

A Survey and Analysis of Various Agricultural Crops Classification Techniques

Surabhi Chouhan
Department of Computer
Science
University Institute of
Technology
Barkatullah University, Bhopal

Divakar Singh
Department of Computer
Science
University Institute of
Technology
Barkatullah University, Bhopal

Anju Singh
Department of Computer
Science
University Teaching
Department
Barkatullah University, Bhopal

ABSTRACT

Mining is way of providing and extracting some meaning information from the data so that the data can be classified and grouped easily and quickly. These mining algorithms can be applied in various fields including classification of agricultural crops production. In the fields of Data Mining various efficient algorithms are implemented for the classification of agricultural crops production. Here in this paper a survey of all the existing techniques as well as their advantages and issues are discussed. Hence by analyzing their various advantages and issues a new and efficient technique for the classification of agricultural crops production is proposed in future such as classification using Fuzzy Conclusion Tree by the Optimizing the Feature Withdrawal using PSO-SVM (Particle Swarm Optimization with Support Vector Machine).

Keywords

Keywords are your own designated keywords which can be used for easy location of the manuscript using any search engines.

1. INTRODUCTION

In recent times, the beginning of data mining systems for discovering usage pattern from Web data (Web Usage Mining) indicates that these techniques can be a viable alternative to traditional decision making tools [1]. Web Usage Mining is the process of applying data mining techniques to the discovery of usage patterns from Web data and is targeted towards applications [2]. Web Usage Mining mines the derived data (Web server access logs, browser logs, cookies, user queries, user reports, registration data, user sessions or transactions, mouse clicks and any other data as the consequence of dealings with the Web) resulting from the communications of the users during definite stage of Web sessions. Application of data mining methods to the World Wide Web submitted as Web mining where this expression has been used in three distinctive ways; Web Content Mining, Web Structure Mining and Web Usage Mining.

This has led to an increasing awareness in the improvement of tools proficient in the automatic extraction of knowledge from data. The expression Data Mining, or Knowledge Discovery in Databases, has been accepted for a field of research agreement with the habitual discovery of understood information or knowledge within databases [3]. The implicit information surrounded by databases, and generally the remarkable association bonding among sets of objects, that lead to association rules, may reveal useful patterns for decision support, financial estimate, marketing strategies, even medical analysis and many additional applications. This actuality fascinated a lot of awareness in current data mining

research [3]. A very dominant association rule mining algorithm, Apriori [4], has been build up for rule mining in large transaction databases. Many other algorithms developed are derived and/or conservators of this algorithm. A major walk forward in improving the presentations of these algorithms was made by the preface of a work of fiction, compact data structure, passed on to as frequent pattern tree, or FP-tree [5], and the associated mining algorithm, FP-growth. The main difference between these two approaches is that the Apriori-like methods are based on bottom-up creation of frequent item set arrangements and the FP-tree based ones are partition-based, divide-and-conquer methods. But in the agriculture area, predictive data mining is mainly used. There are two main techniques namely classification and clustering. Data mining techniques has put forward the guidelines for making recommendations for site specific crop management [28].

DATA MINING

Data mining is the method of discovering or fetching useful information from database tables. Many methods to sequential data mining have been proposed to extract useful information, such as time series analysis, temporal association rules mining, and sequential pattern discovery. Several basic techniques are used in data mining for describing the type of mining and data recovery operation. The rapid growth of the Internet could be largely attributed to the loose governance structure, which beyond some control over domain names, is open to be freely added to by anyone with access to a computer and an Internet connection. Data mining is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to either enhance profits, cuts costs, or both.

