Association rule data mining

Kristian Wichmann

October 31, 2016

1 Transactions and patterns

Let $\mathcal{I}$ be a set of so-called *transactions*. A *pattern* t is a subset of $\mathcal{I}$. So we might consider a dataset $\mathcal{T}$ with n different transactions and m patterns:

$$\mathcal{I} = \{I_1, I_2, \ldots, I_n\}, \qquad \mathcal{T} = \{t_1, t_2, \ldots, t_n\} \tag{1.1}$$

Association rule data mining is an unsupervised learning discipline trying to reveal systematics in the database.

1.1 Support of a pattern

A pattern $X \subseteq \mathcal{I}$ is said to have a *support* equal to the number of elements of $\mathcal{T}$ in which X is a subset. Support may be specified absolutely or relatively.

1.2 Association rules

An *association rule* takes the form $X \Rightarrow Y$, where X and Y are disjoint patterns.