A Study on Bayesian Decision Theoretic Rough Set

Sharmistha Bhattacharya Halder, Department of Mathematics, Tripura University, Bikramnagar, Tripura, India

ABSTRACT

The concept of rough set was first developed by Pawlak (1982). After that it has been successfully applied in many research fields, such as pattern recognition, machine learning, knowledge acquisition, economic forecasting and data mining. But the original rough set model cannot effectively deal with data sets which have noisy data and latent useful knowledge in the boundary region may not be fully captured. In order to overcome such limitations, some extended rough set models have been put forward which combine with other available soft computing technologies. Many researchers were motivated to investigate probabilistic approaches to rough set theory. Variable precision rough set model (VPRSM) is one of the most important extensions. Bayesian rough set model (BRSM) (Slezak & Ziarko, 2002), as the hybrid development between rough set theory and Bayesian reasoning, can deal with many practical problems which could not be effectively handled by original rough set model. Based on Bayesian decision procedure with minimum risk, Yao (1990) puts forward a new model called decision theoretic rough set model (DTRSM) which brings new insights into the probabilistic approaches to rough set theory. Throughout this paper, the concept of decision theoretic rough set is studied and also a new concept of Bayesian decision theoretic rough set is introduced. Lastly a comparative study is done between Bayesian decision theoretic rough set and Rough set defined by Pawlak (1982).

Keywords: Bayesian Decision Theoretic Rough Set, Bayesian Rough Set, Decision Theoretic Rough Set, Probabilistic Rough Set Etc, Rough Set

1. INTRODUCTION

Pawlak (1982) defined the concept of Rough set. And it is aimed at data analysis problems involving uncertain, imprecise or incomplete information, Pattern recognition, machine learning, knowledge acquisition, economic forecasting and data mining. Knowledge classification is a fundamental problem in rough set theory. In Pawlak’s rough set model, the degree of set overlap was not considered, namely, the classification must be totally correct or certain. Therefore, original rough set model cannot effectively deal with data sets which have noisy data and latent useful knowledge in the boundary region may not be fully captured. In order to overcome the limitations, some extended rough set models have been put forward which combine with other available soft computing technologies. Later on Researchers introduced many new concepts of rough set e.g. Probabilistic rough set, 0.5 probabilistic model, symmetric variable precision rough set model, asymmetric variable precision rough set model, statistical rough set model, fuzzy rough set model, covering rough set model, tolerance rough set model, dominance based rough set model. Also depending on the Bayes theorem

DOI: 10.4018/ijrsda.2014010101
in 2007 Yiyu Yao (2007) introduced decision theoretic rough set. In 2002 Sleazak and Ziarko (2002) the concept of Bayesian rough set model was introduced and a detail study was done in 2012 (Zhang, Hangyun et al., 2012). After that lots of applications in attribute reduction were investigated by various Researchers.

Variable precision rough set model (VPRSM) (Ziarko, 1993) is one of the most important extensions. In the model, standard inclusion relation is extended to majority inclusion relation, and the novel notion can be able to allow for some degree of misclassification in the largely correct classification. The strict functional or dependent relations between attributes will be softened.

As a result, more general association decision rules including deterministic and probabilistic ones can be obtained in VPRSM.

Subsequently, Ziarko et al. (1994) put forward an asymmetric variable precision rough set model (AVPRSM), and the model becomes more general and flexible. Variable precision rough set models, symmetric or asymmetric, involve some parameters, \( \beta \). Different parameters will result in different models, and the extracted decision rule sets may be distinct.

In the applications, it is not clear how to find out the optimal parameters and their values are often selected based on the decision makers’ previous knowledge of the domain and their intuition or the proposed criteria.

The connections between rough sets and Bayes’ theory were analyzed by Pawlak. Rough set theory offers a new view on Bayes’ theory, and any decision data set in rough set theory will satisfy the total probability theorem and Bayes’ theorem. Based on Bayesian decision procedure with minimum risk, Yao (1990) put forward a new model called decision theoretic rough set model (DTRSM) which brings new insights into the probabilistic approaches to rough set theory. DTRSM provides a general framework for comparing and synthesizing probabilistic rough set approximations. It not only has good theoretical foundation, but also possesses reasonable semantic interpretation. The Pawlak’s rough set model, VPRSM and AVPRSM can be directly derived from DTRSM under relevant loss functions. If the practical decision problems involve cost or risk environments, the DTRSM will be more beneficial for decision making compared with original rough set model.

Moreover, VPRSM and AVPRSM can be considered as an intermediate step when using the decision theoretic approach for rough analysis.

According to Bayesian reasoning, Sleazak and Ziarko (2002) presented Bayesian rough set model (BRSM), in which the parameters that control the approximation regions in VPRSM are determined by the prior probability of occurrence of the target event under consideration. As the hybrid development, BRSM is reasonable for handling some practical domains in which the prior probability of the assumption will be affected by adding some new evidences, such as medical diagnosis, fault detection, economy forecasting and so on. Since the prior probability can be estimated from data set itself, BRSM is more objective when compared with the parametric versions of variable precision rough set model.

A more general parametric modification of BRSM, called variable precision Bayesian rough set model (VPBRSM), was further proposed by Sleazak (2003) which allows single parameter-controlled degree of \( \varepsilon \)-precision in the approximation region definition. VPBRSM is more applicable to practical data analysis problems where small deviations from prior probability are likely to occur due to noise or measurement inaccuracy. Based on the Bayes factor and the inverse probabilities, Sleazak (1993) also introduced a parameterized extension of rough set model, called rough Bayesian model (RBM) which can be utilized very well if the prior and posterior probabilities derivable from data or background knowledge are not reliable. Under associated parameters, the Pawlak’s rough set model, VPRSM and BRSM can also be derivable from RBM.

Throughout this paper various defined concepts and their possible extensions were discussed. Also the concept of Bayesian decision
Understanding User Experience
Camille Dickson-Deane and Hsin-Liang (Oliver) Chen (2018). Encyclopedia of
Information Science and Technology, Fourth Edition (pp. 7599-7608).
www.igi-global.com/chapter/understanding-user-experience/184455?camid=4v1a