

# DATA MINING TECHNIQUES FOR INCREASING SMART FARMING IN AGRARIAN SECTOR

Vandna, K.L. Bansal

**Abstract :** Agriculture become more fruitful when data mining techniques is used, data mining with agronomy boost up the agribusiness and may accomplish the increasing demand of livestock in the world. Agriculture is art of breeding the food plants where all kind of livestock includes, the agrarian, agronomy & cultivation are the another name for agriculture. So by mixing these two fields and applying various data mining techniques in agronomy it gives so many benefits. The aim of this scrutiny is to disclose peculiar extraction approaches in agrarian field, to find those which gives best result and also try to find out the best data mining tool which can be used in agriculture with these techniques. In this paper, various studies proposed the data eradication technique which gives most appropriate result in the field of cultivation for more clear vision. These eradication approaches can be used to find out the solution for real time problems like weather, temperature, soil, rain, pH, Moisture, pesticides & Disease forewarnings which help farmers to achieve their expected crop result.

**Keywords:** Data mining, Agriculture (agronomy, agrarian), classification, soil science, Fertilizers

## 1 INTRODUCTION

Agribusiness is the soul of Indian financial system. More than fifty percent of Indian population rely on agriculture for its economy and daily livestock, because the population of rural people are more than urban people in India. Although fact that India live in villages and wide sphere in India concealed under watering, there thirty three percent of yield chunk is lay under irrigation. The agronomy rate is awfully crouched in terms of productivity. The farmer, agriculture scientists, researchers and government is putting additional effort and techniques for enhancing production because the demand of the food is increasing in a speedy way. As a result data which is used in the field of agriculture is increasing periodically. As aggregate testimony upturn, it desire number of schemes for extraction of the information when needed. Today, petty number of farmers are really using brand - new techniques, gizmos and mechanisms of cultivation for rendering in a superior way. Data extraction can help out in anticipating the various forthcoming drifts [1]. Data Mining is a procedure of discovering of unfamiliar probably impressive arrangements from large amount of datasets which are previously unexplored. The extracted information from the data can be used for illustrating as a dummy model for the prognosis and classification. The classification of the data mining can be categorized as follows: (a) Descriptive data mining approaches are used to find out those trends and patterns used for illustrate and symbolize the prevailing competence of the data. (b) Predictive data mining approaches are used to detect various new tendencies and patterns which are used for making new predictions about new explicit values which is based on the patterns determine from the noted results.

Marketing and so many are included. The rest of the paper composed of Different data mining techniques there pros and cons, Background Study, Application of data mining techniques in agronomy, Issues and challenges in agronomy, results and conclusion.

## 2 DATA MINING TECHNIQUES

The clustering, association rules, Classification and regression are four major techniques in DM which are used to solve the various agriculture problems are shown as Figure (1). [3]

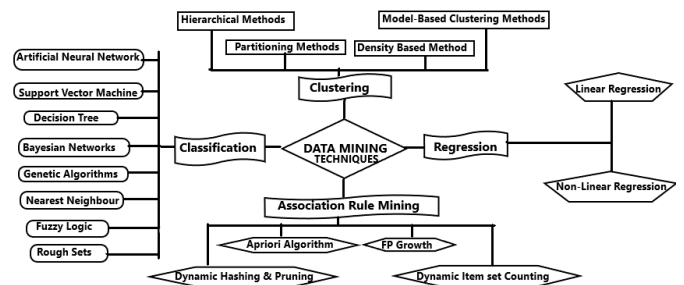


Fig. 1. Various techniques in data mining

### Classification

For predicting the future data patterns or to narrating the data classes classification and prediction are the ways which is used for extracting the patterns. To forecast the distinct class trademark on advanced data fragment, a set of classified samples is used to classify the strange samples. One of the main aims of classification technique is to test uncharted sets during training which helps to increase the predictive accuracy. There is the various classification techniques are available for discovering new knowledge are: [4]

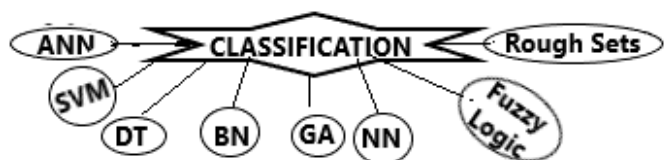


Fig. 2. Classification techniques

• Vandna, K.L. Bansal, Himachal Pradesh University.