CLUSTERING

Clustering is an unsupervised learning technique that takes unlabeled data points (data records) and classifies them into different groups or clusters. This is done in such a way that points assigned to the same cluster have high similarity, while the similarity between points assigned to different clusters is low [8]. Different clustering techniques have been developed [8]. The most important types of clustering techniques are hierarchical agglomerative, partition, robust and unsupervised. Clustering is still an important research issue in the data mining, because there is a continuous research in data mining for optimum clusters on spatial data. There are various types of partition based and hierarchal algorithms implemented for clustering and the clusters which are formed based on the density are easy to understand and it does not limit itself to certain shapes of the clusters. Density-based clustering methods try to find clusters based on the density of points in

regions. Intense or dense regions that are accessible from each other are merged to produce clusters. Density-based clustering methods surpass at finding clusters of arbitrary shapes [9]. We here briefly present, in order, the above stated methods of clustering techniques.

AGGLOMERATIVE HIERARCHICAL CLUSTERING

The Agglomerative Hierarchical Clustering (AHC) [22] is a method of cluster analysis which intends to build a hierarchy of groups. The agglomerative one starts with a similarity matrix in which each element shows a similarity degree regarding the other ones. Then a linkage method is successively joining elements, creating a bunch of clusters called Dendrogram. A cut-off point is selected to identify the most important clusters from the tree plot.

FUZZY CLUSTERING

Mohammad El-Helly, Hoda Onsi, Ahmed Rafea and Salwa El-Gammal in their study used fuzzy clustering in detection of leaf spots in cucumber crop [23]. Leaves spots are indicative of plant diseases; earlier leaf batches are examined manually and are then subjected to expert advice. The experts after the proper investigation, used to announce the disease. In their effort,

Mohammad El-Helly, Hoda Onsi, Ahmed Rafea and Salwa El-Gammal, proposed a segmentation technique for identifying leaf batches in cucumber crop, based on fuzzy clustering algorithm.

The first step of image analysis and pattern recognition is the segmentation of image. Segmentation can be viewed as a clustering problem. It is very critical and inevitable component of image analysis and pattern recognition. This is the task which determines the quality of image analysis. Image segmentation is carried out by partitioning the image into homogeneous disjoint regions pertaining to some criterion as intensity or colour and none of the union of any two adjoining region should be homogeneous.

LINEAR DISCRIMINANT ANALYSIS

Next, we used the LDA to test the a priori classification, to observe the most discriminant variables and to obtain the discriminate functions as well. Since there are three groups to test, the model found two discriminant functions to separate the groups. We have used a stepwise discriminant method to select only those variables that have discriminant power, rejecting the redundant variables or with a low discriminant factor. The Linear Discriminant Analysis (LDA) [24] is a statistical method that comes from the multiple regression analysis and its principal objective is to obtain several classification functions that classify an observation in a set.

CLASSIFICATION

Classification is one of the most frequently occurring tasks of human decision making. A classification problem arises when an object needs to be assigned to a predefined class or group according to its characteristics. A classification task is sometimes also referred to as a supervised learning task since the classes or groups are defined beforehand and can be used to steer the learning process. Many decision problems in a variety of domains such as engineering, medical science, human sciences and management science can be considered as classification problems. Popular examples are speech recognition, character recognition, medical diagnosis, bankruptcy prediction and credit scoring.

BAYESIAN NETWORK

A Bayesian network is a graphical model that encodes probabilistic relationships among variables of interest. When used in conjunction with statistical techniques, the graphical model has several advantages for data analysis. One, because the model encodes dependencies among all variables, it readily handles situations where some data entries are missing. Two, a Bayesian network can be used to learn causal relationships and hence can be used to gain understanding about a problem domain and to predict the consequences of intervention. Three, because the model has both a causal and probabilistic semantics, it is an ideal representation for combining prior knowledge (which often comes in causal form) and data. Four, Bayesian statistical methods in conjunction with Bayesian networks offer an efficient and principled approach for avoiding the over fitting of data. Development of a data mining application for agriculture based on Bayesian networks were studied by [25]. According to him, Bayesian network is a powerful tool for dealing uncertainties and widely used in agriculture data sets. He developed the model for agriculture application based on the Bayesian network learning method. The results indicate that Bayesian Networks are a feasible and efficient.

