A SURVEY ON CROP YIELD RECOMMENDER SYSTEM IN AGRICULTURE

Sangeeta[1], Shruthi G[2]

[1]Student, Department of Information Science & engineering,
Ramaiah Institute of Technology, (India)
[2]Assistant professor, Department of Information Science & engineering,
Ramaiah Institute of Technology, (India)

ABSTRACT

Agriculture is the best utility region especially inside the developing worldwide areas like India. Usage of records age in horticulture can substitute the circumstance of decision making and Farmers can yield in higher manner. About portion of the number of inhabitants in India relies upon on farming for its occupation however its commitment towards the GDP of India is just 14 percent. One suitable explanation behind this is the deficiency of adequate decision making by farmers on yield prediction. There isn't any framework in location to suggest farmer what plants to grow.

The proposed machine learning approach aims at predicting the best yielded crop for a particular region by analyzing various atmospheric factors like rainfall, temperature, humidity etc., and land factors like soil pH, soil type including past records of crops grown. Finally our system is expected to recommend farmers about suitable crops grown in their region which yields best in terms of quantity and quality and also market prices.

Keywords: Crop yield prediction, Demand-based crops, Machine learning techniques, Supervised Learning, Unsupervised learning.

I.   INTRODUCTION

Agriculture is something that individuals have started to finish up moderate on, disregarding that it's miles what is holding us alive. However, there is regardless some driving forward, enthusiastic ranchers whose life continues running on essentially developing. Regardless, there's in addition the pollution that is extending packages these days. The Main intention of the Department of Agricultural Marketing and Agricultural Business is to have a reasonable cost to the cultivating network who are pushed behind the current focused showcasing situation and the mission of accomplishing the reasonable cost is by making the current demonstration and principles solid and progressively compelling by executing new innovations and systems went for lessening pre and post-gather misfortunes through legitimate and sorted out techniques and urge enhancing the market. The vital motivation behind making a managed market is to put off the undesirable exchange work out, to diminish the charges inside the commercial centre and to offer reasonable expenses to the Farmers. A few activities have been taken to advance rural showcasing a decent method to cultivate and keep up the place of country monetary improvement.

To advantage the cultivating from the new worldwide market get admission to potential outcomes, the inward rural promoting device inside the United States of America moreover wishes to be joined and strengthened. In interesting, the commercial centre contraption must be revived to:

- Provide impetuses to Farmer to deliver more.
- Pass on the changing over wishes of the purchasers to the makers to empower producing making arrangements.
- Foster genuine challenge a considerable lot of the market players and
- To improve the offer of Farmers in the last expense of his rural produce.

Today the farmers develop crops dependent on the experience picked up from the past age .Since the customary technique for cultivating is polished there exists an overabundance or shortage of yields without gathering the real necessity. The farmers don't know about the interest that happens in the current horticultural economy. This results in the misfortune to the Farmers. The communicated thought processes arranged by significance in the back of Farmer suicides have been condition, low produce costs, weight and hover of relative's obligations, poor water system, and blast inside the cost of development. The primary reason is the low costs of the items and the expanded expense of development. The expenses of yields are controlled by economic interest and the points of confinement of the creation.

Yield forecast is one of the undertakings that should be possible by bleeding edge ML calculations. Field sensors, satellites, unmanned flying engines (UAVs), and cultivating hardware can give a goliath amount of records on soil circumstances, plant physiology, climate, climate, and several of the procedures taking locale in a homestead. These datasets license the approach of sort and estimate molds that might be very useful to Agriculture generation. India is growing quick in populace. The call for is high and could blast in coming predetermination along these lines, to make certain sustenance security vertical advancement in farming is the need of great importance. For this a blended basic and methodological methodology likes assortment commencement, pesticide and composts the executives, consolidated editing, water collecting, proficient water system techniques and numerous others may be required. Additionally, it transforms into basic to re-enact and expect the harvest yield underneath encompassing circumstances past to the usage arrange for viable yield the executives and favored outcomes, more noteworthy so in the sprinkled locale and when India is inclining nearer to exactness cultivating researchers.

