A Survey on Machine Learning and Deep Learning Techniques used for Agricultural Crop Prediction, Soil Classification and Rainfall Prediction.

[1]Vignesh.K[1], Dr.A.Askarunisa[2]

[1]Assistant Professor, [2]Professor & Head
Department of Computer Science, KLN College of Information Technology, Madurai, India

---***---

Abstract: Machine Learning is the leading technology which enables the machine to think for prediction and forecasting. Deep learning is the subset of Machine Learning which has the capability of learn unsupervised and unconstructed data which gets inspire from human brain. Machine learning and Deep Learning technologies can be combined used to assist in the prediction model for agriculture. Agriculture is an art of cultivating the food crops. Such art needs the technical assistance to avoid the over production and scarcity of the products. This prediction model will save the crops from vilest climatic conditions. Researchers finding many techniques to make the prediction which includes different kind of parameters. By surveying the techniques and parameters, it will be clear to get the right way for the creation of comprehensive prediction model for agriculture.

Introduction:

The prediction Model explains that it will make the decision regarding the future results. Large amount of data in variety of parameters which can influence the results. So computer has to do many functionalities like think, analyze the large range of data and creating the similar patterns and many classification.

Artificial Intelligence (AI) is a technology which make it possible for machines to learn from experiences. It is wide ranging technology which includes many subset techniques. Machine learning (ML) is a subset of AI and it is a method of data analysis to build an analytical model. Machine learning holds 5 supervised and 3 unsupervised learning techniques. Deep Learning is a part or subset of ML. It is a reinforcement learning Technology to process and digest the knowledge better. Artificial Neural Network(ANN), Convolution Neural Network(CNN), Recurrent Neural Network(RNN) are some of the Deep Learning techniques. Among these techniques, the best techniques should be chosen and applied to get the comprehensive prediction value from multi parameters. Lot of researches has been done and going on using such a many algorithms. So it is important to survey the related researches and to analyze the techniques which hold the multiple factors as parameters, following researches explains about its techniques and parameters.

Hsu-Yang Kung, Ting-Huan Kuo, Chi-Hua Chen, Pei-Yu Tsai **"Accuracy Analysis Mechanism for Agriculture Data Using the Ensemble Neural Network Method"** www.mdpi.com/journal/sustainability. Published: 1 August 2016. **[1]**

--Ensemble Neural Network (ENN) had been used to generate agricultural yield predictions. It create multiple network models, each with different numbers of hidden layers and neurons.

--The weighted average of all remaining network models to improve the accuracy of the prediction.

--The method based on ENN had a much lower error rate than traditional back-propagation neural networks.

--Convolutional Neural Network will Produce better result than ENN.

Thayakaran Selvanayagam, Suganya S, Puvipavan Palendrarajah **"Agro-Genius: Crop Prediction using Machine Learning" -International Journal of Innovative Science and Research Technology-2019 [2]**

This paper mainly focusing on vegetable cultivation of particular region. Collected data set is divided into 3 partitions (ie.) 80% for training, 10% for testing, 10% for validation.

--Recurrent Neural Network (RNN) is used for crop yield prediction and optimization

-- Machine Learning algorithms and techniques like LSTM (Long-Short

Term Memory), ARIMA (Auto Regressive Integrated Moving Average) are used for crop price prediction.

N.L. Chourasiya, P. Modi , N. Shaikh3, D. Khandagale, S. Pawar **"Crop Prediction using Machine Learning" IOSR Journal of Engineering (IOSR JEN)- 2019. [3]**

Importance of seed classification is stressed in this paper. Training data model has been created and prediction model gives the fore coming values.

--SVM classifier, Multi Linear Regression and ANN is used to develop the prediction model.

-- Input to these Machine learning techniques are "filtered Data set" is used extracted from raw Seed and soil data.

