Survey on prediction of crop yield by using different techniques

1st Shruti Rajendra Kudagi

Computer Department
BV(DU)COE
Pune, India

2nd Prof. Dr.S. H.Patil

Computer Department
BV(DU)COE
Pune, India

3rd Prof. Dr.Mrs. Sunita Dhotre

Computer Department
BV(DU)COE
Pune, India

Abstract— The Indian economy is mostly based on agricultural occupation. So the environmental factors like temperature, climate, moisture content, irrigation, soil fertility, minerals, sunlight, and variety of crop can affect the agriculture practices. Due to this the crop yield can be less. Farmer uses conventional methods so that they can judge the crop yield. But due to certain climatic changes farmers need to face huge loss. Climates are certain and specific but the weather is unpredictable sometimes this lead to loss in agriculture. Each type of crops need different environmental conditions this is threat to farmer. Machine learning is mainly used for predictive modeling based on existing dataset. In this paper tried to focus on survey on different papers which uses machine learning techniques to predict crop yield. Which helps us to select proper algorithms so that the model can predict crop yield with accuracy and better result.

Keywords— Machine learning, Predictive Model.

I. INTRODUCTION

Sugar factories wants to know the calculated yield of sugarcane in each season for their business planning. Every year the field surveying people of sugar factory collect the data of each plot by visiting there land .But this done conventionally.The survey people use there methods and formulas to estimate sugar production for the year and pricing is done. This helps the sugar factories to keep the historical data of farmers land and on basis of this the future yield prediction can take place.The main problem with the conventional method is human errors. This bridge the gap between the actual value and false value. Various new technologies and methods are used to calculate the crop yield prediction. In this paper , we have done the survey which will help us to analyse the best algorithm so that it fits into solution.We have survey on machine learning techniques that can help us to use algorithm which will provide better accuracy .The main motto of this survey to select best algorithm that would predict the crop yield .

In this paper we surveyed the sugarcane yield predictive methods by studying different papers. Algorithms like random forest classification and gradient boosting tree classification, KNN, ANN, are mainly used . Each algorithm have there own advantage and disadvantage on basis of problem they are going to work on . considering plot features for predicting crop yield temperature, soil attributes , irrigation all matters deeply. The contribution of survey done tends to give machine learning approach so that model can be built to predict the sugarcane yield for plot with accuracy.

II. MOTIVATION

Many farmers are facing the problem of crop yield estimation. There are certain things are made manually , nut the prediction of the crop yield on previous dataset doing manually can result in human error. Machine learning the branch of AI which is one the important branch for predicting the results in accurate manner by using certain machine learning algorithms. In this paper we have used supervied machine learning algorithm. Due to different algorithms the predicting time required is more than the manual estimation.The cost of predicting yield is been reduce. Due to human intervention is not needed on large scale. Managing the dataset from historical records plays important role in predicting the crop yield prediction.

III. RELATED WORK

A broad literature surve has been done considering all different crops .Each paper focus on different attributes , location , weather and crop type. In this paper further research is going to done on sugarcane yield prediction . Sugarcane is basically selected due to its contribution in agriculture in India is more. The survey will help over all the algorithms which can be used for different crop yield prediction . Many of the paper consist of same parameters which actual becomes the base and algorithms used are different according to problem. This actually changes the result between the algorithms.

. Further studying the papers , each represent there contribution towards development to crop yield prediction techniques.[1] In this paper , algorithm used in this is Random forest algorithm . parameters like soil , weather, has been consider .Research mainly focus on Tamil Nadu crops.[2]Climatic parameters are given main focus which uses C4.5 algorithms to calculate yield .Crops like soyabean, Maize, Paddy crops. Location considered is Madya Pradesh. [3]In this paper satellite derived precipitation and soil characteristics are considered and pre-season prediction is done . Crops like soybean and maize yield predicted by using Neural network.

