

# Crop Yield Prediction of Wheat Using Fuzzy C Means Clustering and Neural Network

Akanksha Verma<sup>1</sup>, Aman Jatain<sup>2</sup>, Shalini Bajaj<sup>3</sup>

<sup>1</sup>M.tech Student, Department of Computer Science and Engineering, Amity University, Gurgaon (12001), Haryana, India.

<sup>2</sup>Assistant professor, Department of Computer Science and Engineering, Amity University, Gurgaon (12001), Haryana, India.

<sup>3</sup>Professor, Department of Computer Science and Engineering, Amity University, Gurgaon (12001), Haryana, India.

## Abstract

An efficient framework is required in farming area to anticipate and enhance the product everywhere throughout the world. Due to inaccessibility of appropriate information, the many-sided quality of expectation of products is high by virtue of which nature of forecast gets influenced. In information mining key advance in mining profitable information is grouping. In this paper we proposed trim forecast framework utilizing fuzzy bunching strategies and neural system. Counterfeit neural systems have been shown to be intense apparatuses for expectation and illustrating, to assemble their adequacy. To foresee the reasonable harvest by detecting different parameters like biomass, temperature, rainfall and solar radiation, we utilized approach of yield forecast. Horticultural researchers over the world request the necessity for a compelling part to foresee and upgrade the yield development. The necessity for a fused product improvement control with exact prescient yield organization is exceedingly felt among developing gathering. It deals with a grouping approach over harvest informational collection to foresee the product yield and enhance the basic leadership in accuracy agribusiness. The consequence of the desire models will help agribusiness associations in outfitting agriculturists with beneficial information as to which factors add to high wheat yield. We have evaluated RMSE, PSNR, FAR, FRR factors to analyse the yield prediction.

**Key words:** Crop Yield Prediction, Data Mining Technique, Crop Growth, Precision Agriculture.

## INTRODUCTION

Agribusiness is the foundation of Indian economy. In India, dominant part of the agriculturists is not getting the normal harvest yield because of a few reasons [13]. The agrarian yield is fundamentally relies upon climate conditions like temperature, rainfall, biomass etc. In this unique situation, the agriculturists essentially require an opportune exhortation to anticipate the future yield profitability and an investigation to help the ranchers to amplify harvest creation in their products. Yield expectation is an imperative farming issue. Each rancher is keen on knowing, how much yield he is about to anticipate. Data mining is a notable innovation, creating with database and manmade brainpower [14]. It is a handling suggestion of

activity of removing dependable, novel, helpful and justifiable examples from database. Accomplishing most extreme harvest yield at least cost is one of the objectives of agrarian generation. Early recognition and administration of issues related with trim yield markers can enable increment to return and ensuing benefits [15]. Expectations could be utilized by edit directors to limit misfortunes when ominous conditions may happen. Also, these forecasts could be utilized to amplify trim expectation when potential exists for great developing conditions. Agribusiness now-a-days has turned out to be exceptionally info and cost concentrated zone without sensible utilization of manures and plant assurance measures, horticulture never again stays as beneficial as before on account of vulnerabilities of climate, creation, strategies, costs and so on that regularly prompt misfortunes to the ranchers .

## RELATED WORK

Despite the fact that a great deal of research has been led for building up the choice emotionally supportive network for agriculturists, the greater part of the examination centre around the harvest administration, edit illness administration and product yield estimating. Be that as it may, the agriculturist's harvest determination at the prior stage is a standout amongst the most critical components since proper product choice at the prior stage will help ranchers to enhance edit administration and product yield. A yield forecast is a broad issue that happens.

Ravel et. al [1] discuss about the Knowledge Discovery Process and also a various data mining Techniques such as Association rules, Classification, Clustering, Prediction and Sequential Patterns. Agrawal et. al [2] discuss about different Data Mining apparatuses, for example, Dashboards, Content Mining instruments. They provide a diagram about these instruments and the different situations in which they can be conveyed. Grajales et. al[3] have proposed a web application that utilizes open dataset like local climate conditions, chronicled creation, arrive cover, neighborhood atmosphere conditions, historical production, land cover and the integration method of them so that farmers can easily access that application.