Sneha N, Dr. Jharna Majumdar :"**Big Data Application in Agriculture to Maximize the Rice Yield Crop Production using Data Mining Techniques**" International Journal of Innovative

Research in Computer and Communication Engineering. *Vol. 5, Issue 5, May 2017.* **[4]**

--Big data and Data Mining techniques were applied on huge agriculture data

--Data set ranges from 2005-2013 of rice production. Mining the important patterns from historical data and using those patterns to predict future crop production.

--Chameleon,Random Forest and Regression Techniques were applied for different parameters to optimize the rice production.

Teresa Priyanka,Pratishtha Soni,C. Malathy **"Agricultural Crop Yield Prediction Using Artificial Intelligence and Satellite Imagery" - Eurasian Journal of Analytical Chemistry 2019 [5]**

The motto of this paper is to predict the crop yield well ahead of its harvest time which help the farmers to make necessary steps for selling and storage of crops.

-- ANN(Artificial Neural Network) is used to predict the yields based on diverse soil parameters and different weather conditions.

-- Additionally CNN(Convolution Neural Network) is applied to predict the crop yield with satellite images of remote areas.

Dakshayini Patil , Dr. M .S Shirdhonkar **"Rice Crop Yield Prediction using Data Mining Techniques: An Overview** " International Journal of Advanced Research in Computer Science and Software Engineering – 2017. **[6]**

--This paper showed the overview of rice crop yield prediction.

--Different atmosphere and harvest parameters - Precipitation, least temperature, normal temperature, most extreme temperature, reference trim evapotranspiration, range.

-- The raw data set has been collected and preprocessed, then using WEKA tool prediction model is built and analyzing the outcomes.

Md. Tahmid Shakoor, Karishma Rahman, Sumaiya Nasrin Rayta,"**Agricultural production output prediction using Supervised Machine Learning techniques**" 1st International Conference on MauritiusNext Generation Computing Applications (NextComp), 2017 IEEE. **[7]**

--The researchers suggested area based beneficial crop rank before the cultivation process using Decision Tree Learning-ID3 (Iterative Dichotomiser 3) and K-Nearest Neighbors Regression algorithms.

--Crops' yield per hectare (M. Ton), average of minimum and maximum temperature, rainfall, year range, and region were the parameters that had been used.

Jharna Majumdar* , Sneha Naraseeyappa and Shilpa Ankalaki " **Analysis of agriculture data using data mining techniques: application of big data**" Journal of Big Data Published in Springer Open- 2017 [8]

--This paper focused on the analysis of the agriculture data and finding optimal parameters to maximize the crop production

--PAM, CLARA, DBSCAN and Multiple Linear Regression

--Mining the large amount of existing crop, soil and climatic data, and analyzing new, non-experimental data optimized the production.

Anjela Diana Corraya, Sonia Corraya "**Regression based Price and Yield Prediction of Agricultural Crop**" International Journal of Computer Applications (0975 - 8887) Volume 152 - No.5, October 2016. [9]

--The objective of this research work was predicting future price and yield of agricultural crops considering inflation.

--The Weighted Linear Regression model was used which results 78.75% and 83.55% correct prediction for price and yield respectively.

V.Ramesh,K.Ramar"**Classification of agricultural land soils: A data Mining Approach**" Medwell Journals ISSN:1816-9155, 2019. [10]

--A Large set of soil database was collected from soil science and agricultural Dept. and data mining techniques were applied.

--Naïve-Bayes classifier algorithm had been applied

--Results compared with other classifier algorithms like J48, Random forest, Bayesian Network.

G. Yethiraj, **"Applying Data Mining Techniques in the field of agriculture and allied sciences"** International Journal of Business Intelligentsvol. 1, no. 2, December 2019 [11]

It has been reviewed the applications of data mining techniques and found out that there are several algorithms and techniques being applied in agricultural domain.

-- The data mining techniques could be applied to characterize soil data and found that data mining depends on the amount of data used in the process. An increase in dataset size improves accuracy, which may improve the verification of valid patterns compared to standard statistical analysis.

