



## Information Technology

Elixir Inform. Tech. 67 (2014) 21773-21774

Elixir  
ISSN: 2229-712X

# Excellent agriculture using data mining techniques-a study

G.Sivakumar, C.Parameswari, D.Roja Ramani and M.M.Yamuna Devi

Department of Information Technology, Sethu Institute of Technology, Virudhunagar, India.

### ARTICLE INFO

#### Article history:

Received: 3 August 2013;

Received in revised form:

24 January 2014;

Accepted: 15 February 2014;

#### Keywords

Data mining,  
Agriculture,  
Soil management,  
Eco system.

### ABSTRACT

This paper presents a deep study on the application of Data mining Techniques in the field of Agriculture. Data mining software applications using various methodologies have been developed by both commercial and research centers. Using this software applications Agricultural eco system is also managed. Eco system management includes soil management. Recent technologies are nowadays able to provide a lot of information on agricultural-related activities, which can then be analyzed in order to find important information. Efficient techniques can be developed and tailored for solving complex soil datasets using data mining.

© 2014 Elixir All rights reserved

### Introduction

In many developing countries, hunger is forcing people to cultivate land that is unsuitable for agriculture and which can only be converted to agricultural use through enormous efforts and costs, such as those involved in the construction of terraces. Each country is known for its core competence. India's is agriculture. Yet, it only accounts for 17 per cent of the total Gross Domestic Product. With the pressure of urbanization, it is going to be a challenge to produce food for more people with less land and water. Agriculture or farming forms the backbone of any country economy, since a large population lives in rural areas and is directly or indirectly dependent on agriculture for a living. Income from farming forms the main source for the farming community. The essential requirements for crop harvesting are water resources and capital to buy seeds, fertilizers, pesticides, labor etc. Most farmers raise the required capital by compromising on other necessary expenditures, and when it is still insufficient they resort to credit from sources like banks and private financial institutions. In such a situation, the repayment is dependent on the success of the crop. If the crop fails even once due to several factors, like bad weather pattern; soil type; improper, excessive, and untimely application of both fertilizers and pesticides; adulterated seeds and pesticides etc. then he is pushed into an acute crisis causing severe stress [3]. In addition, the plant growth depends on multiple factors such as soil type, crop type, and weather. Due to lack of plant growth information and expert advice, most of the farmers fail to get a good yield. Data mining is the process of discovering previously unknown and potentially interesting patterns in large datasets (Piatetsky-Shapiro and Frawley, 1991). The 'mined' information is typically represented as a model of the semantic structure of the dataset, where the model may be used on new data for prediction or classification. Alternatively, human domain experts may choose to manually examine the model, in search of portions that explain previously misunderstood or unknown characteristics of the domain under study. Monitoring agricultural crop conditions, weather and climate, and ecosystems, and providing decision support for agricultural

planning and policy-making are critically important for the development of agriculture to support economic and social development [1]. "Agricultural and biological research studies have used various techniques of data analysis including, natural trees, statistical machine learning and other analysis methods" (Cunningham and Holmes, 1999) [2].

### Data Mining Approaches

#### Support Vector Machines

Support Vector Machines (SVM) is binary classifiers (Burgess, 1998; Cortes and Vapnik, 1995). SVM is able to classify data samples in two disjoint classes. The basic idea behind is classifying the sample data into linearly separable. Support Vector Machines (SVMs) are a set of related supervised learning methods used for classification and regression. In simple words given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that predicts whether a new example falls into one category or the other. SVM is used to assess the spatiotemporal characteristics of the soil moisture products [4].

#### Genetic Algorithm

The Genetic Algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of Evolutionary Algorithm (EA), which generates solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection and crossover. Soil liquefaction is a type of ground failure related to earthquakes. It takes place when the effective stress within soil reaches zero as a result of an increase in pore water pressure during earthquake vibration (Youd, 1992). Soil liquefaction can cause major damage to buildings, roads, bridges, dams and lifeline systems, like the earthquakes. Genetic Algorithm approach is used for assessing the liquefaction potential of sandy soils (G. Sen et al. Nat. Hazards Earth Syst. Sci., 2010).

