

Praxisübung zur 7. Übung „Knowledge Discovery“

Wintersemester 2008/2009

Berechnen häufiger Itemsets mit Apriori

Implementieren Sie den Apriori-Algorithmus zum Finden häufiger Itemsets bei einer gegebenen Menge von Transaktionen.

Für diese Aufgabe benutzen wir nicht das RapidMiner-Framework, sondern die eigenständige Klasse `de.unikassel.cs.kde.kdd.Apriori`, die Sie im Web unter <http://www.kde.cs.uni-kassel.de/lehre/ws2008-09/kdd/daten/Apriori.java> finden. Nutzen Sie diese Datei als Grundgerüst für Ihre Implementierung. Behalten Sie Klassen- und Paketnamen bei und implementieren Sie den Apriori-Algorithmus in der Methode

```
apriori(SortedMap<String, SortedSet<String>> transactions, int minSupport)}.
```

Selbstverständlich können Sie ggf. weitere Methoden in der Klasse anlegen. Nutzen Sie zum Testen Ihrer Implementation den Datensatz aus der 7. Übung, welchen Sie als CSV-Datei unter <http://www.kde.cs.uni-kassel.de/lehre/ws2008-09/kdd/daten/warenkorb.csv> finden.

Beachten Sie bitte zusätzlich folgende Formalitäten:

- Abgabe der Praxisübung: bis **13.1.2009, 23:59 Uhr MEZ** per E-Mail an jaeschke@cs.uni-kassel.de. Spätere Einreichungen werden nicht berücksichtigt!
- Alle benötigten Methoden, Attribute, etc. müssen sich in der Java-Datei `Apriori.java` befinden, deren Grundgerüst Sie hier finden: <http://www.kde.cs.uni-kassel.de/lehre/ws2008-09/kdd/daten/Apriori.java>.
- Am 15.1.2009 werden Sie uns Ihren Algorithmus kurz vorführen, erklären und ggf. Fragen beantworten. Termine (pro Teilnehmer ca. 5 Minuten) werden wir vorher vereinbaren.
- Für eine positive Bewertung muss ihr Code problemlos compilieren, einen Durchlauf durch den Warenkorb-Datensatz aus der 7. Übung ohne Fehler bestehen, nachvollziehbare Itemsets liefern und es muss im Gespräch erkennbar sein, dass Sie *Ihre Implementierung* des Algorithmus verstanden haben.

Sollten Sie Fragen haben, wenden Sie sich bitte an Robert Jäschke (jaeschke@cs.uni-kassel.de). Zusätzlich bieten wir Ihnen an, dass sie immer donnerstags vor und nach der Übung kurz Fragen stellen können. Weitere Hinweise und Dateien finden Sie auch auf der Webseite <http://www.kde.cs.uni-kassel.de/lehre/ws2008-09/kdd/uebung/rapidminer.html>.