

# Predictive Analysis in Agriculture to Improve the Crop Productivity using ZeroR algorithm

T.Rajasekaran<sup>1</sup>, P.Jayasheelan<sup>2</sup> and K.R. Sri Preethaa<sup>3</sup>

Assistant Professor <sup>1,2,3</sup>

Department of Computer Science and Engineering<sup>1,2,3</sup>

KPR Institute of Engineering and Technology, India <sup>1,3</sup>

Kristujayanti College, Bangalore, India<sup>2</sup>

rajasekaran30@gmail.com<sup>1</sup>, sheelan.jsp@gmail.com<sup>2</sup>, sripreethaakr@gmail.com<sup>3</sup>

**Abstract:** Agriculture is believed to be as backbone of Indian economic system. For the past few decades, agriculture field has seen lots of technological changes to improve better productivity. The world population grows steadily but the resources for crop production continuously diminish. As per World Trade Organization results in the coming decade, sustainable crop production is caused by environmental degradation. Therefore there have been considerable efforts to develop innovative approaches for sustainable crop production. Using prediction methods, farmers can enhance the productivity of crops. These methods are used to find the required quantity of crops, seeds, humidity, water level and other supplements. So, this may prevent providing too high amount of supplements for cultivation, saves money on pesticides and fertilizers, but also increases yield of crop. Major idea of this concept is higher quantity of yield can be obtained by cultivating required crop at right place and right cost at right moment.

**Keywords:** Precision Agriculture, Weka, ZeroR Algorithm, Matplotlib, Predictive Analysis

## 1. INTRODUCTION

Predictive Analysis Agricultural systems are very successful in increasing productivity and efficiency of crop production. However, population grows steadily, while the resource for crop production diminishes day by day. Traditionally agriculture involves planting or harvesting the crop against a predetermined schedule. Precision agriculture involves in collecting real-time data on weather, air quality, soil, crop maturity, equipment, labor costs and availability of existing data [2]. This predictive analytics can be used to make smarter decisions in agricultural field.

### 1.1. Precision Agriculture

Precision agriculture follows recognizing the previous year yield values, understanding the current requirement of environment and exploiting information that quantifies variations in soil and crops within agricultural fields [3]. Over the past decade, precision agriculture techniques plays vital role in economic system. Although there are multiple techniques available, farmers still figure it out hard to handle the productivity in stable manner [2]. The prediction farming concept is based on the recognition of what happened in past, analyzing what is happening in current situation and controlling the future data need in agriculture [5]. To predict the yield level, previous few year data are believed to be as input for the algorithm and predicted value is shown as graph using matplotlib [6]. Here prediction parameters includes previous year crop yield, cultivated area, irrigation types, humidity in particular area, usage of pesticides, fertilizers and labor cost [9].

### 1.2. Benefits for farmers

The precise crop variety selection, exact types and doses of fertilizers, pesticides and herbicides, and proper irrigation meet the demands of crops for optimum growth and development [12]. It may leads to increase yield, especially in main areas where uniform crop management practices were traditionally practiced. Prediction of advanced technologies, including machinery, tools and information about input, helps the farmers to increase the efficiency of labor, land and farming time [4]. The Prediction of exact quantities of yield at the appropriate time reduces the cost of agrochemical in crop production and also reduces the overall cost per unit [7]. Because of prediction, better decision-making is done at right time. Also agricultural tools helps the farmers to analyze and produce accurate information for appropriate decision-making in land preparation, seed technologies, choosing fertilizer, pesticide and herbicide application, irrigation and drainage, and post-production activities [5].

