

Abstract

Data mining is a process that aims at knowledge discovery from within great extension databases. Although it has been increasingly adopted due to its good performance in several domains, in some cases the results that are generated may be too large or too complex. This is a problem that is specific to the Association Rules technique, which often generates an amount of rules that exceed the limit amount of rules that are humanly viable to manipulate.

In this present paper we propose a new semantic pruning approach for association rules based on previous domain knowledge, which is represented by attribute inter-dependency relations. The proposed method is aimed at facilitating analysis and comprehension of the rules, by means of eliminating redundancy within the mined rules, and by selecting those that have greater impact for the user's needs.

Among the main results of the present study is the proposal and implementation of the DMcut association rules pruning method. The experiments that were conducted on four public domain databases reveal the potential benefits of using the method.

Keywords: Data Mining, Association Rules, Post-processing