Predictive data mining is mainly used for prediction especially in agriculture area to predict future crop, weather forecast estimation, food, fertilizers and pesticides used, yield estimation and so on [2][1]. Different sectors in agriculture field are described as Plant Genetics, Seed Technology, Pests and Pathogens, Horticulture, Engineering, Veterinary, poultry, Dairy, Food Processing, Farm Mechanization, Irrigation, Soil and water conservation, Soil science, Animal fed, Fertilizers and Chemicals, Climate, Green Buildings, Agricultural

**Clustering**

Clustering focuses on segregation of the data directories into clusters where particular cluster is adjacent to one and all. The data with same precedent are grouped together; where data with the different instances lies in the different group and assessment of new cluster is peculiar [5]. There is no preceding expertise around information. The numbers of clustering method are: [6]

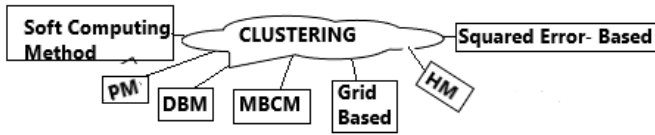


Fig. 3. Clustering techniques

**Association Rule Mining**

Association rules are used to discover undetected or desired pattern among extensive amount of data, it is used to find the link among divergent components in transferable directory. These discovered the ingredient which coexists frequently in the directory and consist of abounding self-reliant selections of items (like purchasing transactions) and to find rule. The elementary problem declaration is: Set of transactions, where individual transaction = set of literals, Association rule = expression of the form  $P \Rightarrow Q$ , P & Q are the sets of objects. The transactions of database which consider P tend to contain Q. [7] Association rule mining is used in the market value analysis, store blueprint, telecommunication alarm prediction and list chart. The algorithms which are used in association rule mining are: [8]

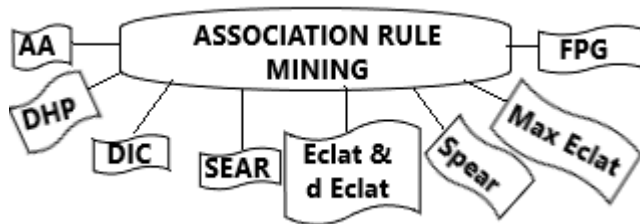


Fig. 4. Association Rule techniques

**Regression**

Learning objective which outline the statistics items into evident value prognosis model known as regression. The application of regression are guessing the probability of survive or die of a patient on the bases of sets of his diagnoses. Estimate the volume of biomass available in a forest, demand of new product in the market. [9] Regression models are also severer as the classification job with numeric or frequent class stub which is used to predict continuous target. The numbers of Prediction methods are shown as:



Fig. 5. Association Rule techniques

Table- I: Data mining techniques: Pros and Cons.

S	Method Name	Techniques	Pros	Cons
1.	Classification	A. Decision Tree	<ul style="list-style-type: none"> <li>Decision trees don't require real knowledge; it can handle either numeric or categorical data.</li> <li>Decision trees makes classification and learning smooth and simple.</li> <li>Good accuracy.</li> </ul>	<ul style="list-style-type: none"> <li>High variance of tree structure and restricted to one output attribute.</li> <li>An unstable classifier i.e. classifier performance depends upon the type of data sets.</li> </ul>
		B. Support Vector Machine	<ul style="list-style-type: none"> <li>For classifying linearly separable data SVM uses maximum marginal hyper plane data can be break down into pieces.</li> <li>SVM extends itself.</li> </ul>	<ul style="list-style-type: none"> <li>These are computationally expensive.</li> <li>Training process takes more time as compare to other methods.</li> </ul>
		C. Neural Network	<ul style="list-style-type: none"> <li>High accuracy and tolerance of noisy data suitable for Continuous-valued data &amp; current data. It is easy to conserve and implemented in parallel hardware.</li> </ul>	<ul style="list-style-type: none"> <li>It requires many parameters and training.</li> <li>Awe full interpretability.</li> </ul>
		D. Bayesian Network	<ul style="list-style-type: none"> <li>Bayesian classifiers are also efficient like other.</li> <li>Accurate high speed.</li> <li>They make the process of computation simple.</li> </ul>	<ul style="list-style-type: none"> <li>Dependency between variables.</li> <li>Totally depend on selection of parameters.</li> <li>Consistent accuracy only gets with large datasets.</li> </ul>
		E. Genetic Algorithm	<ul style="list-style-type: none"> <li>GA is Accurate and can be used</li> </ul>	<ul style="list-style-type: none"> <li>GA is Time-consuming</li> </ul>

