

Decision Tree Algorithm Applied in Suitability Assessment of Temporary Crops Based on Agrometeorological Forecasts

Aleta C. Fabregas*

Polytechnic University of the Philippines, Manila, Philippines: aletfabregas@gmail.com

Abstract

Objectives: To develop Crop Recommender System, using Decision Tree, suggesting suitable temporary crops in Laguna, Philippines based on forecasted agrometeorological variables with Time Series analysis and Autoregressive Integrated Moving Average models. **Methods/Statistical Analysis:** This study develops a tool to assess the suitability of crops based on forecasted variables with higher accuracy using Decision Tree and make the crop more adaptative to a changing weather patterns. **Findings:** Simulation process was done to evaluate and test the accuracy of the system. In forecasting agrometeorological variables: Maximum temperature, minimum temperature, relative humidity, sunshine duration, accumulated rainfall, the system garnered accuracy of 98.23%, 97.71%, 97.03%, 81.96% and 67.51% to each variable respectively. Overall, the accuracy of the agrometeorological forecasts averaged to 91.64%. For the measurement of accuracy of the system in assessing the suitable crops, the model was tested on each crop and resulted to 86.33% overall classification accuracy. In this study, it is determined that high variability of time series affects the forecast accuracy of rainfall, which model was less accurate compare to other agrometeorological variable models. **Application/Improvements:** TANIM, a new web-based decision support system enables decision makers to know the future condition of agrometeorological variables and track which temporary crops are suitable based from the forecast.

Keywords: Agrometeorological Variables Forecasting, Autoregressive Moving Average, Decision Tree Algorithm, Suitability Assessment, Temporary Crops

1. Introduction

Weather and climate have a direct influence on cropping systems and plant yield. Thus, weather fluctuations and climate variability play a significant role in crop growth and yield. Occurrence of abnormal weather episodes during the growing season or during critical development stages may hamper growth processes resulting in yield reduction. This makes climate variability a threat to food production leading to serious social and economic implications¹. Having a direct influence in agriculture, climate and changing weather patterns has been subject to studies that aims to cope up in the said changing element. Numerous studies have suggested that climate variability and climate change can have adverse impacts on global

food production and food security. Climate variability driven by major inter annual-scale climate modes, such as the El Nino Southern Oscillation, has been playing a key role by often leading to droughts and decrease in crop yields that could further result in famine in some food insecure regions²⁻⁴.

Agriculture, in general, has undergone a similar evolution. Technology has become an indispensable part of doing business for every farmer, agriculture retailer and agronomist. In fact, a recent study by Hexa Reports suggests precision agriculture is set to grow to \$43.4 billion by 2025. The increasing adoption rate of technology in agriculture shouldn't be surprising to anyone. According to⁵ farmers are driven to use technology to increase efficiency and manage costs. Precision agriculture is also known as

*Author for correspondence

precision or precision farming. Perhaps the easiest way to understand precision agriculture to think of it as everything that makes the practice of farming more accurate and controlled when it comes to the growing of crops and raising livestock.

Most of precision agriculture studies involves crop yield prediction problem. Based on big-data analysis, crop yield prediction quantifies the yield of crop and predicts its value in the future to know if there will be shortage of production for a certain crop or to know which crop will have the highest amount of yield. Only few studies focuses on recommendation of crop which makes it an interesting area to study. Crop recommenders system solves crop selection problem. Instead of knowing the quantity of yield that a crop could have, it would also be helpful for farmers to know the best selection of crops base from growing condition of environment. Crop yield selection or crop recommender system could be applicable for minimizing losses unfavorable conditions may occur and this could be used to maximize crop yield. In this application, crops or series of crops are outputted as recommendation which basis is dependent on various parameters. These parameters are usually site-specific and are usually listed as main elements that affect crops. Parameters like weather condition (e.g. temperature, rainfall and humidity), soil type (e.g. sandy, silty, clay, saline soil) and soil composition (eg. pH value, nitrogen, phosphate, potassium, organic carbon).

