Nat. Volatiles & Essent. Oils, 2021; 8(5): 10531-10539

A Machine Learning model for Crop and Fertilizer recommendation

Manoj Kumar D P [1], Neelam Malyadri[2], Srikanth M S[3], Dr. Ananda Babu J[4]

[1]Assistant Professor, Department of Computer Science & Engineering, Kalpataru Institute of Technology, Tiptur, India.

[2] Assistant Professor, School of Computer Science and Engineering, Reva University, Bangalore ,India.

[3] Assistant Professor, School of Computer Science and Engineering, Reva University, Bangalore ,India.

[4] Associate Professor, Department of Information Science and Engineering, Malnad College of Engineering, Hassan, India.

Abstract –

India is currently the world's second largest producer of several dry fruits, agriculture-based textile raw materials, roots and tuber crops, pulses, farmed fish, eggs, coconut, sugarcane and numerous vegetables. India is ranked under the world's five largest producers of over 80% of agricultural produce items, including many cash crops such as coffee and cotton. Farmers are growing same crop in the season rather than growing different varieties in various seasons, also applying more quantity of fertilizers without knowing actual contents and quantity. So we have designed a recommendation model based on machine learning , describes the best suitable crop to be grown and fertilizer to be seeded depending on soil and weather conditions. Hence by utilizing our system,farmers can grow new crops in different seasons and benefit a better profit, avoid soil pollution.

Keywords **–**Machine Learning, random-forest,SVM,K-nearest neighbor algorithm

1. INTRODUCTION :

Agriculture is the main occupation for the people of India, covering 60% of the nation land and catering the basic needs of 1.2 billion people [1]. For the benefit of the farmers, modernization of agriculture procedures is carried out today. The crop yield or production majorly depends on the weather conditions, environmental changes, rainfall (which at times is uncertain), water management, and the utilization of pesticides. Therefore farmers are not able accomplish expected yield of crop. Now a days data mining, machine learning as well as deep learning approaches are used by various researchers to enhance and improve the yield of crop and their quality[11,12]. Machine Learning can gain proficiency with the machine without characterized computer programming, so it improves machine execution by distinguishing and portraying the consistency and pattern of drive information. In this work various machine learning approaches such as Linear Regression, Gradient Boosting Regressor, Random Forest Regressor, Decision Tree Regressor, Polynomial Regression, Ridge Regression have been used for yield prediction on crop yield dataset of different states and considering varied crops.

The designed system will recommend the most suitable crop for particular land. Based on weather parameter and soil content such as Rainfall, Temperature, Humidity and pH. They are collected from V C Farm Mandya, Government website and weather department. The system takes the required input from the farmers or sensors such as Temperature, Humidity and pH. This all inputs data applies to machine learning predictive algorithms like Support Vector Machine (SVM) [5] and Decision tree [6] to identify the pattern among data and then process it as per input conditions. The system recommends the crop for the farmer and also recommends the amount of nutrients to be add for the predicted crop. The system has

some other specification like displaying approximated yield in q/acre, required seed for cultivation in kg/acre and the market price of the crop

2. RELATED WORK :

In agriculture, Machine Learning is considered as a novel field, as variety of work has been done with the help of machine learning in the field of agriculture. There are different philosophies made and evaluated by the researchers all through the world in the field of agriculture and related sciences.

CH. Vishnu VardhanChowdary, Dr.K.Venkataramana [2], developed id3 algorithm for getting improved and great quality of crop yield of Tomato and is executed in Php platform and datasets are used as csv. Temperature, area, humidity and the production of tomato crop are the different parameters used in this study.

R. Sujatha and P. Isakki [3], utilizes data mining techniques for prediction. This model worked on different parameters such as crop name, land area, soil type, pH value, seed type, water and also foreseen the boom and diseases of plants and in this way empowered to choose the descent crop dependent on climatic data and required parameters.

N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy [4], proposed the SVM for crop yield prediction of rice. In this method, dataset used consists of different parameters such as place, temperature, precipitation and manufacturing. On this dataset, the implemented classifier is sequential minimal optimization. They prepared the dataset through Weka tool to manufacture the set of rules on current dataset. In python, by using SVM algorithm outcomes were produced.

S. Veenadhari, B. Misra and C. Singh [5], have built up an interactive site for finding the influence of climate and production of crop by utilizing c4.5 algorithm called Crop Advisor. Dependent on c4.5 algorithm, decision tree and ruled have been developed. It gives the idea how crop growth is affected by different climatic parameters. The data with respect to the related years environmental parameters like rainfall, temperature where gathered. The choices were dependent on the zone under the picked crop.

