

## Atividade 7

### Algoritmo Apriori

É um algoritmo de mineração de dados que busca por itens frequentes em banco de dados que ocorrem simultaneamente em transações do banco utilizando regras de associação.

Através da utilização desse algoritmo é possível gerar associações e informar, por exemplo em supermercados, qual produto sempre sai quando determinado produto é comprado.

O núcleo do algoritmo é basicamente esse:

```
Apriori( $T, \epsilon$ )
   $L_1 \leftarrow \{\text{large 1-itemsets}\}$ 
   $k \leftarrow 2$ 
  while  $L_{k-1} \neq \emptyset$ 
     $C_k \leftarrow \{c | c = a \cup \{b\} \wedge a \in L_{k-1} \wedge b \in \bigcup L_{k-1} \wedge b \notin a\}$ 
    for transactions  $t \in T$ 
       $C_t \leftarrow \{c | c \in C_k \wedge c \subseteq t\}$ 
      for candidates  $c \in C_t$ 
         $\text{count}[c] \leftarrow \text{count}[c] + 1$ 
       $L_k \leftarrow \{c | c \in C_k \wedge \text{count}[c] \geq \epsilon\}$ 
     $k \leftarrow k + 1$ 
  return  $\bigcup_k L_k$ 
```

Figura 1. Algoritmo apriori.

Ideia base do algoritmo:

- Obtenção de conjuntos de itens frequentes de cardinalidade, progressivamente superior ao conjunto de itens frequente e se um conjunto de itens é frequente então todos os seus subconjuntos são também frequentes.
- Depois de obtidos todos os conjuntos de itens frequentes transformar cada um numa regra (ou conjunto de regras) com confiança superior ao limiar onde um conjunto de

itens frequente pode dar origem a várias regras, mas também pode não dar origem a nenhuma.