IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 20, Issue 5, Ver. III (Sep - Oct 2018), PP 72-80
www.iosrjournals.org

Survey of Crop Prediction Using Different Classification Analytical Model

M.Saranya[1], Dr. S.Sathappan[2]
[1]Research Scholar, Erode Arts & Science College.
[2]Associate Professor, Erode Arts & Science College.
Corresponding Author: M.Saranya

Abstract— *In this paper, there are different ways has been analyzed to predict the crop yield. Now a day, there's tremendous growth in paper publication and analysis in numerous streams for getting efficient result. The crop yield productions are going to be continuously become terribly crucial in our standard of living. To pursue the quickest and extremely economical thanks to enhance the crop yield the near-term arrange is application and optimizing the present agriculture technologies. This may result in important improvement in agriculture that is suboptimal resolution. The most goal of this analysis is employed to summarize the historical trend of crop production in nominal region to look at the foremost constraint to extend the productivity in future. Except for this paper, to performing comparative study with techniques of LR, KNN, NBC, J.48, MLP, RF, SVM, DT and Artificial Neural Network (ANN) that statistics estimation of crop yield. To boot, here to incorporate the information pre-processing approach to predict the worth while not duplication and efficient time manner that understand high accuracy and prediction capabilities.*

Keywords—*Machine Learning, Crop Prediction, Classification Model, Feature Selection Model, NN Model*

Date of Submission: 13-10-2018 Date of acceptance: 28-10-2018

I. Introduction

Agriculture is that the backbone of Indian Economy. In India, majority of the farmers aren't obtaining the expected crop yield thanks to many reasons [1]. The agricultural yield is primarily depends on weather. Understanding the relative Importance of these Climate factors to crop yield variation might offer valuable data regarding crop planting and management beneath global climate change condition for policymakers and farmers. The quantity of knowledge is gigantic in Indian agriculture. The data once become information is very helpful for several functions. India still depends only on monsoon downfall. The climate variations want to be self-addressed associated an analysis is to be created in order to assist the farmers to maximize the crop productivity [2].

Yield prediction is a crucial agricultural problem. Each farmer is fascinated by knowing, what proportion yield he's concerning expect. Within the past, yield prediction was performed by considering farmer's previous expertise on a specific crop. The degree of information is big in Indian agriculture. The data once become information is extremely helpful for many functions.

Data Mining is wide applied to agricultural issues. Data processing is employed to research massive knowledge sets and establish helpful classifications and patters within the information crop datasets. The general goal of the info Mining method is to extract the knowledge from an information set and remodel it into intelligible structure for additional use [2].