2.	Clustering	A. Hierarchical Methods	<ul style="list-style-type: none"> <li>with other methods.</li> <li>Implementation is easy</li> <li>Good visualization capability.</li> <li>Need to specify no. of clusters in advanced</li> </ul>	<ul style="list-style-type: none"> <li>g iterations.</li> <li>It is slower because of cubic time complexity.</li> </ul>	C. FP Growth	<ul style="list-style-type: none"> <li>There are only two passes over data set, no candidate generation compresses data set and are faster.</li> </ul>	<ul style="list-style-type: none"> <li>FP tree may not fit in memory and also expensive to build.</li> </ul>	
			B. Partitioning Method	<ul style="list-style-type: none"> <li>In PM the use of iterative way to ensure the clusters &amp; are Easy to implement.</li> </ul>		<ul style="list-style-type: none"> <li>No. of clusters should be predefined by user, only spherical shaped clusters can be determined.</li> </ul>	D. Dynamic Item set Counting	<ul style="list-style-type: none"> <li>Data sets are marked in four different ways: Solid box, Solid circle, Dashed box, Dashed circle.</li> </ul>
		C. Density-Based		<ul style="list-style-type: none"> <li>Easily handle cluster with arbitrary shape.</li> <li>Work well in the presence of noise.</li> </ul>	<ul style="list-style-type: none"> <li>Not handle data points with varying densities.</li> <li>Result depends on the distance measures.</li> </ul>	4. Regression		A. Linear Regression
			D. Model-Based	<ul style="list-style-type: none"> <li>Specifying the no. of clusters automatically based on standard statistics also can handle noise.</li> </ul>	<ul style="list-style-type: none"> <li>Depend on the hypothesized model and structure.</li> </ul>		B. Non-Linear Regression	
		E. K-Means		<ul style="list-style-type: none"> <li>Simple clustering approach, Efficient, Less complex method</li> </ul>	<ul style="list-style-type: none"> <li>Require no. of clusters in advance</li> <li>Problem with handling categorical attributes.</li> </ul>	<h3>3 BACKGROUND STUDY</h3> <p>This section basically focused on the various publications in related field. In [10] authors' discussed three algorithms for the comparison of the analysis time and memory for fabrication of clusters in management zones. Management zones are finicky aspect in precision farming environment. So to management of the spatial factors adequately increase the crop productivity. Data mining is the drift technology which can be used aptly in the field of depiction of management zones. Define the Need of more studies there for uncovering the new inclined contributions to boost the agricultural sector. The authors in [11] proposed Multiple Linear Regression techniques and techniques based on the Density based clustering for crop production in particular elected territory i.e. one of the Andhra Pradesh district, East Godavari in India. Compression of the crop produce prognosis contrived with all set of testimony for the improvement of the scheduled techniques for 40 years interval. [12] Anticipated on valuating new data mining techniques for the revenue esteem. The data is used in this paper are proffered by directorate of economics and statistics of India, State Governments Statistical and agriculture department, soil department. This paper includes the crops of the Kolhapur district of Maharashtra State in India for sample design of different crops in different areas of the Kolhapur district. Techniques like K means used to forth the tainting in atmosphere, classification of loam and plants, supervision of wine fermentation, SVM for analyzing different changes of weather. In [13] authors presented the improvement about development and stabilization of the agronomy production as one of the main condition for boosting agronomy. Merits include the aim of the prognosis and its potentiality to equip measure of</p>		
			3.	Association Rule Mining	A. Apriori Algorithm	<ul style="list-style-type: none"> <li>It is easy to use parallelized and uses large item set properties.</li> </ul>	<ul style="list-style-type: none"> <li>It requires many datasets scans, transaction data base is memory resident.</li> </ul>	
B. Dynamic Hashing & Pruning	<ul style="list-style-type: none"> <li>DHP performance doesn't decrease when data grows it increases the size of memory to store the data.</li> </ul>	<ul style="list-style-type: none"> <li>With increase in size storage size also increased where overflow situation also occurs.</li> </ul>						

