International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 1 | ISSN : 2456-3307

A Review on Data Mining Techniques for Fertilizer Recommendation

Jignasha M. Jethva[1], Nikhil Gondaliya[2], Vinita Shah[3]

[1]Research Scholar, Information Technology, G H Patel College of Engineering and Technology, V V Nagar, Gujarat, India

[2]Associate Professor, Information Technology, G H Patel College of Engineering and Technology, V V Nagar, Gujarat, India

[3]Assistant Professor, Information Technology, G H Patel College of Engineering and Technology, V V Nagar, Gujarat, India

ABSTRACT

Agriculture plays a crucial role in the life of an economy. It is the backbone for developing countries like India as more than 70% of population depends on agriculture. To increase crop production many factors are responsible like soil, weather, rain, fertilizers and pesticides. We have used soil parameters to increase crop production because it is an essential key factor of agriculture. To maintain nutrient levels in the soil in case of deficiency, fertilizers are added to soil. The common problem existing among the Indian farmers is that they choose approximate amount of fertilizers and add them manually. Excess or insufficient addition of fertilizer can harm the plant life and reduce the yield. This paper provides review of various data mining techniques used on agriculture soil dataset for fertilizer recommendation. Mainly I focused on various soil parameters like Fe, S, Zn, Cu, N and Ph value etc. In this survey, we also describe some Agriculture problems that can be solved by using data mining techniques.

Keywords : Agriculture, Soil Fertility, Fertilizer Recommendation, Data Mining, Clustering, Classification, Neural Network

I. INTRODUCTION

Extracting meaning full details from large amount of data which can be in different forms or terms is the method known as data mining. These types of terms include binary text, numbers, and texts in the form of consumer data. Data mining is also Known knowledge discovery in Databases (KDD) [1] [6].

There are number of important areas where data mining is widely used such as Future healthcare, Education, Customer Segmentation, Fraud detection, Financial Banking and Agriculture [22].

Data mining in agriculture is a very current research topic because agricultural sector needs more support for its development in developing countries like India. As Mahatma Gandhi said, "India lives in villages and agriculture is the soul of Indian economy" [21]. Agricultural data increases day by day. So there is a need to transform this huge data into technologies and make them available to the farmers [9]. Still recommendations for farmers are based on one to one interaction between farmers and experts and different experts have variety of recommendations. So Recommendation can be provided to farmers using past agricultural data with

help of data mining concepts and they can get optimized results from recommender. [23].The motivation behind this work is to explore data mining techniques, which are suitable for solving complex agricultural problems.

Researchers design and develop applications for solving complex agricultural problems using data mining like Yield Data Prediction, Soil Mapping and Classification, Fertilizer and Pesticide Management, Grading and sorting of agriculture product etc [24].

II. SOIL ANALYSIS FOR FERTILIZER RECOMMENDATION

Among various trust areas of agriculture research as crop cultivation, market analysis and price prediction, rainfall prediction, classification of soil plays an important role because analyzing soil provides major contribution to the support of the farmers. Data mining in agricultural soil datasets is exciting and modern research area [19].

The productive capacity of a soil depends on soil fertility. Soil fertility is the capacity of the soil to support the growth of plants and it is measured by the amount of macro and micronutrients, water, pH etc [10].To maintain nutrient levels in the soil in case of deficiency, fertilizers are added to soil. Better crop yield is highly depends on inputs of fertilizer. Fertilizer not only increases crop yields, but also builds soil organic matter [24].

This paper provides analysis of different data mining techniques to analyze the soil test data sets for fertilizer recommendation.

III. TECHNIQUES USED IN DATA MINING

There are Various Techniques available for data mining which can also be used for agriculture domain.

A. Association rule mining:

Association is a data mining technique in which a pattern is discovered based on a relationship of a particular item on other items in the same transaction. For example, this technique is used in market basket analysis, shopping cart analysis, online advertisement etc [2] [6].

In this method, the focus is on finding relationships between the different items in a transactional database. For example, if there is set of items as I= $\{i_1, i_2, \ldots\ldots, i_n\}$ and the set of transaction as D= $\{t_1, t_2, \ldots..t_n\}$ then the association rule could be,

X => Y, Where X and Y belongs to set I. [19]

B. Clustering

Clustering means identification of common kind of objects. By using clustering method user can additional identify dense and sparse regions in object space and can find overall division pattern among data attributes.