## 2. PREDICTIVE ANALYTICS MODEL

Predictive analytics includes various techniques from machine learning, data mining techniques that are derived from different historical information and current factors to make smarter decision about the future events [8]. By considering previous year data, making smarter decision about the future in present is called predictive analysis model. This model encompasses different patterns from historical facts to analyze the future predictions. This is a type of data mining technique which derives results from previous year data and predicting the behavior using history of data [11]. Predictive analytics includes various machine learning techniques and statistical techniques for classification and pattern matching [9]. This includes analyzing what event had occurred in the past, when it is occurred, and monitoring current terminologies and facts about data, then making smarter decision using machine learning methods. In Prediction, previous year data are referred as training set and the data's are classified based on training set [10]. The first and foremost process in prediction is data collection then data analysis, monitoring and finally result can be plotted as graph using matplotlib function.

## 3. PROBLEMS TO BE SOLVED

Data are collected for different weather conditions, soil type, humidity, air quality, crop maturity, labor costs and statistics of previous few year data have taken under consideration and future will be predicted by using machine learning algorithm [8]. Though previous monitoring techniques gathers the crop conditions properly, prediction results have not yet been optimized. First of all, researchers do not have clear idea about crop condition and crop monitoring methods [6]. They should know how to monitor crop condition on different circumstances. So crop characteristics should be well monitored by researchers to deliver good results in prediction methods. Quantitative models can produce quantitative results in crop monitoring which will help to develop crop growth in different conditions. Basically problems in predictions are finding proper algorithm for prediction methods and assuming different location results for predictions.

## 4. METHODOLOGY

It is of great value to obtain the crop condition information at early stages of cultivation. The same task is more important in sometimes to acquire the exact production after the harvesting time. Predictive analysis in agriculture is a crop management concept to increase the environmental, economical, market and public pressures on arable agriculture [2]. To implement this technique, previous few year yield value, cultivated area, irrigation methods, usage of fertilizer and pesticides, rainfall level at each season, soil maturity and weather condition of particular area details must be progressed in effective manner [6]. The overall concept consists of following activities. First activity is data collection from previous year statistics. Then by processing those data using machine learning algorithm, yield value is predicted. Then the predicted value can be plotted as graph using matplotlib using python language. Final result is shown in graph based on different parameters.

### 4.1. Data collection

| Sl. No. | District           | Paddy Area (Ha.) | % to state total | Gross cropped area | % of paddy area to Gross cropped area |
|---------|--------------------|------------------|------------------|--------------------|---------------------------------------|
| 1       | Thiruvananthapuram | 2395             | 1.15             | 155065             | 1.54                                  |
| 2       | Kollam             | 2097             | 1.01             | 157343             | 1.33                                  |
| 3       | Pathanamthitta     | 2802             | 1.35             | 102385             | 2.74                                  |
| 4       | Alappuzha          | 36251            | 17.41            | 107389             | 33.76                                 |
| 5       | Kottayam           | 21410            | 10.29            | 209452             | 10.22                                 |
| 6       | Idukki             | 1264             | 0.61             | 276493             | 0.46                                  |
| 7       | Ernakulam          | 7731             | 3.71             | 172449             | 4.48                                  |
| 8       | Thrissur           | 21172            | 10.17            | 181287             | 11.68                                 |
| 9       | Palakkad           | 83998            | 40.35            | 302348             | 27.78                                 |
| 10      | Malappuram         | 7528             | 3.62             | 240877             | 3.13                                  |
| 11      | Kozhikode          | 2920             | 1.4              | 206971             | 1.41                                  |
| 12      | Wayanad            | 8995             | 4.32             | 172355             | 5.22                                  |
| 13      | Kannur             | 5740             | 2.76             | 226570             | 2.53                                  |
| 14      | Kasargode          | 3857             | 1.85             | 150773             | 2.56                                  |

Figure.4.1.1. District wise cultivated area

Data collection is the major task in prediction. Since the prediction is based on the previous year statistics, small variations in data collection may cause major effect in the output. So analyzing previous year data and considering original data value for monitoring is important. Data can be of soil type, soil maturity, pesticides, fertilizer, weather conditions, humidity, water level, rain fall, cost and availability of labor of based on previous year values. Using these data, prediction results are calculated based on algorithm.