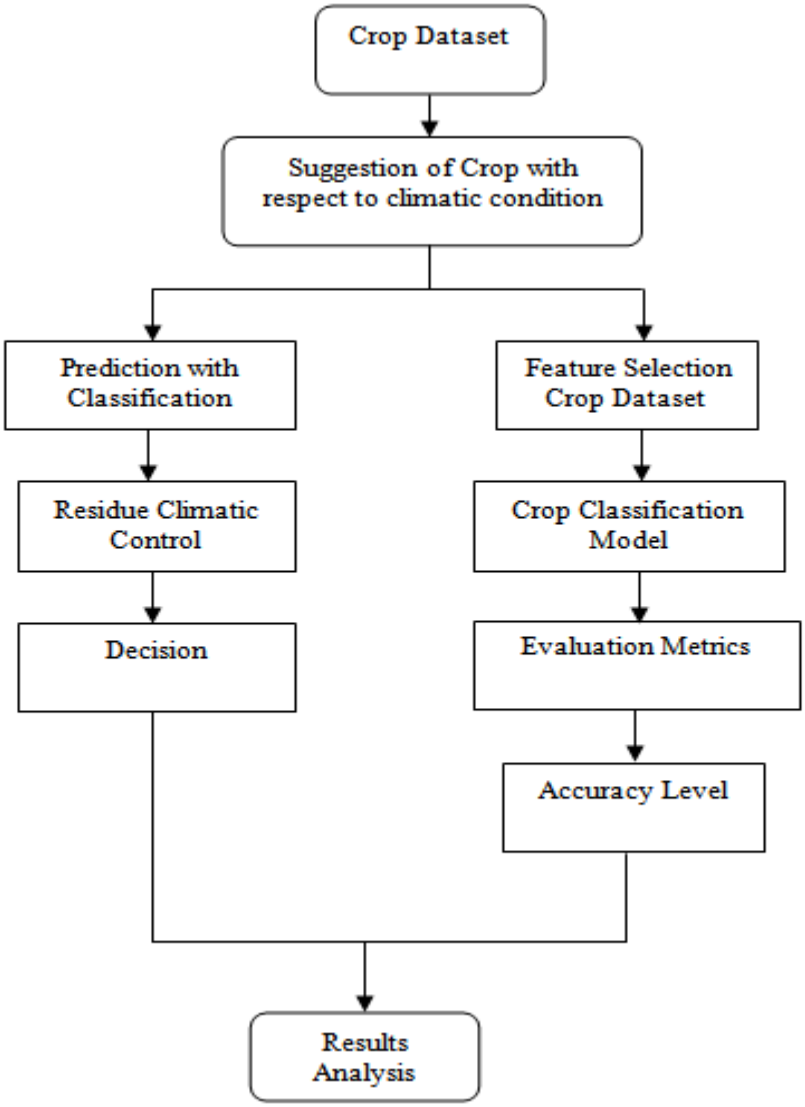

Fig 1.1 Crop Prediction Model

Through this survey paper, predicting totally different crop yield that is correct models for crop yield estimation exploitation data and communication technologies may facilitate the farmers. It won' t to improve prediction of crop yield below totally different climatically eventualities.

- This paper concentrates on classification analytical technique that produces higher results of crop prediction.
- It conjointly considers parameters like space, production rate, mean temperature, mean rainfall, soil, and rainfall PH value and crop disease details. Hence, multiple parameters are considered.
- So the prediction for crop yield can offer appropriate results for the farmers World Health Organization rely on agriculture.
- Predict in step with the posterior results concerning crop yield

In this section presents a characteristic of information mining, want for crop prediction model, issues in crop prediction system and motivation and objectives of the review paper. Section II deals with the literature review on the relevant analysis topics of crop prediction. Recent approaches and methodologies accustomed Suggestion of Crop with relevancy environmental conditions are mentioned. Sections III deals with the literature feature selection model justify the steps to maximize the crop prediction performance. Section IV describes the literature methodology of the crop prediction classification system. This section explains concerning the ways within the classification analytical model that Suggestion of Crop with relevancy environmental condition. Section V discusses concerning the dataset and also the performance of the literature system. Section VI concludes the paper and additionally discusses concerning the additional scope of the analysis.

II. Literature Survey

The analysis of crop yield prediction exploitation Multiple linear regression (MLR) technique and Density primarily based cluster technique for the chosen region to predicting specific crop that adopted for all the districts of Andhra Pradesh to enhance and manifest the validity of yield production [3]. Kidakan Saithanu.et.al exemplified the event of neural network models for statement rice yield within the northern Thailand that support of Multi-Layer Perception (MLP) and normal Radial Basis Function(ORBF)[4] neural spec. Its trained by victimization many earth science variables. This design provides higher performance with smallest RMSE of Validation dataset.

The Agriculture management specialists would like easy and correct estimation techniques to predict rice yields within the designing method. The goal of analysis to research whether or not artificial neural network models might effectively predict Fujian rice yield for typical environmental condition of the mountainous region, evaluated ANN model performance relative to variations of development parameters [5]. This ANN models tried to be superior for accurately predicting rice yields beneath typical Fujian climate condition and also the yield prediction that uses a knowledge mining technique so as to predict the class of the analysed soil datasets. This categorization foretold the crops yielding with classification rule like Naïve Bayes and K-Nearest Neighbor search tested well and satisfies all needs [6].