Most of the existing studies used real-time weather information, but for farm planning and management it is more desirable to know the crops that are suitable to the future weather conditions. Several studies offered methods to forecast or predict the values of the variables that affect crops. These variables are called agrometeorological variables. Agrometeorological variables are vital factors in crop production and plays drastic effects when fluxes of such were observed. In⁶⁻⁸ used ARIMA class of time series models to forecasts values of agrometeorological variables such as relative humidity, minimum temperature, maximum temperature and rainfall.

With available resources of weather and soil information, studies were conducted which focused on suggesting crops based from various independent elements. In⁹⁻¹² proposed different methods in mapping the relation of various independent variables to crops. These studies were utilized for machine learning in determining crops that are suitable for giving information about weather condition, soil and fertilizer. Crops that are found suit-

able will be recommended. Interestingly, in the study of¹³, they proposed a system that utilizes real-time weather information and predicts the crops that are suitable to the current weather condition. They utilized Naïve Bayes classifier and Fuzzy Naïve Bayes in crop-weather prediction task. As the result, they got an accuracy of 54% for Simple Naïve Bayes and 73% for Fuzzy Naïve Bayes. This accuracy rate could possibly not enough for precision in agriculture it is important that the recommendations made are accurate and precise because in case of errors it may lead to heavy material and capital loss.

This study focuses on development of TANIM: Crop Recommender System, a system that suggests suitable temporary crops in Laguna based on forecasted agrometeorological variables of the location. Agrometeorological variables that are forecasted are minimum temperature, maximum temperature, relative humidity, rainfall and sunshine. These variables are forecasted monthly by utilizing Time Series analysis methodology and Autoregressive Integrated Moving Average models. Decision Tree was used to know whether the crop is suitable or not to the weather condition by using the of monthly forecast agro meteorological variables as the independent variables. Temporary crops that are found suitable are added to series of crop as recommendation.

2. Material and Methods

Many studies had demonstrated the use of machine learning specifically classification technique in identifying best selection of crops with agro meteorological data as the influencing variables. One of these studies is¹⁴, where they demonstrated the prediction of rice crop yield by applying one of the machines learning technique, Support Vector Machine (SVM). The study used precipitation (mm), minimum, average and maximum temperature (degree Celsius) as climatic parameters to be able to deliver the objective. In⁶ proposed fuzzy query approach for crops plantation dates selection and optimization. Minimum and maximum temperature and humidity were used as variables to predict for the incoming year. In¹⁵ investigated an ensemble model with majority voting technique using Random Tree, CHAID, K-nearest Neighbor and Naïve Bayes as learners to recommend a crop for the site-specific parameters. In the study of¹⁶, they presented method called Crop Selection Method or CSM which aimed to solve crop selection problem in India. CSM method retrieved all possible crops that were to be sown at a given

time stamp. Yield rate of these crops are evaluated, if yield rate per day of these crops are fair (within tolerance) then those crops are selected for crop sequences. Similarly, in the study of¹⁷, they proposed a method which goal was to automate the decision-making process of the farmers in selecting crops by considering the climatic and soil conditions within the area. The method proposed in the study uses fuzzy logic algorithm. Weather information such as rainfall, temperature, humidity, sunshine, cloud type etc. is considered as the input parameters of the fuzzy logic together with soil information.

To improve crop selection method, it is desirable to combine it with the power of forecasting by predicting the future value of the influencing variables, which in this study are the agrometeorological variables; maximum temperature, minimum temperature, sunshine duration, relative humidity and rainfall. Many studies had also attempted to predict these variables in some parts of the world. One of these is the study of⁶, where they forecasted temperatures in the Sylhet Division of Bangladesh using ARIMA or Auto Regressive Integrated Moving Average models. Models were used to carry out short-term predictions of monthly maximum and minimum temperatures. In⁸, build a model to forecast relative humidity and mean monthly temperature in Ahwaz Station using ARIMA model. While in the study of⁷ experimentation was conducted by modelling and predicting behavioral pattern in rainfall phenomena based on past observations. The study introduced three fundamentally different approaches for designing a model, the statistical method based on Autoregressive Integrated Moving Average (ARIMA), the emerging Fuzzy Time Series (FST) model and the non-parametric method (Theil's regression). In their result, the ARIMA model was better than the Theil's Regression model.