Jun Wu, AnastasiyaOlesnikova, Chi- Hwa Song, Won Don Lee [6], proposed selection tree which is fit for grouping all styles of farming records. A decision tree classifier turned into proposed for information of agriculture. It utilises new facts and can address each and in whole record. 10-fold cross validation method is utilised to check dataset, horse-colic and soyabean dataset.

Kiran Mai,C., Murali Krishna, I.V, A.VenugopalReddy [7], explained in their study that how data mining is incorporated with the other farming data such as meteorological data, usage of pesticides are useful for soothing out of use of pesticides. Topical information related to the business of agriculture which has contiguous properties was represented.

Verheyen, K., Adrianens, M. Hermy and S.Deckers [8], explained statistical mining techniques in their study as they are regularly used to view the characteristics of soil. As kmeans is utilized for sectioning soils in blend with GPS based innovation.

3. PROPOSED WORK :

The Proposed system will predict the most suitable crop and fertilizer for particular land based on soil contents and weather parameters such as Temperature, Humidity, soil PH and Rainfall.

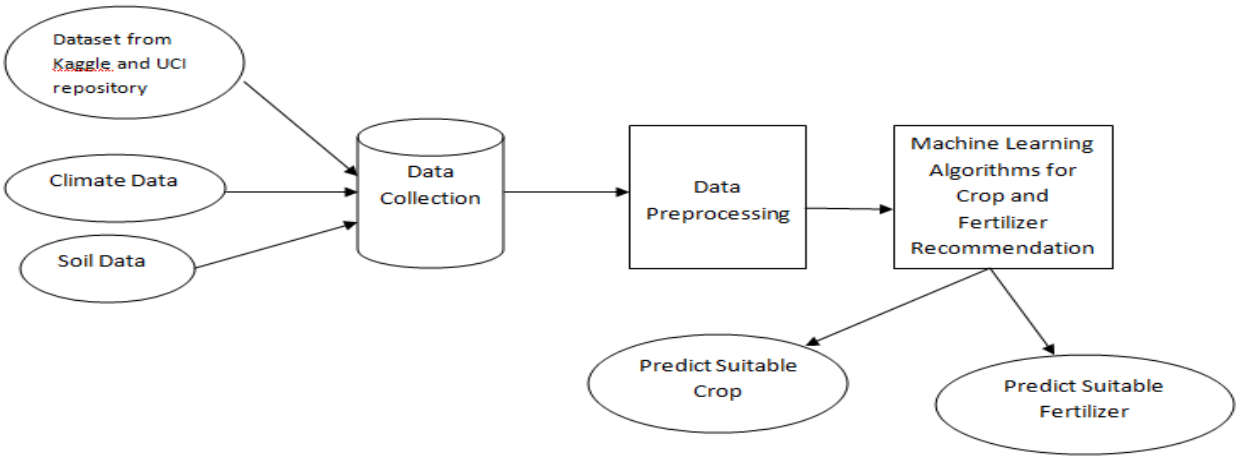

Figure 1: Overall Architecture

The Architecture of the proposed system consists of various blocks as shown in the fig (1) as follows

3.1. **Data Collection**: -

Data collection is the most efficient method for collecting and measure the data from different resources like kaggle and UCI machine learning repository. To get an approximate dataset for the system. This dataset must contain the following attributes i)Soil PH ii) Temperature iii) Humidity iv) Rainfall v) NPK values, those parameters will consider for crop prediction.

Crop Recommendation Dataset can be retrieved using following kaggle dataset link as: https://www.kaggle.com/manojkumardp/crop-recommendation-dataset,

Similarly Fertilizer Recommendation dataset can be retrieved as: https://www.kaggle.com/gdabhishek/fertilizer-prediction are used as training and testing datasets.

3.2. **Data Preprocessing**: -

After collecting datasets from various resources. Dataset must be preprocessing before training to the model. The data preprocessing can be done by various stages, begins with reading the collected dataset the process continues to data cleaning. In data cleaning the datasets contain some redundant attributes, those attributes are not considering for crop prediction. So, we have to drop unwanted attributes and datasets containing some missing values we need to drop these missing values or fill with unwanted nan values in order to get better accuracy. Then define the target for a model. After data cleaning the dataset will be split into training and test set by using sklearn library.