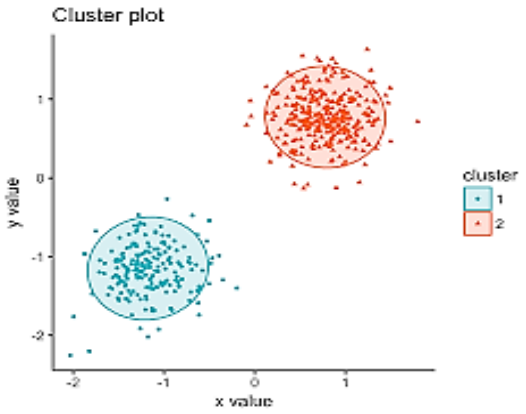

Figure 1: Clustering

Classification way can also be used for productive means of Distinguishing category or category of object but it becomes costly so clustering can be used preprocessing way for attribute subset choice and categorization [2] [3].

C. Classification

Classification is a Technique which forms the common stepping stone for various recursive algorithms and methods designed for data mining. The design aim of a classification algorithm is to divide various data-sets or database which have enough data into different types of sets so that they can be analyzed using a single step exhaustive model. Classification method makes use of statistical techniques like decision trees, linear programming, neural network and statistic-oriented approach, etc. [6]

D. Regression:

Regression is a one type of data mining Method used to fit a formula to a dataset. The process for estimating the relationships among variables is known as regression analysis. Regression analysis can be helpful to understand how the typical value of the dependent variable can be changed. [5][6]

IV. THE SCOPE OF DATA MINING IN AGRICULTURE

There are some research area of agriculture where data mining is useful in many ways. It may provide crucial role in decision making for complex agricultural problems. Some research challenges of agriculture can be solve by data mining algorithm as follow:

A. K-mean in Agriculture: The K-means algorithm is distance based clustering techniques. By applying this algorithm, K cluster are formed. Based on Euclidean distance, object is placed into the respective cluster. [19] The k-means algorithm is used to classify soil in combination with GPS. Classification of plant and soil, grading apples before marketing, Monitoring water quality change, Detecting weeds in precision agriculture, and the prediction of wine fermentation problems can be performed by using a k-means approach. [21]

B. K-nearest neighbor in Agriculture: In pattern recognition, the k-Nearest Neighbors algorithm is a non-parametric method used for classification and regression. In Agriculture, k-NN algorithm is used in simulating daily precipitations and other weather variables and Estimating soil water parameters and Climate forecasting. [21]

C. Neural Networks in Agriculture: In data mining a statistical model known as Artificial Neural Network is a Non-linear predictive model that learns through training and resembles biological neural networks in structure. [2] The neural network is used in Prediction of flowering and maturity dates of soybean and in forecasting of water resources variables in agriculture.

D. SVMs in Agriculture: Support vector machine (SVM) which was originally developed by Vapnik (1998) has been widely applied to many different fields, such as signal process and time series analysis. SVMs are one of the newest supervised machine learning techniques. The current study investigated the applicability of support vector machine in agriculture is in the crop classification and in the analysis of the climate change scenarios. [21][18]

E. Decision Tree in Agriculture: Decision Tree is tree-shaped structures that represent sets of decisions and generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). In agriculture, Decision tree algorithm is used for predicting soil fertility. [2][21]

V. RELATED WORK

P. Jasmine Sheela. Sivaranjani [20] made a comparative study of various classification algorithms such as Random forest, J48, Naïve Bayes and concludes that accuracy level for J48 algorithm is 98.17%, for Random forest algorithm is 90.81%, for Naïve Bayes algorithm is 77.18%, for JRip algorithm is 92.53%, and after using boosting for J48 it increase the accuracy level as 96.73%. This algorithm is used to analyze a large number of datasets and it is used to create a soil fertility prediction in easy manner.

Jay Gholap [7] used Decision Tree algorithm for predicting soil fertility. He also used Selection and Boosting techniques to tune the performance of J48 algorithm. Ad boost is one of the boosting techniques used in this research. After using selection and boosting algorithm it increase the accuracy level to 96.73%.

S.S.Baskar, L.Arockiam, S.Charles [8] classifies the soil dataset using various Classification algorithms such as regression, J48, Naive Bayes. J48 gives the best result and accuracy level is 93.86%. For given soil samples this system will recommend the fertilizer appropriately.

Vrushali Bhuyar [9] focuses on classification of soil sample of Aurangabad District to predict fertility rate. She use J48, Naïve Bayes, and Random forest algorithm to calculate fertility rate of soil data. Her Study shows that among the classifier J48 classifier perform best to predict fertility index. This will help to decision maker to recommend fertilizers accordingly.

