IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676,p-ISSN: 2320-3331, Volume 12, Issue 6 Ver. III (Nov-Dec. 2017), PP 83-86
www.iosrjournals.org

Designing of an Improved Agricultural Classification System

Sudhansu Bisoyi[1], Archana Panda[2]
1 (Department of Computer science & Engineering ,Gandhi Engineering College, India)
2 (Department of Computer science & Engineering ,Gandhi Engineering College, India)

Abstract: *Agriculture is the main source of income for most of the people of African's countries. So there is a need to transform the huge agriculture data into technologies and make them available to the farmers. The aim of this work is to find out the best classification algorithm enhances the classification of the agricultural dataset according to countries, area harvested, yield, production, and seed. Five classification algorithms are used namely J48, PART, Decision Table, IBK, and Naïve Bayes. Real agricultural dataset of the production in African countries is used and applied on WEKA software. The obtained results revealed that J48 algorithm outperformed in terms of error rate and provides slightly better performance than PART and Decision Table. IBK and Naïve Bayes classification algorithms are not suitable for this dataset. This means that trees classifiers and rules classifiers are good for this dataset.:*
Key Word*: Classification of agriculture, WEKA, J48, Decision Table, NaiveBayes, PART, and IBK*

I. Introduction

As it is known that the backbone of the country's economy is agriculture. Hence, the agricultural firms are aimed to computerize their operations in order to increase the productivity [1]. The wide availability of huge amounts of agriculture data has generated an urgent need for the research of data mining. Generating rules with higher accuracy for Agriculture databases can be done using different techniques of data mining [2]. The agricultural classification algorithms enable the farmers to identify the area harvested, the yield, the production, the seed, and etc. The application of agriculture classification is increasing day by day to improve and increase the production of crops. Therefore, the accuracy of agriculture classification systems is very important in agriculture sector. This work aims to find out the best classification algorithm to enhance the accuracy in agriculture classification. Then, recommended the suitable classifier to classify the African's countries and their crops in the harvested area to give the best classification. Five classifiers are applied namely Bayesian, Rules and trees, and Lazy in that Bayesian classifier is NaiveBayes classification algorithm, in rules classifier two classification algorithms are applied Decision Table classification and PART classification algorithms. In trees classifier J48 classification algorithm is examined and in Lazy classifier IKB is applied. These algorithms are examined with 10-fold cross validation test mode by using WEKA software with real agriculture dataset. This paper is organized as: section 2 presents work of other researchers in agricultural classification. Section 3 describes the agricultural dataset which is used. Section 4 implements and discusses the experimental results while section 5 concludes the results.

II. Literature Review

- They used combination of object-based image analysis and advanced machine learning methods to improve the identification of crops. From ASTER satellite images captured in two different dates of nine major summer crops in central California. They evaluated the classification by applying c4.5, LR, SVM, MPL both as single classifiers and combined hierarchical classification. They found that MPL and SVM as single classifiers obtained accuracy slightly higher than LR and notably higher than c4.5. as the hierarchical combination, the best method is (SVM+SVM) which improves the accuracy of classification for all the studied crops.
- They investigated the effect of three different vegetation indices of rapid eye imagery on classification accuracy. They used overall accuracy and kappa coefficient to evaluate the accuracy of the classified images. Their study proved the efficient use of SVM for crop classification. Their results indicated that vegetation indices derived from original spectral band of Rapid Eye imagery could be used for crop classification.
- They used PSO-SVM for the selection of the features then applied the Fuzzy Decision Tree for the classification. From their experimental results they found that their proposed methodology enhanced the accuracy and the execution time as compared to the existing methodologies. They applied their proposed methodology on various datasets and from their experimental results they approved that.
- They assessed the accuracy of support vector machine model and quantified the effect of class imbalance on its accuracy. For their study they used a highly imbalanced dataset of 20 tropical species and one mixed-

species of 24 species. They found that standardizing sample size reduced the accuracy of the model and the level of species under and over prediction.

