ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 9(38), DOI: 10.17485/ijst/2016/v9i38/95032, October 2016*

Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper

Subhadra Mishra[1], Debahuti Mishra[1]* and Gour Hari Santra[2]

[1]Siksha 'O' Anusandhan University, Bhubaneswar - 751030, Odisha, India;
mishra.subhadra@gmail.com, debahutimishra@soauniversity.ac.in
[2]Orissa University of Agriculture and Technology, Bhubaneswar - 751003, Odisha, India;
santragh@yahoo.co.in

Abstract

Objective: This paper has been prepared as an effort to reassess the research studies on the relevance of machine learning techniques in the domain of agricultural crop production. **Methods/Statistical Analysis:** This method is a new approach for production of agricultural crop management. Accurate and timely forecasts of crop production are necessary for important policy decisions like import-export, pricing marketing distribution etc. which are issued by the directorate of economics and statistics. However one has understand that these prior estimates are not the objective estimates as these estimate requires lots of descriptive assessment based on many different qualitative factors. Hence there is a requirement to develop statistically sound objective prediction of crop production. That development in computing and information storage has provided large amount of data. **Findings**: The problem has been to intricate knowledge from this raw data , this has lead to the development of new approach and techniques such as machine learning that can be used to unite the knowledge of the data with crop yield evaluation. This research has been intended to evaluate these innovative techniques such that significant relationship can be found by their applications to the various variables present in the data base. **Application / Improvement**: The few techniques like artificial neural networks, Information Fuzzy Network, Decision Tree, Regression Analysis, Bayesian belief network. Time series analysis, Markov chain model, k-means clustering, k nearest neighbor, and support vector machine are applied in the domain of agriculture were presented.

Keywords: Artificial Neural Network, Decision Tree, Machine Learning, Regression Analysis, Time Series Analysis

1. Introduction

Agriculture gave birth to civilization. India is an agrarian country and its economy largely based upon crop productivity. Thus agriculture is the backbone of all business in India. Now India stands in second rank in worldwide in farm production, Agriculture and allied sectors like forestry and fisheries considered for 14.5% of the GDP in 2015 and about 50% of the total manpower. The economy improvement of agriculture towards India's GDP is strongly declining growth still agriculture is statistically the broadest economic background and plays a significant role in the various socio economic frame work of India.

Indian agriculture is affected by various factors such as climate, due to topography, historical, geographical, biological, political, and institutional and socio economic factors. As time passed there are variations in natural factors and nature of technology so policies also changed. So agriculture production performance also changes in drastic path and large gaps in different geographic loca-

*Author for correspondence

tions of the country. The factors that affect agriculture are independent of one another. So this arise risk and the consistent output of food also affected.

Agricultural production is mostly affected by environmental factors. Weather influences crop growth and development, causing large intra-seasonal yield variability. In addition, spatial variability of soil properties, interacting with the weather, cause spatial yield variability. Crop agronomic management that is planting, fertilizer application, irrigation, tillage etc., can be used to offset the loss in yield due to effects of weather. As a result, yield forecasting represents an important tool for optimizing crop yield and to evaluate the crop-area insurance contracts.

As the climate changes time to time due to the pollution, population, solid waste management, surface and ground water hydrology etc and the impact of climate change in the developing world described by G Yamuna. From that we can get the idea that we have to analyze the production according to the climate change also[1]. In[2] described the rainfall runoff analysis using artificial neural network and they conclude that the artificial neural network model predict the data in accurate manner. In[3] predicted the reservoir water level stage using the neural network.

Machine learning means to give the knowledge to the machine. There are various types of machine learning techniques such as supervised and unsupervised learning. Supervised learning means there is one supervisor to supervise the thing that is the program is trained by training examples and then that can be used to find the accurate conclusion for new data. Artificial neural network, Bayesian network, decision tree, support vector machines, ID3, k-nearest neighbor, hidden markov model etc. are some of the examples of supervised learning. The unsupervised machine learning means a vast amount of data is given to the program and the program will find the patterns and the relations between them. So hidden patterns in the data can be discovered using unsupervised learning. Some examples of unsupervised learning algorithms are k-nearest neighbor, self organizing map, and partial based clustering, hierarchical clustering, k-means clustering etc.

Computer science and statistics together brought by machine learning to improve the prediction power. It is mainly used by data scientists, data analysts and also for them who wants to use the raw data to predict or find trends in data. As there are vast amount of data in the agriculture and also increased day to day, so machine learning techniques can be used for agriculture and agricultural production to find the accurate prediction of crop production.

2. Application of Machine Learning Techniques in Agriculture

Prediction of crop can be performed by using various machines learning algorithms such as mathematical and statistical method etc. Some of the methods those are already studied are presented here.

2.1 Artificial Neural Network

Artificial Neural Network (ANN) is the network of artificial neurons. It is based on the human brain's biological processes. It is one of the examples of supervised learning. Neural network has to be trained once, thereafter similar patterns in future data can be predicted for instance, meaningful solutions to problems can be produced even if the input data is incorrect/ incomplete.

Accuracy of ANN goes on increasing by the addition of more and more data. Also ANNs are capable of adopting their complexity without knowing the underlying principles. ANN can derive relationship between input and output on any process.