A. A. Raoranne and R. V. Kulkarni, **"Data Mining: An effective tool for estimation in the agricultural sector,"** International Journal of Emerging Trends and Technology in Computer Science, vol. 1, no. 2, pp. 75-79, July-August 2019. [12]

--Raoranne A. A., et al., discussed how data mining can bridge knowledge of the data to crop yield estimation. The study assessed new data mining techniques and was applied to various variables to establish if meaningful relationships can be found.

--He observed that efficient techniques can be developed and analyzed using appropriate data to solve complex agricultural problems using data mining techniques

Hooman Fetanat, Leila Mortazavifar, Narsis Zarshenas "The Application of Data Mining Techniques in Agricultural Science" 2016 . [13]
 --Author applied following technique on Agriculture Data --Classification: classifies a data item into one of several predefined categorical classes -- Regression: maps a data item to a real valued prediction variable -- Clustering: maps a data item into one of several clusters, where clusters are natural groupings of data items based on similarity metrics or probability density models .

      -- Rule generation: extracts classification rules from the data -- Discovering association rules: describes association relationship among different attributes

      --Sequence analysis: models sequential patterns, like time-series analysis. The goal is to model the states of the process generating the sequence or to extract and report deviation and trends over time

VibhaMaduskar and Prof. Yashovardhankelkar "Survey on Data Mining"]International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 2, February 2019) [14]

 --surveyed that the data selection is the data relevant to the analysis is decided and retrieved from the various data locations. Data preprocessing is the process of data cleaning and data integration is done. Data cleaning is also known as data cleansing; in this phase noise data and irrelevant data are removed from the collected data.

      --In Data transformation the selected data is transformed into forms appropriate for the mining procedure. It is the crucial step in which clever techniques are applied to extract potentially useful patterns.

Namita Mirjankar , Smitha Hiremath **"Application of Data Mining In Agriculture Field "International Journal of Computer Engineering and Applications,iCCSTAR-2016,Special Issue,May2018 [15]**

      In this paper, author discussed about the problem of predicting yield production and that can be solved by employing data mining techniques.

      --Consider that sensor data are available for some time back to the past, where the corresponding yield productions have been recorded. All this information form a training set of data which can be exploited to learn how to classify future yield productions, once new sensor data are available

V. Rajeswari[*] and K. Arunesh : "**Analysing Soil Data using Data Mining Classification Techniques"** Indian Journal of Science and Technology, *Vol 9(19), DOI: 10.17485/ijst/2016/v9i19/93873, May 2018 .* **[16]**
      --This paper analyzed about various types of soil and classified the soil using 3 classification techniques - JRip, J48 and Naïve Bayes.
      --Results of these 3 algorithms were compared and concluded that JRip classification algorithm produced better result.
      --JRip classification algorithm gives better result of this dataset and was correctly classified into maximum number of instances comparing with the other two.

Vinayak A. Bharadi, Prachi P. Abhyankar, Ravina S. Patil, Sonal S. Patade , Tejaswini U. Nate, Anaya M. Joshi : "**Analysis And Prediction In Agricultural Data Using DataMining Techniques**" International Journal of Research In Science & Engineering. *Special Issue 7-ICEMTE March 2017* **[17]**
      --Sample dataset had been analyzed and compared with K means, Density-based spatial

clustering of applications with noise (DBSCAN) and Expectation–Maximization (EM)

--Production, Rainfall & Temperature were used as attributes .

--K-means clustering algorithm was adopted as base algorithm. DBSCAN and EM algorithms were also applied to data.

Sjaak Wolfert, Lan Ge, Cor Verdouw, Mar cJeroen Bogaardt : "**Big Data in Smart Farming – A review** " Science Direct – Elsevier. *2017* . **[18]**

--The authors reviewed about big data and data mining techniques which could be applied on agriculture data set and mentioned the idea "smart Farming"

--Data chain which include crops and animals related to agri. had been discussed in this paper and It include all the activities that are needed to manage data for farm management.

