Variable Precision Rough Sets Theory and its Application

Student: Jyh-Hwa Hsu   Adviser: Dr. Chao-Ton Su

Department of Industrial Engineering and Management

National Chiao Tung University

Abstract

The Variable Precision Rough Sets (VPRS) theory is a powerful tool for data mining, as it has been widely applied to acquire knowledge. Despite its diverse applications in many domains, the VPRS theory unfortunately cannot be applied to real world classification tasks involving continuous attributes. This requires a discretization method to pre-process the data. Also, the VPRS theory lacks a feasible method to determine a precision parameter (β) value to control the choice of β -reducts. In this study we first propose a new algorithm, named the extended Chi2 algorithm that uses a Chi2 algorithm as a basis, whereby the extended Chi2 algorithm improves the Chi2 algorithm in that the value of pre-defined misclassification rate (δ) is calculated based on the training data itself. In addition, an effective method is proposed to select the β -reducts. First, we calculate a precision parameter value to obtain the subsets of information system that are based on the least upper bound of the data misclassification error. Next, we measure the quality of classification and remove redundant attributes from each subset.

Five numerical examples are analyzed in this study. By running the software of See5, our proposed extended algorithm possesses a better performance than the Chi2 algorithm. To show the effectiveness of the proposed β -reducts selection approach,

a simple example and a real-world medical case are analyzed. Comparing the implementation results from the proposed method with the neural network approach, our proposed approach demonstrates a better performance. Finally, a real example from communication industry is analyzed. The VPRS theory using our proposed procedures is applied to reduce the Radio Frequency (RF) test items in mobile phone manufacturing. Implementation results show that the test items have been significantly reduced. By using these remaining test items, the inspection accuracy is very close to that of the original test procedure. Also, VPRS demonstrates a better performance than that of the decision tree approach.