Applying Data Mining Techniques to Predict Yield of Rice in Humid Subtropical Climatic Zone of India

Niketa Gandhi
Department of Computer Science
University of Mumbai, Mumbai, India
Email Id: niketa@gmail.com

Leisa Armstrong
Department of Computer Science,
University of Mumbai &School of Science,
Edith Cowan University, Perth, Australia
Email Id:l.armstrong@ecu.edu.au

Abstract – **Agricultural crop productivity depends on various factors such as precipitation, climate, soil type and hydrology. Statistical methods and techniques can be used to assess the impact of these factors on crop production. By applying these techniques on historical climate and crop production data it may provide knowledge which can be used by farmers or industry and government stakeholders for strategies which will lead to increased crop production. The present research focuses on application of data mining techniques to extract knowledge from the historical agricultural dataset to predict rice crop yield for Kharif season of Humid Subtropical climatic zone of India. The performance evaluation of different classification techniques has been carried out in free and open source data mining software WEKA (Waikato Environment for Knowledge Analysis) as part of this research on the agricultural dataset. The experimental result provided include sensitivity, specificity, accuracy, mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE). The result on the current dataset shows that J48 and LADTree(Logical Analysis of Data Tree) classifiers provide the best performanceof the classifiers used in this research.**

Keywords – data mining; classifiers; crop analysis;IBk; J48; LADTree; LWL; WEKA; yield prediction.

I. INTRODUCTION

In today's era, decision makers often have access to enormous amount of information from different data sources. Breton et al., argue that this approach is not necessarily the most effective to support the decision maker [1]. All information may not be required to adequately execute the task. Further the limitations in human processing capabilities and the resources required to process all this information might also be problematic. Hall et al., and Salemo et al. also stated that for such large quantities of data, the decision maker would need automated or semi-automated decision support system (DSS)[2,3]. In ever changing agricultural context, a DSS needs to be flexible, easy to use and robust enough to perform well over a large set of agricultural and climatic problems to be useful.

There are various factors which can influenceIndianagricultural crop productivity. This results in Indian farmers having to make an important but difficult tactical decision, which determines the success of their crops and ultimately their livelihood. The exposure to excessive amounts of information may not necessarily help them to make a better decision in terms of increasing the crop productivity for their particular farm. A better way would be the presentation of relevantdata in the processed form and generation of knowledge from the historical datasets which may help farmer see how different climatic agricultural scenarios may affect their particular crop and allow them to make better informed decisions about seeding, fertilizer, pesticides and harvestingtasks.A DSS tool which uses basic inputs from the farmers or decision makers and provides a simple prediction of crop yield under a particular scenarioon the basis of historical climate and crop production data can assist farmer decision making and thereby contribute to improving India's food security and Indian farmer livelihood support.

The present research focuses on predicting the rice crop yield through the use of the historical data of various factors affecting the crop productivity for the Humid Subtropical climatic zone of India by applying data mining techniques. Many research studies have shown the potential of applying data mining techniques in agriculture [4, 5, 6, 7, 8, 9, 10].

II. RELATED WORK

The ability to predict the future crop yield enables farmers or decision makers to take the most appropriate decision for that crop. Many researchstudies showed the potential for machine learning and other techniques to be used for prediction of agricultural production systems.Several crop yield models have been developed in relation to different parameters as influencing factors by applications of data mining techniques, artificial neural networks and statistical techniques. Many researchers have applied data mining techniques for crop yield prediction under a particular scenario and have achieved good results [11, 12, 13, 14, 15].

978-9-3805-4421-2/16/$31.00 ©2016 IEEE

A study by Putch et al., used a backpropagation network to predict rice yield using climatic observation data and predicted with a maximum of 45-60kg/ha [16]. Another study by Ji et al., used neural networks to predict rice yield based on soil parameters and achieved a testing error of 17.3% [17]. Liu et al., used back propagation network to predict rice yield based on weather data [18].