Hemageetha [5] basically concentrates on utilizing the dirt parameters like pH, nitrogen, dampness and so forth for crop yield prediction. The method Apriori calculation is also used

by Fathima et. al [6]. They utilizes the data mining systems on ongoing data that assistance in information revelation. Rub et. al. [7] presents a relative report on the relapse models that could be utilized for foreseeing yield. Raorane et. al [8] also discusses the different data mining systems for enhancing the crop generation in agribusiness. Veenadhari et. al. [9] depicted the motivation behind data mining techniques in the territory of farming.

**PROPOSED METHODOLOGY**

In the proposed approach initially the raw data set is taken and then clustering and classification will be performed in order to get the results in MATLAB. To perform clustering, Fuzzy c means clustering approach is taken.

**Fuzzy C Means Clustering-**

In this research Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1981) is frequently used in pattern recognition. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C \mu_{ij}^m \|x_i - c_j\|^2 \quad (1)$$

$$1 \leq m < \infty$$

Here in equation, *m* is any real number greater than 1, N and C denotes the row and column respectively,  $\mu_{ij}$  is the degree of membership of  $x_i$  in the cluster *j*,  $x_i$  is the *i*<sup>th</sup> of d-dimensional measured data,  $c_j$  is the d-dimension center of the cluster, and  $\|*\|$  is any norm expressing the similarity between any measured data and the centre.

Fuzzy C Means is a précised learning algorithm that provides less error rate probability and arranges the data in the hierarchical manner. Basically it calculates the output in terms of degree of membership that means the data point that have high probability or how one data point is related to other data point and how much correlation will be higher at the time of testing. Fuzzy clustering deals with assigning of data points in such a manner so that it will get the close relationship of similarity between data points as much as possible.

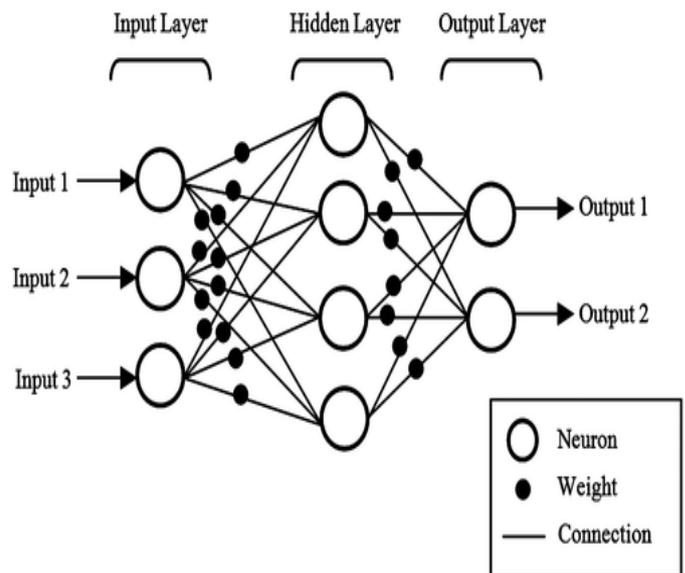
**Neural Network as a Classifier**

In proposed work neural network will be used for the classification purpose. A neural network consists of units (neurons), arranged in layers, which convert an input vector into some output. Each unit takes an input, applies a (often nonlinear) function to it and then passes the output on to the

next layer. Generally the networks are defined to be feed-forward: a unit feeds its output to all the units on the next layer, but there is no feedback to the previous layer. Weightings are applied to the signals passing from one unit to another, and it is these weightings which are tuned in the training phase to adapt a neural network to the particular problem at hand. This is the learning phase.

Neural networks have found application in a wide variety of problems. These range from function representation to pattern recognition-

1. **The Architecture:** The number of layers and the no. of nodes in each of the layers.
2. **The learning mechanism** which has been applied for updating the weights of the connections (Supervised and Unsupervised Learning).
3. **The activation functions** used in various layers. An ANN is having a sequence of layers.



**Figure 3.1.** Neural Network

**GUI to Perform Classification**

A GUI interface is made in MATLAB tool where data set is uploaded in order to perform fuzzy c means clustering.