Menaka et.al evaluated analysis that prove the adaptive Neuro Fuzzy illation System (ANFIS) prediction model provides higher result that alternative approaches. Here, totally different strategies has been analyzed to predict the crop yield by strategies of Artificial neural network (ANN), adaptive Neuro Fuzzy illation system, symbolic logic and Multi Linear Regression[7] to provided the most effective result for prediction. So the process of all attributes can increase the classification accuracy repeatedly than the traditional prediction models. The studies on application of knowledge mining technique within the agriculture field, J48 of C4.5 algorithmic rule used here for analysis. This analysis accustomed choose the foremost glorious crops for the region and its yield thereby rising the values and gain of farming additionally and to allow higher crop yield estimation for the required region. This aids farmers to choose on the crop, wish to plant for forthcoming tenure. Prediction can facilitate the associated industries for designing the logistics [8]. This is often achieved by applying association rule mining on agricultural information. It centred on creation of prediction model which can be accustomed future prediction of crop yield. The transient analysis of crop yield prediction victimisation data processing technique supported association rules for designated region. The experimented result provided the result that planned work expeditiously predict the crop yield production.

Narayana Balakrishnan (2016) developed ensemble model of AdaSVM and AdaNative that project the crop production over amount of your time series dataset exactly than the present model[9]. This model compared to SVM and Naïve Thomas Bayes technique with 2 parameters used one by one for prediction of output area unit the accuracy and also the classification error. The result's smart quantity of fall within the perfection within the accuracy of prediction and conjointly smart quantity of fall within the share of classification error in each the planned techniques. The applying of knowledge mining techniques to wont to predict future production of crops like Rice, Wheat and Maize with reference to varied environmental condition conditions[10]. The survey processed by the parameters worth, like temperature, rainfall, area, space on irrigation, yield and production with data processing formula of Multiple statistical regression, Random forest regression and Support vector regression.

Abirami.et.al targeted on numerous crop yield prediction ways mistreatment data processing techniques. The most goal of the work is to seek out the acceptable knowledge models that attain high accuracy and high generality of crop yield prediction [11]. Here numerous classification approaches evaluated during this study. The survey to check completely different} data processing rules with same parameters on the 10fold cross validation take a look at to predict the crop yield by different data processing classification algorithm like K-Nearest Neighbor, K-Means, Neural Network, Support Vector Machine, Case-based Reasoning, call Tree rule area unit applied of agriculture domain[12]. Finally this work reveals the superior performance of J48 classification rule with 89.33% for crop prediction and applying of various data processing classification techniques within the domain of agriculture for yield prediction. The approach is to forecast the wheat crop yield for various district of Gujarat states. Wheat crop yield, data processing classification rule and step-wise simple regression ways are used for prediction [13]. This rule of knowledge mining was applied on the weather data exploitation weka tool for crop yield forecast. The performance of those rules and applied math forecast model is applied by MLP and AR rule model was quite higher than the opposite algorithm.

Vinayak A.Bharadi(2017) analysed the agricultural knowledge of India exploitation data processing algorithms and to search out some helpful details from the results of those techniques which might be facilitate to enhance the agricultural yield. These are the complete agglomeration rule that utilized in the paper like K-Means, DBSCAN, and EM. Here K-Means cluster rule is adopted as base rule, DBSCAN and EM algorithms are applied to data [14]. Agricultural crop production depends on varied factors like biology, climate, economy and geographics. Further scientific and policy communities have recognized the condition of crop agriculture to

global climate change and talent of farmers to adapt as a result of direct and robust dependence of crop agriculture on climate. many factors have totally different impacts on agriculture, which might be quantified exploitation applicable applied math methodologies. Applying such methodologies and techniques on historical yields of crops, its attainable to get info or data which might be useful to farmers and government organizations for creating higher call and policies that cause increased production. Here, it centred is on the appliance of data mining techniques to extract knowledge from the agricultural data to estimate crop yield for major cereal crops [15]. Advances in computing and knowledge storage hove provided huge a most of information. The challenge has been to extract data from this raw data; this had cause new ways and techniques like data processing that may bridge the data of the info to the crop yield estimation. Here aimed to assess these new data processing techniques and apply them to the varied variables consisting within the info. Finally, its ascertained the economical technique are often developed and analysed victimisation the suitable knowledge, the info that is collected from Kolhapur district to unravel complicated agricultural issues victimisation data processing techniques[16].

