

Übungsblatt 2

12. November 2008

Hinweis: Soweit nicht anders angegeben, gibt es für jede korrekt bearbeitete Teilaufgabe einen Punkt. Die Abgabe der Hausübungen ist bis spätestens zum Beginn der nächsten Vorlesung möglich – entweder persönlich direkt vor der Vorlesung oder per Einwurf in den Briefkasten des Instituts (Informatikzentrum, zweiter Stock, vor Raum 236).

Aufgabe 3 (Fuzzy-Modell)

In dieser Aufgabe betrachten wir eine Dokumentensammlung, die die Dokumente d_1 bis d_3 enthält. Diese Dokumente werden wie folgt durch die Terme t_1 bis t_5 repräsentiert:

$d_1: t_2, t_3, t_4,$

$d_2: t_1, t_3, t_4,$

$d_3: t_1, t_5.$

Ein Benutzer stellt die folgenden Anfragen:

$q_1: t_1 \text{ UND } t_3,$

$q_2: t_1 \text{ ODER } t_3,$

$q_3: \text{NICHT } t_1.$

- Welche Ergebnisse liefert das Boolesche Modell bei diesen Anfragen?
- Welche Ergebnisse liefert das Fuzzy-Modell? Bestimmen Sie dazu zunächst neue Repräsentationen der Dokumente nach dem Ansatz von Ogawa, Morita und Kobayashi.

(3 Punkte)

Aufgabe 4 (Coordination Level Matching und Vektorraummodell)

Wir betrachten noch einmal die in der vorhergehenden Aufgabe beschriebene Dokumentensammlung.

- Welche Ergebnisse liefert das Vektorraummodell bei den folgenden Anfragen, wenn Coordination Level Matching verwendet wird?

$q_4: t_1, t_2, t_4,$

$q_5: t_4, t_5.$

- Welche Ergebnisse liefert das Vektorraummodell für die Anfragen aus dem vorherigen Aufgabenteil? Bestimmen Sie zunächst geeignete vektorielle Repräsentationen dieser Anfragen nach dem TF-IDF-Schema. Verwenden Sie dabei als Maß jeweils das Cosinusmaß und den Euklidischen Abstand.

(2 Punkte)

- c) Wie könnte man folgende Anfragen im Vektorraummodell repräsentieren? Begründen Sie Ihre Antworten kurz.

q_8 : t_1 IST MIT LIEBER ALS t_3 ,

q_9 : t_1 ABER NICHT t_3 .

Aufgabe 5 (Stochastik)

Bestimmen Sie folgende Wahrscheinlichkeiten für das Würfelbeispiel aus der Vorlesung:

- $P(\text{„eine gerade Augenzahl im ersten Wurf und eine ungerade im zweiten Wurf“})$,
- $P(\text{„höchstens 7 Augen in der Summe“} \mid \text{„verloren“})$,
- $P(\text{„mindestens 10 Augen in der Summe“} \mid \text{„keine 6 im ersten Wurf“})$.