sarje.

Under the guidance of S.S.Kare

Abstract—**Agriculture is main the occupation in India but it contributes only 18 percent of GDP to the Indian economy which is not satisfactory. There are many reasons behind India's poor production in the agriculture field is lack of crop planning by farmers or various environmental and economic factors that influence crop production but unpredictable changes in these factors lead to a great loss to farmers. These risks can be measured when appropriate methodologies are applied to data related to soil, weather and past yield. By using data mining, crop yield can be predicted by deriving useful patterns from these agricultural data that helps farmers to decide the crop they would like to plant for the upcoming year for gaining maximum profit. This paper presents a survey of the various techniques used for crop yield prediction.**

Keywords—**Crop yield prediction, Agriculture, Data Mining.**

I. INTRODUCTION

A possible reason for the poor contribution of the agricultural sector to the GDP of India may be the lack of adequate crop planning by farmers as well as by the government. Rapid fluctuations in crop prices are common in the market. In such a scenario, it is difficult for a farmer to make an educated choice of crop to grow in their land or to estimate the yield and price to expect from it. By applying farmer's previous experience for a particular crop, one can make the predictions for crop but For the better crop production, the farmers definitely requires a suitable guidance to predict the future of crop yield and also an analysis is to be made in order to help the farmers to increase crop production for their crops. As every farmer is interested future of crop yield and also

analysis is to be made in order to help the farmers to increase crop production for their crops. As every farmer is interested in knowing how much yield is expected to be grown in their land. Data mining is the process that is used to find out useful patterns from a large amount of data. It can be used for identifying clusters in data and includes a set of rules that describe each category or class in a data set. It is also used to find out hidden, valid, and useful patterns in huge data sets. Data Mining is all about discovering previously unknown relationships amongst the data. Data mining is also called as Knowledge discovery, Knowledge extraction, data/pattern analysis, information harvesting, etc..[9] Data Mining is the method of extraction, transforming, loading and predicting the meaningful information from huge data to extract some patterns and also transform it into an understandable structure for further use.[9]

Data mining procedure is separated into seven methods [15]:

• **Data cleaning**
• **Data integration**
• **Data selection**
• **Data transformation**
• **Data mining**
• **Pattern estimation**
• **Knowledge display**

In this paper, the main aim is analyzing different techniques to maximize and predict the crop yield productivity, one can make use of different data mining techniques.

II. RELATED WORK

Various researchers in the past explain the different kinds of techniques in agriculture sector.

Anshal Savla et al [1] which make a relative study of categorization algorithms and their performance which helps to know the yield and predict it in precision agriculture. These algorithms are

implemented in the various data sets which are collected from various fields and form long back years which helps in yield prediction on ragi crop.

Aakunuri Manjula et al [2] This paper concludes the requirement for crop yield prediction and its major usage and the role in a nation's planned guiding principle which is made in the agriculture development field. A Framework is developed called Extensible Crop Yield Prediction Framework (XCYPF) is developed. It provides facilities to flexible addition of a collection of techniques and methods towards crop yield prediction tool called XCYPF is developed in order to help people to predict more varieties of crops and their yield prediction by using independent and dependent variables.

Yash Sanghvi et al [3] states the procedure of agricultural data through data mining and visual data mining techniques is predicted. It helps to reduce high dimensional agricultural data into smaller size helps to acquire the useful knowledge of the data which are related to the yield, by taking input attributes like fertilizers.

Ramanujam et al [4] state that the recommendation system to be developed for the different types of soils using big data analytics and other various factors.

Paul Monali K. et al [5] acknowledges the significance of crop selection and some of the factors which decide the selection of crops like a market price, production rate and policies of the government that are discussed. In this paper, it also proposes Crop Selection Method (CSM) that is used to solve the yield of the crop and selecting the crop and requires to improve the net yield rate of the crop. It predicts the series of crops which need to be selected according to the seasons including factors like soil type, weather crop type, water type which are all used in the estimation of crop yield for crops like cereal in some major districts of India.

Rakesh Kumar et al [6] aims in solving the problem-related in food problems in the country of Egypt. A framework has been proposed that predicts the production, and it also imports of that meticulous year. Artificial Neural Network and Multi-Layer

Prescription is proposed for perception WEKA to build the prediction.

Xu Y, Zhou W et al [7] In this paper yield is analyzed and category is predicted. Crop yield is identified as Classification rule which is taken from a prediction which depicts Naive Byes and k-Nearest Neighbor algorithms that are used in prediction for yield prediction of soil.

Yogesh gandge et al [8] Discussed various data mining techniques used for predicting the crop yield and describes the importance of accurate data extraction methods of big data analytics.

Kaur et al [12] analyze the different Data Mining techniques to find a suitable data model that helps in achieving high accuracy for price prediction. Dataset of Coimbatore market price of tomato collected and the price is predicted using BP neural network and the result is simulated using MATLAB [12].

III. BASIC MODEL FLOW

India is the producer of many crops. There can be many ways to divide the types of crops they can be based on area, season, economic value, etc. The current study considers the data sets of all crop yield, temperature, rainfall, the humidity of soil and soil type.
 The Farmer's Analytical Assistant system can be used for better planning and decisions to improve the yield.

A. Input

Most of the research papers that were have considered some climatic parameters data related to previous crops like weather data, rainfall, temperature, humidity, soil, and soil type, can be taken as input.

