International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 6, June 2020 ||

Crop Selection and Crop Yield Prediction based on Weather Forecast using Machine Learning

Dr. Kamalakshi Naganna[1], Swapna D[2], Tejaswini C A[2], Rashmi Surya[2], Sushmitha M[2]

H.O.D, Dept. of CSE, Sapthagiri College of Engineering, Visvesvaraya Technological University, Bangalore, Karnataka, India[1]
UG Students, Dept. of CSE, Sapthagiri College of Engineering, Visvesvaraya Technological University, Bangalore, Karnataka, India[2]

ABSTRACT: Agriculture is the mainstay of a developing economy like India. In India agriculture provides employment to large population which lives in villages and fully dependent on it to fulfil their livelihood. In the last decade, an incredible growth has been occurred in the use of electronic infrastructure for agricultural processes and development in India. In the recent time, an impressive development in information technology and electronic infrastructure has brought its usage well in the capacity of common people. The affordability of electronic equipment's and IT based applications resulted to a tremendous development in agriculture sector. Machine learning is an imminent field of computer science which can be applied to the farming sector quite effectively. It can facilitate the up-gradation of conventional farming techniques in the most cost-friendly approach.

KEYWORDS: Decision Tree algorithm; Linear Regression algorithm; Crop selection; Crop Yield prediction

I. INTRODUCTION

Agriculture is the backbone of every economy. In a country like India, which has ever increasing demand of food due to rising population, advances in agriculture sector are required to meet the needs. To add to it, the present economic conditions and government policies of India are such that it necessitates the adoption of Precision farming or smart farming. It will enable the farmers to maximize their crop yields and minimize the input costs as well as the losses due to reasons like uncertain rainfall, droughts etc. The agriculture sector needs a huge up gradation in order to survive the changing conditions of Indian economy. Along with the advances in machines and technologies used in farming, useful and accurate information about different matters also plays a significant role in it. This information is being gathered by the use of remote sensors, satellite images, surveys etc. This information along with the knowledge of subject experts and researchers should be readily available to the farmers in order to exploit its potential worth. Also, as the amount of such information is increasing gradually, there is a direct need to analyse it to extract useful facts and patterns. This is where computer science and technology come into the picture. Many algorithms have been proposed for this reason over time which has yielded good results.

II. RELATED WORK

In [1] states the requirements and planning needed for developing a software model for precision farming is discussed. It deeply studies the basics of precision farming. The authors start from the basics of precision farming and move towards developing a model that would support it. This paper describes a model that applies Precision Agriculture (PA) principles to small, open farms at the individual farmer and crop level, to affect a degree of control over variability. The comprehensive objective of the model is to deliver direct advisory services to even the smallest farmer at the level of his/her smallest plot of crop, using the most accessible technologies. This model has been designed for the scenario in Kerala State where the average holding size is much lower than most of India. Hence this model can be positioned elsewhere in India only with some modifications.

[2] Makes a qualified cogitation of assortment algorithms and their performance in yield prediction in precision husbandry. These algorithms are implemented in a data set collected for several years in yield prediction on soya bean crop. The algorithms used for yield prediction in this paper are Support Vector Machine, Random Forest, Neural Network, REP Tree, Bagging, and Bayes. The conclusion drawn at the end is that bagging is the best algorithm for yield prediction among the above stated algorithms since the error deviation in bagging is minimum with a mean absolute error of 18985.

[3] Shows the importance of crop selection and the factors deciding the crop selection like production rate, market price and government policies are discussed. This paper proposes a Crop Selection Method (CSM) which solves the crop selection problem and improves net yield rate of the crop. It suggests a series of crop to be selected over a season considering factors like weather, soil type, water density, crop type. The predicted value of influential parameters

determines the accuracy of CSM. Hence there is a need to include a prediction method with improved accuracy and performance.

[4] Aims to solve the crucial problem of selecting the classifiers for the ensemble learning. A method to select a best classifier set from a pool of classifiers has been proposed. The proposal aims to achieve higher accuracy and performance. A method called SAD was proposed based on accuracy and classification performance. Using Q statistics, the dependency between most relevant and accurate classifiers is identified. The classifiers which were not chosen were combined to form the ensemble. This measure is supposed to ensure higher performance and diversity of the ensemble. Various methods such as SA (Selection by Accuracy), SAD (Selection by accuracy and Diversity) and NS (No selection) algorithm were identified. Finally it is inferred that SAD works better than others.

[5] Proposes various classification methods to classify the liver disease data set. The paper emphasizes the need for accuracy because it depends on the dataset and the learning algorithm. Classification algorithms such as Naïve Bayes, ANN, Zero R and VFI were used to classify these diseases and compare the effectiveness, correction ate among them. The performance of the models where compared with accuracy and computational time. It was concluded that all the classifiers except naive bayes showed improved predictive performance. Multilayer perceptron show the highest accuracy among the proposed algorithms.

