

Knowledge Discovery in Databases
WS 2009/10
Übungsblatt 12

Aufgabe 12-1 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?

Aufgabe 12-2 Online Mining für Assoziationsregeln

Gegeben sei eine Transaktionsdatenbank DB auf der bereits frequent Itemsets und Assoziationsregeln bestimmt wurden.

Überlegen Sie sich, wie Sie die bereits berechneten Ergebnisse unter folgenden Änderungen mit möglichst wenig Aufwand anpassen und erweitern können:

- Die Datenbank wird um zusätzliche Transaktionen erweitert.
- Der minimale Support wird größer.
- Die minimale Confidence einer Regel wird kleiner.