B. Pre-processing

The data which is collected from past data and resources are in the raw form. It may contain some incomplete data, having some NULL, inconsistent data. In this step, such incomplete data should be altered. Therefore in this step, such redundant data should be filtered i.e., NULL values can be replaced with default values.

C. Attributes Extraction(Feature selection)

The classification of the dataset by identifying and using the most relevant attributes form dataset. The algorithm is applied to predict the crop.

D. Output

The output is the list of crops with respect to its market price per hector.

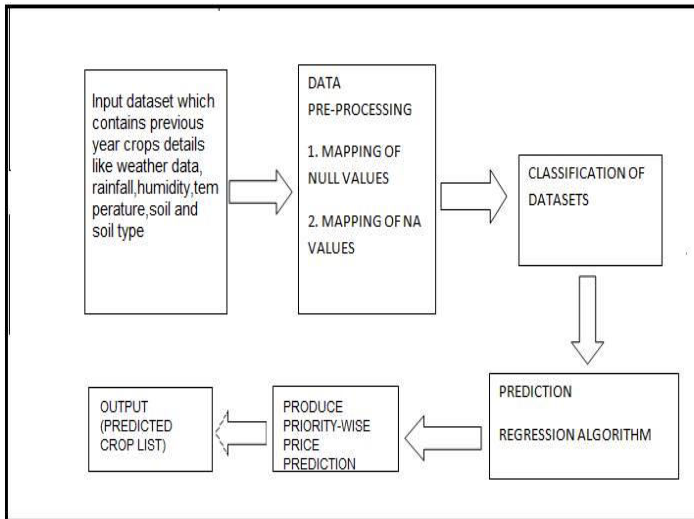**This is an outline of Farmer's Analytical Assistant:**

Figure 1: Schematic Representation of Farmers Analytical Assistant

A brief overview of Farmer's Analytical Assistant :

Step 1: Take the agriculture input dataset which contains the crop details.

Step 2:The data set is pre-processed to remove the improper values for analysis.

Step 3:An appropriate and efficient classification algorithm should be used.

Step 4: The implication is to apply Regression algorithms to predict the crop yield.

Step 5: Prediction or recommendation can be provided to the farmers based on the results of the crop list obtained.

Various techniques are used for prediction of crop yield are discussed below table :

Author and Publications	Techniques used	Description	Limitations
Bendre, 2015 [10]	Map Reduce and Linear Regression algorithm	They are used for weather forecasting	The forecasting is done based on only a weather data.
Yogesh Gandge, 2017 [8]	Support Vector Regression model	To get the excellent accuracy for price forecast in the agriculture products market this method is utilized.	The data which is not stable, the prediction accuracy of GM(1,1) is low.
	K-Nearest Neighbor and Naive Bayes	Support vector regression can serve as a better reference model for yield prediction.	SVM does not perform very well, when the data set has more noise i.e. target classes are overlapping.
Jiajun Zong, 2012	Grey Prediction Method	To get the excellent accuracy for price forecast in the agriculture products market this method is utilized.	The data which is not stable, the prediction accuracy of GM(1,1) is low.
Raorane, 2012 [11]	Arti_cial Neural Network (ANN), Decision	Crop production techniques are discussed.	Requires huge processing time for large

	Tree, Support Vector Machine (SVM) and K means		other methodologies does not give e_cient results in all types of forms
Kaur , 2014[12]	The use of BP neural network and simulation using MATLAB.	Price prediction data model is discussed and high accuracy is achieved.	Only price prediction is achieved.
Sellam , 2016[12]	Regression Analysis (RA), Linear Regression (LR) are discussed.	Describes various factors affect on crop yield.	The linear regression does not give correct results on variable data.
Aakash G Ratkal[13]	Support Vector Regression(SVM), Non-Linear Regression	Suggests best crop choice and income from crop	NIL

IV. CONCLUSION

Farmer's Analytical Assistant tool intends to help farmers to make better choices about the crop which they plans to grow next. Hence features like crop prediction and price prediction which will help the farmer make a reasonable estimate of the price and yield they may get. so the loss can be avoided. Unlike this more crop productions can be predicted. The methodologies can be used with data of other crops to study their relationship with essential climatic parameters.

From the above analysis, SVM and Non-Linear Regression show an average accuracy of less than 10 percent difference between the predicted and actual market price [13]. Non-Linear Regression algorithm increases the accuracy of the system and SVM is used to analyze the past data of various attributes.

References

[1] 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).

[2] Aakunuri Manjula, Dr.G.Narsimha (2015), `XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction' , Conference on Intelligent Systems and Control (ISCO).

[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] Ramanujam, R.Harine Rajashree, C.Kavya, T.Kiruthika, J.Nisha.Pudumalar(2016), `Crop Recommendation System for Precision **Agriculture'(ICoAC).**

[5] Paul Monali K. Vishwakarma Santosh Ashok Verma(2015) "Analysis of Soil Behaviour and Prediction of Crop Yield usingData Mining Approach" International Conference on Computational Intelligence and Communication Network.

[6] 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).

[7] Xu Y, Zhou W, Cui B, Lu L. Research on performance optimization and visualization tool of   hadoop. Proceedings of 10th International Conference on Computer Science and Education; Cambridge University, USA.2015 Jul