2.1 Data Gathering

Two types of datasets were gathered which were used to build the models in the system. The first dataset, the agrometeorological Time Series dataset, was gathered from the National Agrometeorological Station situated in the University of the Philippines – Los Baños, Laguna which were collected and managed by the DOST-PAGASA Climatology and Agrometeorological Division (CAD) located in Diliman, Quezon City. The data gathered were the records of 2004 – August 2018 of the following agrometeorological variables: Maximum Temperature,

Minimum Temperature, Rainfall, Relative Humidity and Sunshine Duration. Second dataset, the Crop-Agromet Dataset, came from the Philippine Statistic Authority – Crop Statistics Division which contains the information regarding the crops being produced in Laguna and the crop production and area harvested respectively. The crop production and area harvested data were used to get the actual crop yield per crop which were produced in the Province of Laguna. Through this dataset, the system was able to generate a classifier model which classify the productivity of a crop depending on the agrometeorological variables forecasted by the ARIMA model through the first dataset. The set of crops included in the study is only limited to those which were usually planted in Laguna and were temporary in nature.

In our experiments, we divided both dataset into two parts. In agrometeorological time series dataset, samples from 2004 to 2014 were used as training set and samples from 2015 to 2018 were used as the testing set to validate the performance of the models. Similarly Crop-agromet dataset was divided into 80% and 20% data samples. 80% was used as the training set to build the classifier and 30% was used as testing set.

The next two sections discuss the algorithms used in this study, namely, Autoregressive Integrated Moving Average and Decision Tree.

2.2 Autoregressive Integrated Moving Average

A general class of univariate models is the Autoregressive Integrated Moving Average (ARIMA) model. An ARIMA model represents current values of a time series in terms of past values of itself (the autoregressive component) and past values of the error term (the moving average terms). The integrated component refers to the number of times a series must be differentiated to induce stationarity¹⁸.

The integrated component of an ARIMA model represents the number of times a time series must be differenced to induce stationarity. A general notation for ARIMA models is ARIMA (p, d, q) (P, D, Q), where p denotes the number of autoregressive terms, q denotes the number of moving average terms and d denotes the number of times a series must be differenced to induce stationarity. P denotes the number of seasonal autoregressive components, Q denotes the number of seasonal moving average terms and D denotes the number of seasonal differences required to induce stationarity¹⁹.

2.3 Decision Tree

A Decision Tree is defined as a classification procedure that recursively partitions a dataset into smaller subdivisions on the basis of a set of tests defined at each branch (or node) in the tree. The tree is composed of a root node (formed from all of the data), a set of internal nodes (splits) and a sets of terminal nodes (leaves). Each node in a Decision Tree has only one parent node and two or more descendant nodes. In this framework, a dataset is classified by sequentially subdividing it according to the decision framework defined by the tree and a class label is assigned to each observation according to the leaf node into which the observation falls²⁰. The Decision Trees are known to produce results of higher accuracies in comparison to traditional approaches such as the “box” and “minimum distance to means” classifiers but the performance of DTs can be affected by a number of factors including: Pruning and boosting methods used and decision thresholds²¹.

TANIM: Crop Recommender System is composed of two modules; 1. The agrometeorological forecasts and 2. Crop Suggestion Module. The agrometeorological forecasts module is used to predict the future values of the four agrometeorological variables, Minimum Temperature, Maximum Temperature, Relative Humidity, Sunshine and Rainfall. This module includes the process of selecting the type of ARIMA models to be used based on the time series data available. The ARIMA Model Selection process includes the identification of parameters for the ARIMA models, evaluation of models selected and lastly, is selecting the best models suited for forecasting. These selected models are used to forecast monthly values of the agrometeorological variables. In the Crop Suggestion module, Crop-agromet dataset was used in training the Decision Tree classifier which determines whether a crop is high yield or low yield to a given weather condition. A set of parameters to be fed to the classifier model is the forecasted minimum temperature, maximum temperature, relative humidity, sunshine and rainfall from the agrometeorological forecasts module and the identity of crop to be checked for weather suitability. After all the crops are tested in the classifier model, all the crops that are identified as suitable crops or high yield are displayed as crop recommendation for the given weather condition. The user input is the monthly time frame then the system's output is the crop recommendation based from the forecasted agrometeorological variables in the given month/s.