Since the relatives between harvest yield and the climate and non-atmosphere factors are non-straight and comprise of fair dimension difficulties, machine examining may demonstrate a triumph elective for yield expectations.
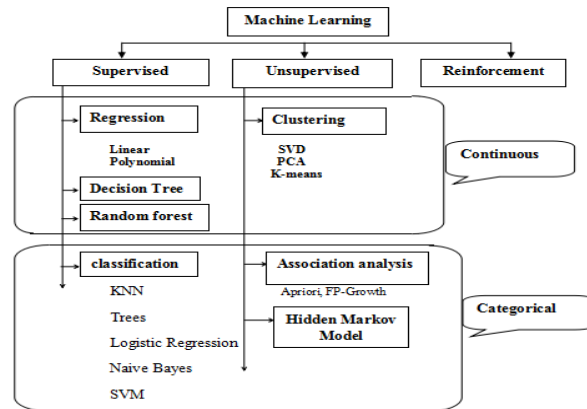
II. METHODS

Fig 1: Variants of machine learning algorithms

Machine learning can be separated into two wide segments: Supervised and Unsupervised. There are numerous different algorithms.

Supervised learning

To learn the relationship between given inputs and outputs the algorithm uses feedback from human resource and trained info. For example, a specialist can use marketing outlay and weather outlook as input info to identify the sales of cans. When the output info is known we can use supervised learning. The algorithm will identify new data.

Supervised learning is further categories into:

- Classification task
- Regression task

Classification

Envision you want to anticipate the gender-specification of a client for a commercial. You will begin gathering information on the stature, weight, work, compensation, purchasing basket, etc. from your client database. You know the gender-specification of every one of your client, it must be male or female. The goal of the classifier will be to allot a likelihood of being a male or a female (i.e., the label) in light of the data (i.e., highlights you have gathered). At the point when the model figured out how to perceive male or female, you can utilize new info to make an expectation. For example, you just got new data from an obscure client, and you need to know whether it is a male or female. In the event that the classifier predicts male = 70%, it implies the calculation is certain at 70% that this client is a male, and 30% it is a female.

The mark can be of at least two classes. The above model has just two classes, yet in the event that a classifier needs to anticipate object, it has many classes (e.g., glass, table, shoes, and so forth each article speaks to a class).

Regression

At the point when the yield is a consistent esteem, the errand is a relapse. For example, money related

expert may need to conjecture the estimation of a stock dependent on a scope of highlight like value, past stock exhibitions, macroeconomics record. The framework will be prepared to assess the cost of the stocks with the most minimal conceivable blunder.

Algorithm Name	Description	Type
Linear regression	Figures out how to relate each element to the yield to help anticipate future qualities.	Regression
Logistic regression	Expansion of linear-regression that is utilized for characterization assignments. The yield variable is paired (e.g., just dark or white) as opposed to consistent (e.g., an endless rundown of potential hues)	Classification
Decision tree	Very interpretable grouping or relapse show that parts information include values into branches at choice hubs (e.g., if a component is a shading, every conceivable shading turns into another branch) until an official choice yield is made	Regression Classification
Naive Bayes	The Bayesian technique is a characterization strategy that utilizes the Bayesian hypothesis. The hypothesis refreshes the earlier information of an occasion with the free likelihood of each element that can influence the occasion.	Regression Classification
Support vector machine	SVM, is ordinarily utilized for the grouping task. SVM calculation finds a hyper plane that ideally partitioned the classes. It is best utilized with a non-linear solver.	Regression (not very common) Classification
Random forest	The calculation is based upon a decision tree to improve the precision definitely. Random forest produces ordinarily basic decision trees and uses the 'majority vote' technique to choose which mark to return. For the arrangement task, the last expectation will be the one with the most votes, while for the relapse task; the normal forecast of the considerable number of trees is the last expectation.	Regression Classification
Gradient-boosting trees	Classification or regression strategy that utilizes a large number of models to think of a choice however gauges them dependent on their exactness in foreseeing the result	Regression Classification
AdaBoost	A gradient-boosting tree is a cutting edge classification/regression strategy. It is concentrating on the blunder submitted by the past trees and endeavors to address it.	Regression Classification

Table I: Supervised machine learning algorithm

Unsupervised learning

In unsupervised learning, a calculation investigates input information without being given an unequivocal yield variable (e.g., investigates client statistic information to recognize designs).