DECISION TREES

A decision tree represents a structure with two types of components [26]:

- leaf nodes that assign class labels to observations;
- Internal nodes that specify tests on individual attributes with one branch and subtree for each outcome of the test.

The tree classifies observations in a top-down manner, starting from the root and working one's way down according to the outcomes of the tests at the internal nodes, until a leaf node has been reached and a class label has been assigned. The tree is then constructed by means of recursive partitioning until the current leaf nodes contain only instances of a single class or until no test offers any improvement. However, since most real-life data sets are noisy, and since in most cases the attributes have limited predictive power, this tree growing strategy often results in a complex tree with many internal nodes that overfits the data.

K-NEAREST NEIGHBOUR

K-nearest neighbour classifiers (KNN) classify a data instance by considering only the k most similar data instances in the training set [27]. The class label is then assigned according to the class of the majority of the k-nearest neighbours.

The choice of k can be made using a cross-validation method. Ties can be avoided by choosing k odd or assigning the observation to the default class. Note that it is important to first normalize the inputs to e.g. mean 0 and standard deviation 1 since they may have been measured in different units.

APRIORI ALGORITHM

Apriori algorithm which is introduced in [6] that is more efficient than AIS by an order of magnitude. The foremost advantage of Apriori is that it incorporates the subset frequency based pruning optimization that means, it only process any item-set whose subsets are frequent also. It utilizes a data structure that is known as hash tree which is used for storing the counters of candidate item-sets. The main drawbacks of Apriori are: i) It performs n passes in excess of the database, where n is the distance end to end of the greatest frequent item-set. The counts of candidate item-sets of length k are obtained in kth pass, ii) it follows a tuple-by-tuple

approach where counters of candidate item-sets are updated after reading individual transaction of whole database so that much redundant work is performed after each individual transaction. Based on this algorithm, lots of new algorithms are deliberated with enhancements and modifications.

FP-TREE ALGORITHM

FP-Tree/Growth Algorithm is proposed in [7] where, after a preprocessing scan over the database, it creates a reduced characterization of the database which is known as FP-tree and then data mining is executed over the FP-tree. Frequent patterns are patterns (such as item sets, subsequences or substructures) that become visible in a data set frequently. In this paper we are interested in using association rule mining for discovering the gene expression patterns those exhibit similar expression levels. One of the oldest, the Apriori algorithm is based upon the Apriori property that for an item set to be frequent; certainly any subset must be frequent. The Apriori algorithm employs an iterative approach known as a level wise search, where k-items are used to explore (k+1) item sets. Apriori algorithm suffers by scanning the database while finding the k-item sets in each and every step. Due to which the processing overhead is drastically increased. FP-growth algorithm which is enhanced version of apriori algorithm gives better performance while scanning the data base.

FP GROWTH ALGORITHM

The FP Growth algorithm operates in the following four modules [10].

- Preprocessing module
- FP Tree an FP Growth Module
- Association Rule Generation
- Results

The preprocessing modules convert the log file, which normally is in ASCII format, into a database like format, which can be processed by the FP Growth algorithm. The 2nd module is performed in two steps.

- FP Tree generation
- Applying FP Growth to generate association rules

2. LITERATURE SURVEY

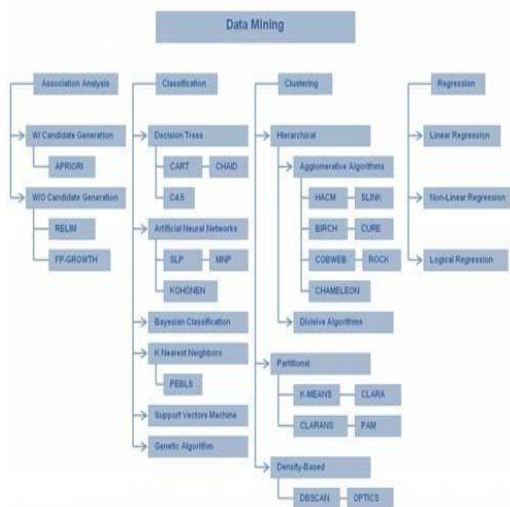


Figure 1. Data Mining Techniques

The main techniques for data mining include Association rules, Classification, Clustering and Regression. The different data mining techniques used for solving different agricultural problem has been discussed [11]. The graphical representation of different data mining techniques is shown in figure 1.