Work by O'Neal et al., described the use of artificial neural networks to predict maize yield based on rainfall, soil and other parameters and obtained a testing error of 14.8% [19]. Mehta et al., did a study on impact of rainfall on crop productivity for Targhadia, Gujarat, India for dataset of 39 years from 1958 to 1996. To derive relation between rainfall and yield, correlation and regression methods were applied. The results were used to develop yield prediction model for major crops [20]. Jim et al., used a neural network to build a hybrid prediction model and performed a prediction experiment on the precipitation in the northern, central and southern regions of the Guangxi province [21]. Nakornphanom et al., conducted research on a principal-component based neural network model and applied it to a water level prediction [22].Dahikar et al., conducted research on agricultural crop yield prediction using artificial neural network using feed forward back propagation[23]. Ahamed et al., applied data mining techniques on environmental variables and biotic input variables to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh [24].

The necessity of the present research were to: (1) identify whether data mining techniques could effectively predict rice yield for typical climatic conditions of the Humid Subtropical climatic zone of India, (2) evaluate the data mining techniques performance with standard evaluators, (3) compare the effectiveness of multiple data mining techniques. Therefore, the present research aims to develop a prediction model for rice yield in the Humid Subtropical climatic zone of India using data mining techniques as an alternate technique for yield prediction.

III. RESEARCH METHODS

A. Study Area

The study area for this research was the India Humid Subtropical climatic zone, based on the Koppen climate classification system [25]. This is one of the climatic zone which covers majority of rice production in the Kharifseason (June to November) of India. All Indian states that fall under this climatic zone were included in this research. There were some states that partially fall under this climatic zone and the remaining part in a different climatic zone. In such cases the districts were selected that fall under Humid Subtropical climatic zone. FifteenIndianstates, which fall under the Humid Subtropical climatic zone,wereselectedfor the present research. Depending on the historic data availability for rice crop produced in Kharifseason, five years data from 1998 to 2002 of various factors viz. precipitation climatic, soil type, physiography and crop yield for 207 districts were selected from these fifteenstates for the present research.

B. Dataset Used

All the datasets used in the present research were sourced from publicly available Indian Government records. Soil type data was acquired from National Bureau of Soil Survey and Land Use Planning (NBSS&LUP), Nagpur, Maharashtra. From the vast initial dataset, only a limited number of important factorswhich have the highest impact on agricultural production were selected for the present research. All the factors were considered for a period of five years from 1998 to 2002. The factors selected are:

i) Precipitation (mm): The total precipitation for Kharifseason (June to November) for each year of every district was calculated from the monthly mean precipitation of that year for a particular district.

ii) Minimum, Average, Maximum Temperature (degree Celsius): Variation in temperaturewould surely have animpact on the crop production. Hence maximum, average and minimum temperature for each year of every district was considered for the present research. The average temperatures for the Kharifseason (June to November) were calculated from the monthly mean temperature for minimum, average and maximum temperatures of that year for every district.

iii) Soil Type: Soil type for each district was identified from the hard copy state maps received from the National Bureau of Soil Survey and Land Use Planning (NBSS&LUP), Nagpur, Maharashtra. These state maps have map ids for each district located on the map and the type of soil is color coded. These soil types are divided in to eight major classes and each class has further subclasses. The major soil classes are Aridisols, Ultisols, Mollisols, Alfisols, Vertisols, Inceptisols, Entisols and Others. Each of the major soil class has been given a unique number from 1 to 8.

iv) Area (Hectares): The rice cultivated area in Kharifseason (June to November) for every year in each selected district of fifteen Indian states was considered for the present research.

v) Production (Tonnes): The rice production for the above cultivated area for Kharifseason (June to November) for every year in each selected district of fifteen Indian states was considered for the present research.

vi) Yield (Tonnes/Hectare): Depending on the rice production and the area cultivated for rice in Kharifseason, for every year of each of the selected district from the fifteen states, the calculated yield was considered for the present research.

The data was then integrated with all the parameters inMicrosoftExcel. The raw dataset consisted of the following columns in Microsoft Excel: sr. no, name of the state, name of the district, year, precipitation, minimum temperature, average temperature, maximum temperature, soil type, area, production and yield.Extensivepre-processing was required to handle the missing values and other data inconsistencies. There were total of 866 records after this pre-processing was completed. For preparing the dataset for applying data mining techniques, unrequired columns were removed. They were sr. no, name of the state, name of the district, year, area and production. Since the present research is based on the Humid Subtropical climatic zone of India, there was no need to individual state and district names. Also the yield data is calculated from the area and production, those columns were not required. The dataset was then sorted on the basis of yield to classify the records in to low, moderate and high yield of rice production in the selected 207 districts of fifteen states within the Humid Subtropical climatic zone of India. Depending on the data set the classes for yield were defined as low with the range of 0.1-1 tonnes/hectare, moderate with the range of 1-2 tonnes/hectare and high with the range of 2-5 tonnes/hectare. There were total of 177 records for class low, 367 records for class moderate and 322 records for class high. This data set was then saved in .csv format for further applying data mining techniques.