3.3. **Machine Learning Algorithm for Prediction**: -

Machine learning predictive algorithms has highly optimized estimation has to be likely outcome based on trained data. Predictive analytics is the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data. The goal is to go beyond

knowing what has happened to providing a best assessment of what will happen in the future. In our system we used supervised machine learning algorithm having subcategories as classification and regression. Classification algorithm will be most suitable for our system.

To predict suitable crop, we use machine learning algorithms like XGBoost, Random Forest and KNN , whereas for prediction of suitable fertilizer, we use machine learning algorithms like SVM and Random forest.

3.3.1 **Crop Prediction:**

To Predict the particular crop to be grown , we use input parameters like N,P,K temperature, humidity and rainfall. Crop prediction process being with the loading the external crop datasets. Once the dataset read then pre-processing will be done by various stages as discussed in Data Pre-processing section. After the data pre-processing, train the models using KNN, Random Forest classifier into training dataset. . For a prediction of the crop, we consider a various factor such as temperature, humidity, soil PH and predicted rainfall. Those are the input parameter for a system that can be entered by manually or taken from the sensors. Predicted rainfall and input parameter values will be appended in a list.

A Summarized Crop prediction model along with its input parameters are tabularized as follows:

N	P	K	Temparature	Humidity	Rainfall	Crop
90	42	43	20.87	82.00	202	rice
85	58	41	21.77	80.31	226	rice
60	55	44	23.00	82.32	263	rice
74	35	40	26.49	80.15	242	rice

Table-1: Given Input Parameters Values with Output for Crop Prediction

Data Analysis on Crop Dataset can be done on Univariate and BiVariate Analysis, here are the some of the examples of crop analysis done.

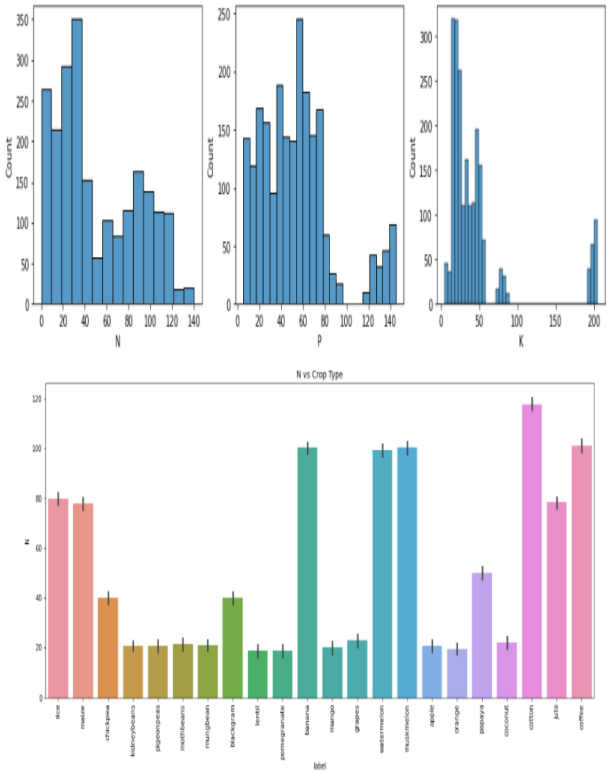

Fig-2: UniVariate and BiVariate Analysis for Parameters N,P,K in Crop prediction

3.3.2 Fertilizer Prediction:

To Predict the particular fertilizer to be used , we use input parameters like N,P,K temperature, humidity,moisture and soil type and also crop to be grown. Fertilizer prediction process being with the loading the external fertilizers datasets. Once the dataset read then pre-processing will be done by various stages as discussed in Data Pre-processing section. After the data pre-processing, train the models usingSVM, Random Forest classifier into training dataset. For a prediction of the fertilizers, we consider a various factor such as temperature, humidity, soil PH and predicted crop to be grown. Those are the input parameter for a system that can be entered by manually or taken from the sensors. Predicted crop and input parameter values will be appended in a list.

A Summarized Fertilizer prediction model along with its input parameters are tabularized as follows:

N	P	K	Temparature	Humidity	Rainfall	Crop
90	42	43	20.87	82.00	202	rice
85	58	41	21.77	80.31	226	rice
60	55	44	23.00	82.32	263	rice
74	35	40	26.49	80.15	242	rice

Table-2: Given Input Parameters Values with Output for Fertilizer Prediction

Data Analysis on Fertilizer Dataset can be done on Univariate and BiVariate Analysis, here are the some of the examples of fertilizer analysis done.