#### 4.2. Prediction based on ZeroR Classifier

Waikato Environment for Knowledge Analysis contains collection of visualization tools, algorithms for data analysis and prediction. It directly provides user friendly graphical user interface. It can be used to implement portability, free availability under GNU, data processing and modeling, ease of use due to GUI. It is fully implemented in Java Programming Language. Thus it can be used as any modern computing platform. TheZeroRclassifier is simplest classifier. This classifier considers all the target attribute and required possible values. More specifically output will always found using target attribute from given data. ZeroR does not include rule that depends on non target attribute. So we can predict the mean for numerical value given in dataset.ZeroR classifier supports for several data processing tasks, classification, regression, visualization tasks. By using ZeroR classifier, data prediction can be done using the data by assuming that the data is available as single file or relation [7]. In that the data may be described by fixed numerals of attributes. Figure 4.2.1 shows the ZeroR classifier processing results.

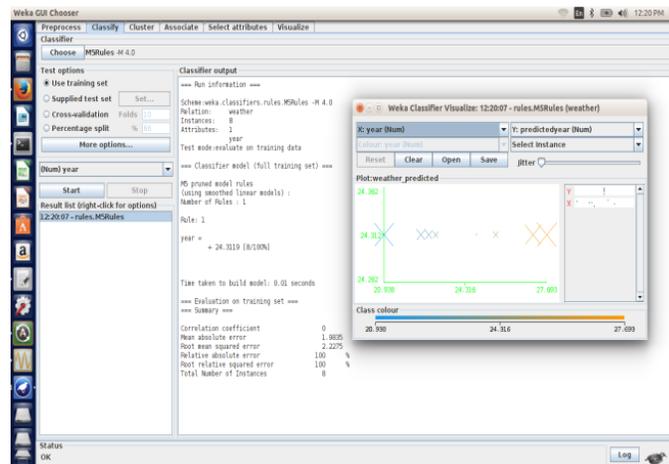


Figure 4.2.1 ZeroR classifier

#### 4.3. Graph based model

Prediction agriculture includes predicting the yield based on previous year data. But the main job is to deliver the prediction result to end user. Basically graph based approach is easy approach to understand the results. So, graph based approach is followed to draw a graph for output. In python language, we follow many built in library function. To plot a graph model, we must invoke matplotlib which is a inbuilt graph plotting library. Just by applying the values to graph parameters such as yield, productivity, area graph can be plotted based on predicted value. This graph based approach is user friendly and easy to analyze the variation in crop production in each year.

### 5. RESULTS AND IMPLEMENTATIONS

The Predictive analysis concept consists of data collection from previous few years. Data may contains production area, crop type, soil type, weather condition, humidity of particular area, rain fall level of particular area, labor cost and availability of existing products are considered as input for prediction tool. Figure 5.1 shows the data classification. Collection of data, processing those data and presenting the data in system guarantees a success in future market[3].ZeroR classifier process the data from mysql and produce classification results as represented in figure 5.1. Figure 5.3 describes the inbuilt graph representation. Real data is mentioned in green color and prediction is represented as red color. This method directly reflects the prediction results into respected text file. By considering prediction results, using python language inbuilt matplotlib library, graph can be plotted as follows;

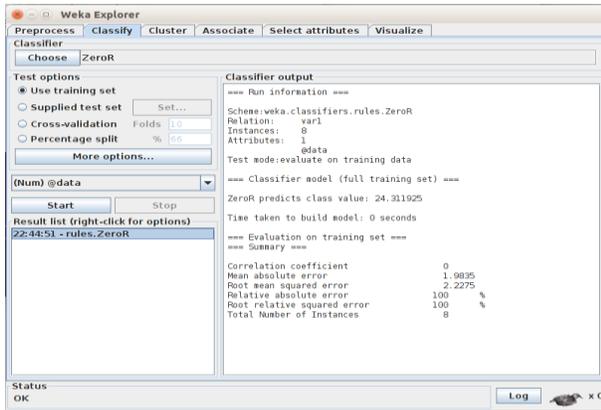


Figure 5.1 Classification model.