The projected new approach that to predict the crop yielding by Machine Learning Approach. This projected approach provides an answer for good agriculture by observation the agricultural field which may assist the farmers in increasing productivity to an excellent extent. forecast information obtained from IMD (Indian scientific discipline Departments) like temperature and rain and soil parameters repository provides insight into that crops square measure appropriate to be cultivated in an exceedingly explicit area[17]. This work presents a system, in from of associate automaton primarily based application, that uses information analytics techniques so as to predict the foremost profitable crop within the weather and soil conditions. The improved approach can integrate the information obtained from repository, weather department and by applying machine learning algorithmic program of multiple statistical regression, a prediction of most fitted crops in line with current environmental conditions is created. This provides a farmer with type of choices of crops that may be cultivated. It develops a system by desegregation information from numerous sources, information analytics, prediction analysis which may improve crop yield productivity and increase the profit margins of farmer serving to them over a extended run. This technique takes into thought the past production of knowledge which is able to facilitate the farmer get insight into the demand and also the value of assorted crops in market.

Aakunuri.et.al delineate the crop yield prediction will facilitate agricultural departments to possess methods for up agriculture. Spacial Hadoop is predicated on the new programming paradigm besides having spacial extensions. Towards this, projected a framework and enforced the extended algorithmic program that's compatible with MapReduce programming [18]. Here, collected five years information of Cotton and Maize crops of Telangana state, India. This enforced framework for crop yield prediction done by MapReduce programming. The disclosed novel system outperformed the present one with reduced error rate. Kanjana Devi(2011) delineated that objective of the work is to check varied data processing techniques which provides the utmost accuracy. data processing is just the manner that assists to convert immense information into technologies and create them on the market to the farmers. the massive quantity of knowledge is used to mine lump of data which will be helpful for farmers and call makes to require effective and prompt call. Here, the main one parameter is employed to extend crop production is soil. Totally different classification algorithms area unit applied to soil information set to predict its fertility and rate victimization K-Means, Random Tree and Apriori [19].

The farmer harvest not solely crops however additionally growing quantity of knowledge. A farmer desires to grasp regarding the applications of recent technologies in agriculture. Such technological demand from the farmer ends up in extracting the information from the on the market information. The information extraction ways in data processing area unit to be explored so as to get the crop yield prediction. Sample data processing techniques were utilized in agriculture. A number of the wide used data processing techniques over agriculture datasets area unit Multiplex simple regression, Density primarily based cluster Technique, K-Means approach, K-Nearest Neighbor, Artificial Neural Networks, Support Vector Machines [20]. The model is employed to optimize the on the market by means that of knowledge mining techniques to predict the crop yield. The wide used data processing techniques used over agriculture datasets and alter Density primarily based agglomeration Technique. Thanushree.et.al [21] planned mobile application wherever the farmers have to be compelled to transfer the app on their sensible phone and enter the desired details regarding their field and therefore the information associated with the crop grow and for the way a lot of space beside the climate, rain and alternative necessary factors and every one these information are compared with the antecedently' s noted information and compared victimization varied data processing techniques then finally the compared results of crop yield would be displayed on their good phone.

III. Feature Selection Model

Feature selection plays an important role in data mining analytical model. It computes an optimal subset of predictive features measured in the original data. It enables to achieve maximum classification performance by reducing the number of features used in classification while maintaining acceptable classification accuracy. Subsets of the original features which retain adequate information to discriminate well among classes are selected. Several search algorithms have been used for feature selection. This work implements Principal Component Analysis, Information gain and Relief-f Attribute Evaluator.

A. Principal Component Analysis (PCA)

Principal component analysis performs a linear mapping of the crop data to a lower dimensional space in such a way that the variance of the data in the low-dimensional representation is maximized. PCA, a non-parametric method builds a set of features by selecting those axes which maximize crop data variance.

PCA can be used to reduce a complex crop data set to a lower dimensionality, to reveal the structures or the dominant types of variations in both the observations and the variables. It is a quantitatively rigorous method that generates a new set of variables, called principal components. Every principal component is a linear combination of the novel variables. The entire principal components are orthogonal to all other, so there is no redundant information. The principal components as a whole form an orthogonal basis for the space of the dataset.

B. Information gain

The information gain of an attribute tells the amount of information an attribute provides with respect to the classification target. Information gain (IG) measures the amount of crop information about the class prediction, if the only information available is the presence of a feature and the corresponding class distribution. In machine learning, information gain can be used to aid ranking the features. A feature with high information gain should be ranked higher than other features because it has stronger power in classifying the crop data. Shannon entropy is the common measure for the information. Information gain is the reduction in the entropy that is archived by learning a variable. Concretely, it measures the expected reduction in entropy.