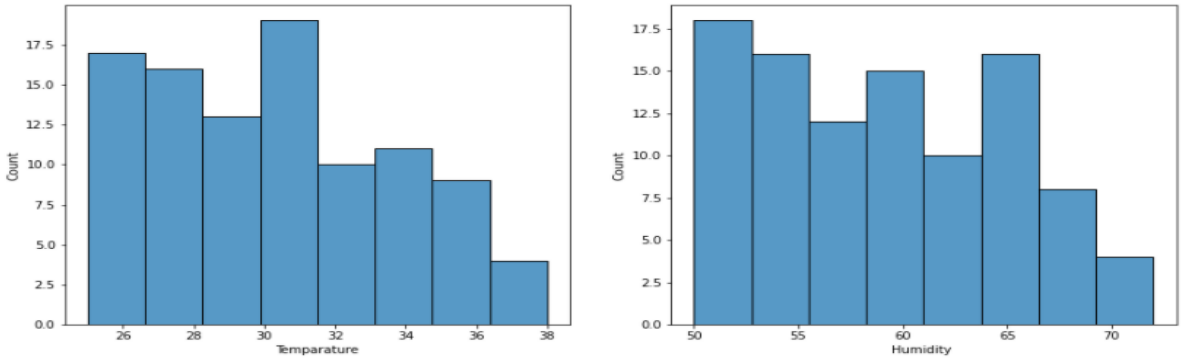

Fig-3: UniVariate and BiVariate Analysis for Parameters N,P,K in Fertilizer prediction

3.4 SUPPORT VECTOR MACHINE:

Support Vector Machine (SVM) is a machine learning algorithm for classification and prediction purposes. In this algorithm, we plot each data item as a point in n-dimensional space with each value as an particular coordinate plotted as a hyper-plane.

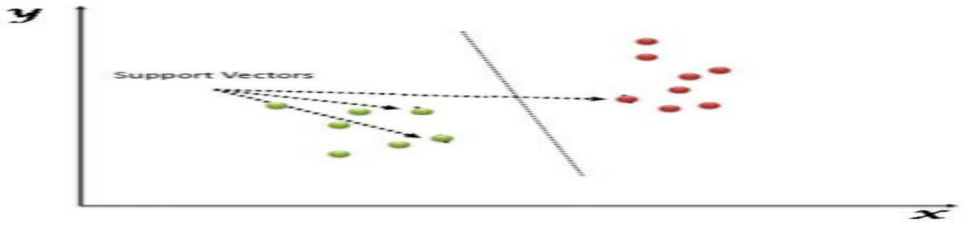

Figure 4: SVM Model

An SVM creates hyperplanes that have the largest margin in a high-dimensional space to separate given data into classes. The margin between the 2 classes represents the longest distance between closest data points of those classes.

STEP1: Select the feature sets from different classes of data

STEP2: Calculate the intersection points of each class of feature and plot, repeat for all the features of data.

STEP3: Remove the features which are intersecting and data of all the classes.

STEP4: Plot the hyper planes for the remaining points.

STEP5: Calculate the distance of the hyper planes in different class of objects.

STEP6: Select the hyper plane which is consistent for each class of data.

4. RESULTS AND DISCUSSION :

The proposed system recommends the best suitable crop for particular land by considering parameters as annual rainfall, temperature, humidity and soil pH. Among these parameters annual rainfall is predicted by system itself by using previous year data with SVM algorithm and other parameters are have to be entered by the user. In the output section the system displays a suitable crop, required seeds/acre, market price and approximate yield of the recommended crop and also the system takes NPK values in the input section to display the required NPK for the recommended crop.

