Nondeterministic Decision Rules in Classification Process

Piotr Paszek and Barbara Marszał-Paszek

Institute of Computer Science, University of Silesia
Będzińska 39, 41-200 Sosnowiec, Poland
{paszek,bpaszek}@us.edu.pl

Abstract. In the paper, we discuss nondeterministic rules in decision tables, called the truncated nondeterministic rules. These rules have on the right hand side a few decisions. We show that the truncated nondeterministic rules can be used for improving the quality of classification.

We propose a greedy algorithm of polynomial time complexity to construct these rules. We use this type of rules, to build up rule-based classifiers. These classifiers, classification algorithms, are used not only nondeterministic rules but also minimal rules in the sense of rough sets. These rule-based classifiers were tested on the group of decision tables from the UCI Machine Learning Repository. The reported results of the experiment show that the proposed classifiers based on nondeterministic rules improve the classification quality but it requires tuning some of their parameters relative to analyzed data.

Keywords: classification, decision tables, nondeterministic decision rules, rough sets, rule based classifier.

1 Introduction

Over the years many methods based on rule induction and rule-based classification systems were developed [10,17]. Some of them are based on rough sets [2,6,14,15,18] and some of them are based on cluster analysis [7]. In this paper we show that exist possibility for improving the rule-based classification systems.

We discuss a method for rule inducing based on searching for strong rules for a union of a few relevant decision classes – nondeterministic decision rules. Because these rules are created by shortening the deterministic rules they are called truncated nondeterministic rules.

In the paper, the following classification problem is considered: for a given decision table T [11,12] and a new object v generate a value of the decision attribute on v using values of conditional attributes on v.

In [16] Skowron and Suraj shown that there exist information systems $S = (U, A)$ [11], where U is a finite set of objects and A is a finite set of attributes, such that the set U can't be described by deterministic rules. In [9] Moshkov shown that for any information system, the set can be described by nondeterministic

P. Herrero et al. (Eds.): OTM 2012 Workshops, LNCS 7567, pp. 485–494, 2012.

(inhibitory) rules. Inhibitory rules [3] are a special case of nondeterministic rules. These results inspired us to use the nondeterministic rules in a classification process [8].

We present an application of (truncated) nondeterministic rules in construction of rule-based classifiers. We also include the results of experiments that shows that by combining the rule-based classifiers based on the minimal decision rules [11,15] with the nondeterministic rules that have sufficiently large support [1], it is possible to improve the classification quality and reduce the classification error.

The paper consists of six sections. In Section 2, we recall the notions of a decision table and deterministic and nondeterministic decision rules. In Sections 3 and 4 we present a greedy algorithm for nondeterministic decision rule construction and main steps in construction of classifiers enhanced by nondeterministic rules. In Section 5 the results of the experiments with real-life data from the UCI Machine Learning Repository [5] are discussed. Section 6 contains short conclusions.

2 Basic Notations

In 1982 Pawlak proposed the rough set theory as an innovative mathematical tool for describing knowledge, including the uncertain and inexact knowledge [11]. In this theory knowledge is based on possibility (capability) of classifying objects. The objects may be for instance real objects, statements, abstract concepts and processes.

Let $T = (U, A, d)$ be a *decision table*, where $U = \{u_1, \ldots, u_n\}$ is a finite nonempty set of *objects*, $A = \{a_1, \ldots, a_m\}$ is a finite nonempty set of *conditional attributes* (functions defined on U), and d is the *decision attribute* (function defined on U).

We assume that for each $u_i \in U$ and each $a_j \in A$ the value $a_j(u_i)$ belong to $V_{a_j}(T)$ and the value $d(u_i)$ belong to $V_d(T)$, where $V_d(T)$ denotes the set of values of the decision attribute d on objects from U.

2.1 Deterministic Decision Rules

In general, the *deterministic decision rule* in T has the following form:

$$(a_{j_1} \in V_1) \wedge \ldots \wedge (a_{j_k} \in V_k) \rightarrow (d = v),$$

where $a_{j_1}, \ldots, a_{j_k} \in A$, $V_j \subseteq V_{a_j}$, for $j \in \{1, \ldots, k\}$ and $v \in V_d(T)$. The predecessor of this rule is a conjunction of generalized descriptors and the successor of this rule is a descriptor.

In the rule-based classifiers, most commonly used are the rules in the form of Horn Clauses

$$(a_{j_1} = b_1) \wedge \ldots \wedge (a_{j_k} = b_k) \rightarrow (d = v)$$

where $k > 0$, $a_{j_1}, \ldots, a_{j_k} \in A$, $b_1, \ldots, b_k \in V_A(T)$, $v \in V_d(T)$ and numbers $j_1, \ldots, j_k$ are pairwise different. The predecessor of this rule (conditional part) is a conjunction of descriptors.