

Information Retrieval

Übungen

Ingo Frommholz

Universität Duisburg-Essen

23.04.2008

UNIVERSITÄT
DUISBURG
ESSEN

Ist IR tatsächlich im Alltag allgegenwärtig?

- **Aufgabe:** Sammelt Gegebenheiten aus Eurem Alltag, bei denen Ihr glaubt, IR-Techniken oder IR-Systeme zu benutzen.

Hausaufgabe: Evaluation

- Wir wollen verschiedene Web-Suchmaschinen evaluieren
- Ist Google wirklich die beste?
- Ca. 3 Gruppen
- Präsentation der Ergebnisse am Ende des Semesters (plus kleine Ausarbeitung, wie vorgegangen wurde (~5 Seiten))
- **Am Ende winken 15 Bonuspunkte für die Klausur!**

Wie sollten wir vorgehen?


Welche Suchmaschinen?

- Google
- Yahoo!
- AskJeeves
- MSN Search
-?

- Beispielanfragen (ca. 5 pro Gruppe)
- Relevanzurteile (ca. 30-40 pro Anfrage)


Recall-Bestimmung?

REL (=Gesamtzahl relevanter Dokumente) kann nur näherungsweise bestimmt werden

- 1 vollständige Relevanzbeurteilung einer Stichprobe 
- 2 Document-source-Methode
- 3 Abgleich mit externen Quellen
- 4 Frageerweiterung
- 5 Pooling-Methode

Recall-Bestimmung?


REL (=Gesamtzahl relevanter Dokumente) kann nur näherungsweise bestimmt werden

- 1 vollständige Relevanzbeurteilung einer Stichprobe 
 - 2 Document-source-Methode
 - 3 Abgleich mit externen Quellen
 - 4 Frageerweiterung
 - 5 Pooling-Methode
- Nur die letzten beiden Methoden sind praktikabel



Recall-Bestimmung?


REL (=Gesamtzahl relevanter Dokumente) kann nur näherungsweise bestimmt werden

- 1 vollständige Relevanzbeurteilung einer Stichprobe 
 - 2 Document-source-Methode
 - 3 Abgleich mit externen Quellen
 - 4 Frageerweiterung
 - 5 Pooling-Methode
- Nur die letzten beiden Methoden sind praktikabel
 - aber: sie liefern nur untere Schranken für $|REL|$



Recall-Bestimmung?

REL (=Gesamtzahl relevanter Dokumente) kann nur näherungsweise bestimmt werden

- 1 vollständige Relevanzbeurteilung einer Stichprobe 
 - 2 Document-source-Methode
 - 3 Abgleich mit externen Quellen
 - 4 Frageerweiterung
 - 5 Pooling-Methode
- Nur die letzten beiden Methoden sind praktikabel
 - aber: sie liefern nur untere Schranken für $|REL|$
 - → Recall-Werte im Allgemeinen zu optimistisch



Mittelwertbildung?

- Makrobewertung
- Mikrobewertung

Bewertungsmaße?

- Recall und Precision, R-P-Graphen
- Abbruchkriterium? Anzahl der relevanten Dokumente, Anzahl der Dokumente?
- *PRECALL_{ceiling}, PRR, EP, Prec@10, MAP*

- RePrec:
`http://search.cpan.org/~goevert/RePrec-0.032/`
- trec_eval: `http://trec.nist.gov/trec_eval/`
- gnuplot: `http://www.gnuplot.info/`

- In 2 Wochen: jede Gruppe präsentiert ein Konzept (10 min + 5 min Diskussion)
 - Anfang Juli: Präsentation/Ausarbeitung
 - Zwischendurch: Fragen und Diskussionen in der Übungsgruppe
 - Gruppeninterne Arbeitsteilung:
 - Wer überlegt Beispielanfragen und gibt Relevanzurteile?
 - Wer implementiert?
 - Wer dokumentiert?
- Letztendlich euch selbst überlassen!