In[4] used artificial neural networks (ANN) in order to design output energy and Green house gas emissions (GHG) for forecasting potato production in Iran on the basis of input energy. They collected data from 260 farmers through face to face approach method. So, various ANN were matured and forecast the efficiency of them was assessed by using the quality aspect. The result depicted that the average total input & output energy of Potato production were 83,723 and 83 059 MJ ha^{-1}, respectively. The most dominant factor was electricity, chemical fertil-

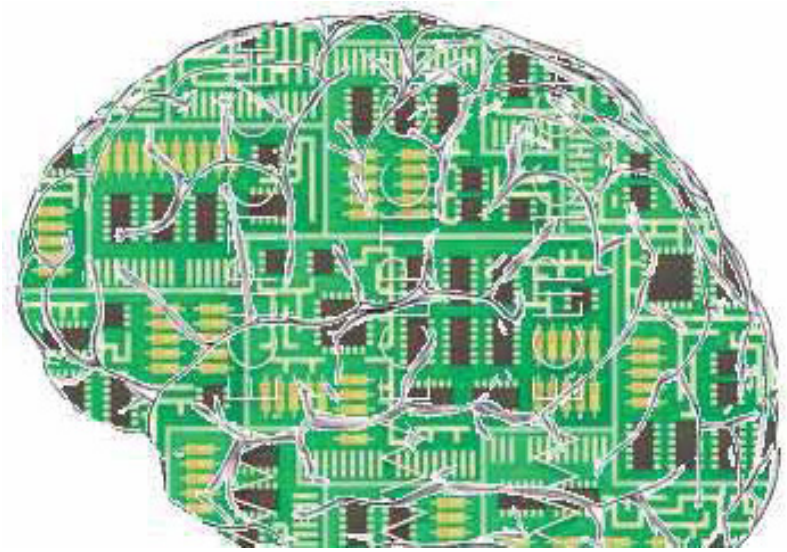

Figure 1. Brain with Artificial Neural Network.
http://en.wikipedia.org/wiki/File:Artifi cialFictionBrain.png

izer and seed in energy utilization. Energy use efficiency and energy productivity were measured. The ANN design with 12-8-2 structure was the best one to assess or forecast the output energy and total GHG emission in potato. The co-efficient of determination (R_2) of the best topology was 0.98 and 0.99 for potato output energy and total GHG emission, respectively as shown in **Figure 1.**

In[5] applied artificial Intelligence and Machine learning Algorithms particularly ID3 Algorithms and few optimization Algorithms to develop a web based expert system with Java as the front and SQL as the backend, to validate the symptoms of the tomato crop. Tomato is the most important vegetable crop which grown most extensively in the world. It is cultivated all over the world in farm garden kitchen garden and by market grown for raw or fresh utilization and processing purpose. This Tomato crop expert conductive system is design at collective venture with distinguished Agriculture scientist and expert in the region of Tomato cultivation with an expert team of computer engineers, programmers and designers. The professional arrangement accord with two essential parts, one is Tomato knowledge systems and other is expert system where in message, the applicant can get all the solid advice regarding various information like varieties, pest and diseases symptoms, cultural practice, mosaic of Tomato Fruits and plant. In expert system, the client is having a communication with the system in on line. Inquire the question to client and client has to give the answer. After obtaining response from the client the expert system will advice and exhibit its control measure of all the hazards. Knowledge expert system associate with recognition with various disease, pest and varieties generally appears in Tomato crop.

In[6] applied machine learning (ML) techniques for Maize breeding as revealed those ML algorithms are promising and can be used in statistical techniques applied in maize, alike the more newly popularize linear mixed models. Among the current technology available for expedite the releasing for new genotypes there is an emerging subject of ML. Several strategic uses of ML in maize breeding, quantitative trait loci mapping heterotic group assignment and the popular genome-wide selections are few of the main areas presently address by the literature. Corn is one of the most important cereals in the world and a primary source of calories for human being along with rice and wheat the evolution of genotypes

adapted to aggravating climate, particularly drought situation which has to be grown in marginal law and changing climatic condition for crop production.

In[7] used extended machine learning and simulation techniques to design crop management strategies and analyzed various methods of a stimulation tool for winter wheat crop management. The simulation tool previous by developed which was outlined the problem encountered by crop production plan engineers to designing the strategies that different from thou presently used. Thus developed newly exploration support system that is originates on reinforcement learning and genetic algorithm techniques that technique to be helpful for user to expedite a broad range of solution which performed to be an advance approach and interesting strategies as an incremental way.

In[8] explained that neural networks can automatically "learn", complex relationships among data. They can be useful in modeling process. They described the neural network advantages over multiple regressions. The neural network can select the independent variable so no need to select independent variable in the data. To propose a model function also not required like multiple regression. More complex variables can be discovered by neural network. Thus a model with maximum precision can be developed. When noise present in the data then it is more resistant.

In[9] studied the presentation of three dissimilar types of multivariate modeling techniques for predicting crop yields. They measured the top soil depth by a soil conductivity meter and on a 30 m grid soil fertility was sampled. They took the data of corn of 1993 and soybeans of 1994. During the harvest time the yield was measured. For the development of the neural network the data was randomly divided into training and testing datasets. They took the topsoil depth, phosphorous, potassium, salt, organic matter and magnesium saturation as input parameters and predict the corn and the same process is for soybean yield. They used the network geometry as shown in the **Figure 2.**

As per[10] ANN is a system inspired by the human brain. In artificial Neural Networks (ANN) each node depicts a neuron whereas each link depicts the way of interaction between two neurons". Very simple tasks are performed by each neuron while the more complex are performed by the network that represents the work of all its neurons. Artificial Neural network is an interconnected set of input and output units where a weight is associated with each connection. During testing phase the network learns how to predict the input sampled by fine tuning the weights. It is a new technique used in flood forecast and is advantageous in modeling rainfall and run off relationship when compared to conventional techniques flood forecast.