C. Methodology Used

*1. WEKA(Waikato Environment for Knowledge Analysis)[26]*isfreely available and open source data mining tool available under the GNU General Public License. It was developed at the University of Waikato, New Zealand and was first implemented in 1997 in its modern form. It is possible to run this software on any modern computing platform as it has been implemented in the Java programming language. It contains a GUI (Graphical User Interface) for easier interaction with data files and producing visual results. Further, the GUI also makes it easier to use the comprehensive collection of data preprocessing and modeling techniques. WEKA can be embedded like any other library, in other applications as it has a general API.

WEKA supports the standard data mining tasks like data preprocessing, clustering, classification, regression, visualization, and feature selection. Assuming that the availability of the data is in a single flat file or relation with each data point is described by a fixed number of attributes, all of WEKA's techniques are predicated.Access to SQL databases is provided using Java Database connectivity which helps to process the result returned by a database query. The dataset prepared for the present research was then opened in WEKA and save in .arff format for further processing.

2. Classification Algorithms – An Overview

An important component of machine learning algorithm is use of a classification technique. From the existing data it helps to extract rules and patterns that can be used for prediction. A technique of mapping data records into one of several predefined classes is classification. Classification is a method of finding a set of models that describe anddistinguish data classes and concepts.The aim is to utilize the model and label the unknown objects [27].

Training and testing phase is followed for building a classifier. A classification model is constructed in the training phase. Individual objects are collectively used as training data. The classification of training data is carried out before building the model. Every object is attached with a class label. In testing phase this model can be used for classification. It provides the estimated value of the predictive accuracy of the classifier. To measure a test set made up of test tuples and their associated class labels is used. The selections of these tuples are random from the general data sets. These tuples are not involved while building the classification model earlier.

Various techniques for classification are used from machine learning, statistics, information retrieval and data mining.Some of them are Support Vector Machines (SVM), Bayesian Methods, Bayesian Belief networks, Decision Trees and Neural Networks.

The present study aims to compare the performance of four classification algorithms such as J48 and LADTree from Trees sub menu and IBK and LWL from lazy sub menu in WEKA [28]. J48 is a decision tree classifier [29, 30]. Decision tree is a predictive machine-learning model. The classifier for binary target variable is LAD (Logical Analysis of Data) Tree [31]. IBK (Instance based) is a K-NN (K- Nearest Neighbour) classifier [32, 33]. It is a supervised learning algorithm. LWL is one of the Lazy learning methods that defer processing of training data until a query needs to be answered[34, 35, 36]. LWL (Locally Weighted Learning) is lazy classifier that uses statistical learning techniques for training and classifying complex tasks.

IV. PERFORMANCE EVALUATION

The standard methods are used for analysis and evaluation of the performance of all the classifiers. Initially sensitivity, specificity and accuracy are used. True Positive (TP, a number of correctly classified that an instances positive), False Positive (FP, a number of incorrectly classified that an instance is positive), False Negative (FN, a number of incorrectly classified that an instance is negative) and True Negative (TN, a number of correctly classified that an instance is negative) are used to calculate all the measures. These values are defined in Table I below.

TABLE I.

True Class	Yes	No	Total
Yes	TP	FN	TP+FN
No	FP	TN	FP+TN
Total	TP+FP	FN+TN	TP+FN+FP +TN

From these quantities, the sensitivity and specificity can be calculated by using Eq. (1) and (2) respectively.

$$\text{Sensitivity} = TP / (TP + FN) \qquad (1)$$

$$\text{Specificity} = TP / (TP + FP) \qquad (2)$$

Sensitivity is defined as percentage of correctly classified instances. Specificity is defined as percentage of incorrectly classified instances.

Accuracy is defined as the overall success rate of the classifier and computed by using Eq. (3).

$$\text{Accuracy} = (TP + TN) / (TP + FP + FN + TN) \qquad (3)$$

Further itshows the relative mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE) and root relative squared error (RRSE) for reference and evaluation.