reliability which a conventional forecast method enable to provide, and introduced vitals of objective methods for pre-gathering forecast of crop outturn in India. In [14] authors have tried to explain different schemas of data mining which are pre-owned in agriculture and gives the comparison of various methods in plant diseases unearthing system, soil moisture system and crop enlargement monitoring system. This paper tells that by using these techniques we can adorn the crop production in farming. The authors in [15] proposed the mixture of precision agriculture and InfoTech leading meliorism in production together for ensuring plants, mammalian and clay snag absolutely what they require for flawless health and profit. Clustering and Regression techniques used for explaining the better outcomes and growth factor of florid plants labeled at various research site across Kish island the Province off the southern coast of Iran. Four clusters are used to show the growth factor and four regression models are used to determine the estimation of leaf number from length of main trunk and primary branch this papers also consider the chattels if chlorophyll contents on the number of flourishing plants out bloom in hot season as future study. In [16] authors proposed the terrific classification technique for the prediction of grass forage damages which may cause disorder in plants. Datasets of 155 records and 8 features have been used to define the classes like low, average, very-high for the estimation of events of distinctive classifiers and Ensemble models for multi-class data & binary class statistics. This paper concludes various classifiers, for binary data RF & GNB performance are finer than others and for primeval dataset NN & RF performance is better. This gives the DT, SVM & RF as the perfect consolidation antiquated among all tested combinations. The authors in [17] focused on the image processing along with the classification and clustering method by applying neural networks to represent a measure for a mechanized classification of various leaf diseases. The comparison of the different techniques are done on the basis of their merits and demerits and identify the problems as leaf color, size, texture are changed with the climate. This paper suggested image processing technique is better than others for identifying the color, shape, and any type of plants in the field of agribusiness. In [18] authors proposed about the data eradication and its various operations used in the range of agriculture area the raw data of the agronomics firm are very copious and distinct so there is the huge need to store data in the sorted way. DM application for farm expert use patterns for measuring plant growth indicators, vintage quality indicators, give the idea about the difficulties of applying data extraction and hurdles faced by data mining in the agriculture. Authors in [19] discussed different survey papers related to the various and recent areas of data extraction. Number of areas where data mining can be applied: web mining, text mining, spatial data mining etc. With the support of different data extraction gizmos named as Rapid miner, WEKA, R-Programming, Orange, KNIME, NLTE by using different data mining techniques like Genetic algorithms, NN, SVM, Association rule mining, Clustering, Classification for the extraction of the data from data marts and warehouses. The authors in [20] are also focused on the estimation of vegetables crop by applying the data extraction methods for finding expertise from various herbaceous plants. WEKA is used as the predictive data mining equip using K-means, Kohonen's Self Organizing Map (KSOM), and EM algorithms. The gathered results find K-means as useful for the prediction of main factors which affect the vegetable production and KSOM