- They made a survey of the classification of agricultural crops production techniques and their advantages. The classification using FDT by the optimizing feature extraction using PSO-SVM is proposed as a new and efficient technique for the classification of agriculture crops production in the future.

III. Dataset Description

The dataset contains the plants that are produced in African countries. This dataset covers quantity produced in tonnes, producer price, value at farm gate, area harvested in hectare, yield per hectare, seed in tonnes, and the African countries where crops are produced. The dataset covers the agricultural production date from 1961 to 2014 with some missing values. This dataset contains 163 crops which are produced in 59 African countries and has the following features:

- Data Set Characteristics: Multivariate,
- Attribute Characteristics: nominal and numeric,
- Associated Tasks: classification,
- Number of Instances: 9919,
- Number of Attributes: 115,
- Missing Values? Yes,
- Area: Life,
- Date Donated: 1961-2014.

The dataset is in .csv extension and is applied on WEKA software. The dataset is downloaded from the following source
fao.org/Portals/_Faostat/Downloads/zip_files/Production_Crops_E_Africa_1.zip.. (10)

IV. Experimental Results and Discussions

The experimental results are shown with 10 fold cross validation to avoid overlapping. Five classifiers are examined namely tree based J48, rules based decision table and PART, Bayesian based naïve bayes, and lazy based IKB. The performance of the classifier is analyzed in terms of correctly classified instances, incorrectly classified instances, Kappa statistic, mean absolute error, root mean squared error, relative absolute error, and root relative squared error. Table (1) and figure (1) show the performance of the five algorithms according to country.

Table 1 Performance Error of Different Classification algorithms According to the Country

Performance of Algorithm	PART	J48	NaiveBayes	DecisionTable	IBK
Mean absolute error	0	0	0.0234	0.0088	0.0257
Root mean squared error	0.0026	0.0026	0.151	0.036	0.1595
Relative absolute error	0.0226 %	0.0224 %	117.1924 %	26.6482 %	77.266 %
Root relative squared error	2.0104 %	2.0008 %	70.5828 %	27.9117 %	123.7828 %
Kappa statistic	0.9998	0.9998	0.2991	0.9998	0.2288

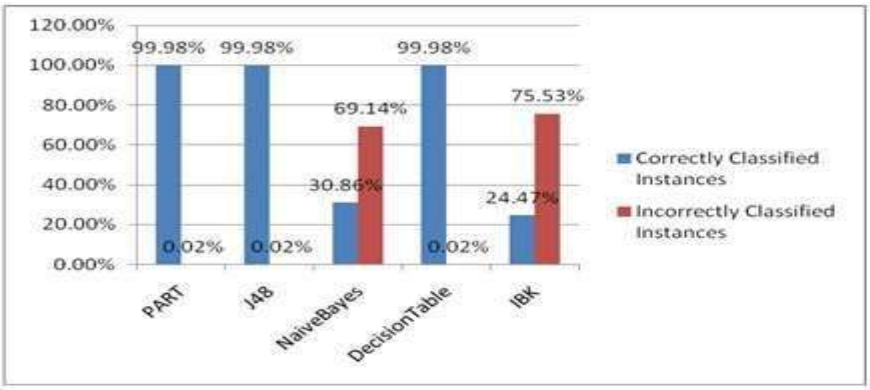

Figure 1 Accuracy of Different Classification algorithms According to the Country

From Table (1), it is observed that the rules and tree classifiers have better performance than the other classifiers. In rules classifiers two algorithms are examined which are PART and Decision Table and from table (1) PART has better performance the Decision Table. Also, J48 is slightly better than PART, therefore we can say that J48 is the best classifier to classify this agriculture dataset. Figure (1) show that J48, PART, and Decision Table are correctly classified with (99.9798%) and Naïve Bayes correctly classified is 30.86 %, and IBK correctly classified is 24.4682% hence Naïve Bayes and IBK give less classification accuracy. Table (2) and figure (2) represent the performance of the five algorithms according to the area harvested, the yield, the seed, and the production.