You can utilize it when you don't have a clue how to order the information, and you need the calculation to discover designs and characterize the information for you.

Algorithm	Description	Type
K-means clustering	Places information into certain gatherings (k) that each contains information with comparable attributes (as controlled by the model, not progress of time by people)	Clustering
Gaussian mixture model	A speculation of k-means clustering that gives greater adaptability in the size and state of gatherings (groups)	Clustering
Hierarchical clustering	A part bunches along a various leveled tree to frame a grouping framework. Can be utilized for Cluster unwaveringness card client	Clustering
Recommender system	Help to characterize the important information for making a recommendation.	Clustering
PCA/T-SNE	Generally used to diminish the dimensionality of the information. The calculations lessen the quantity of highlights to 3 or 4 vectors with the most elevated fluctuations	Clustering

Table II: Unsupervised machine learning algorithm

III. LITERATURE WORK

Grey prediction system which gives an excellent prediction accuracy of price forecast in production market, is made used in this work. This forecast structure is used to predict the market costs of different yields. By implementing demand grade for each crop, the real downside of this framework is destroyed i.e., price of the crops will not be stable all the times. [1]

One of the useful support system in Bangladesh focuses on helping the poor farmers by assisting the demands about the crop through website. The drawback in the system is that the uneducated farmers were not able to use the system, even if the farmer knows about the system they could not able to operate the model. In this model the data will be sent by means of SMS voice message in regional language. [2]

The procedure of crop yield prediction is done by using Data Mining approach which results in prediction of analyzed soil dataset. The interest existing in the rural economy is not considered by the system. This system overcomes the drawback by considering the demands based on the market price crops and it is suggested to the farmers for better growth. [3]

To improvise the value and gain of farming area, data mining techniques are made used which selects the appropriate crops for cultivation and predicting the crop yield. Feasible suggestions to farmers and meeting the current demands are not provided by this system which serves as a drawback here. [4]

A novel system known as extensible Crop Yield prediction framework is built for precision agriculture using data mining techniques. In this paper there is an investigation of requirement for crop yield prediction and different systems have been utilized and finally it results in a framework which is flexible for prediction accuracy. [5]

In this work the author has used various data mining algorithms like Naïve Bayes and KNN to predict the class of analyzed soil dataset. The soil is categorized into high, medium and low. By doing this the farmer and the soil analyst gets the prior knowledge about the land. Meanwhile they can decide which crop best suits to sows. The results in-turn will help in predicting the crop yield. [6]

Agriculture is one of the tedious processes which involve numerous estimations and effecting factors. In this research work author focused on modeling some of the important inputs which plays a major role in the collected dataset and to derive a strong relationship between the variables. SVM and k-means algorithms are used to forward pollution from atmosphere and also to classify between soil and plants. [7]

Fuzzy Cognitive Maps are one of the Soft Computing techniques used for modeling expert's knowledge. In this work the author has made use of this approach to predict yield of cotton plants. Representing the knowledge more visually, effectively, simple and structural makes this technique more advantageous and thus a convenient approach for predicting and improving cotton yield prediction. [8]

The author proposes a model which focuses on removing noise factors to get efficient predictions. By considering various factors like differences in agricultural policies and practices. To accommodate the spatial dependence arising between different regions prior distributions are developed. In addition to this basis expansions and dimension-reduction schemes are incorporated to evaluate the improved predictions. [9]