Sanjay D. Sawaitul et al., focuses the information about weather and are observed and stored. The recorded parameters are used to forecast weather. If there is a change in any one of the recorded parameters like wind speed, wind direction, temperature, rainfall, humidity, then the upcoming climatic condition can be predicted using artificial neural networks, back propagation techniques. The increase in signal range will work in large areas as well [12].

Somvanshi, V.K. et al., deliberate the modeling and prediction of rainfall using artificial neural networks and Box- Jenkins methodology. Other applications of artificial neural networks in hydrology are forecasting daily water hassle and flow forecasting [13].

K. Verheyen et al., Data Mining is the process of discovering meaningful patterns and trends by shifting through huge amount of data, using pattern detection technologies as well as statistical and mathematical techniques. Data Mining techniques are often used to studied soil characteristics. As an example, the K-Mean approach is used for classifying soils in combination with GPS based techniques [14].

Urtubia et al., The prediction of wine fermentation problems can be performed by using a k-means approach. Knowing in advance that the wine fermentation process could get jammed or be slow can help the enologist to correct it and ensure a good fermentation process [15].

I. Jagielska et al., describe applications to agricultural related areas. Such as Yield prediction is a very important agricultural problem. Any farmer might be interested in knowing how much yield is expected. In the past, yield prediction was achieved by considering farmer's experience on particular field, crop and climate condition. We have discussed additional information about data like probability in probability theory, grade of membership in fuzzy set theory [16].

Tellaecche et al., detecting weeds in precision agriculture. The paper summarize an automatic computer vision system for the detection and differential spraying of Avena sterilis, a toxic weed growing in cereal crops. With such purpose it have been designed a hybrid decision making system based on the Bayesian and Fuzzy k-Means classifiers, where the a priori probability required by the Bayes framework is supplied by the Fuzzy k-Means [17].

Veenadhari, S. Influence of climatic factors on major kharif and rabi crops production in Bhopal District of Madhya Pradesh State was considered. The findings of the study revealed that the decision tree analysis indicated that the productivity of soybean crop was mostly influenced by comparative humidity followed by temperature and rainfall. The decision tree analysis shows that the productivity of paddy crop was mostly inclined by Rainfall followed by comparative Evaporation and humidity. For Wheat crop, the analysis shows that the productivity is mostly influenced by Temperature followed by relative humidity and rainfall. The result of decision tree was confirmed from Bayesian classification. The rules formed from the decision tree are useful for identifying the conditions intended for high or low crop productivity [18].

Shalvi D and De Claris N Bayesian network is a powerful tool and broadly used in agriculture datasets. The model developed for agriculture application based on the Bayesian network learning method. The results show that Bayesian Networks are feasible and efficient. Bayesian approach improves hydro geological site characterization even when using low-resolution resistivity surveys [19].

Altannar Chinchulunn et al., The k-nearest neighbor classification algorithmic rule may be divided into 2 phases: coaching section and testing section. Bermejo associated Cabestany urged a reconciling learning algorithmic rule to permit fewer information points to be utilized in coaching information set. Several different techniques are projected to scale back procedure burden of k-nearest neighbor algorithms [20].

B. Rajagopalan and U. Lal A number of studies have been carried out on the application of data mining techniques for agricultural data sets. For example, the K-Nearest Neighbor is applied for simulating daily precipitations and other weather variables [21].

According to **Schupp et al.** [29], imaging technology has the potential to cross multiple areas of tree fruit production, such as crop load assessment, including blossom or green fruit counts, yield estimation, determination of insect presence or disease infection and associated eradication, soil moisture content determination for enhanced irrigation system design, estimation of fertilizer, pesticide, and herbicide application rates, and the development of assisted or automated pruning and harvesting strategies. Schupp et al. [29] state that different technological approaches such as satellite imagery, real-time laser images, and video imagery application will prove successful for fruit and nut crop SSCM.