Table 2 Performance Error of Different Classification algorithms According to Area harvested, Production, Yield, and Seed

Performance of Algorithm	PART	J48	NaiveBayes	DecisionTable	IBK
Mean absolute error	0	0	0.0737	0.0007	0.3182
Root mean squared error	0	0	0.2675	0.0009	0.7157
Relative absolute error	0%	0%	20.4472 %	0.1861 %	28.1257 %
Root relative squared error	0%	0%	63.0162 %	0.2025 %	74.9395 %
Kappa statistic	1	1	0.8023	1	0.1014

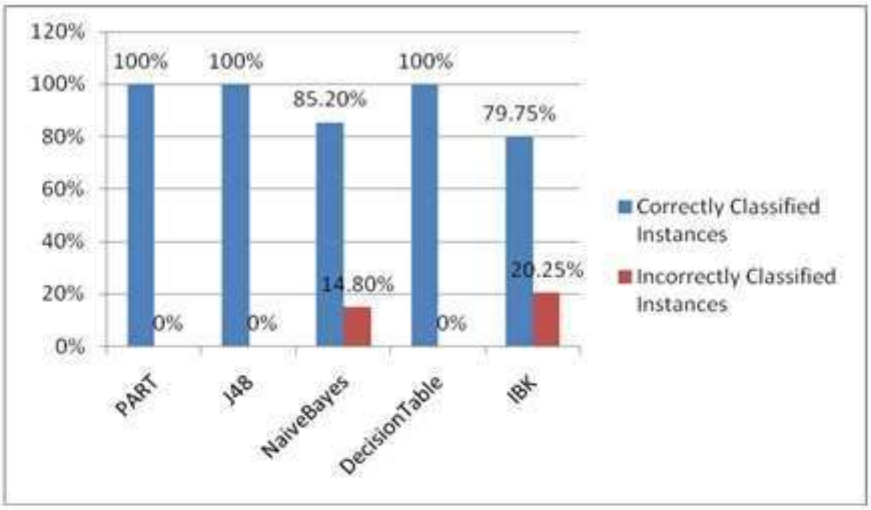

Figure 2 Accuracy of Different Classification algorithms According to Area harvested, Production, Yield, and Seed

From tables (2) it is observed that NaïveBayes and IBK algorithms have higher error rate than the J48, PART, and Decision Table classification algorithms. Figure (2) shows that J48, PART, and Decision Table are correctly classified with no error, while naïve Bayes and IBK are correctly classified with 85.2001% and 79.7459% respectively. It is clear that PART performs slightly better than Decision Table and J48 algorithm is the best one.

V. Conclusion

This research focuses on finding the best algorithm between J48, Decision Table, PART, Naivebayes, and IKB to enhance the classification of African's agriculture dataset. The algorithms are used to classify the area harvested, seed, yield, production, and African countries with their production. From the experimental results the rules and tree classifiers have minimum error rate than the other classifiers. By analyzing the overall experimentation results of the production dataset, it is concluded that J48 algorithm has produced the best classification performance than the IKB and NaïveBayes algorithms and has produced slightly difference performance with Decision Table, PART classifiers. We can say that the performance of the rules and tree classification algorithms are suitable for this dataset to give optimal classification.

References

[1]. Asir, D., Gnana Singh, A., Jebamalar Leavline, E., Priyanka, V., and Swathi, V., "Agriculture Classification System Using Differential Evolution Algorithm", International Advanced Research Journal in Science, Engineering and Technology, 3(5), May 2016, pp. 24-28.

[2]. Thakkar, R., Kayasth, M., and Desai, H. "Rule Based and Association Rule Mining On Agriculture Dataset", International Journal of Innovative Research in Computer and Communication Engineering, 2(11), November 2014, pp. 6381-6384.

[3]. Portmann, F., Siebert, S., Bauer, C., and Döll, P. "Global data set of monthly growing areas of 26 irrigated crops", Frankfurt Hydrology Paper 6 Institute of Physical Geography University of Frankfurt (Main),Germany March 2008, pp.1-400.