[4]This paper is related to our research which focus om sugarcane yield grade prediction . This uses Random forest and Gradient Boosting tree technique. In which Random forest gives better result. Research is carried out in Thailand.[5]Mainly focus on crops in Tamil nadu which uses K-means and modified KNN algorithm for crop yield prediction. [6]Paper focuses on NDVI dataset to predict maize yield .Algorithms like SVM and Gaussian Process Regression , Boosted Regression tree algorithm are used . Research have been done in Iran

[7] The paper focus on Rice crop yield prediction using SVM. But the paper concludes that

the SVM has many errors as compared to other algorithms.[8] The papers focus on bitter melon crop yield prediction using CNN . Images of crop plays important role to identify quality of crop.[9] different algorithms are analysed out of which errors like root mean square error are studied in this paper.[10] Mainly focus on Groundnut crop yield using KNN , where perform good out of other machine learning approach.[11] focus on sugarcane yield prediction using Support vector Regression to predict NDVI . They have done on real time data considering NDVI and good accuracy is achieved. [12] Production output is predicted using Decision and KNN algorithm. They have six major crops of Bangladesh. [13] crop yield prediction is done using multiple linear regression and soil parameters , crop rotation is also considered. [14]Different machine learning algorithms are used to predict wheat yield prediction. NDVI time series have been more focused . The boosted Regression tree algorithm and SVM is used in this.[15]Mainly focus on sorghum yield which uses linear regression and CNN on basis of images of sorghum. All papers have certain drawbacks and futures scope which will be discussed in table below.

IV. TABULAR LITERATURE SURVEY

Table. IV.1

Sr. No	Proposed model	Merits	Demerits	Future Scope
[1]	Maximum yield prediction using random forest	It achieve a largest number of crop yield models with a lowest model	only few crops are selected	Improvement in algorithm
[2]	Crop Advisor" tool is implemented for prediction of yield on basis of climatic factors	User friendly web based interactive software which predicts Yield	Only climatic factors are considered On selected location and crop in MP are considered	Improvement in parameters can increase the yield
[3]	System predicts crop yield using satellite derived precipitates ,Soil features, climatic data	Forecasting yield before beginning of season of crop. Fewer data requirement compare to others	Only for selected location and crop High cost	Improved in data set
[4]	Predict sugarcane yield on individual plot	Improved yield	Need more feature extraction to increase more accuracy	Prediction accuracy can be improved by dataset and feature selection
[5]	Yield prediction on major crops grown in Tamil nadu	Crop yield KNN algorithm gives best result that other two	Data set is not recent	Accuracy in algorithm can be done
[6]	Crop yield prediction using remotely sensed data	GPR out performs better than SVR	Images and data set are limited	More crops can be consider and with different data size
[7]	Prediction of rice yield using SMO classifier. Where SVM is used	Crop yield for rice is more accurate	SMO classifier gives low accuracy Than other like Multilayer perceptron.	Naïve bayes can be used as compared to SVM for implementation
[8]	Prediction of yield using leaf images	No clear results are mentioned	No clear results are mentioned	By computer vision test images of the leaves
[9]	Yield prediction using sensors	Automated system detection and alert system	Completely depend on hardware	Algorithm can be improved
[10]	Yield prediction using Biotic factor	Accuracy is achieved by using algorithm	Only groundnut crop is consider for yield prediction.	More accuracy in algorithm can be done
[11]	Crop yield using SVM algorithm for sugarcane	Accuracy is achieved is 83% in sugarcane yield prediction .SVM outperforms better than others.	Restricted to sugarcane crop only and karanatak region is considered	Improve esemble learning techniques and algorithms
[12]	Proposed model compares two algorithms KNN and decision tree	Six major crops in 10 cities are considered with fixed dataset.	Accuracy failed to meet the accuracy	More algorithms can be used to improve the performance and better precision
[13]	Proposed model compares two algorithm BRT and SVM by considering NDVI	NDVI shows accurate majors in critical situation in wheat crop. BRT paly important role among two.	On wheat crop is consider.	More work can be done for different season
[14]	By using CNN and Linear regression crop yield is predicted Mainly focused on sorghum crop	Accuracy achieved is around 74.5 %. Images plays important role in estimataing weight	Data set consider are very limited	Increasing data set to give more accuracy

V. METHODOLGY

In this paper we have done a survey on different methods given by author who have used different algorithms for prediction of crop yield. Various crop yield are identified . In this paper we have introduce the proposed model with some implementation work done till now.

