

Knowledge Discovery in Databases
WS 2008/09
Übungsblatt 11

Aufgabe 11-1 Herleiten von Assoziationsregeln

Gegeben ist die Menge der Items $I = \{A, B, C, D, E, F, G, H, I, K, L, M\}$.

Weiterhin ist eine Menge von Transaktionen T laut folgender Tabelle gegeben:

Menge der Transaktionen T

Transaktions ID	gekaufte Items
1	B E G H
2	A B C E G H
3	A B C E F H
4	B C D E F G H L
5	A B E K H
6	B E F G H I K
7	A B D G H
8	A B D G
9	B D F G
10	C E F
11	A C E F H
12	A B E G

Bestimmen sie alle Assziationregeln, die man aus dem frequent Itemset $\{B E G H\}$ mit einer Konfidenz von mindestens 60% und dem Support 4 (Frequency 30%) bilden kann. Nutzen sie dabei die in der Vorlesung vorgestellte Monotonie aus.

Aufgabe 11-2 BIRCH

Gegeben seien folgende 1-dimensionale Punktdaten:

1, 2, 5, 10, 4, 7, 15, 9, 3

- (a) Bauen Sie einen CF-Baum mit $B = 2$ und $L = 3$ auf, indem Sie die Punkte der Reihe nach einfügen. Der Schwellwert T ist dabei wie folgt definiert: "Alle Einträge in einem Blattknoten haben eine Euklidische Distanz ≤ 2 zum Centroiden des Blattknotens".
Zeichnen Sie den CF-Baum nach jeder Einfügung.
- (b) Warum ist das Ergebnis abhängig von der Reihenfolge?