3. Results

To measure the performance of the system in forecasting agrometeorological variables and in predicting the productivity of the crop, Forecasting Accuracy was used to evaluate the ARIMA models and Classification Accuracy for the Decision Tree classifier. Forecast Accuracy of the actual percentage value of correct forecast, MAPE is subtracted to 100%. And Classification Accuracy or accuracy is used as a statistical measure of how well a binary classification test correctly identifies or excludes a condition. Table 1 presents the order of ARIMA used to forecast agrometeorological variables and each accuracy of the models. And Table 2 summarizes the accuracy of Decision Tree classifier in predicting the productivity of each crop.

Table 1. Accuracy of ARIMA Models

Agromet Variable	ARIMA Model	MAPE	Forecast Accuracy
Maximum Temperature	(2,0,2)(1,1,2) ₁₂	1.77%	98.23%
Minimum Temperature	(2,0,1)(3,1,0) ₁₂	2.29%	97.71%
Relative Humidity	(2,1,1)(2,1,1) ₁₂	2.97%	97.03%
Sunshine Duration	(2,0,3)(2,1,1) ₁₂	18.04%	81.96%
Rainfall	(2,1,1)(2,1,1) ₁₂	32.49%	67.51%

Table 2. Accuracy of Decision Tree classifier

Crop	Accuracy	Crop	Accuracy
Ampalaya	94.44%	Okra	72.22%
Cabbage	86.11%	Snapbeans	80.55%
Chinese Pechay	77.77%	Squash	83.33%
Eggplant	91.66%	Stringbeans	91.66%
Kangkong	91.66%	Tomato	100.00%
Native Pechay	88.88%	Upo	77.77%

As the Table 1 shows, the accuracy of forecast for agrometeorological variables ranges from 81% to 98%, except 67.51% for the accuracy of rainfall. The system yields highest accuracy for the maximum temperature,

consequently, it had the lowest MAPE. More than 10 years amount of data were used to model the behavior and patterns of agrometeorological variables situated in the province of Laguna. Model selection technique was done to determine the order of hyperparameter PDQ of ARIMA that best fits to model these patterns. In this technique, AIC (Akaike Information Criterion) was used to estimate the relative quality of ARIMA model for the agrometeorological time series data. The Table 1 also shows the order of ARIMA model fitted to the time series data of agrometeorological variables in Laguna.

Order 12 depicts that the time series changes repeats in any 12 times observation, which in this case is yearly. This pattern can be observed in the time plots of the agrometeorological variables in Figures 1-6. It is evident from the time plots that there are yearly pattern and trends present in the data. By doing further analysis on the time plots, it can be observed in maximum and minimum temperature has a downward trend that shows that both variables are on the highest somewhere during the 3rd and 5th months of the year and drops significantly on the last months. While on the time plots of sunshine duration and relative humidity, although there are no linear trend, high values are observed during 2nd and 3rd quarter of the year. By plotting the predicted against the actual values, it shows that the patterns mentioned were captured by ARIMA with a good accuracy. Rainfall, compare to other variables, shows more variability and number of fluctuations. Although, there are a seasonal pattern every 1st quarter of the year, it can be seen that fluctuation or the sudden rise of values are not consistent and differs each year. Because of its high variability, accurate forecast of rainfall is quite difficult to achieve, consequently, among the agrometeorological variables, accuracy for rainfall model is quite low compare to other models.

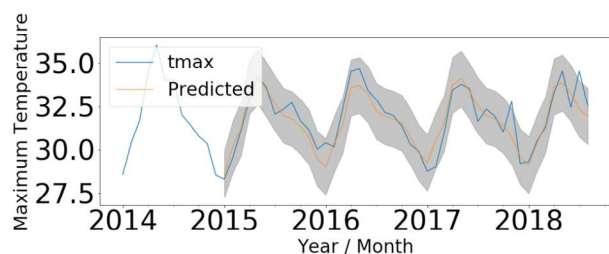


Figure 1. Actual and forecasted maximum temperature from 2015-2018.