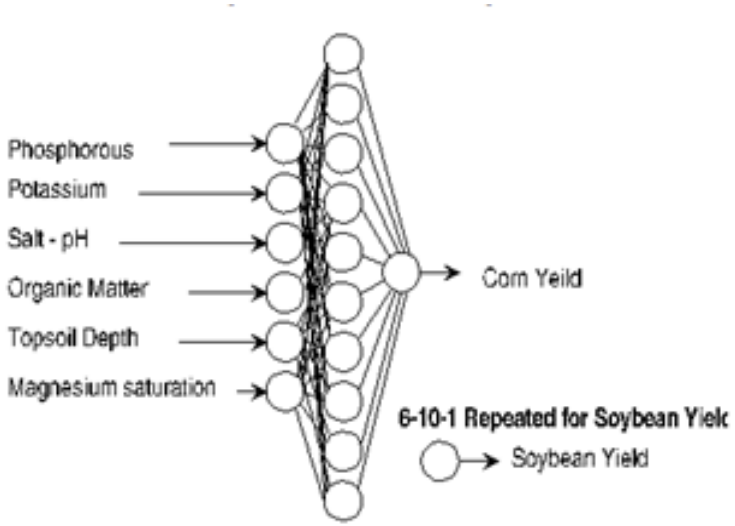

Figure 2. Yield Prediction Model - Drummond et al. 1995.

Neural networks are also better in computing over the conventional methods. ANN is highly suitable for problem having more time for getting solution i.e. Prediction of pest attack incidences can be done in advance, using ANN.

The rice crop monitoring system developed by Chen et al.[11], was established on the change as tested to a neural network classification. For one wet and one dry season the system was delineated rice production areas. On different planting dates it was able to extract information on rice cultivation. For both the seasons a minimum mapping accuracy of 96% was achieved. In a neural network based model this information was used to predict rice yield for the wet season on a regional basis. The predicted yields by neural network were compared with the government statistics and the accuracy of prediction was 94%.

International trade of rice forecasting is very difficult as demand and supply are affected by many unpredictable factors such as trade barriers and subsidies, agricultural and environmental factors, meteorological factors, biophysical factors, changing demographics, etc. those interact in a complex manner.

A comparison study about the performance of ANN's with exponential smoothing and ARIMA models for forecasting rice exports from Thailand was done by[12]. During the validation process of the models they have evaluated various aggregate measures of forecast errors such as MAE, MSE, MAPE, and RMSE to ensure that the models can reproduce acceptable results on hidden future. The satisfactory goodness of fit was shown by the Holt–Winters and the Box–Jenkins models. But during the prediction of unseen data these models were unable to perform well. But as the ANNs were able to track the dynamic non-linear trend and seasonality along with their interactions they showed good performance.

In[13] developed a model for corn and soybean yield forecasting with climatic aspect by applying artificial neural network. They have considered the rainfall, Maryland corn and soybean yield data and predict the corn and soybean yield at state, regional and local levels by applying both the artificial neural network technique and the multiple linear regression model. Lastly they compared both the techniques and conclude that the ANN model gives more accurate yield prediction than the multiple linear regressions.

It[14] studied both the feed forward neural networks and the statistical methods such as linear regression for the prediction of agricultural crop production. They present a brief literature review of both the technology and conclude that we can use the artificial neural network model when the relationship between the variables is unknown to us and complex also it's very difficult to handle statistically. But the statistical linear regression model can be used when the variables are known as it allows interpretation of coefficients of the individual variable s and due to the parametric assumptions of these models.

In[15] suggested weather based model for forecasting crop yield for various crops at selected districts/agro climatic zones/states. The models utilized weekly/fortnightly weather data and, in some cases, agricultural inputs at district level. The techniques included development of suitable weather indices which were used as repressors' in the models, discriminant function analysis and water balance technique. Using these approaches, reliable forecast of crop yield can be provided before harvest - 21/2months (rice and wheat), 1*Y*2month (sorghum), I month (maize) and in middle of September in sugarcane. Pests and diseases, major factors limiting the production, are also influenced by weather conditions. Therefore, weather based models were developed for forewarning of important pests/diseases in rice, mustard, pigeon pea, sugarcane, groundnut, mango, potato and cotton at various locations using the techniques like regression analysis (taking suitable functions of weather variables/indices as repressors'), complex polynomials through GMDH technique and Artificial Neural Network technique. The forewarnings through these models will be very useful in taking timely control measures. These approaches have been successfully used by other workers and organizations for forecasting crop yield and the forewarnings issued enabled the farmers to optimize plant protection measures.

In[16] used the neural network model to study for the development of a model for oil palm yield. They took the percentages of nitrogen, phosphorous, potassium, calcium and magnesium in leave as input variables and fresh fruit bunch as the target variable. Combining the activation function, learning rate, momentum term, number of runs, and number of hidden nodes with all the layer they found that it affects the neural network performance.

They conclude that the numbers of hidden nodes affect significantly and the learning rate, momentum term and number of runs affects insignificantly on the neural network performance. Finally they measured the suitability of the model using R^2values and show the result that the neural network model is better than the regression analysis.