IG= H(Y)-H(Y/X)

H(Y)= -∑ P(Y)LOG(P(Y))

H(Y/X)= -∑P(X)∑ P(Y/X)LOG(P(Y/X))

Where P(Y) is the marginal probability density function for the random variable Y and P(Y |X) is the conditional probability of Y given X.

C. Relief-F Attribute Evaluator

A key idea of the original Relief algorithm (Kira & Rendell, 1992b), is to estimate the quality of attributes according to how well their values distinguish between instances that are near to each other. Given a randomly selected instance Ri, Relief searches for its two nearest neighbors, one from the same class, called nearest hit H, and the other from the different class, called nearest miss M . It updates the quality estimation for all attributes. The ReliefF(Relief-F) algorithm(Kononenko, 1994) is not limited to two class problems, is more robust and can deal with incomplete and noisy data. Similarly to Relief, ReliefF randomly selects an instance Ri, but then searches for k of its nearest neighbors from the similar class, called nearest hits Hj, and also k nearest neighbors from all of the dissimilar classes, called nearest misses Mj (C). It updates the quality estimation WA for all attributes A depending on their values for Ri, hits H.

The selected features from each technique are used for training three classifiers SVM, Naive Bayes and Decision trees. The machine learning algorithms are tested using 10-fold cross validation to obtain better classification result

IV. Classification Techniques

This is a task performed to generalize known structure in data mining to apply to new crop data. It is also the categorization of data for its most effective and efficient use. There are several data mining classification algorithms being considered and implemented in various domains. Some of the most well-liked and well-known' s are modified and presented here in, based on their simplicity, capabilities and robustness.

A. K-nearest neighbor (k-NN)

The principle behind this method is to find predefined numbers of training crop samples closest in the distance to the new point and predict label from these. The numeral of samples can be a user defined constant or diverse based on the local density of points. Here, the distance can be any metric measure. There are various distance measures are implemented in the k-NN, Chebysheb, Euclidean, Edit Distance and Manhattan, but the Euclidean

distance measure is the most frequent preference. However it is successful in huge number of classification problems.

B. J4.8

J4.8 decision trees algorithm is an open source Java implementation of the C4.5 [2]. It grows a tree and uses divide-and-conquer algorithm. J4.8 is a predictive machine-learning model that decides the object value (dependent variable) of a new-fangled sample based on different attribute values of the obtainable data. To classify a new crop item, it creates a decision tree based on the attribute values of the training crop data. When it encounters a set of crop items in a training crop dataset, it identifies the attribute that discriminates. It uses information gain to tell us most about the crop data instances so that it can classify them the best.

C. Naïve Bayes Classification (NBC)

This classifier is based on the Bayes rule of conditional probability. It uses all of the attributes contained in the crop data, and analyses them individually as though they are equally important and independent of each other. The Naïve Bayes classifier works on an easy, but relatively instinctive concept. It makes use of the variables contained in the crop data sample, by observing them individually, independent of each other. It considers every of the attributes individually once classifying a new instance. It assumes that one attribute works independently of the other attributes contained by the crop sample.

D. Multi Layer Perceptron (MLP)

MLP is a feed forward artificial neural network model that maps sets of input crop data onto a set of appropriate outputs. It consists of several layers of nodes, with every layer fully associated to the next one. Each node may be a neuron with a nonlinear activation function. It uses a learning method known as back propagation for training the network.

E. Linear Regression (LR)

It is a statistical measure that can be used to determine the strength of the association between one dependent variable and a sequence of other changing variables known as independent variables (regular attributes). If independent

variable contains several input attributes like in our analysis (rainfall, sunshine hours, humidity, pH etc), then it' s termed as multiple linear regressions. Linear regression provides a model for the relationship between a scalar variable and one or more explanatory variables.

F. Artificial Neural Network (ANN)

One widely used artificial neural network, back propagation neural network (BPNN), was applied to predict rice yield because of its simplicity in structure and robustness in simulation of nonlinear systems. A typical three-layer BPNN comprising one input layer, a hidden layer, and an output layer were used in the current study (Fig.4.1). The neurons of adjacent layers are connected by the nodes' weights. There are two weights in the three-layer BPNN, which are vij between input and hidden layers and ujk between hidden and output layers.