To achieve excellence farming and suggest the farmer to select the best previous agriculture information is used. This research work also paves way to improvised yield prediction and increase in income level of crops. [10]

The author proposed a methodology called crop selection method (CSM), which is utilized to choose the yield for a chose season. This technique amplifies net yield of harvest according to the season. The arranged technique improves yield rate of harvest. There's a need for expectation technique of high exactness and execution because of the CSM approach relies upon predicted estimations of impacted parameters. [11]

In agriculture, ranchers are as yet confronting a trouble of harvest yield forecast. The objective of this paper is to propose and realize a standard essentially based system to anticipate the harvest yield creation from the get-together of past enlightening list. To achieve this, affiliation rule is utilized on the horticulture informational collection from 2000 to 2012. [12]

In proposed demonstrate the data investigation is finished and prescient models were planned. Together with these procedures they utilized relapse models like linear, non-linear and multiple linear models. Regression models are tried for viable expectation or figure the harvest yield in Andhra Pradesh and Telangana states. [13]

This paper presents survey on different crop yield prediction methods. Prediction will helps the farmers for better yield as per the season, crop selection, methods for agriculture etc. These methods results for better yield and profit. [14]

In this investigation the author has built up a website page called 'Crop-Advisor' by utilizing C4.5 algorithm that centers around hunting down the premier affecting area conditions on crop yields of hand-picked crops in Madhya Pradesh. Upheld area input parameters potential information handling strategies are utilized to anticipate the harvest yield. The exactness of this arranged framework yields over 75% on every one of the harvests and areas hand-picked. [15]

Sl. No	AUTHOR	TECHNOLOGY	RESULT
1	Raorane A.A. Kulkarni R.V.	ANN, Decision tree, Bayesian network, SVM	**K-means** is used to forward atmospheric pollution and to classify soil and plants. **SVN** is used to analyze variation in whether and Wine Fermentation process.
2	E.I. Papageorgiou, A.T. Markinos, T.A. Gemtos	Fuzzy set	**Fuzzy Cognitive Mapping approach** is used to make decision making tasks in yield prediction (cotton crop)
3	Luke Bornn, James V. Zidek	Bayesian model	Inefficiency of averaging SI over the entire season which doesn't improve the crop yield is solved by implementing **Bayesian model** which smoothens and stabilizes the prediction
4	Monali Paul, Santosh K.Vishwakarma, Ashok Verma	Naïve Bayes and KNN	**KNN and Naïve Bayes** algorithms are implemented to classify soil into low, medium and high to help farmers sow better crops and yield more.
5	Yogesh Gandge, Sandhya	Classification algorithm	A **Novel classifier** is used to improvise the prediction and helps farmer to sow water tolerant seeds to withstand floods.
6	R.Sujatha, Dr.P.Isakki Devi	Classification algorithms like Naïve bayes, Random forest, ANN, Decision tree, SVM etc.	Various classification algorithms are implemented to improvise crop prediction and to select best crop suitable for the season.
7	S.Veenadhari, Dr. Bharat Misra, Dr. CD Singh	C 4.5 algorithm	The proposed methodology predicts the crop yield based on the climatic input parameters. The system Accuracy is above 75%.
8	Aakunuri Manjula, Dr.G.Narsimha	Decision support system(DSS)	A framework that facilitates flexible inclusion of various techniques towards crop yield prediction is proposed which helps in precision agriculture.
9	Rakesh Kumar, M.P. Singh, Prabhat Kumar, and J.P. Singh	Crop selection method (CSM),Gradient boosted decision tree(GBDT)	A method called CSM is used which resolves the selection of crop according to season. Performance and accuracy of CSM is approach is high.
10	E. Manjula, S. Djodiltachoumy	K- Means clustering algorithm, association rule	Implementing rule based prediction by collecting past data is achieved by applying Association rule mining.
11	S.Nagini, Dr. T. V. Rajini Kanth, B.V.Kiranmayee	Regression models like Linear, multiple linear, non-linear	This work proposes that if area of crop increases yield increases by applying two-three predictor formulas and highest correlation is achieved between Area-in-hectares and production in tons. The lm and glm contributes for effective crop prediction.
12	Yung-Hsing Peng, Chin-Shun Hsu, Po-Chuang Huang	Autoregressive integrated moving average(ARIMA), partial least square(PLS), ANN, Response surface methodology(RSM)	Based on market prices a crop price forecasting service is built using ARIMA, PLS, ANN, and RSMPLS algorithms to get accurate predictions of price forecasting.