In[17] used the artificial neural network to predict the crop by using the soil parameters such as types of soil, pH, nitrogen, phosphate, potassium, organic carbon, calcium, magnesium, sulphur, manganese, copper, iron, depth and climate parameters such as temperature, rainfall, humidity. They did the experiment on the crops such as Cotton, Sugarcane, Jawar, Bajara, Soyabean, Corn, Wheat, Rice and Groundnut.

2.2 Information Fuzzy Network

In[18] analyzed the remote sensing and other parameter for predicting crop yield through aggressive of neural network. They applied the flexible Neuro-fuzzy Inference system (ANFIS). Soil moisture content, ground biomass and repository organ are the inputs to ANFIS. It has only a single number is sought or otherwise one output node, i.e. yield. The other problem in forecasting yield in that remote sensing data do not go long behind in time. Hence any forecasting effort is compelled to apply a very finite number of past years in order to build up a design to predict future values. The arrangement is disciplined by leaving one year out and using all the other data. They calculate the deviation of our estimate compared to the yield of the year that is left out. The procedure is used to all the years and average prediction efficiency was given.

In[19] used FINKNN to the effect of forecasting sugar production established on population of assessment furnished by Hellenic sugar Industry. FINKNN a K nearest-neighbor classifier performing over the metric lattice of traditional interruption supported convex fuzzy sets. They observed that FINKNN give better forecast efficiency on this effect and consider the extensive opportunity and possible service of these techniques. They also observed that for complication effect. Population of measurement, data can be interpreted by fuzzy interval numbers (FIN) and they started an algorithms for buildup FINs from such populations. They commenced a lattice theoretic metric gap amongst FINs with discretionary. Shaped membership activity, which developed the basis for FINKNN's analogy measurements.

In[20] studied to group the aspect of incomplete information, used fuzzy modeling advancement to conquer to the confusion during the advancement of an agriculture web based decision support system, because the effect of climate development is accepted to consequence high productivity and crop arrangement in agricultural field of Malaysia. So, data synthesis in the progeny tests needs decision of combat issuing from the heterogeneity of data, they conclude decision algorithms classifier from incomplete knowledge to administer transparent and more acceptable information during user communication. The detailed data about the rainfall arrangement, soil structure of the region will allotment data by taking full asset of the incomplete information to accomplish better results. The improvements of this study are newly advanced algorithms and investigation for planting material classification. So, the researcher addressed the experimental results which may afford high recovery in planting material breeders in agriculture management through achieve of policies of decision making.

In[21] applied aggressive effective group to of Fuzzy Cognitive Maps (FCMS) for design and forecast method in apples (cv. Red Chief). For designing FCMS are excellent causal cognition device and affecting influential process. Clarity, elasticity properties and ability to various conditions and comfort of use they achieve strength. They categories apple production by using In common, they design the attitude of complicated system have interpretation effectiveness and that can be applied to forecast new science. They have selected a data driver non-linear FCM ability access to demarcate yield in apples. To perform skillful knowledge for yield forecast and crop management by using the recommended technology they modeled and developed the FCM model. The nodes of advanced FCM design linked with the supervised edges and the nodes produce the main soils factors influencing yield such as soil texture like clay and sand content, Organic Matter (OM), potassium (K), calcium (Ca), Electrical Conductivity (EC), phosphorous (P), zinc (Zn) contents and other edges display the cause effect or weighted affinity among the soil characteristics and pro-

duction. They implement all the algorithms in the same dataset and they found superiority of the FCM learning model in the yield forecast.

A fuzzy logic information network and a decision-support system were developed by for cultivation of Olive in Andalusia[22].

In[23] used prototype tool to help the farmers of Canadian farmers for their crop selections. It has been very problematic to collect together the information data of maximize crop performance during growing season. They were advocating for extending prototype system which allow submitting reports of crop performance along with describe date for growing conditions. These experience reports was submitting by many farmers and the data in these reports could be mined to facilitate generalized information about the performance of various crops and their conditions which best support each.

2.3 Decision Tree

The decision tree models include the concepts as nodes, branches, terminal values, strategy, payoff distribution, certain equivalent, and the rollback method. There are three kinds of nodes and two kinds of branches. The decision node which is represented as square is a point where a choice must be made. The decision branches are extending from a decision node. Each terminal node has an associated terminal value, sometimes called a payoff value, outcome value, or endpoint value. The result of a scenario or the sequences of decisions are measured by each terminal value. There are two step processes for the construction of a decision tree algorithm- first, growth of large decision tree then reduction of size and over fitting the data, in the second step, and tree is pruned. The pruned decision tree that is used for classification purposes is called the classification tree described[24].

There are various factors influencing the prediction. In[25] used agronomic variables, nitrogen application and weed control using the machine learning algorithm such as artificial neural network and Decision tree to develop the yield mapping and to forecast yield. They conclude that high prediction accuracies are obtained by using ANNs.