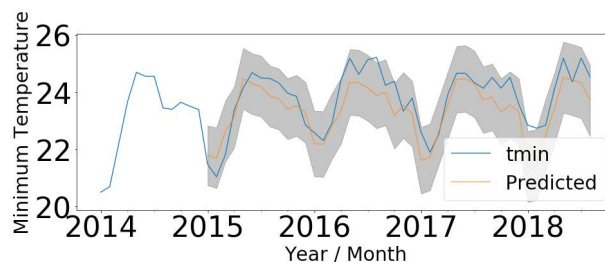


Figure 2. Actual and forecasted minimum temperature from 2015-2018.

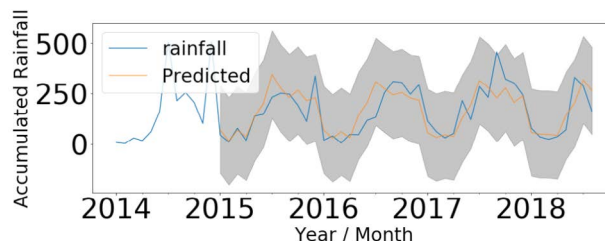


Figure 3. Actual and forecasted rainfall from 2015-2018.

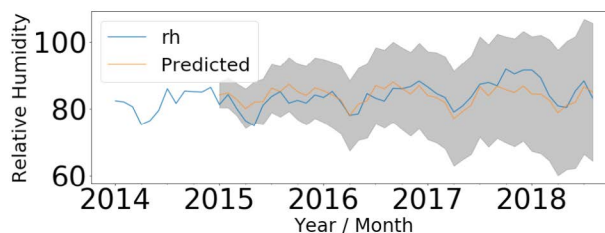


Figure 4. Actual and forecasted relative humidity from 2015-2018.

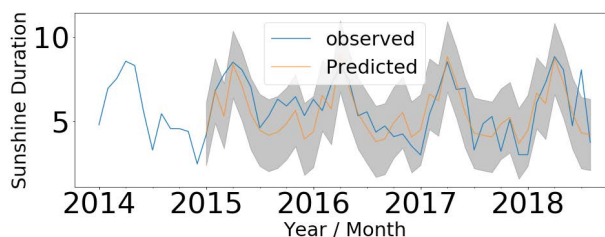


Figure 5. Actual and forecasted sunshine Duration from 2015-2018.

The Table 2 depicts the classification accuracy for each crop ranges from 77% to 98%, except for the crop Okra which resulted to 72.22%. Tomato garnered the highest accuracy with 98% accuracy, and along with it, Ampalaya, Eggplant, Kangkong and Stringbeans had resulted to high accuracy. In classifying the suitability of Okra, the system

got the lowest accuracy. This was due to the high false negatives which means some observations of Snapbeans should have been classified as Low Yield but the model predicted High Yield. Possibly, this result occurred because of skewness of data present in both training and testing data for Okra. Overall, the average accuracy of the system in assessing the suitability of crops was 86.33%.

To further study the relation of the five agrometeorological variables to the productivity of temporary crops, Extra Tree Classifier is used. Extra Tree Classifier was used to estimate the value of importance of each independent variables to the true outcome, in this case, the five agrometeorological variables and crop productivity. It is a feature selection technique where the importance of a feature is measured by calculating the increase in the model's prediction error after permuting the feature. A feature is "important" if shuffling its values increases the model error, because in this case the model relied on the feature for the prediction. A feature is "unimportant" if shuffling its values leaves the model error unchanged, because in this case the model ignored the feature for the prediction²². Table 3 shows the result from Extra Tree Classifier.

Each crop differs on its agrometeorological requirements and one variable may have more influence than other variables in terms of the productivity of the crop. As shown in Table 3, among the five agrometeorological variables, productivity of temporary crop was mostly influenced by maximum temperature is said to have more influence followed by minimum temperature and rainfall.