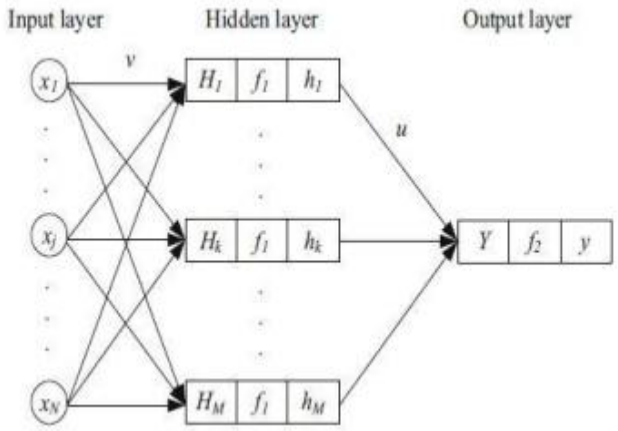

Fig 4.1 Three-Layer Artificial Neural Network Model

The aim of BPNN is to constantly modify the weights of connections between contiguous layers based on the deviation between actual values and outputs until the accuracy of the model meets the requirement of forecasting. The BPNN model can be used to forecast with new data when the weights are determined after numerous modifications. The ANN-Marquardt algorithm [18] combined with Newtonian gradient descent algorithm was used to adjust the connection weights and biases to minimize the error. The number of neurons in the hidden layer was usually determined by trial and error.

G. Support Vector Machine (SVM)

The current study investigated the applicability of support vector machines (SVMs) in determining the relative importance of climate factors (mean temperature, rainfall, relative humidity, sunshine hours, daily temperature range, and rainy days) to yield variation of paddy rice in India.[4] Support vector machine (SVM) which was originally developed by Vapnik (1998) has been generally applied to several different fields, such as time series analysis and signal process. Based on the statistical learning hypothesis and Structural risk reduction principle, SVM is less Vulnerable to over fitting problem and it uses a hypothesis area of linear functions in a very higher dimensional feature area. Studies have demonstrated that SVMs are superior to traditional artificial neural networks in solving classification and regression problems due to their good generalization ability.

H. Decision Trees (DT)

A decision tree represents a structure with two forms of parts:

- Leaf nodes that assign class labels to observations
- Internal nodes that specify tests on individual attributes with one branch and sub tree for each outcome of the test.

The tree classifies observations in a very top-down manner, ranging from the basis and working one's method down according to the outcomes of the tests at the inner nodes, till class label has been assigned and a leaf node has been reached. The tree is then constructed by means that of algorithmic partitioning till this leaf nodes contain only instances of a single class or till no analysis offers any enhancement. However, since most crop data sets are noisy, and since in most cases the attributes have limited predictive power, this tree growing strategy often results in a complex tree with many internal nodes that over-fits the crop data.

I. Random Forest (RF)

The Random Forest models were trained to predict crop yield using numerous biophysical variables as predictors. Environmental variables integrated soil, water, climate, photoperiod and fertilization data. The equivalent data were used for training MLR models for benchmarking purposes. The RF algorithm as such set aside partial data for its own internal validation, referred to as out-of-bag (OOB) data. However, to ensure a reasonable and conservative comparisons between MLR and RF, we tend to used only a random half of every dataset (i.e., potato, wheat, silage maize, maize grain) for training (‘ training dataset’) both RF and MLR models. The other half that was not intended for training was then used because the ‘ test dataset’ to validate and evaluate performances between the RF and MLR models. This method ensured that identical data points were obtainable for training and independent testing with the data points not included in training.

