ABSTRACT

Bayesian Decision theoretic rough set has been invented by the author. In this paper the attribute reduction by the aid of Bayesian decision theoretic rough set has been studied. Lot of other methods are there for attribute reduction such as Variable precision method, probabilistic approach, Bayesian method, Pawlaks rough set method using Boolean function. But with the help of some example it is shown that Bayesian decision theoretic rough set model gives better result than other method. Lastly an example of HIV/AIDS is taken and attribute reduction is done by this new method and various other method. It is shown that this method gives better result than the previously defined methods. By this method the authors get only the reduced attribute age which is the best significant attribute. Though in Pawlak model age sex or age living status are the reduced attribute and variable precision method fails to work here. In this paper attribute reduction is done by the help of discernibility matrix after determining the positive, boundary and negative region. This model is a hybrid model of Bayesian rough set model and decision theory. So this technique gives better result than Bayesian method and decision theoretic rough set method.

Keywords: Bayesian Decision Theoretic Rough Set Model, Bayesian Method, Bayesian Rough Set, Decision Theoretic Rough Set, Pawlaks Rough Set Method Using Boolean Function, Probabilistic Approach, Variable Precision Method

1. INTRODUCTION AND PRELIMINARIES

Pawlak (1982) defined the concept of Rough set. Most of the preliminaries are defined in the paper ‘A study on Bayesian decision theoretic rough set method’. So they are not cited twice.

1.1. A BAYESIAN DECISION THEORETIC ROUGH SET MODEL

The concept of Bayesian decision theoretic rough set model is as follows:

\[ \text{apr}^*_D([x]_c) = \text{Pos}^*_D([x]_c) = \bigcup \{ [x]_c : \text{[x]}_c \cap D_i = \emptyset \} \cap \text{P}(D) \}
\]
\[ \text{Neg}^*_D([x]_c) = \bigcup \{ [x]_c : \text{[x]}_c \cap D_i \neq \emptyset \} \cap \text{P}(D) \}
\]

DOI: 10.4018/ijrsda.2014010102
The main aim of this paper is to show the process of attribute reduction using Bayesian decision theoretic rough set model. Also by an example of HIV AIDs a comparative study is shown between various types of rough sets and the Bayesian decision theoretic rough set model.

Attribute reduction is an important problem of rough set. For classification tasks we consider two possible interpretations of the concept of a reduct. The first interpretation views a reduct as a minimal subset of attributes that has the same classification power as the entire set of condition attributes. The second interpretation views a reduct as a minimal subset of attributes that produces positive and boundary decision rules with precision over certain tolerance levels. Studies on attributes reduction can therefore be divided into two groups.

The first group concentrates on the description of the classification power i.e. the decision class or classes to which an equivalence class belongs. An object indiscernible within an equivalence class is classified by one decision class in consistent decision tables and may be classified by more than one decision class in inconsistent decision tables. In general a membership distribution function over decision classes may be used to indicate the degree to which an equivalence class belongs. Zhang et al. (2012) propose the maximum distribution criterion based on the membership distribution function, which pay more attention to the decision classes that take the maximum distribution over the set.

A reduct can be defined as a minimal subset of attributes that has the same classification power for all objects in the universe in terms of the types of decision classes such as generalized decision and majority decision or the proportion of the decision classes such as decision distribution and maximum distribution.

The second group concentrates on the evaluation of the classification power i.e. the positive, boundary and/or negative regions of decision classes to which an equivalence class belongs. In the Pawlak rough set model (Pawlak, 1982), each equivalence class may belong to one of the two regions, positive or boundary. The positive region is the union of equivalence classes that induce certain classification rule. The boundary region is the union of equivalence classes that induce uncertain classification rules. The negative region is in fact the empty set. The positive and boundary region induce two different types of decision rules called the positive rules and boundary rules respectively. While a positive rule leads to a definite decision, a boundary rule leads to a “wait and see” decision.

In rough set theory by a reduct we usually mean a compact yet informative subset of the available attributes.

1.2. DEFINITION (DECISION REDUCT)

Let $S = (U, A, d)$ be a decision table with a decision attribute $d$ indicating belongingness of objects to investigated concepts. A subset of attributes $DR \subseteq A$ will be called a decision reduct iff the following conditions are met:

1. For any pair $u, u' \in U$ of objects belonging to different decision classes (i.e. $d(u) \neq d(u')$) if $u$ and $u'$ are discerned by $A$ (i.e. there is $a \in A$ such that $a(u) \neq a(u')$) then they are also discerned by $DR$.
2. There is no proper subset $DR' \subseteq DR$ for which the first condition holds.

A decision reduct can be interpreted as a set of attributes that are sufficient to discriminate all objects from different decision classes. At the same time this set has to be minimal in a sense that no further attributes can be removed from DR without losing the discernibility property. For example $\{a_3, a_4\}$ and $\{a_3, a_5\}$ are decision reducts of the decision table $S_4$ from the table. The first condition in the above definition is often replaced by some other requirements for preserving information about decision while reducing attributes. In this paper for simplicity we restrict ourselves to the above discernibility based criterion which is a very well documented in the rough set literature.
Olympics Big Data Prognostications

[www.igi-global.com/article/olympics-big-data-prognostications/163102?camid=4v1a](www.igi-global.com/article/olympics-big-data-prognostications/163102?camid=4v1a)