

Subjective Evaluation of Labeling Methods for Association Rule Clustering

Renan de Padua¹, Fabiano Fernandes dos Santos¹, Merley da Silva Conrado¹,
Veronica Oliveira de Carvalho², and Solange Oliveira Rezende¹

¹ Instituto de Ciências Matemáticas e de Computação,
USP - Universidade de São Paulo, São Carlos, Brazil
{padua,fabianof,merleyc,solange}@icmc.usp.br

² Instituto de Geociências e Ciências Exatas,
UNESP - Univ Estadual Paulista, Rio Claro, Brazil
veronica@rc.unesp.br

Abstract. Among the post-processing association rule approaches, clustering is an interesting one. When an association rule set is clustered, the user is provided with an improved presentation of the mined patterns. The domain to be explored is structured aiming to join association rules with similar knowledge. To take advantage of this organization, it is essential that good labels be assigned to the groups, in order to guide the user during the association rule exploration process. Few works have explored and proposed labeling methods for this context. Moreover, these methods have not been explored through subjective evaluations in order to measure their quality; usually, only objective evaluations are used. This paper subjectively evaluates five labeling methods used on association rule clustering. The evaluation aims to find out the methods that presents the best results based on the analysis of the domain experts. The experimental results demonstrate that there is a disagreement between objective and subjective evaluations as reported in other works from literature.

1 Introduction

Association rule mining (ARM), introduced in [1], is an important task of data mining. ARM aims to “find all co-occurrence relationships, called associations, among data items” [11].

Association rules have been successfully applied for decision support (such as the cross-marketing, attached mailing applications, catalog design, add-on sales, store layout, and customer segmentation based on buying patterns) [3], for applications of telecommunications alarm diagnosis and prediction [2], for inter-disciplinary domains beyond data mining (such as indexing and similarity search of complex structured data, spatio-temporal and multimedia data mining, stream data mining, web mining, software bug mining, and page-fetch prediction) [8], and for disease prediction [17].

When generating association rules, it is necessary to deal with a huge amount of rules since the number of rules grows exponentially with the number of items in

the data set [9]. Many algorithms have been developed to overcome the problem of dealing with these generated rules. These algorithms follow one of these post-processing approaches: *Querying (Q)*, *Evaluation Measures (EM)*, *Pruning (P)*, *Summarizing (S)*, or *Grouping (G)* [5,22,14,10]. The algorithms that belong to the approaches of *Q*, *P*, and *S* aid the exploration process by reducing the exploration space (*RES*); the ones that belong to *EM* approach explore the process by directing the user to what is potentially interesting (*DUPI*); and, finally, the algorithms of *G* approach explore the process by structuring the domain (*SD*).

Grouping is a relevant approach related to *SD*, since it organizes the rules in groups that contain, somehow, similar knowledge. These groups improve the presentation of the mined rules, providing the user a view of the domain to be explored [18,19]. A methodology was found in the literature for post-processing association rules that utilizes the grouping approach. This methodology, called PAR-COM [5], combines clustering and objective measures to direct the user to what is potentially interesting and, consequently, reduces the association rule exploration space. Thus, the user only needs to explore a small subset of the groups that contain the potentially interesting knowledge. However, it is essential that groups be represented by labels that may provide the user a view of the subjects contained in the exploration space, helping to guide its search.

Although some methods have been proposed to label document clusters in Text Mining (TM) and Information Retrieval (IR) [13,12,16], there are few researches in the literature that deal with selecting labels for association rule clustering. Padua et al. [15] and Carvalho et al. [4] assess some labeling methods using objective evaluations. Chang et al. [7] discuss about a disagreement between objective and subjective evaluation results in a topic extraction context. The latter found that some results of objective measures are not always a good predictor of human judgments regarding the terms selected as labels for the topic extraction task. The same problem is found here since the label selection task is similar to topic extraction and association rule clustering approaches.

Considering that, we use a subjective methodology to evaluate label sets obtained by labeling methods for association rule clustering. For that, this paper presents an adapted version of the subjective evaluation methodology proposed in [7] (details in Section 3). The evaluation was applied in five labeling methods for association rule clustering in order to identify which one obtains suitable label sets according to the

The proposal of an evaluation methodology adapted from [7] is introduced and adjusted for an environment that considers clusters of association rules obtained from structured data. Specifically, the proposed evaluation methodology is based on a task named *word intrusion*. The word intrusion task, proposed in [7], consists of identifying a spurious word inserted into a set of words¹ that represent the extracted topic. The *word intrusion* task was initially proposed to evaluate whether an extracted topic has human-identifiable semantic coherence.

¹ In this work, a set of words represents the labels of a group.