as best suitable for predicting results of yield production. The extending this model by considering the earlier attributes for the improvement of prediction accuracy as future scope. In [21] proposed the identification of convenient fact that attains high definiteness and generality in yield prediction. The Knowledge Discovery in Database process used to analyze new and convenient patterns from large data houses. Statistical Analysis of crops is gathered from 2004 to 2010 in Amravati district of Maharashtra which gives useful statistical information about soil, crop yield investigation, weather, fertilizers and also introduces GPS technique for finding the useful pattern in soil identification. In [22] authors presented exhaustive investigation of crop yield prognosis by applying MLR (Multiple Linear Regression) which mainly spotlight on estimation of the agronomy investigation of organic cultivation & inorganic cultivation, also gives the architecture for agriculture analysis the production of the plants on time, gain & defeat of knowledge for determining the actual demesne land in a particular area. The real demesne data sets are used to consider the discrepancy in performance and prediction between the organic and inorganic cultivation. In [23] authors attempted to select the correct arrangement in food safety for the forecast of cultivation amount and adopting need of strategic crops for Egyptian settlers belonged to public extension. The introduced framework focus on building a data extraction device for fore forth needs in agribusiness to defeat food uncertainty by constructing process of ANN, WEKA using Multilayer Perception (MLP) function as predictive method of data eradication. Outcome tells that framework achieve to predict yearly need of amount of Wheat, Rice, Beans up to coming years. In [24] a background of the application of data eradication methods as ANN, Bayesian networks and support vector machines in the discipline of agriculture. Different kind of studies has marked on the functions of data extraction manner. It determined the amount of encouraging ways used to figure out the relationship of various weather conditions and factors facts on crop yielding productivity. Inspections find out how these methods used with difficult agricultural datasets and used for the yielding forecast with assimilation of both periodically and spatially using GIS automation. In [25] authors described the data mining process which are applied on the data sets which comprises the attribute of soil as type of soil and properties as in soil available percentage of nutrients. Extraction of the input data can be done with data eradication algorithms where neural network used to autumn data from datasets.

#### **4 DATA MINING TECHNIQUES: APPLICATION IN THE AREA OF AGRONOMY**

There are huge amount of studies on application of data extraction techniques are available for agribusiness data storage. For classification of soils that is used to analyze large soil experimental data set based on Naive Bayes technique of data mining [26]. For predicting soil fertility Decision tree algorithm is used [27]. Neural Networks are used for Prediction of flourishing and full growth dates of soybean [28], information about weather forecasting using recorded parameters. [9] K-Nearest is practiced in simulation of everyday precipitations for alternative weather variables [29]. SVM is used in classification of crops and analysis of different scenarios for climate change [30] [31]. SVMs for differentiate between birds sound and other sound [32]. K-means apply for soil classification, grading apples and precision agriculture for detecting weeds [33] [34] [35].

### 5 ISSUES & CHALLENGES IN AGRONOMY

There are many challenges occur in applying data mining techniques Agriculture. These challenges are like the storage of the large amount of the data, extraction of useful data from huge amount of data, security of the data, management of quality in agriculture like data, standards plans treatmentsThe problem of prediction in agriculture can be dividing into phases as learning phase and decision making phase where the large amount of dataset is transformed into extracted dataset where number of characteristics and objects in new set are much lesser than the original data. Are these data and predictive characteristics sufficient for making predictive model of acceptable performances? [18]

#### Other challenges like leaf disease identification techniques are as follow:

- i. In cultivation image data like:
  - a) Consumption of maximum storage for jpeg, png and gif.
  - b) Use of maximum radio bandwidth for transmission capacity that affects image quality.
- ii. The old knowledge influences the output image.
- iii. In the exiting field conditions, optimization  
Approaches used for a peculiar crop disorder and frequent computerized thorough care of plant can be done by automated gizmos.
- iv. Changes with climatic and environment conditions:
  - a) Leaf colour
  - b) Size and texture.
  - c) The area expertise and continuous observations are demanded well in time. [17]

### 6 RESULT

Table 2 gives the review of research papers using different data mining tools, techniques, languages, and the results in the form of inference of those research papers. The techniques used are Decision tree, K-Nearest, Clustering, Classification, Bagging, Boosting, Ensemble Models, Association rule mining, EM, Regression, Multiple Linear Regression, tools/languages used are WEKA, DIONE data miner, and Rapid miner.