Goel et al. [30] reported a strong correlation between the digital information, e.g., spectral data of the aerial image, and soybean crop physiological parameters such as chlorophyll fluorescence, leaf greenness, leaf area index (LAI), photosynthesis rate, and plant height. In another application, Goel et al. [30] used multi-spectral (24 wave band) airborne optical remote sensing for detecting weed infestation in site-specific managed field crops. They used the multi-spectral band, with a range of 475.12 nm to 910.01 nm, and their results indicated that the wave band centered on 675.98 and 685.17 nm in the red region and 743.93 to 830.43 nm in the NIR region, had good potential for weed classification in a maize field. The Schiffes multiple range tests provided a p-value that was less than 0.05 to support their findings.

3. METHODOLOGY USED IN FUTURE

The main steps that may be implemented are as follows:

1. Initialization of Input Parameters such as No. of Particles and Dimension and Cross Folds validation. Since PSO is applied here for the optimization of selection of features, hence the parameters of PSO need to be initialized such as number of particles and their position velocity and the cross folds used by SVM.
2. For each of the particle and objects choose a random instance values from the dataset and their respective moving position. As soon as the initialization step is over the particles or features of the PIMA Indian dataset is chosen so that the best or dependent attributes are selected.
3. The Particle Current Position can be considered as the Best Position means as the best instance value. The

random feature selected assumes to be the best attribute of the dataset and so is the fitness value as best and selection of features starts from this feature of the dataset.

4. For Each of the particle for max_iterations Now compute for all the features of the PIMA Indian Diabetes Dataset.
5. Selection of the particle position in the dataset as an instance values as (x, y). The feature to be selected (Particle) moves along 'X' and 'Y' axis for the next best feature from the dataset depending upon the fitness value.
6. Now Initialize the Input Parameters of SVM Here initialization of parameters such as type of kernel used and margin width of SVM and number of iterations is done.
7. ker = '@linearKernel' OR 'GaussianKernel' Selection of kernel as linear or Gaussian (RBF) on the basis of which SVM iterates and select features.

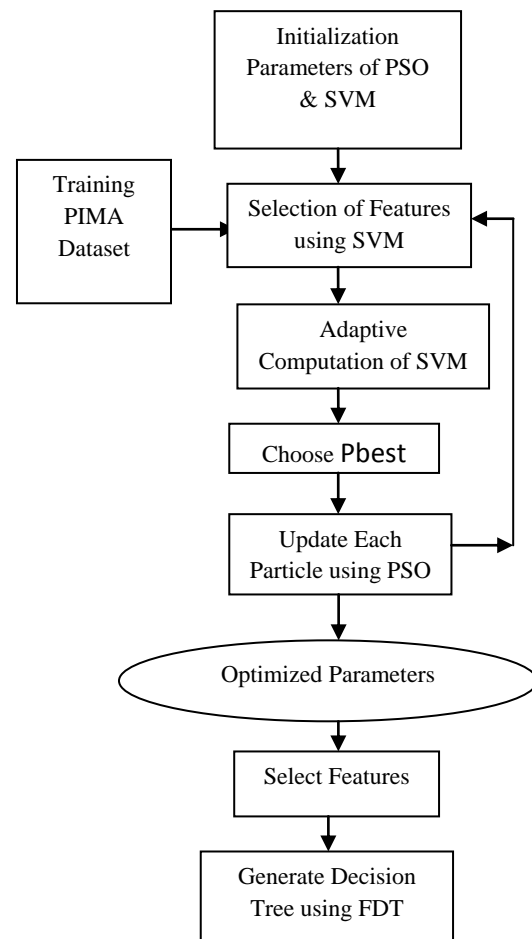


Figure 2 Flow of the Proposed Methodology

8. $p1 = x$; Here 'x' is selected as the particle or feature of the dataset.
9. $C = y$; 'y' is assumed to be the class index of SVM.
10. $trnX=X$; Select the 'X' as the training value of the SVM.
11. $trn Y=Y$; Select Y as the training classes of SVM.
12. $tstX=X'$;