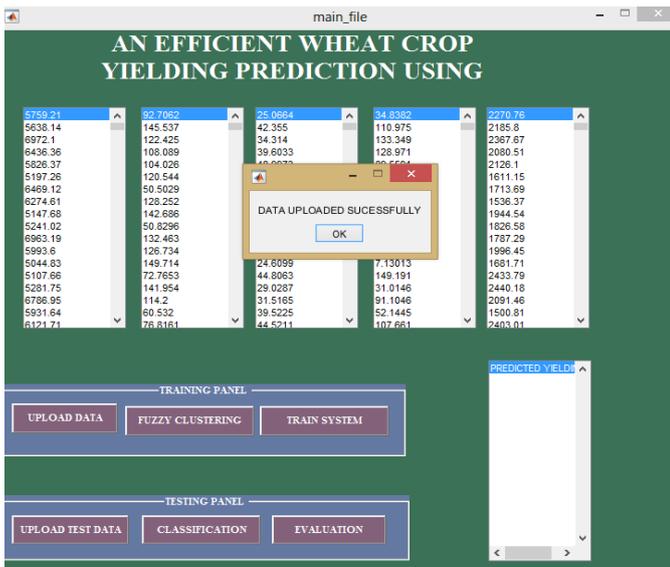


Figure 4.1. A GUI to upload Data-set

Fuzzy C Means algorithm is used for clustering, where it shows (Fig: 4.2) the results in terms of degree of membership means which data points have higher probability or how much similarity is there in between the data points.

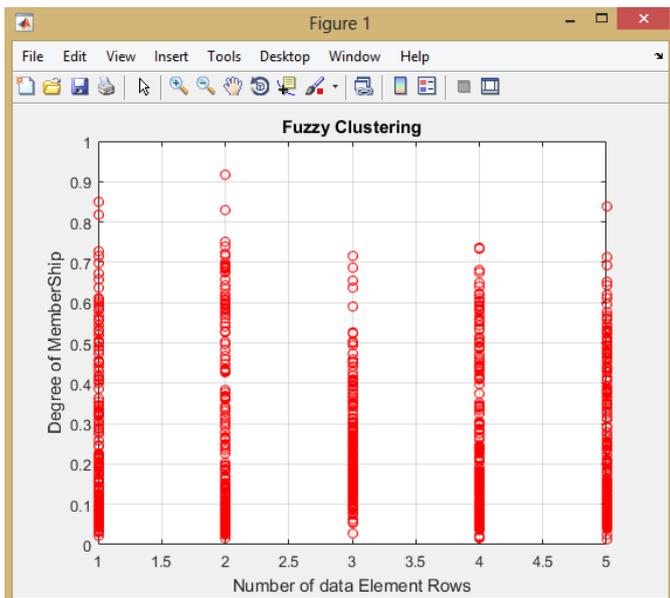


Figure 4.2. Fuzzy C Means clustering (DOM)

Fig 4.3 shows, how neural network is trained with the help of train command already inbuilt in MATLAB and then in fig 4.4 shows the processing of test data to complete the classification.