I. Proposed model

a. Algorithms used

The below section gives a brief idea about proposed idea and algorithm used. We have selected machine learning methods used for sugarcane yield prediction system by selecting best algorithm among three algorithm to be used. For this previous existing dataset are used by selecting some predefined attributes . The output expected is crop yield prediction. By using three different model, In this paper we have introduce initial work further can be done in next paper.

1. Random forest algorithm

Random forest is one of the prediction method which can be used for the regression and classification work . Mainly Random forest is used where esembling technique can be introduce which can be used to increase accuracy . In this dependent varianle and independent variable are used the attribute as independent or dependent . Multiple independent attribute is called as multilinear regression properly to determine relation . Before going in to detail we need to study decision tree by selecting different attributes. Decision tree algorithm is technique to classify large data and determine patterns among it.

2. K- nearest neighbor

This algorithm is based on proximity which includes near values. It includes close points, proximity, distance is also considered. It just club together similar data point which are near to each other. Selecting K is one one important task by calculating previous distances.

3. Naïve Bayes

It is collection of different algorithm . Where it uses Bayes theorem which work on probability. It completely work on indepent assumptions , which in practical could not give perfect results. The dataset are mainly split into feature matrix and response vector. The feature matrix is complete collection of all dependent variable. The response vector consist of class variables.

II. Proposed flow

This consist of various modules Like
   a. New registration and Login
   b. Form filling for crop yield prediction
   c. Rainfall , soil attribute selecting .
   d. Dataset
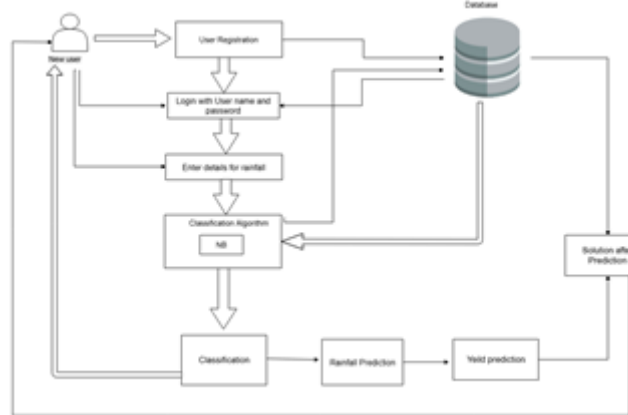   e. Result with crop yield prediction value.

Fig. V. II

III. Implementation

At this time we have done initial work which includes
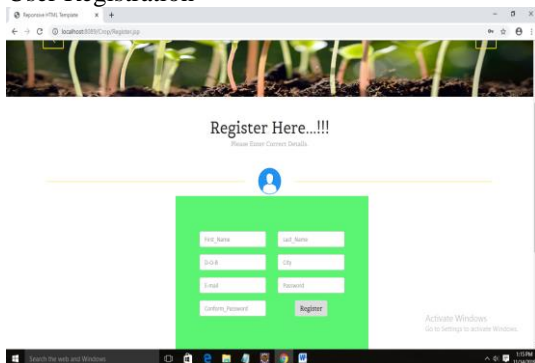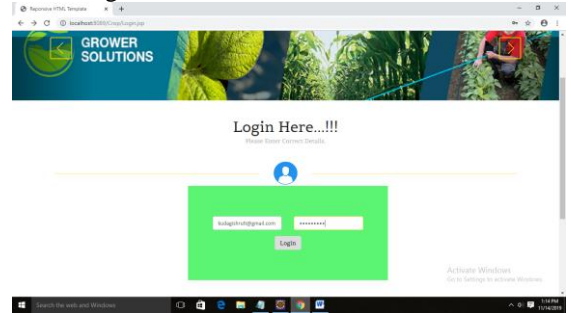   a. User Registration