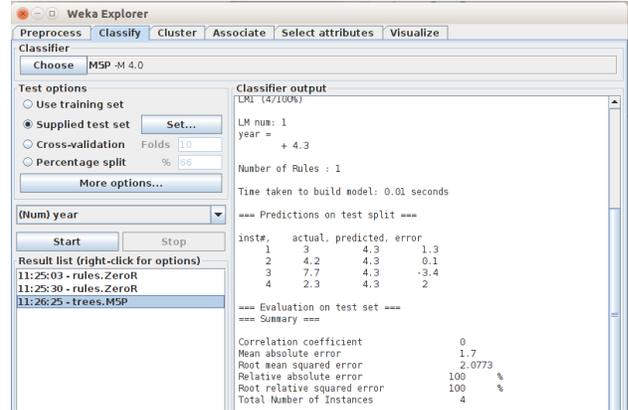


Figure 5.2 Prediction Value

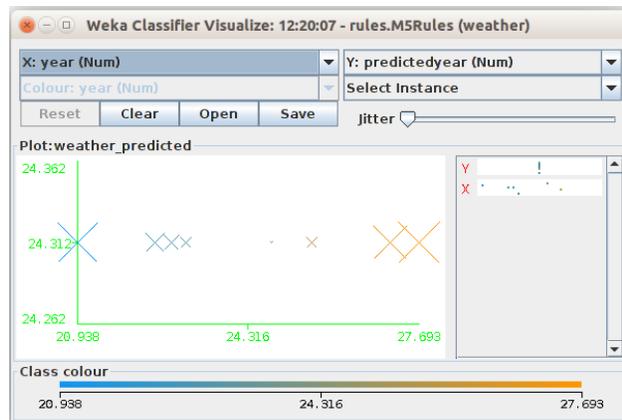


Figure 5.3 Prediction result.

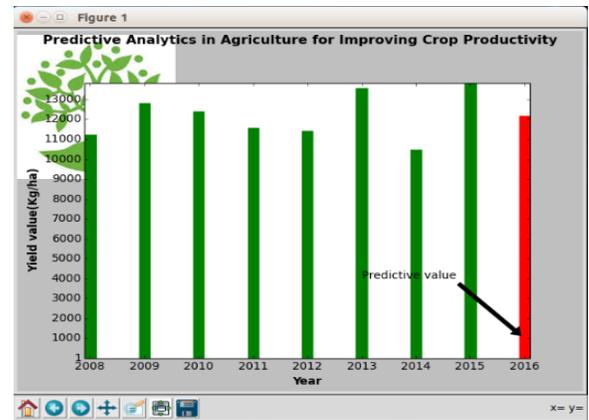


Figure 5.4 Matplotlib graph representation

Figure 5.4. shows the matplotlib bar chart graph representation. Here the green color represents previous year data collection and red color line represents prediction about the future. X axis represent year representation and y axis represents yield value. By using same method, we can able to calculate and predict the results based on different type of soil, different type of crops, different weather conditions. The respected results can be plotted as different graph representation. In order to increase crop productivity of particular crop, farmer can see the results and can cultivate the crops under best productivity seasons. So that farmers can get best profit and yield will be high. In this aspect, productivity can be increased.

## 6. CONCLUSION

Predictive analysis in agriculture is crop yield monitoring concept. Implementation of this concept should help farmer to produce higher yield. With this concept, input factors can be minimized and output can be maximized in precise way. In order to increase production, the system manages crop diseases also. System can find required amount of land, water, fertilizer, pesticides to maximize the crop production. This will help the farmers to plough exact resources for production and minimize the cost in proper manner. This technique significantly reduces cost and chemical wages. So that farmers obtain a return on their investment by saving on fertilizer cost. Another benefit of this concept is adjusting with environmental factors. Based on different environmental structures farmer can plough respected crop on respected field at right time increases the productivity to larger manner. Therefore precision agriculture seeks important role in economy field as well as to achieve environmental goal.

