

Aufgabenblatt 7: Ähnlichkeitsmaße und Evaluierung

Aufgabe 1 Vergleich von Ähnlichkeitsmaßen

Berechnen Sie das Cosinusmaß, das Pseudocosinusmaß, das Dice-Maß und das Overlap-Maß für den Queryvektor $(3, 1, 0)$ und die Dokumentvektoren

1 0.1 0	1 0.5 0	1 2 0	1 20 0
1 0.2 0	1 0.7 0	1 3 0	
1 0.3 0	1 1 0	1 5 0	
1 0.4 0	1 1.5 0	1 10 0	

Vergleichen Sie die Ergebnisse.

Aufgabe 2 Precision und Recall

In einer Literaturdatenbank mit 100 000 Dokumenten befinden sich zu einer Anfrage 30 relevante Dokumente. Ein zu bewertendes Retrieval System liefere die folgende Rangfolge von Dokumenten, in der jedes einzelne "R" ein relevantes Dokument und jedes "U" ein nicht relevantes Dokument bezeichne:

```
RRURRURUR UURURRUUUU URUUURUUUR RUUUURUUUU URUUUUUUUR UUUUUUUUUU
UUUUUUUUUU UUUUUUUUUU UUUUUUUUUU UUUUUUUUUU UUUURUUUUU UUUUUUUUUU
UUUUUUUUURU UUUUUUUUUU UUUUUUUUUU UUUUUUUUUU . . .
```

- a) Berechnen und zeichnen Sie ein Precision-Recall Diagramm.
- b) Wie würde das Precision-Recall Diagramm aussehen, wenn das System optimal arbeiten würde?
- c) Wie würde das Precision-Recall Diagramm im Mittel aussehen, wenn die Auswahl der Dokumente durch das System zufällig wäre?
- d) Gibt es andere Möglichkeiten zu den Recallwerten gehörende Precisionwerte für ein Precision-Recall Diagramm zu bestimmen, als die in der Vorlesung angegebene? Wie könnten die aussehen? Wie verändert sich dadurch das Diagramm?
- e) Worauf muss man achten, wenn der Ähnlichkeitswert zur Anfrage für mehrere Dokumente gleich ist?

Aufgabe 3 Testkollektionen

In der Tabelle in Abbildung 28 des Vorlesungsskripts sind verschiedene statistische Werte zu verschiedenen Testkollektionen angegeben. Wie könnte man daraus Maße konstruieren, die

- a) die Breite der einzelnen Testkollektionen
- b) die Abdeckung der einzelnen Testkollektionen durch die zugehörigen Anfragen abschätzen?

Aufgabe 4 TREC Evaluierung

Beschreiben Sie die Evaluationsmethode, die in TREC verwendet wurde.

- a) Welche Aufgaben wurden gestellt?
- b) Wie wurden Ergebnisse der Teilnehmenden ausgewertet?
- c) Wie wurde die Relevanz bestimmt?
- d) Wie wurde das Verfahren zur Bestimmung der Relevanz überprüft? Wie war das Ergebnis dieser Überprüfung?