Analysis:

By surveying these papers, it is come to know that predictions are done with various parameters. Mainly parameters related to Soil, weather plays major role. Technologies used by the researchers which gives appropriate data based on their requirements. CNN, ANN, RNN, ARIMA, LSTM, SVM and so many Artificial Intelligence and Machine learning algorithms have been used. **Classification** is used to predict the outcome of a given sample when the output variable is in the form of various distinct types. **Regression** is used to predict the outcome of a given sample inputs is in the form of real values. For example, a regression model might process input data to predict the amount of rainfall. **Ensembling** is another type of supervised learning. It means combining the predictions of multiple machine learning models that are individually produce a more accurate prediction on a new sample.

Survey Paper Nos.	Technologies	Parameters	Conclusion
[1]	Ensemble Neural Network (ENN)- for Yield prediction	Meteorological Factors, Environmental Factors	Soil factors have not taken. ENN results(lower error rate) are compared with BPN and Regression Analysis
[2]	Recurrent Neural Network (RNN), Long-Short Term Memory (LSTM), Auto Regressive Integrated Moving Average (ARIMA)	Weather factors, soil classification	Crop price prediction has been done. Without calculating overall planting area, the predicted price will be inaccurate
[3]	SVM classifier, Multi Linear Regression and ANN	Filtered data set of Raw Seed and Soil classification Factors	Using ANN the crop yield is predicted. Less number of inputs involved.
[4]	Chameleon, Random Forest and Regression Techniques	Soil types, temperature , rainfall	Multiple linear regression to forecast the crop yield

[5]	ANN,CNN, Back Propagation	Satellite images, past crop production data	Rule based system to predict the crop yield production
[6]	Multiple linear regression, Bayesian Network	Precipitation, least and high temperature	Survey results conclude that many data mining techniques have been applied.
[7]	Decision Tree Learning, K-NN Regression	Area(Hectares), Weather factors	Area based crop prediction results
[8]	Partition around medoids (PAM), KNN, DBSCAN	Soil Type, Rainfall range	This model produces optimal temperature for different crops and prediction for the current scenarios.
[9]	Auto Regressive Moving Average Model, Linear regression	Rainfall, Past price range	The Weighted Linear Regression model
[10]	Data Mining and classification techniques	Soil types	Soil classification done using data mining techniques and framed the patterns which will helpful for making the soil profiles
[11]	k-means. KNN, ANN, SVM	Climate and environment factors	Initial step out applications test on agricultural field which performs crop as well as animal farm prediction.
[12]	Decision Tree, ANN, SVM	Cropped area, past crop production data	Classification and clustering techniques involved for solving agricultural crops estimation problems
[13]	Regression and clustering techniques	Growth factors of ornamental plants.	Based on the clusters , many useful patterns have been made using regression techniques.
[15]	Classification and other Data Mining Techniques	past crop production data, pesticide data	Predicting yield production and assist with suitable pesticide
[16]	JRIP, Naïve Bayes	Soil data	Classifiers applied on soil types and extract the required information
[17]	K means, DBSCAN, Expectation–Maximization (EM)	Precipitation, Soil data	Creating useful patterns for optimizing the prediction value
[18]	Smart sensing and analyzing, smart control	Agricultural farming data	Smart way to manage the agricultural activities using big data techniques.

Conclusion: From the survey it is clear to understand that the parameters involved in each paper will impact the accuracy of prediction values. This paper describes the role of machine learning and deep learning techniques in prediction on agriculture factors. A comprehensive prediction model has to be created which predicts the rainfall and water availability, soil classification, crop yield prediction that based on acreage and

Crop market value prediction which base on predicted values such as rainfall, soil classification and crop productivity. So supervised machine learning techniques, efficient classification algorithm such as Naïve-Bayesian, Deep learning techniques such Artificial Neural Network(ANN) and Recurrent Neural Network (RNN) will be the precise algorithms which will produce more accuracy on prediction values.