## 7. FUTURE ENHANCEMENTS

Currently, larger number of companies focuses on predictive analytic techniques to enhance their business. Since agriculture is backbone of any country, taking smarted decision in agriculture becoming more important task in day to day life. As long as the population increases, the demand for food increases and arable land level diminishes. These factor forcing the farmers

to do high productivity with fewer amounts of resources in a sustainable manner. In the future, sensors can be fixed to predict the yield of any crop. A farmer can take a photo of crop in mobile phone and upload the picture in the system where the system could predict the properties of crop and can judge the result. Crop properties can be matched with earlier image database, and by image processing techniques, crop maturity can be identified based on color and weather factors. As long as the technology develops, prediction techniques may enhance its efficiency. A device called quad copter can be fixed in the agriculture field to calculate the humidity and weather condition of particular area. These small rotary and fixed wing aircraft can fly at low altitudes and be programmed to fly a certain pattern. The quad copter can take high definition images with its camera and produce properties of the crops from database. Frequently, quad copter provides sensor reading and based in the real time results, to make more efficient predictions among the crops. Therefore predictive analysis plays major role in future agriculture.

#### REFERENCES

- [1] The concept and implementation of precision farming and rice integrated crop management systems for sustainable production in the twenty-first century- D.V. Tran a and N.V. Nguyen.
- [2] Implementing Precision Agriculture in the 21st Century. V. John Stafford, Silsoe Solutions, Ampthill, UK. doi: 10.1006/jaer.200
- [3] Using technology of data collection and data processing in precision farming. Sber a vyhodnoceni.
- [4] Precision Agriculture M&A, Investment, and startups on the Rise, Special market report 2014.
- [5] Predicting Maize Phenology: Intercomparison of Functions for Developmental Response to Temperature S. Kumudini, F. H. Andrade, K. J. Boote, G. A. Brown, K. A. Dzotsi, G. O. Edmeades, T. Gocken, Vol. 106 No. 6, p. 2087-2097( 2014)
- [6] Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison - vol. 111 no. 9, 3268-3273(2013)
- [7] A Stochastic Approach for Predicting the Profitability of Bioenergy Grasses Joseph Dolginowa, Raymond E. Massey, Kenneth A. Sudduth Vol. 106 No. 6, p. 2137-2145( 2014)
- [8] Predicting climate change effects on wetland ecosystem services using species distribution modeling and plant functional traits Helon Moor, Kristoffer Hylander, Jon Norberg(2014) DOI 10.1007/s 13280-014-0593-9
- [9] Agricultural water management- Simulation and Yield of wheat crop under irrigated saline regimes P.Kumar, A.Sarang Vol 148, 31 jan 2015(2014)
- [10] Climate variation explains a third of global crop yield variability Deepak K. Ray, James S. Gerber, Graham K. MacDonald & Paul C. West
- [11] P.Kumar, A.Sarangi, D.K.Singh, S.S.Panhar (2015) - Simulation of salt dynamics in the root zone and yield of wheat crop under irrigated saline regimes using SWAP model
- [12] Samir Kumar Sarangi and Dr. Vivek Jaglan ,ISSN : 2279-0047- IJETCAS (International Journal of Emerging Technologies in Computational and Applied Sciences)- Performance Comparison of Machine Learning Algorithms on Integration of Clustering and Classification Techniques
- [13] Ritika, Aman Paul - Prediction of Blood Donors Population using Data Mining Classification Technique, ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering Research Paper
- [14] ChitraNasa, Evaluation of Different Classification Techniques for WEB Data - International Journal of Computer Applications (0975 – 8887) Volume 52– No.9, August 2012