V. EXPERIMENTAL RESULTS

This section discusses the results achieved on rice crop yield data set of Humid Subtropical zone of India after executing the four classification techniques. The algorithms were executed in WEKA. The result obtained for each classification algorithms are shown and described below.

A. Result for classification using J48 algorithm

Using J48 algorithm in WEKA, classification for generating a pruned or unpruneddecision tree was carried with the following parameters: binary splits on nominal attributes when building the trees = false; the confidence factor used for pruning = 0.25; debug = false; the minimum number of instances per leaf = 2; the amount of data used for reduced error pruning = 3, one fold is used for pruning, the rest for growing the tree; reduced pruning error is used = false; to save the training data for visualization = false; the seed used for randomizing the data when reduced error pruning is used = 1; to consider the subtree raising operation when pruning = true; unpruned = false; use Laplace = false. J48 algorithm achieves accuracy of 100%, sensitivity of 100% and specificity of 100%. It gives a mean absolute error and root mean squared error of 0 and relative absolute error and root relative squared error of 0%.

B. Result for classification using LADTree algorithm

Classification by generating a multi-class alternating decision tree using the LogitBooststrategywas executed in WEKA using the LADTree algorithm with the following parameters: the number of boosting iterations to use which determines the size of the tree = 10; debug = false. LADTree algorithm achieves accuracy of 100%, sensitivity of 100% and specificity of 100%. It gives a mean absolute error and root mean squared error of 0.0001, relative absolute error of 0.0188% and root relative squared error of 0.0233%.

C. Result for classification using IBk algorithm

K-nearest neighbour classifier can be used to select appropriate value of K based on cross validation and can be used for distance weighing. The IBk algorithm was executed in WEKA with the following parameters: kNN, number of neighbours to use = 1, whether hold-one-out cross validation will be used to select the best k value between 1 and the value specified as the kNN parameter = false; debug = false; use of distance weighing method = no; mean squared error is used rather than mean absolute error when doing cross validation for regression problems = false; the nearest neighbour search algorithm to use with distancefunction = EuclideanDistance-R first-last; the maximum number of instances allowed in the training pool = 0, A value of 0 signifies no limit to the number of training instances. The addition of new instances above this value will result in old instances being removed. IBk algorithm achieves accuracy of 95.612%, sensitivity of 93.01% and specificity of 96.51%. It gives a mean absolute error of 0.0454, root mean squared error of 0.2091, relative absolute error of 10.6359% and root relative squared error of 45.2543%.

D. Result for classification using LWL algorithm

Classification using locally weighted learning can be carried out by using an instance based algorithm to assign instance weights which are then used by a specified weighted instance handler. The LWL algorithm was executed in WEKA with the following parameters: kNN, number of neighbours used to determine the width of the weighting function = -1, less than equal to 0 means all neighbours; the base classifier used = decision stump; debug = false; the nearest neighbour search algorithm to use with distancefunction = EuclideanDistance-R first-last; weighting kernel which determines weighting function = 0, 0 means linear. LWL algorithm achieves accuracy of 89.299%, sensitivity of 73.82% and specificity of 90.71%. It gives a mean absolute error of 0.1636, root mean squared error of 0.2727, relative absolute error of 38.3222% and root relative squared error of 59.0183%.

VI. DISCUSSION ANDCONCLUSIONS

One of the goals of agricultural production is achieving maximum crop yield at minimum cost. Rice is one of the most important food crops in the world and a primary source of food for more than half of the world's population. Over 90% of the rice is produced and consumed in Asian region [37]. To meet this demand, it becomes increasingly important to study the factors impacting the rice production. Variations inclimaticfactors,precipitation, soilandhydrology can determine the success of rice crop production in a particular region or climatic zone. Indian farmers are faced with the challenges in making decisions about which crop and agricultural practices are best suited to their situation. These challenges may be overcome with a good prediction model that can predict rice crop yield under different climatic and geographic regions.

In the present research, four classifiers namely J48, LADTree, IBk and LWL have been used. Among the classifiers used, our experimental results show that J48 and LADTree achieve highest accuracy, sensitivity and specificity. On the other hand, the worst classification was performed by LWL classifier with lowest accuracy, sensitivity and specificity. The summary of all the results of the classifiers compared is shown in the Table II below. The summary of all the error results of the classifiers compared is shown in Table III below.

TABLE II.