In[26] described the soybean productivity modeling using decision tree algorithms. They had collected the climate data of Bhopal district for the period 1984-2003. They considered the climatic factors such as evaporation, maximum temperature, maximum relative humidity, rainfall and the crop was soybean yield. They have applied the Interactive Dichotomizer3 (ID3) algorithm which is information based method and based on two assumptions. Using the induction tree analysis it was found that the relative humidity is a major influencing parameter on the soybean crop yield. Decision tree formed for influence

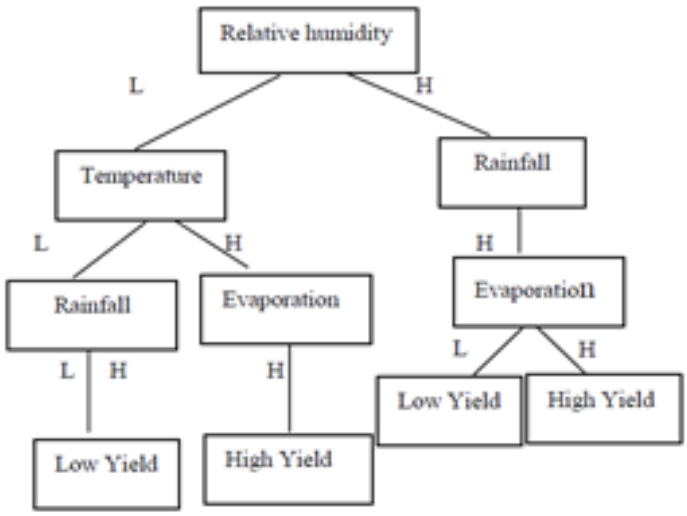

Figure 3. Decision tree for influence of climatic factors on soybean yield, S. Veenadhari et al.

of climatic factors on soybean yield. Using the if-then-else rules the decision tree is formulated to classification rules which are shown in **Figure 3** Relative humidity affects much on the production of soybean and some rules generated which help to in the low and high prediction of soybean. One of the drawbacks was only the low or high yield can be predicted but the amount of yield production cannot be predicted[27].

With the consequence of climate diversity in India, bulk of the agriculture crops are being poorly impressed in terms of their achievement over a period of past two decades. Forecasting the crop production well advanced of its yield might be helpful to policy inventor and farmers to take convenient decision for marketing and storage. Such forecasting will also advice concurrent industries for planning the coordination of their business as these don't take into account component of the climate and are chiefly factual. In the present course a software tool titled 'Crop Advisor' has been advanced as a client friendly webpage for forecasting the effect of weather parameters on the crop yields. C4.5 algorithm is applied ascertain the most effective climatic parameter on the crop yields of specified crops in preferred district of Madhya Pradesh. The software will be helpful for advice the effect of various weather parameters on the the crop yield. Other agro –input parameters liable for crop yield are not accommodating in this tool. Since the application of these input parameters differ with individual fields in space and time.

2.4 Regression Analysis

The crop production is very important for agriculture related organizations, consultants, producers etc. Crop forecasting can be done by collecting data from different sources and taking different types of data such as metrological data, agro metrological, soil, remote sensing data, agricultural statistics etc. Accurate and timely forecast is required for marketing, storage and transportation decisions.

In[28] investigated the development of a crop prediction model framework and concluded that climate related variables were not the main determinants of corn yield, rather yield was greatly affected by planting practices, particularly by the application right amount of fertilization.

Forecast the influence of temperature on the Jowar crop yield was done by[29]. The experiment was conducted by taking the Jowar production and both maximum and minimum temperature of India during the period from 1950-2011. Pearson correlation coefficients were applied the test of significance was 2-tailed and their result were significant at 0.01 level. Then regression analysis was done between crops yield and temperature. It was concluded that the Jowar yields were very less reliant on the temperature but the another factors affect the yield at a higher level and when the temperature was decreased, then increased the yield of Jowar crops.

A model was developed by[30] for forecasting the yield of the sugarcane in Coimbatore district by using the fortnightly weather variable such as average daily maximum and minimum temperature, relative humidity in the morning and evening and total fortnightly rainfall and the yield data. They took the data from 1981-2004 and developed the model by the data for a period of (1981-2001). The model was validated by taking the data from (2002-2004). Their forecast model was able to explain 87% of variation in the sugarcane yield and they conclude that the sugarcane yield can be forecasted using the regression technique successfully two months before harvest.

There are various statistical techniques for crop production but the regression analysis is widely used statistical techniques[31.]

In[32] developed a regression models and techniques to predict the response variable that is yield and the explanatory variables such as weather, soil properties.

In[33] found that parametric regression model is used in many of the yield forecasting method. The functional forms of the predictor variables are known, taking these assumption they have used the parametric regression.

The commonly used models are linear regression models described by[34], the and Wilcox et al.[35,] found the use of polynomial regression models, and House et al.[36] found about nonlinear regression models.

During the crop's growing season developed a model by taking two climatic parameters such as rainfall and temperature to predict the crop yield. They have used a functional linear regression analysis to find the relationship between the yield which is the response variable and temperature which is the predictor variables[37].

In[38] used the correlation analysis taking the yield as response variable and temperature and precipitation as predictors. For selecting the best predictors they have used a stepwise regression technique.

The effect of climate change on corn production under the anticipated seasonal climate change conditions was found by To produce the daily weather data they have been used a weather generator and to predict the corn yield they have used the multiple linear regression models taking the climate variables such as temperature, rainfall, humidity, wind speed and solar radiation. They conclude that climate variability *significantly* affects crop yields[39].

2.5 Clustering

Cluster analysis or clustering is the process of identifying objects that are similar to each other but different from individuals in other groups. It is mainly used for data analysis. Clustering is used in many fields such as machine learning, pattern recognition, image analysis, information retrieval, and agriculture etc. There are various clustering algorithms are there such as k-means, k-medoid etc but the common and important clustering algorithm is k-means.

In[40] demonstrates an evaluation of modified k-Means clustering algorithm in crop prediction. Their results and evaluation showed the comparison of modified k-Means over k-Means and-Means++ clustering algorithm and found that the modified k-Means has achieved the maximum number of high quality clusters, correct prediction of crop and maximum accuracy count.