**Table- II:** is a comparison-based table on data mining tools, techniques, and inference.

S. No.	Author	Year	Paper	Technique	Tool/Language	Inference
1.	D. Ramesh [11].	2015	Analysis of crop yield prediction using data mining techniques	Multiple Linear Regression and Density Based Clustering		The result shows density based clustering gives maximum production (estimation) as compare to MLR ranging between -13% And +8%
2.	Silly Jo	1999	Developing	Association clustering,	WEK	Result shows that
3.	D. Ramesh [37].	2013				
4.	Michael Stadler [38].	2005				
5.	Dr. You	2017				

	Cunningham [36].		innovative application in agriculture using data mining	Iterative approach, NaiveBayes, j4.8, bagging, boosting	A	Longit boosting stumps for 10 or 100 repetitions are the excellent among the other schedules present in WEKA and j4.8 model is useful in developing more objective standard for quality classification and marketing pricing for mushrooms as compare to others.
	D. Ramesh [37].	2013	Data mining techniques and applications to agriculture yield data	Multiple linear regression, K-means		The assessment of average manufacturing by using MLR method is given as 98% accuracy and using K-means algorithm is given as 96% accuracy.
	Michael Stadler [38].	2005	Using data mining techniques for exploring the key features of plant dynamics upon a newly built plant trait database	Decision tree, Centre based clustering, Association Rules, Decision rules	DIONE data miner, J2EE	Experiment shows that the values plant traits life interval, breeding structure, seed farming, are each mapped to a no. Between 0.1 and 0.9 indicating the vulnerability of a species for extinction and DIONE data miner will be used in terms of additional algorithms, and putting extra statistical appraisal possibilities for extracting results.
	Dr. You	2017	Predict the Main	K- Means, Kohonen's	WEK	K-means is useful for

	sef Abu zir [20].		Factors that Affect the Veg. Production In Palestine Using WEKA Data Mining Tool	Self Organizing Map (KSOM), Expectation Maximization (EM)	A	predicting the main factor which affects the vegetables cultivation. KSOM is more accurate to predict yield production. It gives the accuracy result as 60% (72997.1667) for increasing the yield production compared to actual average actual data is 34144.25.		n [40].	tools/ K-Means methods on clustering of rice crop by province as efforts to stabilize food crop in Indonesia	K-means	miner	the K-means used for increasing rice crop production. K-means accuracy % performance based on the Davies-Bouldin index of rice crop are -0.392.
6.	Monali Paul [39].	2015	Analysis of Soil Behavior and Prediction of Crop Yield using Data Mining Approach	K-Nearest Neighbors (KNN) and Naive Bayes (NB)	Rapid Miner 5.3.	The result shows the demographic result of KNN, which shows that the low class holds 30 land soil patterns provide awful yielding of crops, 45 lands of medium class soil will provide good cropping and 25 areas of high class soil will give moderate yielding of crops. And statistic result of the Naive Bayes algorithm, 15 land areas which has low class soil will provide awful yielding of crops, 40 lands of medium class soil will provide good cropping and 45 areas of high class soil will provide moderate yielding of crops.						
7.	Sudirna	2018	Data mining	Clusterin g,	Rapid	The result shows that						

### 7 CONCLUSION

In this paper, number of data extraction techniques in favour of agronomy and its various fields are reviewed in terms of certain parameters. Data mining techniques like k-means for soil fertility, prediction analysis for wine fermentation, bi clustering, high performing system in data mining helps in solving highly complex problem. this study show that large number of algorithms and extraction approaches have been used in agronomy but very less number of data mining gizmos have been used with these techniques. So there can be various number of tools can be used with data mining technique in agrarian field which improves the lifetime of agronomy.

### 8 REFERENCES

- [1] M. C. S. Geetha, "A Survey on Data Mining Techniques in Agriculture", International Journal of Innovative Research in Computer and Communication Engineering ISSN (Print):2320-9798, Vol. 3, Issue 2, 2015.
- [2] Dr. P. Jaganathan, S. Vinothini, P. Backialakshmi, "A Study of Data Mining Techniques to Agriculture", International Journal of Research in Information Technology (IJRIT) ISSN: 2001-5569, Vol. 2, Issue 4, pp.306-313, April 2014.
- [3] Hatel Patel, Dharmendra Patel, "A Brief survey of Data Mining Techniques Applied to Agriculture Data", International Journal of Computer Applications (0975-8887), Vol. 95-No. 9, June 2014.
- [4] Beniwal, S. & Arora, J. (2012), "Classification and feature selection techniques in data mining", International Journal of Engineering Research & Technology (IJERT), 1(6), 1-6.
- [5] Xu, R & Wunsch, D (2005), "Survey of clustering algorithms", Neural Networks, IEEE Transactions on, 16(3), 645-678.
- [6] Fayyad, U., Piatetsky - Shapiro, G., & Smyth, P. (1996), "From data mining to knowledge discovery in databases", AI magazine, 17(3), 37.
- [7] Srikant, R V Q & Agrawal, R (1997, August), "Mining Association Rules with Item Constraints. In KDD" (Vol. 97, pp. 67-73).
- [8] Zaki, M J (1999), "Parallel and distributed association mining: A Survey". IEEE concurrency, 7(4), 14-25.
- [9] 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.

