A Consolidated Approach to Rule Induction

Hwalsik Chang

Pusan National University

ABSTRACT

One of the most critical aspects of developing knowledge-based systems is to acquire domain knowledge from human experts. The knowledge acquisition phase, however, is often considered to be very difficult and time-consuming.

To solve this bottleneck problem in knowledge-based system construction, much effort has been devoted to the development of machine learning capabilities, with rule induction algorithms coming under the scrutiny of many researchers. The crux of rule induction algorithms is to uncover a set of hidden rules for decision making from examples of the experts' decisions.

Like any other inductive modeling technique such as statistical modeling, rule induction algorithms must be evaluated in terms of their predictive precision. A basic premise of this research is that different rule induction algorithms may have different levels of predictive precision depending on the characteristics of the data set from which the rules are derived.

This research examined variants of Quinlan's ID3 algorithm and Cendrowska's PRISM and evaluated their performances by using two sets of data -- bond rating problem and burn feasibility diagnosis problem. The results showed that each algorithm was better for one problem, but not for the other problem.

This paper indicates that it may be necessary to apply multiple rule induction algorithms to a given problem for the purpose of improving the predictive precision of the induced rules. That is, model builders are advised to examine and

comparatively evaluate multiple modeling approaches for a given problem rather than to select one modeling approach over others without justifiable reasons. This paper presents a consolidated approach to rule induction that "guarantees" the quality of Lockean rule formulation systems.