Two Decision Supporting Techniques—Decision Tree and Rough Set based on SCILAB

Wang Guanjue, Wang Yuhua, Zheng Qi,Jia Haitao,Chu Yanbo,Duan Peikun
Beijing University of Posts and Telecommunications,Beijing,China,*100876*
lellewang@gmail.com

Abstract

SCILAB is an open source platform for numerical computation providing a powerful open computing environment for engineering and scientific applications.

In this paper, we will analyze the merits of SCILAB language as well as four reasons why we chose SCILAB as our development platform for this discrete mathematics toolbox, and how we came up with the topic Decision Tree and Rough Set.

A decision tree is a decision supporting tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. In general, there are two types of decision tree induction, i.e., crisp decision trees and fuzzy decision trees.

Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal. Another use of decision trees is as a descriptive means for calculating conditional probabilities. Various process parameters exert different effects in different processes. In order to study the relationships among the process parameters, we provide several examples introducing the application of decision trees, such as the data mining in business administration, the analysis of the examination results, the model of predicting smokers' quit intentions, etc. We will give a brief explanation of the decision trees in the thesis and realize several algorithms of decision tree based on SCILAB.

Unlike decision tree or other DSS method, rough set method doesn't need us to provide any other prior information except the target data sets, as a result, the description of uncertain problem is more objective.

Rough set is a famous researching made by Z. Pawlak in 1982, which focuses on the uncertain and imperfect information processing. The main idea of rough set is to make an approximate characterization of uncertain or imperfect information in use of knowledge base. Rough set can get knowledge from experiential learning, reduce information without lost, recognition and evaluate dependence between data. Nowadays, rough set has wide applications especially in decision support system (DSS) and knowledge discovery in database (KDD, database mining), in which core and reduction are two key concepts. Rough set is an important method of soft computing.

Key words: SCILAB, decision tree, rough set