E

Evaluation of Decision Rules by Qualities for Decision-Making Systems

Ivan Bruha
McMaster University, Canada

INTRODUCTION

A 'traditional' learning algorithm that can induce a set of decision rules usually represents a robust and comprehensive system that discovers a knowledge from usually large datasets. We call this discipline Data Mining (DM). Any classifier, expert system, or generally a decision-supporting system can then utilize this decision set to derive a decision (prediction) about given problems, observations, diagnostics. DM can be defined as a nontrivial process of identifying valid, novel, and ultimately understandable knowledge in data. It is understood that DM as a multidisciplinary activity points to the overall process of determining a useful knowledge from databases, i.e. extracting high-level knowledge from low-level data in the context of large databases.

A rule-inducing learning algorithm may yield either an ordered or unordered set of *decision rules*. The latter seems to be more understandable by humans and directly applicable in most expert systems or decision-supporting ones. However, classification utilizing the unordered-mode decision rules may be accompanied by some conflict situations, particularly when several rules belonging to different classes match ('fire' for) an input to-be-classified (unseen) object. One of the possible solutions to this conflict is to associate each decision rule induced by a learning algorithm with a numerical factor which is commonly called the *rule quality*.

The chapter first surveys empirical and statistical formulas of the rule quality and compares their characteristics. Statistical tools such as contingency tables, rule consistency, completeness, quality, measures of association, measures of agreement are introduced as suitable vehicles for depicting a behaviour of a decision rule.

After that, a very brief theoretical methodology for defining rule qualities is acquainted. The chapter then concludes by analysis of the formulas for rule qualities, and exhibits a list of future trends in this discipline.

BACKGROUND

Machine Learning (ML) or Data Mining (DM) utilize several paradigms for extracting a knowledge that can be then exploited as a decision scenario (architecture) within an expert system, classification (prediction) one, or any decision-supporting one. One commonly used paradigm in Machine Learning is *divide-and-conquer* that induces decision trees (Quinlan, 1994). Another widely used *covering* paradigm generates sets of decision rules, e.g., the CNx family (Clark & Boswell, 1991; Bruha, 1997), C4.5Rules and Ripper. However, the rule-based classification systems are faced by an important deficiency that is to be solved in order to improve the predictive power of such systems; this issue is discussed in the next section.

Also, it should be mentioned that they are two types of agents in the multistrategy decision-supporting architecture. The simpler one yields a single decision; the more sophisticated one induces a list of several decisions. In both types, each decision should be accompanied by the agent's confidence (belief) into it. These functional measurements are mostly supported by statistical analysis that is based on both the certainty (accuracy, predictability) of the agent itself as well as consistency of its decision. There have been quite a few research enquiries to define formally such statistics; some, however, have yielded in quite complex and hardly enumerable formulas so that they have never been used.

One of the possible solutions to solve the above problems is to associate each decision rule induced by a learning algorithm with a numerical factor: a *rule quality*. The issue fo the rule quality was discussed in many papers; here we introduce just the most essential ones: (Bergadano et al., 1988; Mingers, 1989) were evidently one of the first papers introducing this problematic. (Kononenko, 1992; Bruha, 1997) were the followers; particularly the latter paper presented a methodological insight to this discipline. (An & Cercone,

2001) just extended some of the techniques introduces by (Bruha,1997). (Tkadlec & Bruha, 2003) presents a theoretical methodology and general definitions of the notions of a Designer, Learner, and Classifier in a formal matter, including parameters that are usually attached to these concepts such as rule consistency, completeness, quality, matching rate, etc. That paper also provides the minimum-requirement definitions as necessary conditions for the above concepts. Any designer (decision-system builder) of a new multiple-rule system may start with these minimum requirements.

RULE QUALITY

A rule-inducing algorithm may yield either an ordered or unordered set of decision rules. The latter seems to be more understandable by humans and directly applicable in most decision-supporting systems. However, the classification utilizing an unordered set of decision rules exhibits a significant deficiency, not immediately apparent. Three cases are possible:

1. If an input unseen (to-be-classified) object satisfies (matches, 'fires' for) one or more rules of the same class, then the object is categorized to the class assigned to the rule(s).
2. If the unseen object is not covered by any rule, then either the classifier informs the user about its inability to decide ('I do not know'), or the object is assigned by default to the majority class in the training set, or some similar techniques are invoked.
3. Difficulty arises if the input object satisfies more rules assigned to different classes. Then some schemes have to be applied to assign the unseen input object to the most appropriate class.

One possibility to clarify the conflict situation (case 3) of multiple-rule systems is to associate each rule in the decision-supporting scheme with a numerical factor that can express its properties and characterize a measure of belief in the rule, its power, predictability, reliability, likelihood, and so forth. A collection of these properties is symbolized by a function commonly called the *rule quality*. After choosing a formula for the rule quality, we also have to select a scheme for combining these qualities (*quality combination*).

Quality of rules, its methodology as well as appropriate formulas have been discussed for many years. (Bergadano et al., 1992) is one the first papers that introduces various definitions and formulas for the rule quality; besides rule's power and predictability it measures its size, understandability, and other factors. A survey of the rule combinations can be found, e.g. in (Kohavi & Kunz, 1997). Comprehensive analysis and empirical expertise of formulas of rule qualities and their combining schemes has been published in (Bruha & Tkadlec, 2003), its theoretical methodology in (Tkadlec & Bruha, 2003).

We now discuss the general characteristics of a formula of the rule quality. The first feature required for the rule quality is its *monotony* (or, more precisely, *nondecreasibility*) towards its arguments. Its common arguments are the *consistency* and *completeness* factors of decision rules. Consistency of a decision rule exhibits its 'purity' or reliability, i.e., a rule with high consistency should cover the minimum of the objects that do not belong to the class of the given rule. A rule with high completeness factor, on the other hand, should cover the maximum of objects belonging to the rule's class.

The reason for exploiting the above characteristics is obvious. Any DM algorithm dealing with real-world noisy data is to induce decision rules that cover larger numbers of training examples (objects) even with a few negative ones (not belonging to the class of the rule). In other words, the decision set induced must be not only reliable but also powerful. Its reliability is characterized by a consistency factor and its power by a completeness factor.

Besides the rule quality discussed above there exist other rule measures such as its size (e.g., the size of its condition, usually the number of attribute pairs forming the condition), computational complexity, comprehensibility ('Is the rule telling humans something interesting about the application domain?'), understandability, redundancy (measured within the entire decision set of rules), and similar characteristics (Tan et al., 2002; Srivastava, 2005). However, some of these characteristics are subjective; on contrary, formulas of rule quality are supported by theoretical sources or profound empirical expertise.

Here we just briefly survey the most important characteristics and definitions used by the formulas of rule qualities follow. Let a given task to be classified be characterized by a set of training examples that belong

Related Content