

Übung zur Vorlesung „Web Data Mining“

# Übungsblatt 1

Abgabe am 18.4.2007

**Aufgabe 1** Implementieren Sie den MSapriory-Algorithmus zum Finden von häufigen Artikelmenigen mit mehreren unteren Schranken für Unterstützung. Definieren die untere Schranke für Unterstützung eines Artikels  $A$  als  $MIS(A) = \lambda \cdot sup(A)$  mit  $\lambda \in [0, 1]$ . Der Parameter  $\lambda$  soll vom Anwender eingegeben werden können. Sie können bei der Implementierung auf die Perl-Implementierung des einfachen Apriory-Algorithmus von der Vorlesungswebseite<sup>1</sup> zurückgreifen. Als überschaubare Testdaten stehen Ihnen die Wetterdaten zur Verfügung. Anfang der Woche können Sie auch eine größere Datenmenge, die Dokumente beschreibt herunter laden. Führen Sie Ihren MS-Apriory Algorithmus auch auf diesen Daten aus.

**Aufgabe 2** Beweisen Sie, dass bei der Erzeugung von Assoziationsregeln mit mehreren unteren Schranken für Unterstützung es ausreicht, den zusätzlichen Zähler `c.tailcount` (siehe Folien 39 und 47) einzurichten.

---

<sup>1</sup>[http://users.informatik.uni-halle.de/~hinnebur/Lehre/WebKDD\\_SS07\\_web/](http://users.informatik.uni-halle.de/~hinnebur/Lehre/WebKDD_SS07_web/)