Usage of Fuzzy, Rough, and Soft Set Approach in Association Rule Mining

Satya Ranjan Dash, School of Computer Application, KIIT University, Bhubaneswar, Odisha, India

Satchidananda Dehuri, Department of Systems Engineering, Ajou University, Suwon, South Korea

Uma kant Sahoo, School of Computer Application, KIIT University, Bhubaneswar, Odisha, India

ABSTRACT

This paper is two folded. In first fold, the authors have illustrated the interplay among fuzzy, rough, and soft set theory and their way of handling vagueness. In second fold, the authors have studied their individual strengths to discover association rules. The performance of these three approaches in discovering comprehensible rules are presented.

Keywords: Association Rule Mining, Data Mining, Fuzzy Sets, Rough Sets, Soft Sets

1. INTRODUCTION

An investigation of the Laws of Thought George (1958) is a variant of ordinary elementary logic differing in its values, operations, and laws. Instead of the usual logic of numbers, Boolean logic is the algebra of binary values \( \{0, 1\} \), have attracted many researchers from diverse field to a common platform of digital world. The operations defined on Boolean algebra are combinations of conjunction \( \wedge \), disjunction \( \vee \) and negation \( \neg \), and with constants 0 and 1. The laws defined here, are those propositions that hold for all values of their variables, for example \( x \lor (y \land x) = x \). This has high applicability in digital logic and computer programming. It assumes that the values attained by a Boolean variable is distinct, and is either true or false.

However, the society we live in is not homogeneous. Any study of the interplay between empirical material and theoretical concepts has to take into account the issue of vagueness. We use the term vague, rather than a generic term, since it encompasses various nuances of being imprecise, uncertain, ambiguous, indeterminate, fuzzy, etc. There is a striking contrast between

DOI: 10.4018/jalr.2012070105
our ability to successfully recognize objects and our inability to give complete descriptions about them. In Skala and Canaria (1982), the author says that from experience any attempt at characterizing any event completely leads to an uncertainty as to the adequacy of the identification criteria suggested. Hence, Boolean logic ceases to explain even simple problems with vague parameters.

Hence, we cannot successfully use classical methods because of the various uncertainties involved in real life problems. However there are mathematical theories of probability and interval mathematics which deal with uncertainties, but these have their own difficulties.

Probability theory assumes that all events occur randomly. This limits its uses to random events only. However real life events depend on previously occurring events hence cannot be explained by probability theory. Also to apply probability theory we need to have a large number of trials which is always not feasible specially while dealing with economic, environmental or social problems. On the other hand interval mathematics takes into account the errors of calculation by constructing an interval estimate for the exact solution. For cases where different uncertainties are involved, this technique ceases to be adaptable.

Hence, to address the problem of uncertainties various approaches like fuzzy Zadeh(1965), rough Pawlak, Z.(1982), and soft set theory Molodtsov, D.(1999) are developed under the umbrella of soft computing. In Section 2, we discuss the basic preliminaries of these approaches, in addition to association rule mining.

Association rule mining is one of the important tasks of data mining. Association rule mining maps a given problem into the popular transactions versus items domain which was the first application in the realm of market basket analysis. For instance, documents could be treated like transactions and the words within the documents could be the items. The technique allows for capturing in the given transactions a possible trend between two disjoint sets of items, i.e., all possible rules that reflect the presence of some items according to the presence of other items in the same transaction. An association rule is defined as a correlation \( X \rightarrow Y \), where both \( X \) and \( Y \) are defined as sets of items, interchangeably called objects or attributes. More will be discussed in Subsection 2.4.

The interplay among fuzzy, rough, and soft set theory are discussed in Section 3. Association rule mining using fuzzy, rough, and soft set approaches are demonstrated in Section 4 through an example. The conclusions are derived in Section 5.

2. PRELIMINARIES

In Section 1 we saw that the classical mathematical techniques fail to explain certain problems concerning uncertainties. Now we shall see some of the mathematical theories developed to tackle this problem.

2.1. Fuzzy Sets

Fuzzy set theory is an excellent mathematical tool to handle uncertainty occurred due to vagueness Zadeh (1965). Unlike crisp situations (dealt by boolean logic), fuzzy situations cannot be well-defined as there is always some uncertainty involved in them. The impetus for the transition from a crisp theory to a fuzzy one derives from the fact that both the generality of a theory and its applicability to real-world problems are substantially enhanced by replacing the concept of a set with that of a fuzzy set.

The notion of fuzzy set provides a convenient tool for representing vague concepts by allowing partial memberships. Let \( U \) be a finite and nonempty set called universe. A fuzzy set \( A \) of \( U \) is defined by a membership function:

\[
\mu_A : U \rightarrow [0,1]
\]

There are many definitions for fuzzy set complement, intersection, and union. With the min-max system proposed by Zadeh (1965), fuzzy set operators are defined component wise as:
12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:
www.igi-global.com/article/usage-of-fuzzy-rough-and-soft-set-approach-in-association-rule-mining/81214?camid=4v1

This title is available in InfoSci-Journals, InfoSci-Journal Disciplines Medicine, Healthcare, and Life Science. Recommend this product to your librarian:
www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

Head Motion Stabilization During Quadruped Robot Locomotion: Combining CPGs and Stochastic Optimization Methods
www.igi-global.com/article/head-motion-stabilization-during-quadruped/55448?camid=4v1a

Wave Propagation in Filamental Cellular Automata
www.igi-global.com/article/wave-propagation-filamental-cellular-automata/41944?camid=4v1a

Correctness of Self-Stabilizing Algorithms Under the Dolev Model When Adapted to Composite Atomicity Models
Chih-Yuan Chen, Cheng-Pin Wang, Tetz C. Huang and Ji-Cherng Lin (2012). International Journal of Artificial Life Research (pp. 16-31).
www.igi-global.com/article/correctness-of-self-stabilizing-algorithms-under-the-dolev-model-when-adapted-to-composite-atomicity-models/101292?camid=4v1a
Evolutionary Turing Machines: The Quest for Busy Beavers
www.igi-global.com/chapter/evolutionary-turing-machines/28322?camid=4v1a