Chapter 1
Critical Parameters for Fuzzy Data Mining

Sinchan Bhattacharya
Maharaja Agrasen Institute of Technology, India

Vishal Bhatnagar
Ambedkar Institute of Advanced, India

ABSTRACT

Research on data mining is increasing at an incessant rate and to improve its effectiveness other techniques have been applied such as fuzzy sets, rough set theory, knowledge representation, inductive logic programming, or high-performance computing. Fuzzy logic due to its proficiency in handling uncertainty has gained its importance in a variety of applications in combination with the use of data mining techniques. In this chapter we take this association a notch further by examining the parameters which allow fuzzy sets and data mining to be combined into what has come to be known as fuzzy data mining. Analyzing and understanding these critical parameters is the main purpose of this chapter, so as to acquire maximum efficiency in applying the same which impelled the authors to work extensively and find out the crucial parameters essential to the application of fuzzy data mining.

INTRODUCTION

"Data mining is the analysis of observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner" (Hand et al., 2001). Ranjan (2008) provided an alternative view to this concept as he dined data mining as a technology that has been developed for investigation and analysis of large quantities of data to discover meaningful patterns and rules, facilitating the selection procedure.

Fuzzy logic is a multivalued logic that allows intermediate values to be defined between predictable values such as true and false, yes and no or high and low, concepts which are unclear or have value out of index or do not have a distinct value, can be formulated mathematically and processed by computers, so as to facilitate programming in computers (Zadeh, 1984). The main characteristic of fuzzy logic is that it allows us to define values without specifying a precise value which is not possible with classical logic, upon which computer development has been, based so far (Aroba et al., 2007).

DOI: 10.4018/978-1-4666-7456-1.ch001

Data mining is applied in various fields where hidden information or a pattern is to be extracted or where the data is insufficient. Data mining is of a more explanatory nature, and patterns discovered in a data set are usually of a descriptive rather than of a predictive nature (Hüllermeier, 2005). This provides the scope for application of fuzzy logic in the mining system. Fuzzy logic imparts a skeleton to handle uncertainty. Application of the data mining techniques is a fast emerging trend in the modern world. This amalgamation gives wings to the concept of data mining extending its utility and flexibility.

It is quite discernable that the combination of these two concepts is not uncomplicated. There have to be certain rules and restrictions while applying the concepts of fuzzy sets in data mining systems and vice versa.

The chapter is organized as follows: Section I, presents a brief review of literature on fuzzy data mining. Section II provides a glimpse of all the application areas that have been exploited so far. Section III discusses the need for these critical parameters which essentially elaborates on the problems faced by the application of fuzzy data mining and the facility these parameters can bring. Section IV discusses the research methodology adopted in the formulation of the work. Section V presents the critical parameters for the application of fuzzy data mining. In Section VI discussion is done on the research in this area of fuzzy data mining. It also suggests the prospects of fuzzy data systems in the future while section VII presents the limitations of the study carried out by the authors. Finally, the chapter is concluded in section VIII by urging the researchers to use fuzzy data mining in various business purviews.

BACKGROUND

The concept of data mining was introduced in the late 1980's and since then it has been a blooming field of research. In his work, Kruse et al. (1999) defined, data mining as a set of tasks (Fayyad et al., 1996; Nakhaeizadeh, 1998) which include segmentation, classification, concept description, prediction, deviation analysis, and dependency analysis.

The foundations of fuzzy logic were laid in the year 1965 by Lofti, A. Zadeh. He stated that, "In a narrow sense, fuzzy logic is a logical system which is an extension of multivalued logic and is intended to serve as logic of approximate reasoning but in a wider sense, fuzzy logic is more or less synonymous with the theory of fuzzy sets, that is, a theory of classes with unsharp boundaries" (Zadeh, 1994).

Fuzzy data mining methods denote the approaches to analyze fuzzy data based on the data mining techniques available in order to predict a trend or a pattern from the available fuzzy data (Feil & Abonyi, 2008). The fuzzy logic theory brings a paradigm in work with the graduation, uncertainty and ambiguity described by linguistic expressions derived from the operations of data mining which uses knowledge that does not have clearly defined boundaries.

Research has been done in surfeit in the separate fields of fuzzy logic and data mining, and combination of these two fields is gradually gaining recognition. It was Hüllermeier (2005), who discovered the potential of combining fuzzy sets and data mining techniques. In his work in the year 2005, he stated the potential contributions of fuzzy logic in data mining. Pedrycz in 1998 and Baraldi and Blonda in 1999 independently established that data mining and fuzzy logic can be used to find various patterns only if the data is such that it can be of assistance in formation of such patterns.

Many other discoveries were made related to the application areas of fuzzy data mining which gradually led to certain rules. Au Keith and Chan in 1999 and Simha and Iyengar in 2006 realized that the linguistic variables used in fuzzy should not lose its significance while it is applied in the data mining systems. The data fed in the fuzzy