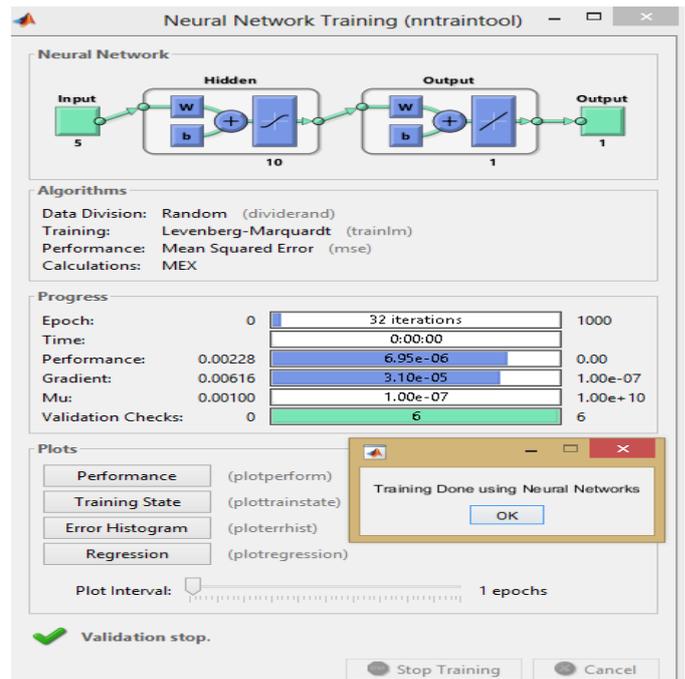


Figure 4.3. Training of Neural Network

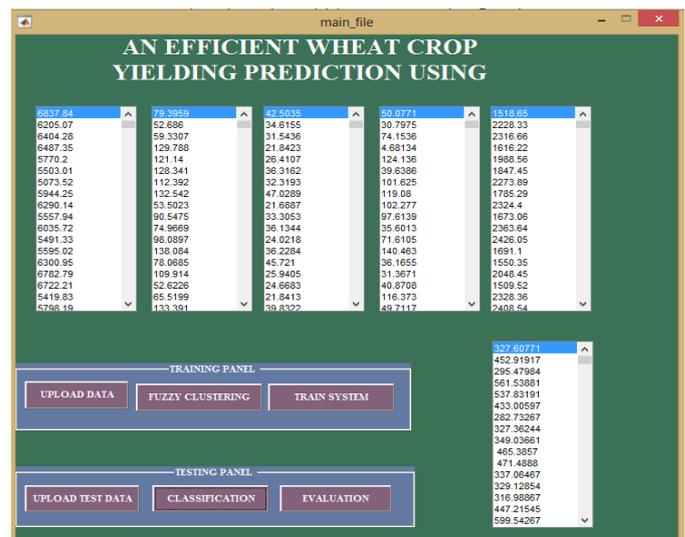


Figure 4.4. Classification performed to predict yield

## DISCUSSION AND RESULTS

Each of the parameter discussed below is calculated and presented in the form of graph as:

1. **RMSE**- Root Mean Square Error tells that how much error rate is taken in predicting the system. It should be less which is a desired output, as shown in Fig. 5.1.
2. **PSNR**- Peak Signal to Noise Ratio is the performance parameter which is calculates to remove the noise and to see to performance of signal. It is inversely proportional to RMSE. So it should be high as shown in Fig. 5.2.

$$PSNR \propto \frac{1}{RMSE} \quad (2)$$

3. **FAR**- False Acceptance Rate show that how many right things are wrongly accepted during prediction of yielding. It should be low as shown in Fig. 5.3.

4. **FRR**- False Rejection Rate shows that how many right things are wrongly rejected during prediction of yielding. It should also be low as shown in Fig. 5.4.

