IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.7, Special Issue.1, Jan 2019 E-ISSN: 2347-2693

A Study on Data mining techniques used in Agriculture

Biswajit Mondal

Department of computer Science, Burdwan University, West Bengal, India

Corresponding Author: biswajitpanagarh85@gmail.com

Available online at: www.ijcseonline.org

***Abstract*—** Agriculture plays a vital role in India . Most of the people in India are involved for cultivation. This paper presents different techniques of data mining used in Agriculture sector. Several methodologies has been used in data mining techniques such as Neural Network, K-Means, Fuzzy-set , Bayesian Network, K-nearest neighbor, Decision tree analysis etc.

***Keywords*—**Data Mining Techniques, Agriculture

I. INTRODUCTION

In agriculture field data mining techniques has been used for collecting several data to help farmers understand and uses data mining results. To process the data in agriculture field may be used data so that any farmer may gets his/her desired query in very short period of time in efficient ways.

II. DATA MINING TECHNIQUES

we can use several method of data mining techniques such as i) Classification Techniques ii) Association Techniques iii) Clustering Techniques iv) Regression Techniques.
In data mining techniques such as data classification and data clustering can be used for data analysis. Data classification is supervised learning where training data set is used to classify the farther data. Data clustering is unsupervised learning where training set in unavailable. We are used multiple data mining techniques for analyzing agriculture data. Classification is one kind of analytical modeling. The different data mining techniques used in agriculture given below.

i) Classification Techniques:
Classification is the method of generalizing known structure to perform new structure. It is a techniques for prediction of class levels and on the basis of this class level to classify available new data. In agriculture soil plays a vital role. we can predict soil type using data mining classification techniques.
Various Classification tools that may be applied in agriculture are briefly outlined below:
1. Decession Tree 2. SVM(Support vector Machine) 3.Aritifical Neural 4.Bayesian Networks 5.Genetic Algorithm.

1. Decision Tree : A Decision tree is a structure that includes a root node , branches and left node. Each internal node denote a test on attribute, each branch denotes an out come of a test and each leaf node holds a class level. The top most node in the tree is the root node[1].

The decision tree classifier is widely used techniques for classification. As the name suggests, decision tree classifier use a tree ;each left node is associate class , and each internal node has a predicate associated with it[2].

2. SVM(Support vector Machine):-Support vector Machine is a best classifier to give very accurate classification across a range of application. A support vector machine allow us to classifying data that's separable. It has main advantages for higher speed and better performance in comparison with neural network. It is very useful for text classification problems. Classifications of images can also be performed using SVM. We can also recognized hand written character by SVM.

3. Artificial neural network(ANN):-Neural network or Artificial networks are most vital class of tools for quantative modeling. Neural networks are computing models for information processing and are particular useful for identifying the fundamental relationship among a set of variables or pattern in the data[3].In more practical terms Neural networks are non linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to patterns in data [4].

In neural network basic characteristics are exhibits mapping capabilities that is ,they can map input pattern to their patterns to their associated output patterns. It is robust systems and faults tolerant and they can recall full patterns from incomplete ,partial or noisy patterns.[5]

4. Bayesian Network : Bayesian Network classification are called supervised learning method which explain rainfall and temperature from the crop data for using classification and probability values of potato, cabbage and cauliflower etc. Bayesian Network which is a classification analysis techniques for using to explore data set.

In Bayesian Network incomplete data set can be handle. It is allow one to learn about causal relationship. Bayesian networks in conjunction with Bayesian statistical techniques facility the combination of domain Knowledge and data. Bayesian Network can also be used as influence diagram instead of decision trees. In comparison to Decision Tree ,Bayesian Network are usually more compact , easier to built and easier to modify .In this Network direct probabilities has been used. The limitation of Bayesian networks require prior Probability distribution and despite innocuous choice, this can have misleading effect on the result.