Table 3. Results of Extra Tree Classifier

Crop	Max. Temp	Min. Temp	Rel Hum	Sun Dur	Rain
Ampalaya	0.4282	0.2893	0.1114	0.0913	0.0796
Cabbage	0.2995	0.1754	0.1765	0.1736	0.1748
Chinese Pechay	0.2102	0.1757	0.1992	0.1754	0.2392
Eggplant	0.4178	0.2877	0.1024	0.0843	0.1075
Kangkong	0.1986	0.2929	0.1790	0.1318	0.1975
Native Pechay	0.3384	0.1899	0.1729	0.1433	0.1551
Okra	0.2664	0.2074	0.2109	0.1926	0.1226
Snapbeans	0.2373	0.1403	0.1772	0.2526	0.1923
Squash	0.3476	0.1521	0.1422	0.1793	0.1785
Stringbeans	0.2171	0.2856	0.1675	0.1446	0.1850
Tomato	0.1683	0.4559	0.0986	0.0723	0.2046

Upo	0.2188	0.1952	0.2069	0.2045	0.1744
Total	3.3487	2.8479	1.9453	1.8461	2.0117

4. Conclusions

The study presented a method which aimed to utilize ARIMA in forecasting agrometeorological variables in Laguna; maximum temperature, minimum temperature, relative humidity, sunshine duration and rainfall and based from this forecast, the suitability of crops in the region were assessed by modelling its production using Decision Tree. TANIM, web-based decision support system enables decision makers to know the future condition of agrometeorological variables and track which temporary crops are suitable based from the forecast. This could provide a wealth of information about crops' potential for adaptation to a changing weather patterns.

Simulation process was done to evaluate and test the accuracy of the system. In forecasting agrometeorological variables: Maximum temperature, minimum temperature, relative humidity, sunshine duration, accumulated rainfall, the system garnered accuracy of 98.23%, 97.71%, 97.03%, 81.96% and 67.51% to each variable respectively. Overall, the accuracy of the agrometeorological forecasts averaged to 91.64%. For the measurement of accuracy of the system in assessing the suitable crops, the model was tested on each crop and resulted to 86.33% overall classification accuracy. In this study, we identify that high variability of time series affects the forecast accuracy of rainfall, which model was less accurate compared to other agrometeorological variable models. Similarly, due to the skewness of dataset used for Okra, both training and testing set, it produced less accurate result.

Future works could be focus in using other agrometeorological variables and consider other factors such as soil to predict the production of crops.

5. Acknowledgements

To our experts, Dr. Marcelino Q. Villafuerte and his good office, PAG ASA – Climatology and Agrometeorological Division for sharing their expertise in order to develop our research. Greatest thanks as well to Ms. May Lyn J. Reyes, Mr. Leo A. Gelisan, Mr. Marino Mendoza and their humble office, Calamba City Agriculturist Office and to my student Adriana Capua, a Bachelor of Computer Science graduate from Polytechnic University of the Philippines (PUP) for

sharing their valuable time and efforts with us to localize our research make it more feasible and possible.