Fig. V.II. a
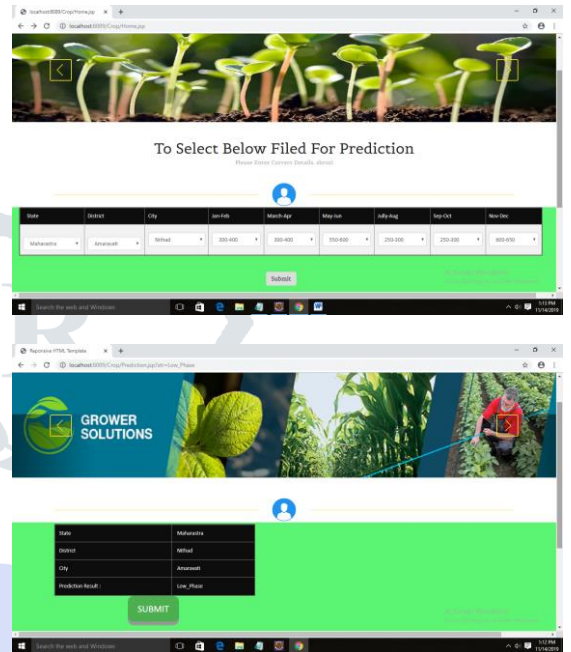
b. Home login

Fig. V. II. b

c. Rainfall Prediction

Fig. V.II. c

IV. Conclusion

The papers concludes literature survey on various papers which mainly falls in crop yield prediction . Every papers introduces new methodology which gives idea to identify research gap and give way to introduce new methodology. According to survey we are going to use Random forest, KNN and naïve Bayes classification techniques of machine learning for predicting crop yield. Sugarcane is the crop is given more focus. Further enhancement can be discuss in next paper.

References

[1] P.Priya et.al ,"Predicting yield ,of the crop using machinelearning algorithm", IJESRT, 7(4): April

[2] S.Veenadhari et.al,"Machine learning approach for forecasting crop yield based on climatic parameter" ,ICCCI-2014.

[3] Ignor Oliveira et.al,"A Scalable Machine learning system for Preseason Agriculture yield forecast", 2018 IEEE 14 International conference on e –Science.

[4] Phusania Charoen – Ung et.al,"sugarcane yield grade Prediction using Random Forest and Gradient Boosting Tree Techniques", JCSSE, 2018.

[5] Mr A Suresh et.al,"Prediction of major crop yield of Tamil Nadu using K means and modified KNN", ICCES 2018.

[6] Hossein Aghighi et.al ."Machine learning Regression technique for the silage Maize Prediction using Time series of Landset 8 OLI", 2018 IEEE.

[7] Niketa Gandhi et.al,"Rice crop yield prediction in India using Support Vector Machines", IEEE

[8] Marizel B. Villanueva et.al, "Bitter melon crop yield prediction using machine learning algorithm", IJACSA, Volume 9,No. 3, 2018

[9] Kodimalar Palanivel et.al ,"An approach for prediction of crop yield using machine learning and Big data techniques", IJCET, Volume 10, Issue 3, May 2019.

[10]Vinita Shah et.al ,"Ground crop yield Prediction using Machine learning Techniques", 2018

[11] Ramesh A. Medar at.al", Sugarcane crop yield forecasting model using supervised machine learning ", 2019 IJISA, 2019,8 11-20.

[12]Md. Tamid Shakoor et.al,"Agriculture production output prediction using supervised Machine learning technique",2017 IEEE

[13] Michiel Stas, Jos Van et.al, "A comparison of machine learning algorithms for regional wheat yield prediction using NDVI time series of SPOT – VGT", 2018 IEEE J-STARS

[14] Judicael Geraud N. ZANNOU,"Sorgham yield prediction using machine learning ",13 June 2019, IEEE.

[15]https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761

[16] https://www.geeksforgeeks.org/naive-bayes-classifiers/