#### Artificial neural networks

Artificial neural networks are models inspired by animal central nervous systems (in particular the brain) that are capable

of machine learning and pattern recognition. They are usually presented as systems of interconnected "neurons" that can compute values from inputs by feeding information through the network. Artificial neural networks are instead used for predicting the total necessary power of agricultural machinery[7]

### K-Means Algorithm

The algorithm for partitioning, where each cluster's center is represented by mean value of objects in the cluster.



Fig. 2. Data mining process.

### Recent Researches In Agriculture

Since spatial association mining needs to evaluate multiple spatial relationships among a large number of spatial objects, the process could be quite costly. An interesting mining optimization method called progressive refinement can be adopted in spatial association analysis. The method first mines large data sets roughly using a fast algorithm and then improves the quality of mining in a pruned data set. The k-means algorithm randomly selects k number of objects, each of which initially represents a cluster mean or center. For each of the remaining objects, an object is assigned to the cluster to which it is most similar, based on the distance between the object and the cluster mean. Then it computes new mean for each cluster. This process iterates until the criterion function converges. The above concept is applied in the area of agriculture where giving the temperature and the rainfall as the initial spatial data and then by analyzing the agricultural meteorology for the enhancement of crop yields and also reduce the crop losses. Recent studies by agriculture researchers in Pakistan (one of the top four cotton producers of the world) showed that attempts of cotton crop yield maximization through pro-pesticide state policies have led to a dangerously high pesticide usage. These studies have reported a negative correlation between pesticide usage and crop yield in Pakistan. Hence excessive use (or abuse) of pesticides is harming the farmers with adverse financial, environmental and social impacts. By data mining the cotton Pest Scouting data along with the meteorological recordings it was shown that how pesticide usage can be optimized (reduced). Clustering of data revealed interesting patterns of farmer practices along with pesticide usage dynamics and hence help identify the reasons for this pesticide abuse[8].

A research has been done on the soil data sets. The research compares the different classifiers and the outcome of this research could improve the management and systems of soil uses throughout a large number of fields that include agriculture, horticulture, environmental and land use management [9]. Data

mining and the various methodologies associated with it can reduce the complexity of the data enabling farmers to make decisions more easily[10].



### Conclusion

From the study it is very clear that data mining techniques are very much useful in the field of agriculture and there are lot of rooms for research in the field of agriculture using data mining techniques.

### References

- [1] "Fuzzy Logic", Stanford Encyclopedia of Philosophy, Stanford University, 2006-07-23. Retrieved 2008-09-29.
- [2] Cunningham, S. J., and Holmes, G. (1999), "Developing innovative applications in agriculture using data mining", In the Proceedings of the Southeast Asia Regional Computer Confederation Conference, 1999.
- [3] Sudarshan Reddy S, Vedantha S, Venkateshwar Rao B, Sundar Ram Reddy and Venkat Reddy, "Gathering Agrarian Crisis Farmers Suicides in Warangal district", Citizens Report, 1998.
- [4] Anish C. Turlapaty, Valentine Anantharaj, Nicolas H. Younan, Spatiotemporal.
- [5] "Consistency analysis of AMSR-E soil moisture data using wavelet-based feature extraction and one-class SVM", In the Proceedings of the Annual Conference Baltimore, Maryland, March 9-13, 2009.
- [6] Dr. D. Ashok Kumar, N. Kannathasan, "A Survey on Data Mining and Pattern Recognition Techniques for Soil Data Mining".
- [7] Q. Lai, H. Yu, H. Chen, Y. Sun, "Prediction of Total Power of Agricultural Machinery Using Artificial Neural Networks", IEEE Conference Proceedings, International Conference on Computing, Control and Industrial Engineering, Volume 2, Wuhan, China, 394-396, 2010.
- [8] Abdullah, Ahsan; Brobst, Stephen; Pervaiz, Ijaz; Umar, Muhammad; Nisar, Azhar (2004), "Australasian Workshop on Data Mining and Web Intelligence", Dunedin, New Zealand.
- [9] Ramesh and K. Ramar, "Classification of Agricultural Land Soils: A Data Mining Approach", Agricultural Journal Year: 2011, Volume: 6, Issue: 3, Page No.: 82-86.
- [10] D. Diepeveen and L. Armstrong, "Identifying key crop performance traits using data mining", World Conference on Agricultural Information and IT.