

Datenbanksysteme II
SS 2006

Übungsblatt 8: Grundlagen der Ähnlichkeitssuche

Besprechung: 03.07. – 04.07.2006

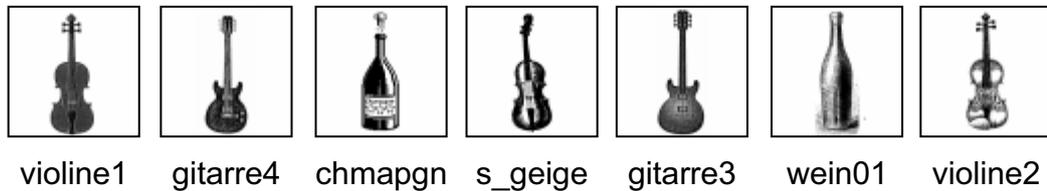
Aufgabe 8-1 *Recall/Precision und Sensitivität/Spezifität*

Gegeben sei das folgende gewünschte Anfrageergebnis aus einer Datenbank von 10.000 Bildern:

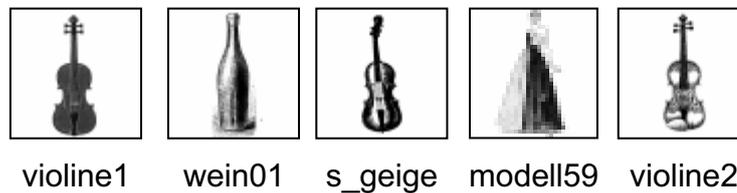


(a) Berechnen Sie jeweils Recall und Precision sowie Sensitivität und Spezifität für die folgenden beiden Anfrageergebnisse:

(i) Anfrageergebnis 1:



(ii) Anfrageergebnis 2:



(b) Wie ändern sich die Werte, wenn nur jeweils die ersten k der angegebenen Ergebnisse ausgegeben worden wären ($k = 1, \dots$)?

Aufgabe 8-2 *Distanzfunktionen*

(a) Zeigen Sie:

Für eine Metrik d gilt: $\forall o, q \in O : d(o, q) \geq 0$

(b) Sind die folgenden Distanzfunktionen für Punkte $o, q \in \mathbb{R}^n$ Metriken?

- $d_1(o, q) = \sum_{i=1}^n (o_i - q_i)$
- $d_2(o, q) = \max\{|o_i - q_i|, i = 1 \dots n\}$
- $d_3(o, q) = \sum_{i=1}^n (o_i - q_i)^2$
- $d_4(o, q) = \sum_{i=1}^n |o_i - q_i|$
- $d_5(o, q) = \sqrt{\sum_{i=1}^{n-1} (o_i - q_i)^2}$
- $d_6(o, q) = \sqrt{\sum_{i=1}^n (o_i - q_i)^2}$
- $d_7(o, q) = \sum_{i=1}^{n-1} \begin{cases} 1, & \text{falls } o_i = q_i \\ 0, & \text{sonst} \end{cases}$
- $d_8(o, q) = \sum_{i=1}^{n-1} \begin{cases} 1, & \text{falls } o \neq q \\ 0, & \text{sonst} \end{cases}$