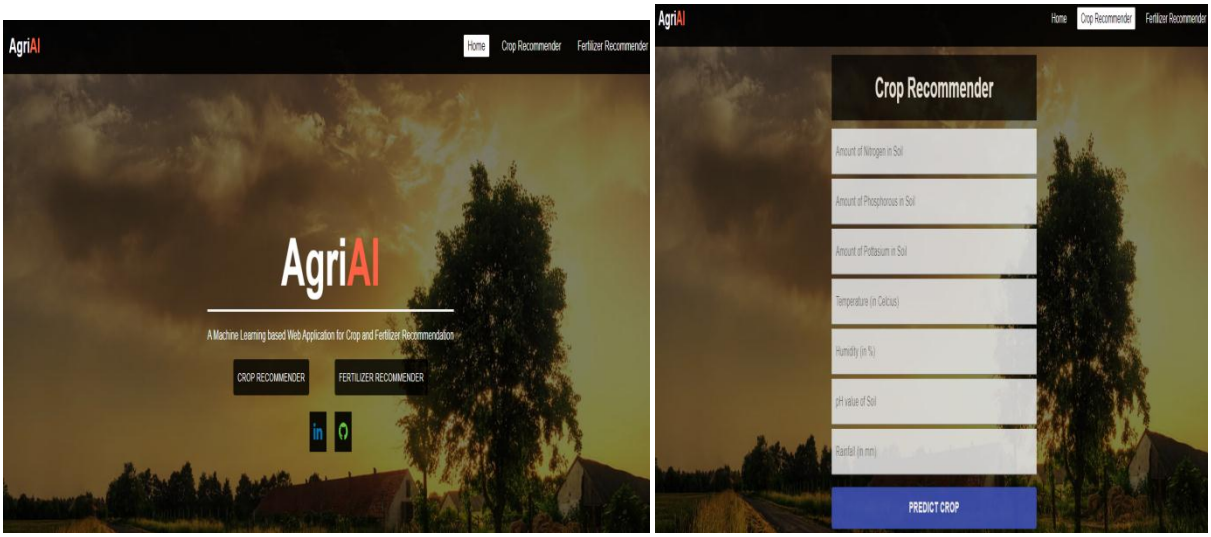

Fig.5: A Home Page displaying Crop and Fertilizer Recommendation

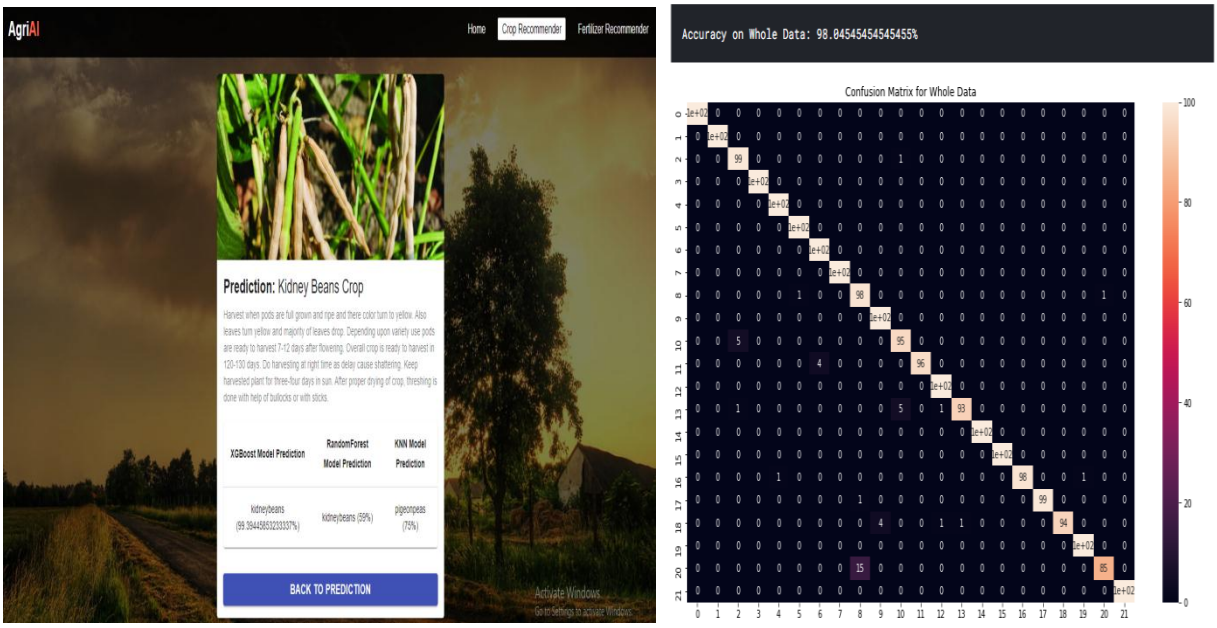

Fig.6: Crop Prediction Settings and Crop Accuracy Result

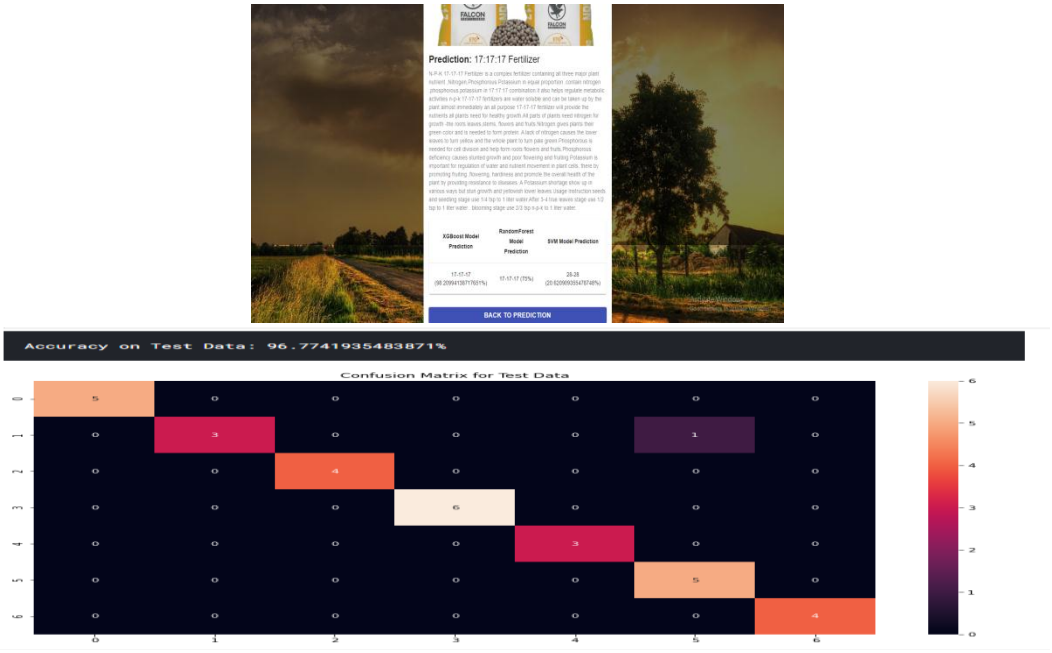

Fig.7:Fertilizer Prediction Settings Accuracy Result

5. CONCLUSION :

Presently our farmers are not effectively using technology and analysis, so there may be a chance of wrong selection of crop for cultivation that will reduce their income. To reduce those type of loses we have developed a farmer friendly system with GUI, that will predict which would be the best suitable crop for particular land and this system will also provide information about required nutrients to add up, required seeds for cultivation, expected yield and market price. So, this makes the farmers to take right decision in selecting the crop for cultivation such that agricultural sector will be developed by innovative idea..

REFERENCES:

[1]https://www.hilarispublisher.com/open-access/agriculture-roleon-indian-economy-2151-6219-1000176.pdf.

[2] CH. Vishnu Vardhanchowdary, Dr.K.Venkataramana, “Tomato Crop Yield Prediction using ID3”, March 2018,IJIRT Volume 4 Issue 10 pp,663-62.

[3] R. Sujatha and P. Isakki, “A study on crop yield forecasting using classification techniques” 2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16), Kovilpatti, 2016, pp. 1-4.