$$FRR = 1 - FAR$$

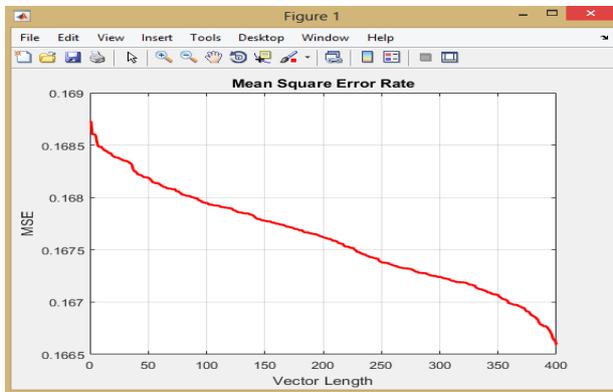


Figure 5.1. Root Mean Square Error

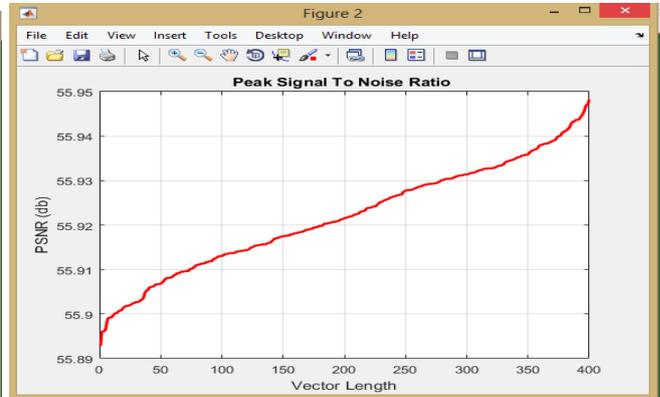


Figure 5.2. Peak Signal to Noise Ratio

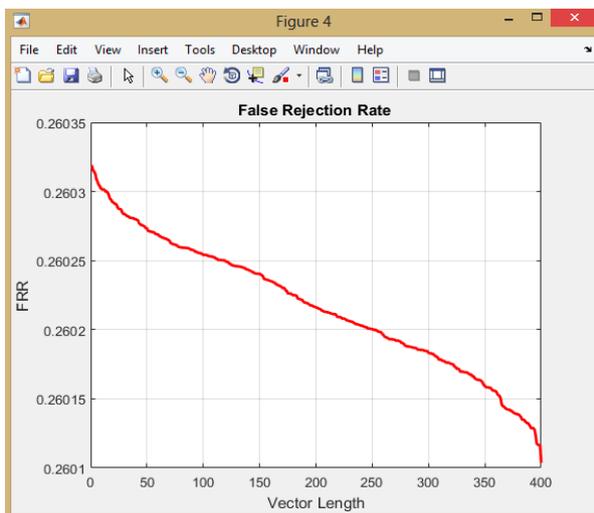


Figure 5.3. False Acceptance Rate

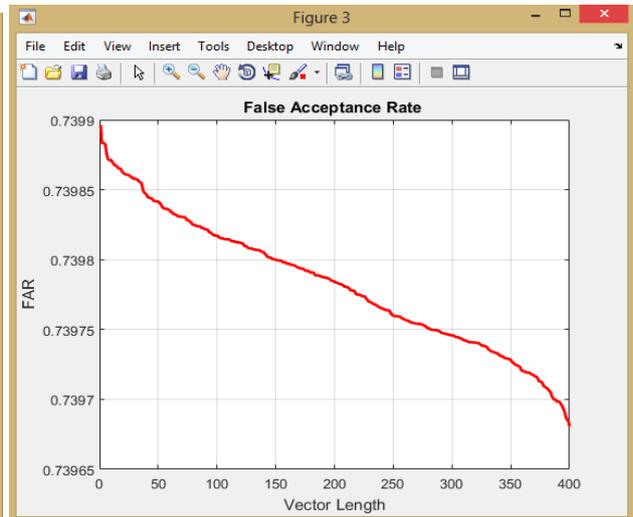


Figure 5.4. False Rejection rate

### CONCLUSION AND FUTURE SCOPE

Agribusiness is spine to each nation on the planet either specifically and in a roundabout way. Humankind proceeds with its excursion with sustenance created by methods for farming. From the audit of writing it is comprehended that the best in class innovation utilization in agribusiness has been disillusioning. The adjustment of innovation towards better efficiency in horticulture has been moderate and further research is expected to speed up it towards accuracy farming. Product yield expectation is a critical segment of national sustenance security appraisal and nourishment approach making. Product development and yield information are basic for controlling farming development framework, and horticultural task and administration. The proposed edit yield expectation comprises of three stages to be specific, pre-preparing, highlight lessening and forecast. Here the proposed

strategy utilizes input information as genuine information. Genuine information is regularly deficient, conflicting, or potentially ailing in certain conducts or slants, and is probably going to contain numerous mistakes. Horticulture in the nation can be enhanced with programmed forecast of harvest yield in view of the solid factors. The framework encourages ranchers to do right things at opportune time. The profitability gets enhanced in farming with managed examine in the field of spatial information mining to acknowledge accuracy horticulture. The model will help the ranchers in expanding their profitability by choosing the proper harvest for their property and climatic conditions. The agrarian structure for edit determination at the prior stage with all the prepared data helps the ranchers in an exceptionally valuable way. The ranchers can get all the data at only a tick of the mouse, and they require not to movement to Agricultural Universities for that.