[6] Tries to solve the problem of food in security in Egypt. It proposes a framework which would predict the production, and import for that particular year. It uses Artificial Neural Networks along with Multi-layer perceptron in WEKA to build the prediction. At the end of the process we would be able to visualize the amount of production import, need and availability. Therefore It would help to make decisions on whether food has to be further imported or not.

[7] The soil datasets are analyzed and acategory is predicted. From the predicted soil category the crop yield is identified as a Classification rule. Naïve Bayesand KNN algorithms are used for crop yield prediction. The future work stated is to create efficient models us ingvarious classification techniques such as support vector machine, principal component analysis.

III. PROPOSED METHODOLOGY

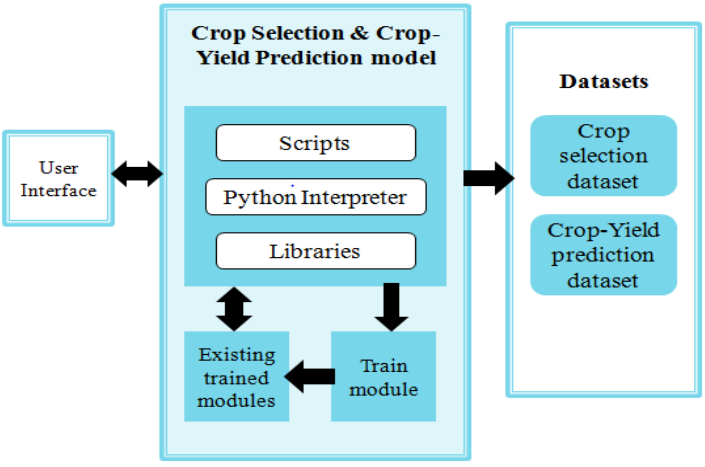

Fig.3.1. System Design

Fig.3.1 shows the System design where the system predicts the values by collecting the datasets of previous data. The proposed methodology contains two phases: Training Phase and Test Phase. In the training phase the data is collected and pre-processed. The pre-processed data is clustered using algorithms. In the testing phase, the yield value is predicted based on the generated rules. On training of data of both algorithms, the values are predicted based on trained data. The various applications of machine learning techniques in agriculture have been listed in this section. These techniques will enhance the productivity of fields along with a reduction in the input efforts of the farmers.

1. **Crop Selection:** Crop Selection is to select the appropriate crop that will be sown plays a vital role. It depends on various factors like the type of soil and its composition, climate, geography of the region etc. Machine learning provides many effective algorithms which can identify the input and output relationship in crop selection Techniques like Decision Trees.

Decision Tree:

A decision tree is a graph that uses a branching method to illustrate every possible outcome of a decision.

Decision trees can be drawn by hand or created with a graphics program or specialized software. Informally, decision trees are useful for focusing discussion when a group must make a decision. Programmatically, they can be used to assign monetary/time or other values to possible outcomes so that decisions can be automated.

2. **Crop Yield Prediction:** Crop Yield Prediction to maximize the crop yield, selection of the appropriate crop that will be sown plays a vital role. It depends on various factors like the type of soil and its composition, state, crop yield, market prices etc. Machine learning provides many effective algorithms which can identify the input and output relationship in yield prediction. Techniques like Linear Regression.

Liner Regression:

Linear regression is a linear approach to modelling the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables). The case of one explanatory variable is called simple linear regression. Linear regression is a basic and commonly used type of predictive analysis. These regression estimates are used to explain the relationship between one dependent variable and one or more independent variables.

IV. ALGORITHM

Crop Selection Algorithm:

Step 1: Start
Step 2: Input: Type of soil, temperature, rainfall as testing dataset.
Step 3: Get the datasets as training dataset.
Step 4: Calculate the gain for training dataset and build decision tree.
Step 5: test input variables with decision tree and get the crop
Step 6: predict the output (crop)
Step 7: Stop

Crop-Yield Prediction Algorithm:

Step 1: Start
Step 2: Get the input
Step 3: merging all the inputs into one value by creating list.
Step 4: From dataset import state, district, crop, area in hectors.
Step 5: we train the model
regr.fit(x_train,y_train) where the training value is stored in regr
Step 6: we then test the model
y_perd=regr.predict(x_test)
my_pred=regr.predict(mydata)
output=int(my_pred)
Step 7: Stop

V. SIMULATION RESULTS

Fig 5.1 shows crop yield prediction home page, user should give input like state, district, crop-for which yield to be predicted and area in hectors. After the input, these inputs are given to model where linear regression gets the input compares with previous year datasets and predicts the result by plotting graph as shown in Fig 5.2. Fig 5.3 shows the resultant output in tons. For crop selection, the home page is shown in Fig 5.4, where the user gives input type of soil, rain and temperature. Taking these as input parameters, it is given to decision tree algorithm. Further the Decision tree algorithm select the crop by taking previous year datasets and user input as shown in Fig 5.5.