5. Genetic algorithm : Genetic algorithm is a nice approach of optimization techniques using evolution ,more specific it is used in inheritance, mutation and selection. Genetic algorithm has been used for decision making purpose to select best crops for growing which are profitable.

ii)Association Techniques :
Association Techniques plays a vital role in the agriculture field . It is a main components of data mining .Association rules has been used for finding different pattern in data. In agriculture field association techniques has been used to discover knowledge from agriculture data bases such as various soil type and cultivation .It has been used different data for taking decision about selection of crops in the suitable area under geographical condition. There are so many different association rule mining algorithm are apriority algorithm ,Hashing and pruning ,Dynamic item set counting and FP growth. Apriority algorithm is easy to execute all frequent item set in the data base.

iii)Clustering Techniques :
Clustering is the process of grouping or making sets of similar or nearly similarly type of physical or abstract objects. The groups are known as clusters[6].There are various types of clustering such as hierarchical methods, density based method ,partitioning method ,grid-based method ,model based method etc. In agriculture field fuzzy clustering,hierarchical device clustering has been used in different purpose . There so many problem in fuzzy nature and solve this problem through fuzzy clustering methods. Fuzzy clustering analysis in data mining improve the accuracy of the decision making system in agriculture. The uses of fuzzy clustering in detection of leaf sports in cucumber crop.

iv) Regression Techniques :
In 1997 Osman and et al has been used in Regression Techniques in Agriculture. Regression techniques represent a straight line (y=mx+b) where m takes appropriate values and b to predict value of y based upon a given value of x[7]. There are two types of prediction such as Linear regression (LR) and nonlinear(NLR).

Multiple linear regression(MLR) which consists of statistical method least squares and applying climatology for preparing models to reuse climate variables from tree ring services. It is crop yield prediction model which represent with the help of multiple regression techniques whose production is used as predict ant and there is lot of predictors such as year, rain fall ,area of showing ,Yield and fertilizers etc.

III. LITERATURE SURVEY

I have gone through many research paper in agriculture domain during 1999 to 2017.The observation is highlighted briefly with the help of following Table(1).

Table(1)

Author and publication	**Techniques used**	**Parameters achieved**
Cuningham and Holmes, 1999[8]	They have used various data mining techniques such as natural trees ,statistical machine learning ,and other analysis method. various data mining techniques uses several agriculture data may yields out comes in the agriculture field.	This techniques has been used to classify soils which explain large soil profile experimental data sets.
K.Verheyn et al, 2001[9]	He has used pattern detection technology as well as statistical and mathematical techniques.	It has been focused data mining techniques used to studied soil characteristics. The K-Mean approach is used for classifying soils in combination with GPS based techniques.
RTNg, JHan,2002	He has been used Spatial data mining.	A new clustering method called CLARANS, whose aim is to identifies spatial structures that may be present in the data. This method is used to different district of state which have similar crop

		production.
Ramirez , Mishra, 2003.	He has studied yield distribution of these crops and conclude that they are non normal and left skewed.	Yield distribution of crop is predicted and they are non normal and left screwed. As a result making use of data often leads to considerable gains in efficiency and therefore economic advantage.
Norwood B. Roborts, Lusk 2004.	He has been observed that the semi parametric model ranked highest for forecasting purposes.	It has been used forecasting Crop yield model. As a result farmer is benefited.
L.N.Na, G.H.Hua and L.B.Ying, 2006	He has been invented a new genetic algorithm based on negative selection.	Genetic Algorithm (GA) is an effective tool to use in data mining and pattern recognition. The negative selection method is used to show promising results.
Barghavi.pandJyothi.S,2009 [10].	They were applying dataminig techniques for classification agriculture land soils.Naive Bayes data mining techniques is used to classify soils.	Naïve Bayes data mining technique is used to explain different type of soils which is helpful for farmers. As a result farmers are choose different soil for best crops production.
S.veenadhari et al, 2011.	Data mining Techniques such as ANN,ID3,the K-means, the K-NN and support vector machine applied in the field of agriculture were used.	Data mining application in agriculture is a relatively new approach for forecasting or predicting of agricultural crop or animal managementAnimal feed, seed, chemical, poultry, fertilizer pesticides, seed, paper and many other industries use agricultural intergradientent products as their production processes.
Raorane et al,2012	They have used data mining techniques with the help of Decision Tree Algorithm, Regression tree, Artifical Neural Network (ANN), Bayesian Network ,Support Vector Machine(SVM), and K-means.	It is a data mining techniques used for crop production in agriculture field.
Alvaro calzadilla et al, 2014[11].	He assessed that impact of climate change in agriculture.	Climatic parameters such as Average solar radiation, Average minimum temperature, Average maximum temperature, Water deficit and Phase length were encountered As a result farmer is benefited in this survey.
Swatihira at al, 2015[11].	He suggested that generally the agriculture data is spatio-Temporal data. This data has to be analyzed by multidimensional analysis.	These data has to be analyzed by multidimensional analysis, Statistical analysis and Data Mining Techniques (Association Rule Mining) for obtaining a useful pattern which helps to analyze the agriculture productivity.
Aruna.P.et al, 2016[12].	He has been explain an application of a wireless sensor network for low cost wireless monitored and control irrigation.	The project describes two Microcontroller units, one unit is placed in agricultural field and the other unit is placed in main control unit which is interfaced with motor unit.
Shriyash Thawali et. al, 2017[12].	He represent a robot capable of performing operations like automatic plugging ,seed dispensing and pesticide spraying control of this agro-bot will be wireless.	Design and analyze a real time system for these robot give a solution and proposed a model which can be used in real time field. the robot Analyzed the design of plough tool and developed for real time system.

