

4.3 Evaluation von IR Systeme

4.3.1 Precision und Recall

Wir haben gesehen, wie wir umgehen, um eine Menge von Dokumente zu finden. Was passiert, falls diese Menge 100, 1000, oder 10,000 Dokumente enthält? Wie können wir messen, ob unser IR System wirklich funktioniert, ob die Ähnlichkeit, die mit der Cosinus-Funktion angegeben wird, wirklich uns die relevante Dokumente liefert?

Um diese Frage zu beantworten brauchen wir weitere Messungen. Wir brauchen einen Weg, um die Genauigkeit des Systems zu messen. Deshalb führen wir einen Relevanzwert ein, wo die „Wirkung“ der Ordnung getestet sein kann. Das wird mit Precision und Recall gemacht.

Wir betrachten eine Testmenge von Dokumenten, die benutzt werden, um den Recall und Precision zu messen.

- **Recall:** Anzahl relevante Dokumente, die von einer Abfrage gefunden worden sind, dividiert durch die gesamte Anzahl für diese Abfrage relevante Dokumente in der Testmenge;
- **Precision:** Anzahl relevante Dokumente, die von einer Abfrage gefunden worden sind, dividiert durch die totale Anzahl gefundene Dokumente.

4.3.2 Recall Precision Graphs

Eine perfekte Precision von 1 bedeutet, dass alle Resultate aus einer Abfrage relevant sind, sagt aber nichts über die gesamte Anzahl relevante Dokumente, die zurückgegeben hätte sein sollen. Im Gegensatz, ein perfekter Recall von 1 bedeutet, dass alle relevante Dokumente zurückgegeben worden sind, sagt aber nichts über die Anzahl irrelevante Dokumente, die auch mit-zurückgegeben worden sind.

Wir sagen „im Gegensatz“, weil es eine inverse Beziehung gibt zwischen Recall und Precision. Es ist möglich Recall zu erhöhen, aber auf den Kosten der Precision. Zum Beispiel kann ich Recall erhöhen in dem ich mehr Dokumente zurückgebe, aber ich werde dann auch mehr irrelevant Dokumente haben, was die Precision vertiefen wird. Deshalb sind normalerweise Precision und Recall zusammen betrachtet, nicht einzeln. Dafür wurden auch Precision-Recall Maps entwickelt, wobei die Precision für ein gegebenes Recall berechnet wird, wie in Abbildung 23 gezeigt.

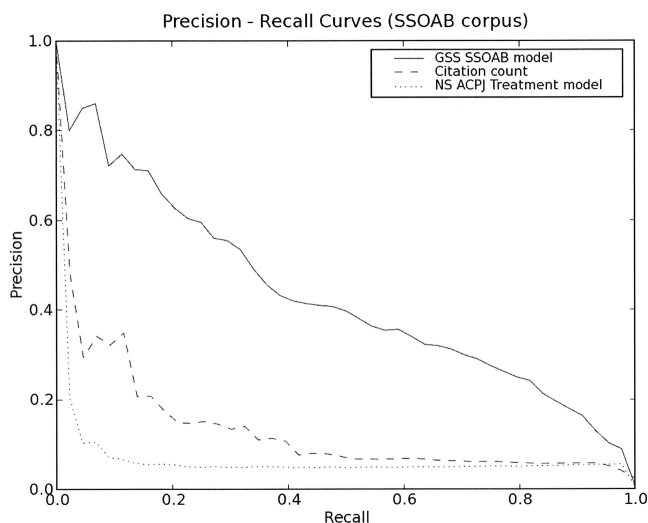


Abbildung 23: Recall Precision Graph (aus: <http://jamia.bmj.com/content/13/4/446.abstract>)