V. Review Analyzer

Sno	Author name	Techniques	Advantages	Disadvantages
1	D.Ramesh and B.Vishnu Vardhan	Multiple Linear Regression(MLR) technique and Density based clustering technique.	To create a user friendly interface for farmers which gives the analysis of rice production based on available data.	To improve and authenticate the validity of yield prediction which are useful for the farmers of Andhra Pradesh for the prediction of a specific crop only.
2.	Kidakan saithanu and Jatupat Mekpary up	Multi-Layer Perception (MLP) and Ordinary Radial Basis Function(ORBF) neural network architecture.	Training network model with the Simple architecture provided better performance than advanced one.	This research can’ t able to adjusting the network parameters like hidden nodes and hidden layers.
3.	Biaojun Ji	Analysis of investigate about ANN, evaluate ANN model, Comparison of Multiple Regression Model with ANN Models.	Accurately predicting rice yields under Fujian climate condition.	The models reported here are appropriate for predicting rice yields in Fujian province in China for average climatic conditions and for the specific soil types only.
4.	Supriya DM	Data mining Technology used such as classification rule like Naïve bayes and K-Nearest neighbors methods are used.	Simplified and reduced the manual work, Large volumes of data can be stored and Smooth workflow.	Values are not supposed to get directly from soil testing lab to server. The direct interaction with administrator is tough.

5.	K.Menaka and N.Yuvaraj	Survey and analysis done on Crop yield prediction model using Adaptive Neuro-Fuzzy Inference System(ANFIS)	Analysis will increase the classification accuracy compare than other models.	More complex to predict the optimized number of input parameter. ANN and MLP methods don't achieve better accuracy.
6.	Pooja MC, Sangeetha M, Shreyaswi J Salian.	J48 of C4.5 algorithm used	It makes very simple classifier to make a decision tree.	It' s not contrived to build Fertilizer Recommendation system which can be utilized effectively by the Soil Testing.
7.	Narayanan Balakrishnan, Dr.Govindarajan MuthuKumarasamy	AdaSVM and AdaNaive ensemble model used.	Achieved accuracy and classification error of prediction.	Enhanced methods needs to implement for better accuracy for rainfall methodologies.
8.	Suvidha Jambekar, Shikha Nema and Zia Saquib	Multiple Linear Regression, Random forest regression and Support vector regression	Data pre-processing efficiently reduced data duplication.	This prediction not analysed for Rice, wheat and Maize crops.
9.	Shital H. Bhojani1, Dr. Nirav Bhatt2	Different classification algorithms of data mining were applied on the weather data using WEKA tool for crop yield forecast.	The statistical model and MLP and AR algorithms model was quite better than the other algorithms.	Better results can be generated by changing the training and validation year for both statistical model and data mining classification technique
10	Aakunuri Manjula1 and Dr. G. Narsimha2	Map reduce program, makes use of parallel processing power of modern computing resources available through cloud computing	solution not only focuses on the crop yield prediction but also throws light into security while processing the data.	No investigation of the relationships and tradeoffs between the variables and crop yield prediction values towards precision agriculture.

VI. Conclusion

This paper is a beginning for further research in forecasting the crop using data mining algorithms. This paper presents new research possibilities for the methodologies to the problem of yield prediction. In agriculture, there are an emergent number of applications of data mining techniques are available and a growing number of data that are presently available from several resources. Future research can explain to the study whether by varying the technique produces enhanced results or by growing the input data set for the identical technique results change in the findings.

References

[1]. Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Paper presented at International conference on knowledge discovery and data mining. 1996

[2]. Ramesh D, Vishnu Vardhan B. Data mining techniques and applications to agricultural yield data. In: International journal of advanced research in computer and communication engineering. 2013; 2(9).

[3]. D Ramesh and B Vishnu Vardhan(2015) " Analysis Of Crop Yield Prediction Using Data Mining Techniques" . IEEE.

[4]. Thailand K. Saithanu1 and J. Mekparyup(2015), " Development Of Neural Network Models For Forecasting Rice Yield In The Northern" . Global Journal of Pure and Applied Mathematics. ISSN 0973-1768 Volume 11, Number 4 (2015), pp. 2191-2194 © Research India Publications http://www.ripublication.com

[5]. B. JI, Y. Sun, S.Yang And J. Wan (2007), " Artificial neural networks for rice yield prediction in mountainous regions" . Fujian Agriculture and Forestry University, Fuzhou 350002, China (Revised MS received 22 October 2006; First published online 16 January 2007)

[6]. Supriya D.M (2017), " Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining Approach" Vol. 5, Issue 5, May 2017 Copyright to IJIRCCE DOI: 10.15680/IJIRCCE.2017. 0505067 9648.

