Evolutionary Fuzzy Rules for Ordinal Binary Classification with Monotonicity Constraints

Christian Moewes and Rudolf Kruse

Faculty of Computer Science, Otto-von-Guericke University of Magdeburg, Germany
cmoewes@ovgu.de, kruse@iws.cs.uni-magdeburg.de

Abstract. We present an approach to learn fuzzy binary decision rules from ordinal temporal data where the task is to classify every instance at each point in time. We assume that one class is preferred to the other, e.g. the undesirable class must not be misclassified. Hence it is appealing to use the Variable Consistency Dominance-based Rough Set Approach (VC-DRSA) to exploit preference information about the problem. In this framework, the VC-DomLEM algorithm has been used to generate the minimal set of consistent rules. Every attribute is then fuzzified by first applying a crisp clustering to the rules' antecedent thresholds and second using the cluster centroids as indicator for the overlap of neighboring trapezoidal normal membership functions. The widths of the neighboring fuzzy sets are finally tuned by an evolutionary algorithm trying to minimize the specificity of the current fuzzy rule base.

1 Introduction

In many real-world problems, complex systems need to be described in a highly comprehensible way. The explaining descriptions are usually based on observations of several variables describing the state of the system. For many systems, it is quite common that the number of variables is large, say around 50. Thus, it can be quite difficult for human experts to build a model that describes the system in an adequate way. Usually, a trade-off has to be found to balance both the correctness and the complexity of the model.

Complex models are naturally regarded skeptically since a proof of correctness is hard to obtain. In our highly technologized information society, rapidly increasing quantities of data coming from complicated systems are stored without any possibility of being analyzed manually. Therefore, it is extremely desirable to have machines that learn from empirical data *and* guarantee both interpretable and correct models.

Here, we will restrict ourselves to machines that solve supervised classification problems based on numerical data. A Fuzzy Rule-based Classifier (FRBC) is such kind of machine. Nowadays, an FRBC is the state of the art in many real-world applications, e.g. automobile controllers (Schröder et al, 1997). It has nice properties that are demanded and appreciated by safety experts. That is, an FRBC is linguistically interpretable and its behavior is easy to approve by these experts. An open research problem is the question how such an FRBC can be found and tuned automatically. We want to

R.R. Yager et al. (Eds.): *Soft Computing: State of the Art Theory*, STUDFUZZ 291, pp. 105–112.
DOI: 10.1007/978-3-642-34922-5_8

develop machine learning tools that come up with interpretable fuzzy rule bases for such systems.

The online discrimination of vehicle crashes to deploy certain stages of a restraint system is such complex problem (Moewes, 2007; Moewes et al, 2008, 2010). Fuzzy binary decision rules shall be obtained from ordinal temporal data. Every instance is classified at each point in time. The so-called fire class is preferred to the no-fire class, e.g. we must not deploy in a no-fire case. To the best of our knowledge, no suitable (fuzzy) rule learner exists for this safety-related problem. An exhaustive discussion and recent recent approaches to tackle safety-related problems can be found in (Nusser, 2009).

In this paper, we show that the crash discrimination problem corresponds to ordinal binary classification with monotonicity constraints. Hence, the idea is to use a rule-based ordinal classifier that exploits monotonicity. This important property is implicitly handled by the Variable Consistency Dominance-based Rough Set Approach (VC-DRSA) (Greco et al, 2001). In this framework, the minimal set of consistent rules has been generated by the VC-DomLEM algorithm (Błaszczyński et al, 2011).

Another novelty in this paper is a proposal to fuzzify the crisp classifier to an FRBC. Every attribute is fuzzy partitioned in three steps. First, a crisp clustering is applied to the rules' antecedent thresholds. Second, the cluster centroids are used as indicator for the overlap of neighboring trapezoidal normal membership functions. Finally, the widths of the neighboring fuzzy sets are tuned by an evolutionary algorithm that minimizes the specificity of the current fuzzy rule base. The fuzzification step is basically performed to compress the original rule base for interpretation issues.

2 Fundamentals

2.1 Ordinal Binary Classification

Given an object x which can be described by attribute values $x = (x_1, \ldots, x_n)$, the aim in ordinal classification is to predict an unknown class label ω from an ordered set $\Omega = \{0, \ldots, k-1\}$. A meaningful order between classes is assumed which corresponds to a natural order between the labels $\omega \in \Omega$. When dealing with $|\Omega| > 2$, ordinal classification problems are found in many real-world applications, e.g. recommender systems. There, users can rate items on a finite scale, e.g. a school grade from 1 to 5. The ultimate goal is to predict the rates of a new user given the known rates. For $|\Omega| = 2$, we say that the ordinal classification problem is binary.

2.2 Rough Set Approach

The rough set approach is performed in two steps to extract knowledge from observations (Błaszczyński et al, 2011). First, all inconsistencies in the data are found by computation of lower and upper approximations of all observations. In the indiscernibility-based rough set approach (IRSA) (Pawlak, 1991), these sets correspond to decision classes. Data points do not obey any order and can be therefore compared by the indiscernibility relation.