References:

[1]. Hsu-Yang Kung, Ting-Huan Kuo, Chi-Hua Chen, Pei-Yu Tsai **"Accuracy Analysis Mechanism for Agriculture Data Using the Ensemble Neural Network Method"** www.mdpi.com/journal/ sustainability. Published: 1 August 2016.

[2]. Thayakaran Selvanayagam, Suganya S, Puvipavan Palendrarajah **"Agro-Genius: Crop Prediction using Machine Learning" -International Journal of Innovative Science and Research Technology- 2019**

[3]. N.L. Chourasiya, P. Modi , N. Shaikh3, D. Khandagale, S. Pawar **"Crop Prediction using Machine Learning" IOSR Journal of Engineering (IOSR JEN)- 2019**

[4].Sneha N, Dr. Jharna Majumdar :"**Big Data Application in Agriculture to Maximize the Rice Yield Crop Production using Data Mining Techniques**" International Journal of Innovative Research in Computer and Communication Engineering. *Vol. 5, Issue 5, May 2017.*

[5].Teresa Priyanka,Pratishtha Soni,C. Malathy **"Agricultural Crop Yield Prediction Using Artificial Intelligence and Satellite Imagery" - Eurasian Journal of Analytical Chemistry 2019**

[6].Dakshayini Patil , Dr. M .S Shirdhonkar **"Rice Crop Yield Prediction using Data Mining Techniques: An Overview** " International Journal of Advanced Research in Computer Science and Software Engineering – 2017.

[7]. Md. Tahmid Shakoor, Karishma Rahman, Sumaiya Nasrin Rayta,"**Agricultural production output prediction using Supervised Machine Learning techniques"** 1st International Conference on MauritiusNext Generation Computing Applications (NextComp), 2017 IEEE.

[8].Jharna Majumdar* , Sneha Naraseeyappa and Shilpa Ankalaki " **Analysis of agriculture data using data mining techniques: application of big data**" Journal of Big Data Published in Springer Open- 2017

[9].Anjela Diana Corraya, Sonia Corraya "**Regression based Price and Yield Prediction of Agricultural Crop**" **International Journal of Computer Applications (0975 - 8887) Volume 152 - No.5, October 2016.**

[10].V.Ramesh,K.Ramar"**Classification of agricultural land soils: A data Mining Approach**" **Medwell Journals ISSN:1816-9155, 2018.**

[11]. G. Yethiraj**, "Applying Data Mining Techniques in the field of agriculture and allied sciences"** International Journal of Business Intelligentsvol. 1, no. 2, December 2019

[12]. A. A. Raoranne and R. V. Kulkarni, **"Data Mining: An effective tool for estimation in the agricultural sector," International Journal of Emerging Trends and Technologyin Computer Science, vol. 1, no. 2, pp. 75-79, July-August 2019.**

[13]. Hooman Fetanat, Leila Mortazavifar, Narsis Zarshenas "The Application of Data Mining Techniques in Agricultural Science" 2016

[14]. VibhaMaduskar and Prof. yashovardhankelkar**"Survey on Data Mining"]International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 2, February 2019)**

[15]. Namita Mirjankar , Smitha Hiremath **"Application of Data Mining In Agriculture Field "International Journal of Computer Engineering and Applications,iCCSTAR-2016,Special Issue,May2018**

[16]. V. Rajeswari[*] and K. Arunesh : "**Analysing Soil Data using Data Mining Classification Techniques"** Indian Journal of Science and Technology, *Vol 9(19), DOI: 10.17485/ijst/2016/v9i19/93873, May 2018*

[17]. Vinayak A. Bharadi, Prachi P. Abhyankar, Ravina S. Patil, Sonal S. Patade , Tejaswini U. Nate, Anaya M. Joshi : "**Analysis And Prediction In Agricultural Data Using DataMining Techniques"** International Journal of Research In Science & Engineering. *Special Issue 7-ICEMTE March 2017*

[18]. Sjaak Wolfert, Lan Ge, Cor Verdouw, Mar cJeroen Bogaardt : "**Big Data in Smart Farming – A review** " Science Direct – Elsevier. *2017*