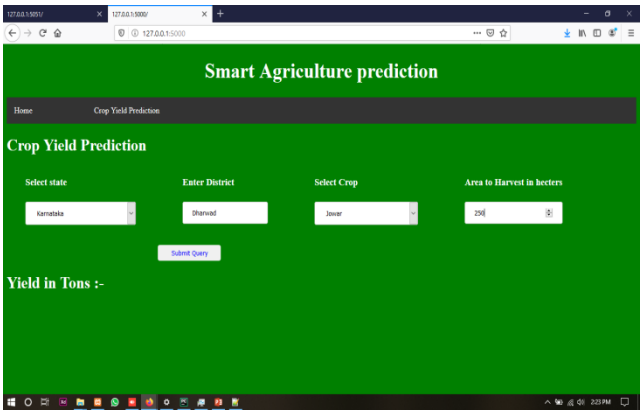

Fig. 5.1. Crop yield prediction home page

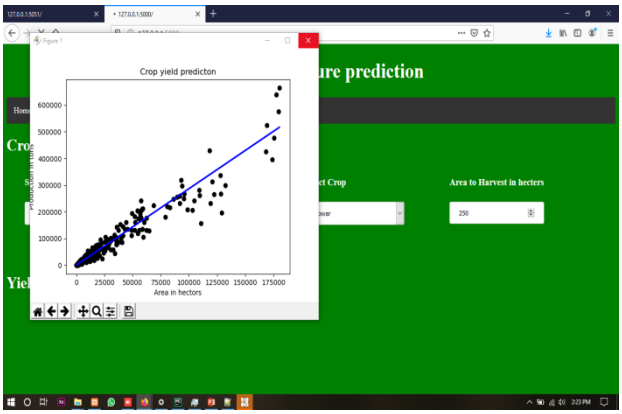

Fig. 5.2. Linear regression graph.

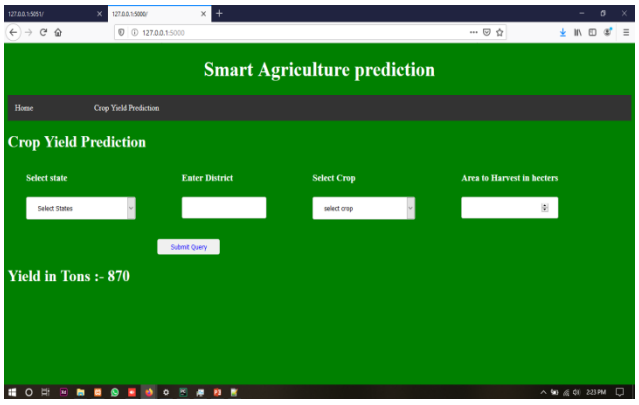

Fig. 5.3. Crop yield prediction in tons.

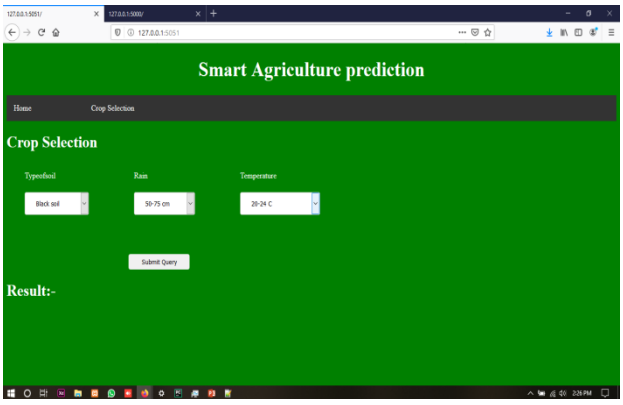

Fig. 5.4. Crop Selection home page.

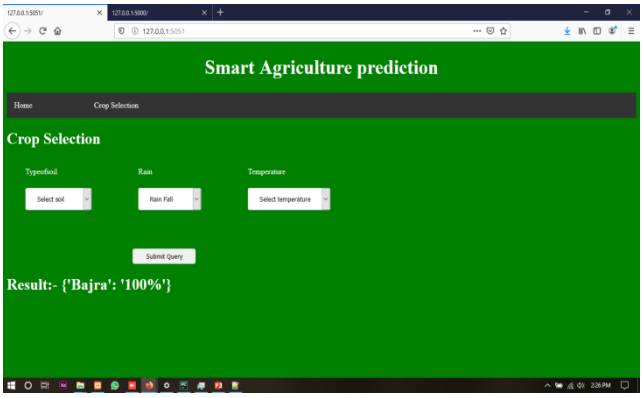

Fig. 5.5. Crop Selected result.