Based on the frequency of variables available by the weather forecast model by[41] classifies the metrological data. They have identified the patterns that are associated to severe convective activity. For some selected mini-regions of Brazil during summer of 2007 their result showed good classification performance. They thought that their metrological model Eta serve as a support tool for meteorologists to identify patterns in advance.

2.6 Principal Component Analysis

In[42] adopted method through data mining process based on principal component analysis to ascertain the arrangement and familiarity in data used for correct forecasting. Rainfall is achievable and the main case variable in the evidence of monsoon the rainfall amount may alter from time to time depending on the place. Rainfall outlining is important for a catchment area where storing of water or rainfall had been carried out, particularly for flood observant methods. The current method based on statistical method for broad range prediction of Indian monsoon rainfall has issues clear cut application, for decade. The limited variation in the periodic rainfall can have destructive impact on economy of India. Component derivation method has been used generally by weather and water assets analyzed to evaluate high spatial datasets such as overall sea surface temperature & rainfall periods. The principal component analysis for obtaining the time based development of monsoon rainfall in India is studied.

2.7 Bayesian Belief Network

A Bayesian network or Bayes network or belief network or Bayesian model or probabilistic directed acyclic graphical models a type of statistical model.

A belief network to assess the effect of climate change on potato production was formulated by[43]. They have shown a belief network combining the uncertainty of future climate change, considering the variability of current weather parameters such as temperature, radiation, rainfall and the knowledge about potato development. They thought that their network give support for policy makers in agriculture. They test their model by using synthetic weather scenarios and then the results are compared with the conventional mathematical model and conclude that the efficiency is more for the belief network.

2.8 Time Series Analysis

Time series analysis is method to analyze time on parametric, series data to extract meaningful statistics and other characteristics of the data. Time series forecasting is a model to predict future values based on previously observed values. There are various types of time series analysis methods such as frequency domain and time domain, parametric or non-parametric, may be linear or nonlinear, univariate and multivariate. Frequency domain includes spectral analysis and wavelet analysis, time domain includes auto-correlation and cross-correlation,

parametric approaches includes autoregressive or moving average model, non-parametric approaches includes covariance or spectrum of the process. Time series analysis is the important tool for crop forecasting. Taking the dependent variable yield as a function of time we can find the relation between yield and time.

In[44] described the new concept of crop yield under average climate conditions and used the time series techniques on the past yield data to set up a forecasting model. They tested the model in Liaoning province and used the data grain per unit yield obtained by adding a variety of crop per unit yields from 1949 to 2005. They used the moving average method first then applied regression equation and finally find the difference and find the impact of climate on yield. So they conclude that the moving average model is regarded as the potential yield forecasting model. The strong point is it needed a relatively small amount of data.

2.9 Markov Chain Model

Markov chain is a stochastic process which is mathematical model in a probabilistic manner. In case of markov chain the outcomes of an experiment depends only on the outcome of the previous experiments. In other words we can say the next state of the system depends on the present state. Markov chains are named after the Russian mathematician (1856-1922) who started the theory of stochastic processes.

For forecasting cotton yield from pre-harvest crop data applied the Markov chain approach[45]. They have investigated the utility of the markov chain approach in predicting crop yields. Separately they analyze cotton yield data from two key producing states, California and Texas taking the four year period 1981-1984. They estimate the probability distribution using markov chain. To assist in selecting the key variables within each period for the baseline 1981-1983 data they have used two regression models such as multiple linear regressions and multiple rank regressions. From each of the state California and Texas four transition matrices were calculated. Then the predicted yield distribution is calculated by multiply consecutive transition matrices. For yield forecast they have used the means of these predicted yield distributions.

In[46] applied the second order Markov chain model for forecasting of sugarcane yield through which, it was possible to use data from two stages simultaneously. They found that this model is better than the models in use i.e. first order Markov chain model and the regression model. Thus they conclude that the second order Markov chain model can be used for crop yield forecasting in preference over regression model and first order Markov chain model.

In[47] appraised Markov logic model of crop rotations for early crop mapping. Furnished and up to date knowledge on crop area, production and productivity is essential for the understanding of environmental condition in agriculture, as the regulation of land use, management method and food insurance in early direction systems. A machine learning technique is suggested to commendable crop rotations which can predict with good ability, at the outset of agricultural season, the crops all most achievable to be presented in a given field by implementing the cropping system of coming 3-5 years. The access capable to analyze from data and acquire professional awareness as mentioned by first order logic rules.

Its ability was assessed by applied the French Land Information System attained in the frame of the EU's common Agricultural policy. This estimated was done by applying different location in terms of temporal depth and dimensional coverage. The estimated results advertised expected approach is capable to forecast the nature of crop in each and every field, before sowing of the crop season, with an efficiency as high as 60%. That is an exceptional than the results attained with present approaches due to remote sensing remote sensing imagery.