Table III: Machine learning methodologies used in agriculture

IV. CONCLUSION

This paper describes, the difference types of algorithms such as K-means algorithm, ANN, Decision tree, Bayesian network, SVM, Naïve Bayes and K-Nearest Neighbor are compared and we can see that all algorithms has different advantages and limitations and furthermore we can conclude that compared to other algorithm SVM, Naive Bayes and KNN has better prediction. This paper combines the work of different authors in a single spot so it is valuable for specialists to get data of current circumstance of the methods and application in substance to agricultural field.

V. ACKNOWLEDGEMENTS

REFERENCES

[1] Jiajun Zong, Quanyin Zhu, "Apply Grey Prediction in the Agriculture Production Price", Fourth International Conference on Multimedia Information Networking and Security, 2012.

[2] Michael gurstein,"a decision support system to assist the rural poor in Bangladesh", IEEE TECHNOLOGY AND SOCIETY MAGAZINES, September 2013.

[3] Monali Paul, Santhosh K. Vishwakarma, Ashok Verma, "Prediction of crop yield using Data Mining approach" Computational Intelligence and Communication Networks (CICN), International Conference 12-14 Dec. 2015.

[4] Tng Zhang, "Solving large scale linear prediction problems using stochastic gradient descent algorithms", Proceedings of the twenty-first international conference on Machine Learning Aakunuri Manjula, G. Narsimha, "XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction", IEEE Sponsored 9th ISCO, 2015.

[5] Aakunuri Manjula, G. Narsimha, "XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction", IEEE Sponsored 9th ISCO, 2015.

[6] D. Ramesh, B. Vishnu Vardhan, "Analysis Of Crop Yield Prediction Using Data Mining Techniques", 2015.

[7] A.A. Raorane, R. V. Kulkarni, "Data Mining: An effective tool for yield estimation in the agricultural sector", IJETTCS, vol. 1, no. 2, pp. 75-79, 2012.

[8] E. I. Papageorgioua, A. T. Markinosb, T. A. Gemtosb, "Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application", Elsevier, vol. 0, no. 11, pp. 3643-3657, 2011.

[9] Luke Bornn, James V. Zidek, "Efficient stabilization of crop yield prediction in the Canadian Prairies", Elsevier, vol. 0, no. 0, pp. 223-232, 2012.

[10] R.Sujatha and P.Isakki, "Crop yield estimation using classification techniques" Computational Intelligence and Communication Networks (CICN), International Conference 7-9 Jan. 2016.

[11] Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh, "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique" International Conference 6-8 May. 2015.

[12] E. Manjula, S. Djodiltachoumy, " A Model for Prediction of Crop Yield" International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017.

[13] S.Nagini, Dr. T. V. Rajini Kanth, B.V.Kiranmayee, "Agriculture Yield Prediction Using Predictive Analytic Techniques"International Conference on Contemporary Computing and Informatics 2016.

[14] Yung-Hsing Peng, Chin-Shun Hsu, and Po-Chuang Huang, "Developing Crop Price Forecasting Service Using Open Data from Taiwan Markets"Nov. 20-22, 2015.

[15] S.Veenadhari, Dr. Bharat Misra, Dr. CD Singh, "Machine learning approach for forecasting crop yield based on climatic parameters" 2014 International Conference on Computer Communication and Informatics (ICCCI -2014), Jan. 03 –05, 2014, Coimbatore, INDIA