VI. CONCLUSION AND FUTURE WORK

Being dependent on agriculture for a long time, our country has not seen much collaboration between technology and agriculture so far. There are some websites and also a few mobile applications already in use, for agriculture in our country. Scope of the system is to assist the farmer in selecting the crops based on predicted values. The system predicts the values by collecting the datasets of weather condition parameters data and according to the predicted values crop selection and yields are predicted. The system intended to make cropping prediction and cultivation procedures digital. Agricultural Extension of our country also welcomes this initiative and thinks that this will be a standard to be followed in the future, by other countries as well. Our step is very little but we hope that this is the beginning to something big.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 6, June 2020 ||

REFERENCES

1. Washington Okori, Joseph Obua,"Machine Learning Classification Technique for Famine Prediction". Proceedings of the World Congress on Engineering 2011 Vol II WCE 2011, July 6 - 8, London, U.K, 2011.
2. Miss.Snehal, S.Dahikar, Dr.SandeepV.Rode, "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach". International Journal of Innovative Research in Electrical, Electronic, Instrumentation and Control Engineering, Vol. 2, Issue 1, January 2014.
3. Shivnath Ghosh, SantanuKoley, "Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks". International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue: 2, 292 297 ISSN: 2321-8169
4. Kumar, Rakesh, M.p. Singh, Prabhat Kumar, and J.p. Singh. "Crop Selection Method to Maximize Crop Yield Rate Using Machine Learning Technique." 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM) (2015).
5. Y.Radhika and M.Shashi, "Atmospheric Temperature Prediction using Support Vector Machines", International Journal of Computer Theory and Engineering, Vol. 1, No. 1, April 2009.
6. Gupta, Aditya et al. "Need Of Smart Water Systems In India". International Journal of Applied Engineering Research 11.4 (2016): 2216-2223. Print.
7. Gutierrez D. D. (2015). Machine Learning and Data Science: An Introduction to Statistical Learning Methods with R. Basking Ridge, NJ: Technics Publications.
8. Mehra, Lucky K. et al. "Predicting Pre-Planting Risk of StagonosporaNodorum Blotch in Winter Wheat Using Machine Learning Models." Frontiers in Plant Science 7 (2016): 390. PMC. Web. 16 Apr. 2016.
9. Development of a disease risk prediction model for downy mildew (Peronosporasparsa) in boysenberry. Kim KS, Beresford RM, Walter M Phytopathology. 2014 Jan; 104(1):50-6.
10. Rumpf, T., A.-K. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer. "Early Detection and Classification of Plant Diseases with Support Vector Machines Based on Hyperspectral Reflectance." Computers and Electronics in Agriculture 74.1 (2010): 91-99. Web.]
11. M.P.Raj, P,R,Swaminarayan, J.R.Saini, D.K.Parmar "Application Of Pattern Recognition Algorithm In Agriculture : A Review" Int.J.Advanced Networking and Application, vol:6, issue:5, 2015
12. http://farmer.gov.in/mspdet.html
13. Nandyala, Chandra Sukanya, and Haeng-Kon Kim. "Big and Meta Data Management for U-Agriculture Mobile Services." International Journal of Software Engineering and Its Applications IJSEIA 10.2 (2016): 257-70. Web.
14. Bhatt, Navtej and Dr. P.V. Virparia. "A Survey Based Research For Data Mining Techniques To Forecast Water Demand In Irrigation". International Journal of Computer Science and Mobile Applications 3.8 (2015): 14-18. Print.
15. Szepesvári, Csaba. "Algorithms for Reinforcement Learning." Synthesis Lectures on Artificial Intelligence and Machine Learning 4.1 (2010): 1-103. Web.
16. Oladipupo, Taiwo. "Types of Machine Learning Algorithms." New Advances in Machine Learning (2010). Web.
17. Sperschneider, Jana, Donald M. Gardiner, Peter N. Dodds, Francesco Tini, Lorenzo Covarelli, Karam B. Singh, John M. Manners, and Jennifer M. Taylor. "E Ffector P: Predicting Fungal Effector Proteins from Secretomes Using Machine Learning." New Phytologist New Phytol 210.2 (2015): 743-61. Web.
18. Mehta, Parth, Hetasha Shah, Vineet Kori, VivekVikani, Soumya Shukla, and MihirShenoy. "Survey of Unsupervised Machine Learning Algorithms on Precision Agricultural Data." 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS) (2015). Web.