Classifier	Accuracy	Sensitivity	Specificity
J48	100	100%	100%
LADTree	100	100%	100%
IBk	95.612	93.01%	96.51%
LWL	89.299	73.82%	90.71%

TABLE III. .

Classifier	MAE	RMSE	RAE	RRSE
J48	0	0	0%	0%
LADTree	0.0001	0.0001	0.0188%	0.0233%
IBk	0.0454	0.2091	10.6359%	45.2543%
LWL	0.1636	0.2727	38.3222%	59.0183%

To help increase the crop yield and subsequent profit, it is important for early detection and management of problems associated with crop yield. Predictions can help the farmers or decision makers to minimize losses when unfavorable conditions may occur. Additionally, these predictions can help maximize crop production when potential exists for favorable growing conditions. These decisions can be easily taken with the help of decision support system (DSS) if developed for a particular crop under a specific climatic scenario; hence the need and importance of DSS in the agricultural domain is increasing.

REFERENCES

[1] R. Breton, S. Paradis and J. Roy, "Command Decision Support Interface (CODSI) for human factors and display concept validation", *International Conference on Information Fusion*, pp. 1284 - 1291, 2002.

[2] D. L. Hall and J. Llinas - Handbook of multisensor data fusion; CRC Press, Boca Raton, Fla., 2001.

[3] J. Salemo, M. Hinman, D. Boulware and P. Bello, "Information Fusion for Situational Awareness", *International Conference of Information Fusion*, pp. 507 - 513, 2003.

[4] M.C.S.Geetha, "A survey on data mining techniques in Agriculture", *International Journal of Innovative Research in Computer and Communication Engineering",* Vol. 3, No. 2, pp. 887-892, February 2015.

[5] A.B.Mankar and M.S.Burange, "Data Mining – An evolutionary view of Agriculture", *International Journal of Application or Innovation in Engineering and Management,* Vol. 3, No. 3, pp. 102-105, March 2014.

[6] R.Kalpana, N.Shanthi and S.Arumugam, "A survey on data mining techniques in agriculture", *International Journal of Advances in Computer Science and Technology,* Vol. 3, No. 8, pp. 426-431, August 2014.

[7] A.A.Raorane and R.V.Kulkarni, "Review-Role of data mining in Agriculture", *International Journal of Computer Science and Information Technologies,*Vol. 4, No. 2, pp. 270-272, 2013.

[8] G.N.Fathima and R.Geetha, "Agriculture crop pattern using data mining techniques", *International Journal of Advanced Research in Computer Science and Software Engineering,* Vol. 4, No.5, pp. 781-786, May 2014.

[9] H.Patel and D.Patel, "A brief survey of data mining techniques applied to agricultural data", *International Journal of Computer Applications,* Vol. 95, No. 9, pp. 6-8, June 2014.

[10] M.Kaur, H.Gulati and H.Kundra, "Data mining in Agriculture on crop price prediction: Techniques and Applications", *International Journal of Computer Applications,* Vol. 99, No. 12, August 2014.

[11] D.Ramesh and B.V.Vardhan, "Analysis of crop yield prediction using data mining techniques", *International Journal of Research in Engineering and Technology,* Vol. 4, No. 1, pp. 47-473, January 2015.

[12] A.A.Raorane and R.V.Kulkarni, "Data Mining: An effective tool for yield estimation in the agricultural sector", *International Journal of Emerging Trends and Technology in Computer Science,* Vol. 1, No. 2, pp. 75-79, July-August 2012.

[13] D.Ramesh and B.V.Vardhan, "Data mining techniques and applications to agricultural yield data", *International Journal of Advanced Research in Computer and Communication Engineering,* Vol. 2, No. 9,pp. 3477-3480, September 2013.

[14] S.Veenadhari, B.Misra and C.D.Singh, "Data mining techniques for predicting crop productivity – A review article", *International Journal of Computer Science and Technology,* Vol. 2, No. 1, pp. 98-100, March 2011.

[15] R.A.Medar and V.S.Rajpurohit, "A survey on data mining techniques for crop yield prediction", *International Journal of Advance Research in Computer Science and Management Studies,* Vol. 2, No. 9, pp. 59-64, September 2014.

[16] S. Puteh, M. Rizon, M. Juhari, J.Nor Khairah, S. Siti Kamarudin, B.Aryati, and R. Nursalasawati, "Backpropagation algorithm for riceyield prediction",*Proceedings of9th of the Ninth International Symposium on Artificial Lifeand Robotics,* Beppu, Japan, Oita, pp. 586-589, 2004.

