Chapter 10
Utility and Significance of Vague Set Theory and Advanced Optimization Mechanisms for Uncertainty Management

Sowkarthika B
Madhav Institute of Technology and Science, India

Akhilesh Tiwari
Madhav Institute of Technology and Science, India

R. K. Gupta
Madhav Institute of Technology and Science, India

Uday Pratap Singh
Madhav Institute of Technology and Science, India

ABSTRACT
In this digital world, tremendous data are generated in every field. Useful information is inferred out of this data, which is valuable for effective decision making. Data mining extracts the interesting information from huge volumes of data. Association rule (AR) mining is one of the core areas of data mining where interesting information is extracted in the form of rules. Traditional AR mining is incapable of handling uncertain situations. In order to handle uncertainty, mathematical tools like vague theory can be utilized with AR mining methodologies for the development of novel vague theory based algorithms, which will be more suitable in effectively handling vague situations that helps framing effective selling strategy. Since an organization can’t analyze the huge rule set obtained from these algorithms, every resultant rule should have a certain ratio of factors customized to the interest of the organization that can be achieved through optimization algorithms. This chapter explores the significance of vague theory and optimization means for effective uncertainty management.

DOI: 10.4018/978-1-5225-3531-7.ch010
INTRODUCTION

Data is being generated and gathered in every field which results in an increase in the size of the data. Any organization is not in need of data. It needs information that is inferred out of data. That is, the data is useless until and unless it is converted to useful information. Since manual examination is entirely incomprehensible with huge volume of data, automated tools came into the picture. Data Mining is an automated tool which helps in extracting useful or interesting patterns or information from huge volumes of data like database or other data repositories. There are various tasks pertaining to data mining. Some of the essential tasks are Association Rule mining (AR mining), Classification and Clustering. Since classification and clustering need AR mining to be performed, and AR mining has direct relevance with real-life scenarios, the authors focus more on AR mining.

AR mining is useful in examining and analyzing the customer’s behavior. Association in retail store refers to the patterns that contain items that are frequently purchased together. It gives the frequently occurring patterns in the transaction data. This basically gives the relationship between the products that is useful in strategic business decisions and product marketing. The problem of AR mining is defined by Agrawal (Han & Kamber, 2006) as:

Let \( I = \{i_1, i_2, \ldots, i_n\} \) be a set of \( n \) binary attributes called items.\( I \)

Let \( D = \{t_1, t_2, \ldots, t_m\} \) be a set of transactions called the database.\( D \)

Each transaction in \( D \) has a unique transaction identifier \( TID \) and contains a subset of the items in \( I \). A rule is defined as an implication of the form \( A \Rightarrow B \) where \( A, B \subseteq I \). Also, \( A \cap B = \emptyset \). For example, from \( \text{Bread} \Rightarrow \text{Jam} \), the retailer could get an idea that whenever \( \text{Bread} \) is purchased, \( \text{Jam} \) is also purchased. Based on the rule generated, some discount could be offered or the store layout could be changed accordingly thereby boosting the sales. Association rules are represented by ‘if-then’ rules. In \( A \Rightarrow B \), \( A \) is the antecedent and \( B \) is the consequent and it implies ‘if \( A \) is purchased, then \( B \) is more likely to be purchased’. Antecedent is the set of items in the database and consequent is the set of items that are in relationship with the antecedent. The two important criteria in identifying a relationship and making a rule interesting are support and confidence. AR is considered interesting only when it satisfies minimum support and minimum confidence thresholds. The thresholds are set by the domain experts or users who have rich knowledge of the system. These play a key role in making any business decisions. A strong association rule is generated if the support and confidence of a rule are greater than minimum support and minimum confidence thresholds.

Support gives how frequent \( A \) and \( B \) occur in the database of transactions \( D \). Support of the rule is the probability of the occurrence of items \( A \) and \( B \) in the database \( D \).

\[
\text{support}(A \rightarrow B) = P(A \cup B) = \frac{|A \cup B|}{|D|}
\]