- [10] Mrs. Krishna Priya C B, Dr. S. Venkateswari, "Delineation of Management Zones in Precision Agriculture using Different Clustering Algorithms", *International Journal of Applied Engineering Research* ISSN 0973-4562, Vol. 13, No. 22 (2018), pp. 15951-15955.
- [11] D Ramesh, B Vishnu Vardhan, "ANALYSIS OF CROP YIELD PREDICTION USING DATA MINING TECHNIQUES", *International Journal of Research in Engineering and Technology* p ISSN: 2321-7308, Vol. 04, Issue: 01, Jan-2015.
- [12] Raorane A.A., Kulkarni R.V., "Data Mining: An effective tool for yield estimation in the agricultural sector", *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)* ISSN 2278-6856, Vol. 1, Issue 2, July – August 2012.
- [13] Raorane A. A., Kulkarni R.V., "Review- Role of Data Mining in Agriculture", *International Journal of Computer Science and Information Technologies (IJCSIT)*, Vol. 4 (2) , 2013, 270 – 272.
- [14] D. Sabareeswaran, A. Edwin Robert, "A survey on data mining techniques in agriculture", *Indian Journal of Innovations and Developments* ISSN (Print): 2277-5382, Vol. 5 (8), August, 2016.
- [15] Hooman Fetanat, Leila Mortazavifar, Narsis Zarshenas, "The Application of Data Mining Techniques in Agricultural Science", *Revista do Centro de Ciências Naturais e Exatas – UFSM (Ciência e Natura)* ISSN on-line: 2179-460X, v. 37 Part 2 2015, p. 108–116.
- [16] Umair Ayub, Syed Atif Moqurrab, "Predicting Crop Diseases Using Data Mining Approaches: Classification" in *Conf. Power, Energy and Smart Grid (ICPESG)*, 1-6, April 2018.
- [17] Subhadra k, Dr. N. Kavitha, "PLANT LEAF DISEASES DETECTION FOR DATAMINING - A SURVEY", *International Journal of Research and Analytical Reviews (IJRAR) P- ISSN 2349-5138*, Vol. 5, Issue 4, December 2018.
- [18] B. Milovic, V. Radojevic, "APPLICATION OF DATA MINING IN AGRICULTURE", *Bulgarian Journal of Agricultural Science*, 21 (No 1), 26-34, 2015.
- [19] Sadiq Hussain, "Survey on Current Trends and Techniques of Data Mining Research", *London Journal of Research in Computer Science and Technology*, Vol. 17, Issue 1, Compilation 1.0, March 2017.
- [20] Dr. Yousef Abuzir, "Predict the Main Factors that Affect the Vegetable Production in Palestine Using WEKA Data Mining Tool", *Palestinian Journal of Technology & Applied Sciences - Vol. 1 - No. 1- January 2018*.
- [21] Mr. Abhishek B. Mankar, Mr. Mayur S. Burange, "Data Mining – An Evolutionary View of Agriculture", *International Journal of Application or Innovation in Engineering & Management (IJAEM)* ISSN: 2319-4847, Vol. 3, Issue 3, March 2014.
- [22] P. Vinciya, Dr. A. Valarmathi, "Agriculture Analysis for Next Generation High Tech Farming in Data Mining", *International Journal of Advanced Research in Computer Science and Software Engineering* ISSN:2277 128X, Vol. 6, Issue 5, May 2016.
- [23] Ayman E. Khedr, Mona Kadry, Ghada Walid, "Proposed framework for implementing data mining techniques to enhance decision in agriculture sector", in *Conf. Communication, Management and Information Technology (ICCMIT)*, *Procedia Computer Science* 65 (2015), 633-642, Available online at [www.sciencedirect.com](http://www.sciencedirect.com).
- [24] Niketa Gandhi, "A review of the application of data mining techniques for decision making in agriculture", in *conf. Contemporary Computing and Informatics*, 978-1-5090-5256-1/16, 2016.