[4]. Peña, J., Gutiérrez, P., Martínez, C., Six, J., Richard E. and Granados, F." Object-Based Image Classification of Summer Crops with Machine Learning Methods" Remote Sening, 2014, 6, pp. 5019-5041; doi:10.3390/rs6065019

[5]. Ustunera, M., Sanli, F., Abdikan, S., Esetlili, M., and Kurucu, Y. " Crop Type Classification Using Vegetation Indices Of Rapideye Imagery", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-7, 2014, ISPRS Technical Commission VII Symposium, 29 September – 2 October 2014, Istanbul, Turkey, pp. 195-198.

[6]. Chouhan, S., Singh, D., and Singh, A. "An Improved Feature Selection and Classification using Decision Tree for Crop Datasets", International Journal of Computer Applications (0975 – 8887) 142(13), May 2016, pp.5-8.

[7]. Graves, S., Asner, G., Martin, R., Anderson, C., Colgan, M., Kalantari, L., and Bohlman, S. "Tree Species Abundance Predictions in a Tropical Agricultural Landscape with a Supervised Classification Model and Imbalanced Data", Remote Sensing, 2016, 8(161) doi:10.3390/rs8020161, pp.1-22.

[8]. Chouhan, S., Singh, D., and Singh, A. "A Survey and Analysis of Various Agricultural Crops Classification Techniques", International Journal of Computer Applications (0975 – 8887), 136(11), February 2016, pp. 25-30.

[9]. Parthiban, C. and Balakrishnan, M. "Expert System for Land Suitability Evaluation

[10]. using Data mining's Classification Techniques: a Comparative Study", International Journal of Computer Trends and Technology (IJCTT), 33(2) March 2016, pp. 87-92.

[11]. Babu, S., Ananthanarayanan, N., and Ramesh, V. "A Study on Efficiency of Decision Tree and Multi Layer Perceptron to Predict the Customer Churn in Telecommunication using WEKA", International Journal of Computer Applications (0975 – 8887), 140(4), April 2016, pp. 26-30.

[12]. Faimida M. Sayyad, Exploring Novel Method of Using ICT for Rural Agriculture Development. International Journal of Computer Engineering and Technology (IJCET), 4(1), 2013, pp. 430–437.

[13]. Waldner, F., Lambert, M., Li, W., Weiss, M., Demarez, V., Morin, D., Sicre, C., Hagolle, O., Baret, F., and Defourny, P."Land Cover and Crop Type Classification along the Season Based on Biophysical Variables Retrieved from Multi-Sensor High-Resolution Time Series", Remote Sensing, 2015, 7, pp.10400-10424.

[14]. Vamanan, R. and Ramar, K. "Classification of Agricultural Land Soils A Data Mining Approach", International Journal on Computer Science and Engineering (IJCSE), 3(1) Jan 2011, pp.379-384.

[15]. Mader,S., Vohland, M., Jarmer, T., and Casper, M. "Crop Classification With Hyperspectral Data of the Hymap Sensor Using Different Feature Extraction Techniques", Proceedings of the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover, Center for Remote Sensing of Land Surfaces, Bonn, 28-30 September 2006, pp.96-101.

[16]. Shrinivas R. Zanwar and Prof R. D. Kokate, Advanced Agriculture System. International Journal of Advanced Research in Engineering and Technology (IJARET), 3(2), 2012, pp. 111–118.

[17]. Mishra, V., Kumar, P., Gupta, D., and Prasad, R."Classification Of Various Land Features Using RISAT-1 Dual Polari metric Data", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-8, 2014, ISPRS Technical Commission VIII Symposium, 09 – 12 December 2014, Hyderabad, India, pp. 833-837.

[18]. Bhuyar, V. "Comparative Analysis of Classification Techniques on Soil Data to Predict Fertility Rate for Aurangabad District", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 3(2), March – April 2014, pp. 200-203.