

**Einführung in IR - Winter 2006/07**  
Dipl.-Inform. Sascha Kriewel, LF 137  
sascha.kriewel@uni-due.de

### Übungsblatt 3

**Evaluierung  
keine Abgabe**

---

#### **Aufgabe 6: Methoden zur Recallabschätzung**

In Aufgabe 4 wurde angenommen, dass unter den kombinierten ersten 10 Ergebnissen der beiden Suchmaschinen alle relevanten Dokumente zu finden sind. Welcher Methode zur Recallabschätzung entspricht dies? Wie müsste man praktisch vorgehen, um tatsächlich den Recall abzuschätzen, und welche anderen Methoden gibt es?

#### **Aufgabe 7: Recall und Precision bei schwachen Ordnungen**

Gegeben seien die folgenden Distributionen mit schwacher Ordnung. In der Dokumentenkollektion seien bezüglich der Anfrage genau 12 relevante Dokumente enthalten. Die Zahl in Klammern gibt die Anzahl der irrelevanten Dokumente im letzten Rang an.

$$\Delta_1 = (+ + + - | + + + - - - - - | + + - - | + + + + - - - - | - - - - | - (73))$$

$$\Delta_2 = (+ + | + - | + + + + - - - - - - | + + - - - - - - | + + + - - - - - | - (72))$$

Bestimme die Werte für Recall und Precision an den einzelnen Rängen und zeichne diese in einen Graphen ein. Wie ist das Diagramm in Bezug auf die gegebene Anfrage zu bewerten? Gibt es sinnvolle Möglichkeiten zur Interpolation?

#### **Aufgabe 8: Erweiterungen für Recall und Precision**

In der Vorlesung wurden verschiedene formale Konzepte zur Erweiterung von Recall und Precision eingeführt. Zeichne das RP-Diagramm zu den beiden Distributionen

$$\Delta_1 = (+ + - | + | + + - | + - - - | + + - | + + - - - - - -)$$

$$\Delta_2 = (+ - | + + - | + + - | + - - | + + - - | + + - - - - - -)$$

unter Verwendung von  $PRR_{NR}$ :

$$PRR_{NR} := \frac{NR}{NR+j+\frac{s-1}{r+1}}$$

Bestimme dazu  $PRR_{NR}$  für  $NR = 1, \dots, 10$  und benutze lineare Interpolation. Welcher Benutzerstandpunkt wird hier vertreten und wie lässt sich das

Diagramm interpretieren?