3. Conclusions

Now-a-days a growing number of applications of machine learning techniques in agriculture are required for which a large amount of data currently available from many resources can be analyzed to find the hidden knowledge. This is an advanced researched field and is expected to grow in the future. The integration of computer science with agriculture helps in forecasting agricultural crops. It is required to build on objective methodology for pre-harvest crop forecasting. Building up a suitable model will have certain merits over the traditional forecasting method. A detailed about the study is shown in the **Table 1.**

Table 1. List of important fusion of machine learning techniques in agriculture

Fusion type	**Application area**
Nonlinear Regression[36]	Forecasting Corn Yields
Markov Chain Approach[45]	Forecasting Cotton Yields from Surveys
Linear Regression[34]	Estimating Grain Yield of Maturing Rice
Belief Networks [43]	Future Crop Production
Neuro-Fuzzy Modeling[18]	For Crop Yield Prediction
Secpnd Order Markov Chains[46]	Forecasting of Crop Yields
Polynomial Regression[35]	Factors Affecting the Yield of Winter Cereals in CropMargins
Deterministic and Probabilistic Prediction Approaches[38]	Season to Inter-annual Climate Forecasting
FINkNN: a fuzzy interval number k-nearest neighbour classifier[19]	for prediction of sugar production from populations of samples
Artificial neural networks[13]	corn and soybean yield prediction
Neural Network[11]	Rice Crop Monitoring
Artificial Neural Networks[12]	Forecasting Thailands Rice Export
Building a fuzzy logic information network and a decision-support system[8]	for olive cultivation in Andalusia
Regression[30]	sugarcane yield using climatic variables
Decision Tree Algorithms[26]	Soybean Productivity Modelling
A Fuzzy Modelling of Decision Support System[20]	for Crop Selection
Time Series Techniques[44]	Crop Yield Forecast
Datamining with climate variable[29]	Jowar Crop Yield in India

Table 1 Continued

Fuzzy Cognitive Map learning approach[21]	Yield prediction in apples
Regression and Neural Networks Models[14]	for Prediction of Crop Production
Modified K-Means Clustering[40]	Crop Prediction
Artificial Neural Network Approach[17]	Agricultural Crop Yield Prediction
Markov logic model[47]	Crop rotations for early crop mapping

4. References

1. Yamuna G. Impacts of climate change in the developing world. Indian Journal of Science and Technology. 2016 Feb; 9(6):1–1.
2. Chakravarti A, Joshi N, Panjiar H. Rainfall runoff analysis using artificial neural network. Indian Journal of Science and Technology. 2015 Jul; 8(14):1–7.
3. Ashaary N, Ishak WHW, Ku-Mahamud KR. Neural network application in the change of reservoir water level stage forecasting. Indian Journal of Science and Technology. 2015 Jul; 8(13):1–6.
4. Khoshnevisan B, Shahin Rafiee S, Omid M, Mousazadeh H, Rajaeifar MA. Application of artificial neural networks for prediction of output energy and GHG emissions in potato production in Iran, Elsevier. Agricultural Systems. 2014; 123:120–7.
5. Babu MSP, Ramana Murty NV, Narayana SVNL. A web based tomato crop expert information system based on artificial intelligence and machine learning algorithms. International Journal of Computer Science and Information Technologies. 2010; 1(3):1–5.
6. Ornella L, Cervigni G, Tapia E. Applications of Machine Learning for Maize Breeding. In: Venkateswarlu B, Shanker AK, Shanker C. Book chapter of Crop stress and its management: Perspectives and Strategies, Springer, New York, USA. 2012; 1–29.
7. Attonaty JM, Chatelin MH, Garcia F, Ndiaye N. Using extended machine learning and simulation technics to design crop management strategies. First European Information Technology in Agriculture, Copenhagen. 1997; 1–500.
8. Gorni G, Augusto A. The Application of Neural Networks in the Modeling of Plate Rolling Processes, JOM. The Minerals, Metals and Material Society. 2008; 49(4):1–4.
9. Drummond ST, Suddeth KA, Birrell SJ. Analysis and Correlation Methods for Spatial Data. ASAE. 1995; 1–6.
10. Baskar SS, Arockiam L, Arul Kumar V, Jeyasimman L. Brief survey of application of data mining techniques to agriculture. Agricultural Journal. 2010; 5(2):116–8.
11. Chen C, Mcnairn H. A neural network integrated approach for rice crop monitoring. International Journal of Remote Sensing. 2006; 27(7):1367–93.
12. Co HC, Boosarawongse R. Forecasting Thailand's Rice Export: Statistical Techniques vs. Artificial Neural Networks, Computers and Industrial Engineering. 2007; 53(4):610–27.
13. Monisha Kaul M, Robert L, Hill H, Walthall C. Artificial neural networks for corn and Soybean yield prediction, Elsevier. Agricultural System. 2005; 85(1):1–18.
14. Prasad PR, Begum SA. Regression and neural networks models for prediction of crop production. International Journal of Scientific and Engineering Research. 2013 Sep; 4(9):98–108.
15. Agrawal R, Mehta SC. Weather Based Forecasting of Crop Yields, Pests and Diseases - IASRI Models. Journal of Ind Soc Agril Statistics. 2007; 62(2):1–12.
16. Azme Khamis A, Ismail Z, Haron K, Mohammed AT. Neural network model for oil palm yield modeling, Asian network of scientific information. Journal of Applied Sciences. 2006; 6(53):1–9.
17. Dahikar MSS, Rode SV. Agricultural crop yield prediction using artificial neural network approach. International Journal of Innovative Research in Electrical, Electronics,

