Iterative bayesian network implementation by using annotated association rules

Clément Fauré[1,2], Sylvie Delprat[1], Jean-François Boulicaut[2], and Alain Mille[3]

[1] EADS CCR, Learning Systems Department, Centreda 1, F-31700 Blagnac,
{clement.faure, sylvie.delprat}@eads.net
[2] LIRIS UMR 5205, INSA Lyon, Bâtiment Blaise Pascal, F-69621 Villeurbanne
[3] LIRIS UMR 5205, Université Lyon 1, Nautibus, F-69622 Villeurbanne
{amille, jboulica}@liris.cnrs.fr

Abstract. This paper concerns the iterative implementation of a knowledge model in a data mining context. The proposed approach relies on coupling a bayesian network design with an association rule discovery technique. First, discovered association rule relevancy is enhanced by exploiting the expert knowledge encoded within a bayesian network, i.e., avoiding to provide trivial rules w.r.t. the available expertise. Moreover, the bayesian network can be updated thanks to an expert-driven annotation process on computed association rules. Our approach is experimentally validated on both synthetic and real data. We sketch a practical case study for which the data report on operational interruptions in the aeronautic industry.

1 Introduction

One major goal of the knowledge discovery from databases (KDD) community is to support the discovery of valuable information or patterns within the data. In the so-called transactional data sets (say 0/1 data), the association rule mining technique is quite popular. It has been studied extensively from the computational perspective. Many researchers have been considered the relevancy as well. Clearly, valuable patterns have to be valid statements (e.g., w.r.t. some objective interestingness criteria like confidence or lift). It is also needed that we support the discovery of useful ones w.r.t. expert expectation, i.e., the so-called subjective interestingness criteria. In this research, we consider that expert expectation is related to novelty, i.e., patterns like association rules are valuable if they provide some information which is somehow new given the encoded domain knowledge. Furthermore, the encoded knowledge for Process n might be updated for Process $n+1$ by using the expert-driven annotation of the relevancy for the patterns extracted during Process n. More concretely, we focus on association rule mining when a bayesian network (BN) captures domain knowledge. Given the BN, some extracted rules can be filtered out. Then, we suggest to perform an expert-driven annotation of the presented rules and these annotations are then used to perform updates on the BN. Doing so, we propose a methodology which iteratively