13. $tstY=Y'$;
14. Training is performed on the basis of (trnX, trnY, C)The selection of features starts with the basic input to SVM as the training values and class index.
15. On the basis of Training Parameters as (trnX, trnY, tstX, ker, alpha, bias, actfunc); A predefined function is defined which computes the features on the basis of above parameters.
16. Selection of 'Y' as the features values can be predicted.

4. CONCLUSION

Here in this paper a complete survey and analysis of various agricultural crops classification is done and find their various limitations and issues. The various technique implemented for the analysis of crops production using classification and clustering and association rule mining is analyzed here so that on the basis of their various limitations a new and efficient technique is implemented in future such as classification using Fuzzy Decision Tree by the Optimizing the Feature Extraction using PSO-SVM (Particle Swarm Optimization with Support Vector Machine).

5. REFERENCES

- [1] Geraldin B. Dela Cruz, Bobby D. Gerardo, Bartolome T. Tanguilig III, "Agricultural Crops Classification Models Based on PCA_GA Implementation in Data Mining", International Journal of Modeling & Optimization, Vol. 4, No. 5, 2014.
- [2] Srivastava, Jaideep, Robert Cooley, Mukund Deshpande, and Pang-Ning Tan. "Web usage mining: Discovery and applications of usage patterns from web data." ACM SIGKDD Explorations Newsletter, vol. 1, no. 2, pp. 12-23, 2000.
- [3] U.M. Fayyad, et al.: "From Data Mining to Knowledge Discovery: An Overview", Advances in Knowledge Discovery and Data Mining: 1-34, AAAI Press/ MIT Press, 1996, ISBN 0-262-56097-6.
- [4] Agrawal, Rakesh, and Ramakrishnan Srikant. "Fast algorithms for mining association rules." In Proceedings of 20th international conference on very large data bases, VLDB, vol. 1215, pp. 487-499, 1994.
- [5] Han, Jiawei, Jian Pei, and Yiwen Yin. "Mining frequent patterns without candidate generation." In ACM SIGMOD Record, vol. 29, no. 2, pp. 1-12. ACM, 2000.
- [6] Toivonen, Hannu "Sampling large databases for association rules." In Proceedings of International Conference on very Large Databases (VLDB), vol. 96, pp. 134-145, 1996.
- [7] Han J, Pei J, Yin Y, Mao R, Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach, Data Mining and Knowledge Discovery, Vol 8, Issue 1, Jan 2004, pp 53 – 87, 2004.
- [8] J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques", 3rd edition, The Morgan Kaufmann Series in Data Management Systems, USA, 2006.
- [9] Xin Wang and Howard J. Hamilton "A Comparative Study of Two Density-Based Spatial Clustering Algorithms for Very Large Datasets", Proceedings of the 18th Canadian Society conference on Advances in Artificial Intelligence, pp. 120-132, 2005.
- [10] Santhosh Kumar, B., and K. V. Rukmani. "Implementation of Web Usage Mining Using APRIORI and FP Growth Algorithms." International Journal of Advanced Networking & Applications 1, no. 6 (2010).
- [11] Yethiraj N G, "Applying Data Mining Techniques in the Field Of Agriculture and Allied Sciences", International Journal of Business Intelligents ISSN: 2278-2400, Vol 01, Issue 02, and December 2012.
- [12] Sanjay D. Sawaitul, Prof. K.P. Wagh, Dr. P.N. Chatur, "Classification and Prediction of Future Weather by using Back Propagation Algorithm-An Approach", International Journal of Emerging Technology and Advanced Engineering, Vol. 2, Issue 1, January 2012, pp. 110-113.
- [13] K. SOMVANSHI, ET AL., "MODELING AND PREDICTION OF RAINFALL USING ARTIFICIAL NEURAL NETWORK AND ARIMA TECHNIQUES", J. IND. GEOPHYS. UNION, VOL. 10, NO. 2, PP. 141-151, 2006.
- [14] K. Verheyen, D. Adriaens, M. Hermy, and S. Deckers, "High resolution continuous soil classification using morphological soil profile descriptions", *Geoderma*, vol. 101, pp. 31-48, 2001.
- [15] Urtubia, A., Pérez-Correa, J. R., Soto, A., & Pszczolkowski, P. (2007, "Using data mining techniques to predict industrial wine problem fermentations", *Food Control*, 18(12), 1512-1517.
- [16] I. Jagielska, C. Mattheews, T. Whitfort, "An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems", *Neurocomputing*, Vol. 24, pp. 37-54, 1999.
- [17] Tellaeché, A., BurgosArtizzu, X. P., Pajares, G., & Ribeiro, A. (2007), "A vision-based hybrid classifier for weeds detection in precision agriculture through the Bayesian and Fuzzy k-Means paradigms", In *Innovations in Hybrid Intelligent Systems* (pp. 72-79). Springer Berlin Heidelberg.
- [18] Veenadhari, S. 2007, "Crop productivity mapping based on decision tree and Bayesian classification". Unpublished M.Tech Thesis submitted to Makhlanlal Chaturvedi National University of Journalism and Communication, Bhopal.
- [19] Shalvi D and De Claris N., "Unsupervised neural network approach to medical data mining techniques", in Proceedings of IEEE International Joint Conference on Neural Networks, (Alaska), pp. 171-176, May 1998.
- [20] B. Rajagopalan and U. Lal, "A K-nearest neighbor simulator for daily precipitation and other weather variable", *Water Resources*, vol. 35, pp. 3089-3101, 1999.
- [21] S.Veenadhari, Dr. Bharat Misra, Dr. CD Singh, "Data mining Techniques for Predicting Crop Productivity – A review article", International Journal of Computer Science and Technology IJCST Vol. 2, Issue 1, March 2011.
- [22] Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical Taxonomy*. Freeman: San Francisco.