[17] B. Ji, Y. Sun, S. Yang, and J. Wan, "Artificial neural networks for riceyield prediction in mountainous regions",*Journal of AgriculturalScience*, Vol. 145, pp. 249-261, May 2007.

[18] J. Liu, C. E. Goering, and L. Tian, "A neural network for setting targetcorn yields", *Transaction of the ASAE*, Vol. 44, No. 3, pp. 705-713, July2001.

[19] M. R. O'Neal, B. A. Engel, D. R. Ess, and J. R. Frankenberger, "Neuralnetwork prediction of maize yield using alternative data codingalgorithms",*Biosystems Engineering*, Vol. 83, No.1, pp. 31-45, May2002.

[20] D.R.Mehta, A.D.Kalola, A.D.Saradava, A.S.Yusufzai, "Rainfall variability analysis and its impact on crop productivity – A case study", *Indian Journal of Agricultural Research,* Vol. 36, No. 1, pp. 29-33, 2002.

[21] L. Jim, G. Randolph, and M. Mary, "On the construction and trainingof reformulated radial basis function neural networks",*IEEE TransactionsonNeural Networks*, Vol. 14, July 2003.

[22] K. Nakornphanom, C. Lursinsap, and A. Rugchatjaroen, "Faultimmunization model for elliptic radial basis function neuron", *Proceedings of 9thInternational Conference on Neural Information Processing*,pp.1027-1031, *2002*.

[23] S.S.Dahikar and S.V.Rode, "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach", *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering,* Vol. 2, No. 1, January 2014.

[24] A.T.M.S.Ahamed, N.T.Mahmood, N.Hossain, M.T.Kabir, K.Das, F.Rahman, R.M.Rahman, "Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh", *16th IEEE/ACIS International*

Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing(SNPD), Japan, June 2015.

[25] M. C.Peel, B.L.Finlayson and T.A.McMahon, “Updated world map of the Köppen-Geiger climate classification”, *Hydrology Earth System Sciences*, Vol. 11, 2007.

[26] Weka 3: Data Mining Software in Java, *Machine Learning Group at the University of Waikato*, Official Website: http://www.cs.waikato.ac.nz/ml/weka/index.html, accessed on 6th November 2015.

[27] J. Han, M. Kamber, J. Pei, “Data Mining: Concepts andTechniques”, Third edition, Elsevier and Morgan Kauffman Publishers, 2012.

[28] I.H.Witten, “Data Mining with Weka”, *Class Lesson,* Department of Computer Science, University of Waikato, New Zealand.

[29] F. Kiamarzpour, R. Dianat, M.Bahrani, M. Sadeghzadeh, “Improving the methods of email classificationbased on words ontology”,*arXiv preprint,* arXiv:1310.5963, 2013.

[30] J. Ross, Quinlan, “C4.5: programs for machine learning” *Elsevier*, 2014.

[31] D. Buhmann, “Radial Basis Functions: Theory and Implementations”,*Cambridge Monographs on Applied and Computational Mathematics*, Cambridge University Press, Cambridge, 2003.

[32] G. Antonion, G. Potamias, C. Spyropoulos, D.Plexousakis, “Advances in Aritificial Intelligence”, *4thHelenic Conference on AI*, Greece, Springer Proceedings, May 2006.

[33] D.W.Aha, D. Kibler, M.K. Albert, “Instance-based learning algorithms”, *Journal of Machine Learning*, Vol. 6, No. 1, pp. 37, 1991.

[34] C. G. Atkeson, A. W. Moore, and S. Schaal, “Locally weighted learning”, *Artificial Intelligence Review*, Vol. 11, pp. 11-73, 1997.

[35] C. G. Atkeson, A. W. Moore, and S. Schaal, “Locally weighted learning for control”,*Artificial Intelligence Review*, Vol. 3, pp. 913-918,1997.

[36] R. Beichel and M. Sonka, “Computer vision approaches to medical image analysis”, *Lecture Notes in Computer Science*, Springer, 4241, 2006.

[37] M.K.Papademetriou, “Rice production in the Asia-Pacific region: Issues and Perspectives”, *Food and Agriculture Organization of the United Nations,* October 2000.