[7]. K. Menaka and N.Yuvaraj, " A survey on crop yield prediction models" . Indian Journal of Innovations and Developments Vol 5(12), October, 2016 ISSN (online) : 2277-5390 ISSN (Print): 2277-5382

[8]. Pooja M C, Sangeetha M, Shreyaswi J Salian, Veena Kamath and Mithun Naik(2018), " Implementation of Crop Yield Forecasting Using Data Mining" . © 2018, IRJET | Impact Factor value: 6.171 | ISO 9001:2008 Certified Journal | Page 2057

[9]. Narayanan Balakrishnan1 and Dr.Govindarajan Muthukumarasamy(2016), Crop Production-Ensemble Machine Learning Model for Prediction International Journal of Computer Science and Software Engineering (IJCSSE), Volume 5, Issue 7, July 2016 ISSN (Online): 2409-4285 www.IJCSSE.org Page: 148-153

[10]. Suvidha Jambekar, Shikha Nema and Zia Saquib(2018), " Application of Data Mining Techniques for Prediction of Crop Production in India" . International Journal of Innovations & Advancement in Computer Science ISSN 2347 – 8616 Volume 7, Issue 4 April 2018

[11]. L. Abirami & Dr. R. Rathipriya(2017), " Survey on Data Mining Techniques for Croppredictiion" , International Journal of Computational Research and Development, Special Issue, January, Page Number 43-46, 2017.

[12]. Hetal Patel and Dharmendra Patel (2016), " A Comparative Study on Various Data Mining Algorithms with Special Reference to Crop Yield Prediction" . Indian Journal of Science and Technology, *Vol 9(22), DOI:* 10.17485/ijst/2016/v9i22/92713, June 2016

[13]. Shital H. Bhojani and Dr. Nirav Bhatt(2018), Application of Data Mining Technique for Wheat Crop Yield Forecasting for Districts of Gujarat State. International Journal of Scientific and Research Publications, Volume 8, Issue 7, July 2018 302 ISSN 2250-3153 .

[14]. Vinayak A. Bharadi, Prachi P. Abhyankar, Ravina S. Patil, Sonal S. Patade , Tejaswini U. Nate and Anaya M. Joshi(2017), " Analysis And Prediction In Agricultural Data Using Data Mining Techniques" . International Journal of Research In Science &

Engineering e-ISSN: 2394-8299 Special Issue 7-ICEMTE March 2017 p-ISSN: 2394-8280 IJRISE| www.ijrise.org|editor@ijrise.org [386-393]

[15]. Aishwarya.B.R (2016), " A Literature Study on Application of Data Mining Tools for Rice Yield Prediction" (IJITR) International Journal Of Innovative Technology And Research Volume No.4, Issue No.1, December - January 2016, 2757 – 2759, @ 2013-2016 http://www.ijitr.com All rights Reserved. Page | 2757

[16]. Raorane A.A and Kulkarni R.V. (2012) " Data Mining: An effective tool for yield estimation in the agricultural sector" . IEEE.

[17]. D.S. Zingade, Omkar Buchade, Nilesh Mehta, Shubham Ghodekar and Chandan Mehta(2017), " Crop Prediction System using Machine Learning" . Special Issue on Recent Trends in Data Engineering *Volume 4, Special Issue 5, Dec.-2017* @IJAERD-2017,

[18]. Aakunuri Manjula and Dr. G. Narsimha(2016). Spatial temporal data mining for crop yield prediction. Research article

[19]. Ms.P. Kanjana Devi and Ms.S.Shenbagavadivu(2011), " Enhanced Crop Yield Prediction and Soil Data Analysis Using Data Mining" . IEEE.

[20]. B Vishnu Vardhan and D. Ramesh(2014), " Density Based Clustering Technique on Crop Yield Prediction" . RESEARCH ARTICLE.

[21]. Thanushree C, Harshitha .P and Manjunath C R(2018), " A Mobile Application Through Crop Yield is Predicted using Data Mining Techniques" . Volume 3, Issue 4, April – 2018 International Journal of Innovative Science and Research Technology ISSN No:-2456-2165 IJISRT18AP602 www.ijisrt.com 431.

M.Saranya. " Survey of Crop Prediction Using Different Classification Analytical Model." IOSR Journal of Computer Engineering (IOSR-JCE) 20.5 (2018): 72-80.