- [23] Detecting Leaf Spots in Cucumber Crop Using Fuzzy Clustering Algorithm: Mohammed El Helly, Hoda Onsi, Ahmed Rafea, Salwa El-Gamma.
- [24] Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems, *Annals of Eugenics*, 7: 179–188.
- [25] Huang, J., Y. Yuan, W. Cui and Y. Zhan, 2008. IFIP international federation for information processing. http://en.wikipedia.org/wiki/International_Federation_for_Information_Processing.
- [26] J.R. Quinlan. C4.5 programs for machine learning. Morgan Kaufmann, 1993.
- [27] T. Hastie, R. Tibshirani, and J. Friedman. The Elements of Statistical Learning, Data Mining, Inference, and Prediction. Springer, 2001.
- [28] Chosa, T., Y. Shibata, M. Omine, K. Kobayashi, K. Toriyama and R. Sasaki, 2003. Map based variable control system for granule applicator. *J. Jap. Soc. Agric. Machinery*, 65: 128-135.
- [29] Baugher, T.; Schupp, J.; Travis, J.; Hull, L.; Ngugi, H.; Krawczyk, G.; Harsh, M.; Reichard, K.; Ellis, K.; Remcheck, J.; Crassweller, R.; Marini, R.; Harper, J.; Kime, L.; Heinemann, P.; Liu, J.; Lewis, K.; Hoheisel, G.; Jones, V.; Glenn, M.; Miller, S.; Tabb, A.; Park, J.; Slaughter, D.; Johnson, S.; Landers, A.; Reichard, G.; Singh, S.; Bergerman, M.; Kantor, G.; Messner, W. *Speciality Crop Innovations: Progress and Future Directions; Specialty Crop Innovations Progress Report; College of Agricultural Sciences, Penn State University: University Park, PA, USA, 2009; pp. 1-16.*