Amrutha A, Lekha R, A Sreedevi [10] says that the time taken for soil testing by chemical methods in a laboratory takes a few days, while in the proposed system the results are obtained within 30 minutes. The results obtained from the soil test are fed to sensors and the results are analyzed using a microcontroller which in turn needs a few seconds. The system has been developed for the controlled addition of fertilizers in order to avoid excess/ deficient fertilizers in the soil.

Navneet, Nasib Singh Gill [12] propose a technique that produces a compact decision tree having increased classification accuracy and also Schwartz criterion is used to get the optimal number of clusters. The experimental results show that incorporating both clustering and classification algorithm is fast and more accurate for recommendation of fertilizers of real world soil dataset.

VI. CONCLUSION

There are a lot of opportunities for research on data mining in agriculture to develop practical applications to solve difficult problems in agriculture sector. This paper provides an analysis of data mining techniques used in agriculture soil data and also compares the J48, Naïve Bayes, JRip, K-Means Classifier algorithm. Also suggest that an Artificial Neural Network is best other than other classification algorithms in data mining. It concludes that decision tree algorithm performs best to analyze soil fertility condition and the results and finding motivate us to do further research work on soil data for fertilizer recommendation. This survey helps to the Research Scholar for getting an efficient algorithm selection in according to the fertilizer recommended based on soil fertility.

VII. REFERENCES

[1]. https://en.wikipedia.org/wiki/Data_mining

[2]. Neha Patel, Rakesh Patel, Shruti Upadhyay, "A Survey Paper of Study about Data Mining" International Journal for Research in Computer Science,2016

[3]. Simmi Bagga, Dr. G. N. Singh, "Applications of Data mining" ,international journal for Science and Emerging Technology with latest Trends, 2012

[4]. https://www.tutorialspoint.com/data_mining/dm_classification_prediction.htm

[5]. J. Han and M. Kamber, jianpei "Data mining: Concepts and techniques (3rd edition)", Morgan Kaufman Publishes, 2012

[6]. Aashay kumar Dubey, Saurabh Kamath, Dhruv Kanakia "Learning Data Mining Techniques" International Journal of Computer Applications, 2016.

[7]. Jay Gholap, "Performance tuning of j48 Algorithm for prediction of soil Fertility", AJCSIT, 2012.

[8]. S. S. Baskar, L. Arockiam, S.Charles "Applying Data Mining Techniques on Soil Fertility Prediction", IJCAT, 2013.

[9]. Vrushali Bhuyar, "Comparative Analysis of Classification Techniques on Soil Data to Predict Fertility Rate for Aurangabad District", IJETTCS, 2014.

[10]. Amrutha A, Lekha R, A Sreedevi, "Automatic Soil Nutrient Detection and Fertilizer Dispensary System", IEEE ,2016

[11]. Prof. NileshPatil, Apoorav Singh, Piyush Patel, Nisarg Joshi, "Soil Classification support for Farming Decisions", IJRITCC, 2017.

[12]. Navneet, Nasib Singh Gill , "Algorithm for Producing Compact Decision Trees for

Enhancing Classification Accuracy in Fertilizer Recommendation of Soil",IJCA,2014

[13]. Navneet and Nasib Singh Gill, "Classification Using the Compact Rule Generation", OJCST, 2015.

[14]. Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki, "Analysis of agriculture data using data mining techniques: application of big data",Springer,2017

[15]. http://mpkrishi.org/EngDocs/AgriLeft/Soilt/soiltHome.aspx#SS

[16]. Soil Test Interpretations and Recommendations Handbook, Daryl D. Buchholz

[17]. A. Mucherino, Petraq Papajorgji, P. M. Pardalos," A survey of data mining techniques applied to agriculture", Springer, 2009

[18]. Niketa Gandhi, Leisa J. Armstrong, "A review of the application of data mining techniques for decision making in agriculture ",IEEE,2016

[19]. N. Hemageetha," A Survey on Application of data mining techniques to analyze the soil for agriculture purpose", IEEE,2016

[20]. P. Jasmine Sheela, M.Sc., K. Sivaranjani, M.Sc., M.Phil., "A Brief Survey of Classification Techniques Applied To Soil Fertility Prediction", ICETSH-2015

[21]. Hetal Patel, Dharmendra Patel," A Brief survey of Data Mining Techniques Applied to Agricultural Data", IJCA, 2014

[22]. http://bigdata-madesimple.com/14-useful-applications-of-data-mining/

[23]. Kiran Shinde, Jerrin Andrei, Amey Oke," Web Based Recommendation System for Farmers", IJRITCC,2015

[24]. Tejas S. Mehta, Dr. Dhaval R. Kathiriya, "Survey of Data Mining Techniques in Precision Agriculture", IJSR, 2015