## REFERENCES

- [1] Kalyani and M.R., 2012, "Data Mining Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, 10(2), pp. 439-442.
- [2] Agrawal, H. and Agrawal, P, 2014, "Review on Data Mining Tools", International Journal of Innovative Science, Engineering & Technology, 2(1) pp. 52-56.
- [3] Grajales, D.F.P., Mosquera, G.J.A, Mejia, F., Piedrahita, L.C and Basurto, C, 2015, "Crop-Planning, Making Smarter Agriculture With Climate Data" International Conference on Agro-Geo Informatics, pp. 240-244.
- [4] Bendre, M. R., Thool, R.C., Thool and V. R, 2015, "Big Data in Precision Agriculture: Weather Forecasting for Future Farming", International Conference on Next Generation Computing Technologies, pp. 744-750.
- [5] HemaGeetha, N, 2016, "A survey on application of data mining techniques to analyze the soil for agricultural purpose", International Conference on Computing for Sustainable Global Development (INDIACom), pp. 3112-3117.
- [6] Fathima, G.N and Geetha, R, 2014, "Agriculture Crop Pattern Using Data Mining Techniques", International Journal of Advanced Research in Computer Science and Engineering, 5(4), pp. 781-786.
- [7] Rub G., 2014, "Data Mining of Agricultural Yield Data: A Comparison of Regression Models", Industrial Conference, pp. 24-37.
- [8] Raorane A.A. and Kulkarni R.V, 2012, "Data Mining: An effective tool for yield estimation in the agricultural sector", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 1(2), pp. 75-79.
- [9] Veenadhari, S., Misra, B and Singh, C.D, 2014, "Machine learning approach for forecasting crop yield based on climatic parameters", International Conference on Computer Communication and Informatics, pp.1-5.
- [10] Sujatha R. and Isakki, P 2016 "A study on crop yield forecasting using classification techniques", International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE), pp. 1-4.
- [11] Kushwaha A.K. and SwetaBhattacharya, 2015, "Crop yield prediction using Agro Algorithm in Hadoop", International Journal of Computer Science and Information Technology & Security (IJCSITS), 5(2), pp. 271-274.
- [12] Sellam,V. and Poovammal, 2016, "Prediction of Crop Yield using Regression Analysis", Indian Journal of Science and Technology, 9(38), pp.1-5.
- [13] Kaur M., Gulati H. and Kundra, H, 2014, "Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications", International Journal of Computer Applications, 99(12), pp.1-3.
- [14] Ramesh and Vijay S., 2014, "A survey on Data Mining Techniques for Crop Yield Prediction "International Journal of Advance Research", 2( 9), pp. 2321- 7782.
- [15] S.Shenbagavadiyu and P. Kanjana Dev, 2016, "Enhanced Crop Yield Prediction and Soil Data Analysis Using Data Mining "International Journal of Modern Computer Science (IJMCS) 4(6), pp. 2320-7868.
- [16] Patil D. and P.G. Halakatti, 2017, "Rice Crop Yield Prediction using Data Mining Techniques", International Journal of Advanced Research in Computer Science and Software Engineering 7(5), pp. 427-431.
- [17] Singh R. and Singh G., 2017, "Wheat Crop Yield Assessment Using Decision Tree Algorithms", International Journal of Advanced Research in Computer Science, 8 (5),pp. 1809-1817.
- [18] E. Manjula, S. Djodiltachoumy, 2017, "A Model for Prediction of Crop Yield" International Journal of Computational Intelligence and Informatics, 6( 4), pp.2349-6363.
- [19] Marzieh Mokarram, Mohammad Jaafar Mokarram and Behrouz Safarianejadian, 2017, "Using Adaptive Neuro Fuzzy Inference System (ANFIS) for Prediction of Soil Fertility for Wheat Cultivation "Biological Forum – An International Journal 9(1), pp. 37-44.
- [20] Menaka K and N. Yuvraj, 2017, "ANFIS Based Crop Yield Prediction Model" International Journal of Science, Engineering and Technology Research (IJSETR) ,6(5), pp. 2278 -7798
- [21] Marizel B. Villanueva, and Louella M. Salenga, 2018, "Bitter Melon Crop yield Prediction" International Journal of Advanced Computer Science and Applications (IJACSA), 9(3).
- [22] M.C.S. Geetha, I.Elizabeth and S. Seetha Raman, 2018, "A survey or analysis on regression data mining techniques in agriculture" International Journal of Pure and Applied Mathematics (IJAM), 118(8), pp. 341-346.
- [23] Dakshayini Patil, M.S. Shirdhonkar, 2017, "Rice Crop Yield Prediction using Data Mining Techniques: An Overview" IJARCSSE, 7(15), pp. 5-13.
- [24] Aman Jatain and Deepti Gaur, "Estimation of component reusability by identifying quality attributes of component: a fuzzy approach", Proceeding of International Conference on Computational Science, Engineering and Information Technology, pp. 738-742, 2012.
- [25] Varisha Ashraf, Ankit Jain, Manjunath C. R and

Sahana Shet, 2018, "Prediction model of crop yield for food crop grown above ground level through big data analytics" International Journal of Advance Research, Ideas and Innovations in Technology (IJARIIT), 4(3), pp. 13-18.