IV. CONCLUSION AND FUTURE WORK

Agriculture is back bone of India. Most of the people in India directly or indirectly involve in Agriculture. Different types of data mining techniques have been used in agriculture sector. Crop growth and Crop yield production has been discovered through data mining techniques such as regression analysis, clustering techniques etc.

It is hoped that web mining techniques may be implemented in Agriculture in near future. It is also essential for clouding computing technology has been introduced in agriculture field exploring the data from clouds.

REFERENCES

[1]. "Data Mining - Decision Tree Induction ", https://www.tutorialspoint.com/data_mining/dm_dti.htm.

[2]. Abraham Silberschaartz, Henry F.Korth,S.Sudarshan, "Data base System Concepts", Mcgraw Hill Education(India) Private Limited,New delhi

[3]. Zhang G.Peter," . Nerual Network for data mining", Georgia State University,Department of managerial Science, gpzhang@gsu.edu

[4]. Dr. Yashpal Sing,Alok Sing Chauhan," Nerual Network for data mining" Journal of Theoretical and Applied InformationTechnology ,© 2005 -2009 JATIT,www.jatit.org.

[5]. Amrender Kumer ,"Forecasting of crops using data mining techniques", Indian Agriculture statistic research Institute,Newdelhi-110012.

[6]. Mamta Tiwari, Dr. Bharat Misra, " Application of cluster analysis in Agriculture – A review article", International Journal of Computer Applications (0975 –8887)Volume 36–No.4, December 2011

[7]. Hooman Fetant, Leila Mortazavifarr , NarsisZarshenas, "The analysis of agriculture data with regression data mining techniques"- https://periodicos.ufsm.br/cienciaenatura/article/view/20759

[8]. Hooman Fetant, Leila Mortazavifarr, Narsis Zarshenas ," The analysis of a agriculture data with regression data mining techniques", DOI: http://dx.doi.org/10.5902/2179460X20759

[9]. Surabhi chouhan, Divaker Singh, Anju Sing . "A Survey and analysis of various agriculture crops classification Techniques"- International Journal of Computer Applications (0975–8887)Volume 136 No.11, February 2016.

[10]. E. Manjula , S.Djodilta Choumy ,"Analysis of data mining techniques for agriculture data", International Journal of Computer Science and Engineering , Communications Vol.4, Issue.2, Page.1311-1313, (2016) www.scientistlink.com

[11]. S. Revanthi, M. Brindha,"Impact of climate change in agriculture with data mining concept", International Research Journal of Engineering and Technology (IRJET) Volume: 03 Issue:03 | Mar-2016 www.irjet.net

[12]. Pratiki Srichandan , Ashis Kumer Mishra, Harkishen Singh "Data Science and Analytical Technology in agriculture", International Journal of Computer Applications (0975 –8887)Volume 179–No.37, April 2018

Authors Profile

Mr. B Mondal Bachelor of Science from Burdwan University, West Bengal and Master of Computer Application from Jorhat Engineering College under Dibrugrah University, Assam. His main research work focuses on Datamining, Artificial Intelligence and Soft Computting.