[4] N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy, “Rice crop yield prediction in India using support vector machines” 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), KhonKaen, 2016, pp. 1-5.

[5] S. Veenadhari, B. Misra and C. Singh, “Machine learning approach for forecasting crop yield based on climatic parameters” 2014 International Conference on Computer Communication and Informatics, Coimbatore, 2014, pp. 1-5..

[6] Jun Wu, AnastasiyaOlesnikova, Chi- Hwa Song, Won Don Lee (2009), “The Development and Application of Decision Tree for Agriculture Data” IITSI, pp 16-20

[7] KiranMai,C., Murali Krishna, I.V, an A.VenugopalReddy, “Data Mining o f Geospatial Database for Agriculture Related Application”, Proceedings of Map India,New Delhi, 2006,pp 83-96.

[8] Verheyen, K., Adrianens, M. Hermy and S.Deckers(2001),“High resolution continuous soil classification using morphological soil profile descriptions”Geoderma, 101:31-48.

[9] https://www.kaggle.com/

[10] https://data.gov.in/

[11] P.Priya, U.Muthaiah, M.Balamurugan, “Predicting yield of the crop using machine learning Algorithm”, IJESRT et al., 7(4): April-2018, pp 2277-2284.

[12] Georg Ruß, Rudolf Kruse, Martin Schneider, and Peter Wagner. “Estimation of neural network parameters for wheat yield prediction” In Max Bramer, editor, Artificial Intelligence in Theory and Practice II, volume 276 of IFIP International Federation for Information Processing, pages 109–118. Springer, July 2008.