- [25] H.D.Gadade, Riddhi Singh, Vaishali Chaudhari, "A Survey on Predictive Analysis in Agricultural Soil Health Data to Predict the Best Fitting Crop", *International Journal of Computer Science and Mobile Computing (IJCSMC)* ISSN 2320–088X, Vol. 8, Issue. 3, March 2019, pg.161 – 165.
- [26] Bhargavi, P, & Jyothis, S. (2009), "Applying Naive Bayes data mining technique for classification of agriculture land soils", *International Journal of Computer Science and Network Security*, 9(8), 117-122.
- [27] Jay Gholap. (2012), "Performance tuning of j48 algorithm for prediction of soil fertility", *Asian Journal of Computer Science and Information Technology* 2: 8 (2012) 251-252.
- [28] Elizondo, D. A., McClendon, R. W., & Hoogenboom, G. (1994), "Neural network model for predicting flowering and physiological maturity of soybean", *Transactions of the ASAE(USA)*, 37(3): 981-988.
- [29] Rajagopalan, B., & Lall, U. (1999), "A k-nearest neighbour simulator for daily precipitation and other weather variables", *WATER RESOURCES RESEARCH*, 35(10), 3089-3101.
- [30] Camps-Valls, G., Gomez-Chova, L., Calpe-Mara villa, J., Soria-Olivas, E., Martin-Guerrero, J.D., & Moreno, J. (2003), "Support vector machine for crop classification using hyper spectral data", In *Pattern recognition and image analysis* (pp. 134-141). Springer Berlin Heidelberg.
- [31] Tripathi, S., Srinivas, V. V., & Nanjundiah, R. S. (2006), "Downscaling of precipitation for climate change scenarios: a support vector machine approach", *Journal of Hydrology*, 330(3), 621-640.
- [32] Fagerlund, s. (2007), "Bird species recognition using Support Vector Machine", *EURASIP Journal of Advanced Signal Processing*, Article ID 8637: 1-8.
- [33] K. Verheyen, D. Adriaens, M. Hermy, & S. Deckers, "High-resolution continuous soil classification using morphological soil profile descriptions", *Geoderma*, Vol. 101(3), pp. 31-48, 2001.
- [34] Leemans, V., & Destain, M. F. (2004), "A real-time grading method of apples based on features extracted from defects", *Journal of Food Engineering*, 61(1), 83-89.
- [35] Tellaache, A., Burgos Artizzu, 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.
- [36] Sally Jo Cunningham, Geoffrey Holmes, "Developing innovative application in agriculture using data

mining”, The proceedings of the Southeast Asia regional computer confederation conference, 25-29, 1999.

- [37] D Ramesh, B Vishnu Vardhan, “Data Mining Techniques and Application to Agriculture Yield Data”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 9, September 2013.
- [38] Michael Stadler, Renee M. Bekker, Jens Finke, Dierk Kunzmann, Michael Sonnenschein, “Using Data Mining Techniques for Exploring the key features of Plant Dynamics Upon a Newly Built Plant Trait Database”, Information for Environmental Protection – Enviroinfo Brno 2005, ISBN 80-210-3780-6, pp.90-94.
- [39] Monali Paul, “Analysis of Soil Behaviour and Prediction of Crop Yield Using Data Mining Approach”, CICN- International Conference on Computational Intelligence and Communication Networks, Available at: [www.ieeexplore.ieee.org](http://www.ieeexplore.ieee.org), 766-771, 2015.
- [40] Sudirman, Agus Perdana Windarto, Anjar Wanto, “Data mining tools| K-Means methods on clustering of rice crop by province as efforts to stabilize food crop in Indonesia”, 2nd Nommensen International Conference on Technology and Engineering, IOP Conf. Ser.: Mater. Sci. and Eng. 420 (2018) 012089.