6. References

1. Lansigan FP, De los Santos WL, Coladilla JO. Agronomic impacts of climate variability on rice production in the Philippines. *Agriculture, Ecosystems and Environment*. 2000; 82(1-3):129–37. [https://doi.org/10.1016/S0167-8809\(00\)00222-X](https://doi.org/10.1016/S0167-8809(00)00222-X)
2. Barnard PL, Allan J, Hansen JE, Kaminsky GM, Ruggiero P, Doria A. The impact of the 2009-10 El Nino Modoki on US west coast beaches. *Geophysical Research Letters*. 2011; 38(13):102–10. <https://doi.org/10.1029/2011GL047707>
3. Maxwell D, Fitzpatrick M. The 2011 Somalia famine: Context, causes and complications. *Global Food Security*. 2012; 1(1):5–12. <https://doi.org/10.1016/j.gfs.2012.07.002>
4. Iizumi T, Sakuma H, Yokozawa M, Luo JJ, Challinor AJ, Brown ME, Yamagata, T. Prediction of seasonal climate-induced variations in global food production. *Nature Climate Change*. 2013; 3(10):904–8. <https://doi.org/10.1038/nclimate1945>
5. What is Precision Agriculture and How is Technology Enabling it? 2018. agfundernews.com/what-is-precision-agriculture.html/
6. Nury AH, Koch M, Alam MJB. Time series analysis and forecasting of temperatures in the sylhet division of Bangladesh. 4th International Conference on Environmental Aspects of Bangladesh; 2013. p. 24–6.
7. Olatayo TO, Taiwo AI. Statistical modelling and prediction of rainfall time series data. *Global Journal of Computer Science and Technology*. 2014; 14(1-G):1–11.
8. Sarraf A, Vahdat SF, Behbahaninia A. Relative humidity and mean monthly temperature forecasts in Ahwaz station with ARIMA model in time series analysis. *International Conference on Environment and Industrial Innovation. IPCBEE, Singapore*. 2011; 12:51–62.
9. Pudumalar S, Ramanujam E, Rajashree RH, Kavya C, Kiruthika T, Nisha J. Crop recommendation system for precision agriculture. 2016 Advanced Computing (ICoAC) Eighth International Conference on Advanced Computingl 2016. p.32–6. <https://doi.org/10.1109/ICoAC.2017.7951740>
10. Rajak RK, Pawar A, Pendke M, Shinde P, Rathod S, Devare A. Crop recommendation system to maximize crop yield using machine learning technique. *International Research Journal Engineering Technology*. 2017; 4(12):950–3.
11. Kumar R, Singh MP, Kumar P, Singh JP. Crop selection method to maximize crop yield rate using machine learning technique. *Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials*; 2015. p. 138–45. <https://doi.org/10.1109/ICSTM.2015.7225403>
12. Shobha N, Asha T. Monitoring weather based meteorological data: Clustering approach for analysis. *Innovative Mechanisms for Industry Applications*; 2017. p. 75–81. <https://doi.org/10.1109/ICIMIA.2017.7975575>
13. Rageena PM, AbyAbahai T, Varghese SM. Agriculture crop prediction system based on meteorological information. *International Journal of Control Theory and Applications*. 2017; 10:23–31.
14. Gandhi N, Armstrong LJ, Petkar O, Tripathy AK. Rice crop yield prediction in India using Support Vector Machines. 2016 13th International Joint Conference on Computer Science and Software Engineering; 2016. p. 1–5. <https://doi.org/10.1109/JCSSE.2016.7748856>
15. Rajashree RH, Kavya C, Kiruthika T, Nisha J. Crop recommendation system for precision agriculture. 2016 Eighth International Conference on Advanced Computing; 2016. p. 32–6.
16. Kumar R, Singh MP, Kumar P, Singh JP. Crop selection method to maximize crop yield rate using machine learning technique. *Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials*; 2015. p. 138–45. <https://doi.org/10.1109/ICSTM.2015.7225403>
17. Kapoor A, Verma AK. Crop selection using fuzzy logic-based expert system. *Applications of Soft Computing for the Web*. Springer; Singapore. 2017. p. 121–37. PMCid: PMC5395383. https://doi.org/10.1007/978-981-10-7098-3_8
18. Ljung G, Box G. On a measure of lack of fit in time series models. *Biometrika*. 1978; 66:67–72. <https://doi.org/10.2307/2335657>
19. Meyler A, Kenny G, Quinn T. Arima modelling. *Forecasting Irish inflation using ARIMA models*. 1998. p. 4–5.
20. Friedl MA, Brodley CE. Decision Tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*. 1997; 61(3):399–409. [https://doi.org/10.1016/S0034-4257\(97\)00049-7](https://doi.org/10.1016/S0034-4257(97)00049-7)
21. Mahesh P, Mather PM. An assessment of the effectiveness of the Decision Tree method for land covers classification. *Remote Sensing of the Environment*. 2003; 86:554–65. [https://doi.org/10.1016/S0034-4257\(03\)00132-9](https://doi.org/10.1016/S0034-4257(03)00132-9)
22. Breiman L. Random forests. *Machine Learning*. 2001; 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>