Instrumentation and Control Engineering (IJIREEICE). 2014 Jan; 2(1):1–4.
18. Stathakis D, Savin I, Negre T. Neuro-fuzzy modeling for crop yield prediction. The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences. 1994; 34:1–4.
19. Petridis V, Kaburlasos VG. FINk NN: a fuzzy interval number k-nearest neighbour classifier for prediction of sugar production from populations of samples. Journal on Machine Learning Research. 2003; 4:17–37.
20. Salleh MNM. A Fuzzy Modelling of Decision Support System for Crop Selection. IEEE Symposium on Industrial Electronics and Applications (ISIEA2012), Bandung, Indonesia. 2012; 17–22.
21. Papageorgiou EI, Aggelopoulou KD, Gemtos TA, Nanos GD. Yield prediction in apples using Fuzzy Cognitive Map learning approach. Computers and Electronics in Agriculture, Elsevier. 2013; 91:19–21.
22. Delgado G, Aranda V, Calero J, Sanchez-Maranon M, Serrano J M, Sanchez D, Vila MA. Building a fuzzy logic information network and a decision-support system for olive cultivation in Andalusia. Spanish Journal of Agricultural Research. 2008; 6(2):252–63.
23. Daryl H, Hepting H, Maciag T, Hill H. Web-Based Support of Crop Selection for Climate Adaptation. 45th Hawaii International Conference on System Sciences, IEEE, Maui, HI. 2012. p. 1227–605.
24. Vijaysinh A, Solanki S. Data mining techniques using WEKA classification for Sickle Cell Disease. (IJCSIT) International Journal of Computer Science and Information Technologies. 2014; 5(1):1–26.
25. Uno Y. Artificial Neural Networks to Predict Corn Yield from Compact Airborne Spectographic Imager Data. Computers and Electronics in Agriculture. 2005; 47(2):149–61.
26. Veenadhari S, Mishra B, Singh CD. Soybean Productivity Modelling using Decision Tree Algorithms. International Journal of Computer Applications. 2011; 27(7):975–8887.
27. Veenadhari S, Misra B, Singh CD. Machine learning approach for forecasting crop yield based on climatic parameters. IEEE International Conference on Computer Communication and Informatics, Coimbatore. 2014. p. 1–16.
28. Rossana MC, Leon D, Rex EL, Jalao J. A Prediction Model Framework for Crop Yield Prediction. Asia Pacific Industrial Engineering and Management Society Conference Proceedings Cebu, Phillipines. 2013. p. 185.
29. Kumar AVTV, Rajini Kanth R. A data mining approach for the estimation of climate change on the jowar crop yield in India. International Journal of Emerging Science and Engineering (IJESE). 2013; 2(2):16–20.
30. Priya SRK, Suresh KK. A study on pre-harvest forecast of sugarcane yield using climatic variables, Statistics and Applications. 2009; 8(2):1–8.
31. Horie T, Yajima M, Nakagawa H. Yield Forecasting. Agricultural Systems. 1992; 40(1):211–36.
32. Rosa DLD, Cardona F, Almorza J. Crop Yield Prediction Based on Properties of Soils in Sellilla, Spain, Geoderma. 1981; 25(3-4):267–74.
33. Kaspar TC. Relationship between six years of corn yields and terrain attributes. Precision Agriculture. 2003; 4(1):87–101.
34. Shibayama M. Estimating grain yield of maturing rice canopies using high spectral resolution reflectance measurements. Remote Sensing of Environment. 1991; 36(1):45–53.
35. Wilcox A. Factors Affecting the Yield of Winter Cereals in Crop Margins. Journal of Agricultural Science. 2000; 135(4):335–46.
36. House CC. Forecasting Corn Yields: A comparison Study using Missouri Data, Statistical Research Division, United States Department of Agriculture. 1979; 17(16):3189–200.
37. Kantanantha K, Nantachai N. Crop Decision Planning Under Yield and Price Uncertainties. 2007.
38. Oludhe O, Christopher C. Deterministic and Probabilistic Prediction Approaches in Season to Inter-annual Climate Forecasting. 2002.
39. Lansigan L. Analysis of Climatic Risk and Coping Strategies in Two Major Corn Growing Areas in the Philippines. 2010.
40. Utkarsha P, Narkhede N, Adhiya KP. Evaluation of Modified K-Means Clustering Algorithm in Crop Prediction. International Journal of Advanced Computer Research. 2014; 4(3):1–1.
41. Glauston R, Liman TD, Stephany S. A new classification approach for detecting severe weather patterns, Computers and Geosciences, ELSEVIER. 2013; 57:158–65.
42. Brindha KP. Data mining based on principal component analysis for rainfall forecasting in India. International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE). 2013 Sep; 3(9):1–11.
43. Yiqun Gu Y, James W, McNicol M. An Application of Belief Networks to Future Crop Production. IEEE Conference on Artificial Intelligence for Applications, San Antonia, TX. 1994. p. 305–9.

44. Hong-Ying L, Yan-Lin H, Yong-Juan Y, Hui-Ming Z. Crop yield forecasted model based on time series techniques. Journal of Northeast Agricultural University (English Edition). 2012; 19(1):73–7.
45. Matis JH, Birkett T, Boudreaux D. An Application of the Markov Chain Approach to Forecasting Cotton Yields from Surveys. Agricultural Systems. 1989; 29(4):357–70.
46. Jain RC, Ramasubramalliall V. Forecasting of Crop Yields using Second Order Markov Chains. Journal of the Indian Society of Agricultural Statistics. 1998; 51:61–72.
47. Osman J, Inglada J , Dejoux JF. Assessment of a Markov logic model of crop rotations for early crop mapping, Elsevier. Computers and Electronics in Agriculture. 2015; 113:234–43.