

Knowledge Discovery in Databases II  
SoSe 2008

Übungsblatt 9: Multi-Instanz Data Mining  
Besprechung am 20.6.2008/23.6.2008

**Aufgabe 9-1** Hausdorff-Distanz

Zeigen Sie, dass die Hausdorff-Distanz alle drei Eigenschaften einer Metrik erfüllt.

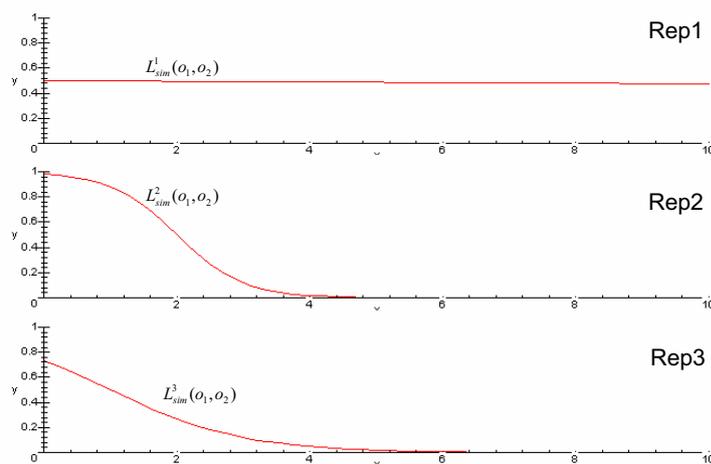
**Aufgabe 9-2** Kuhn-Munkres Algorithmus

Gegeben sei folgende Kostenmatrix K. Führen Sie ein Hungarian Matching auf dieser Matrix durch.

$$K = \begin{bmatrix} 90 & 75 & 75 & 80 \\ 35 & 85 & 55 & 65 \\ 125 & 95 & 90 & 105 \\ 45 & 110 & 95 & 115 \end{bmatrix};$$

**Aufgabe 9-3** Multirepräsentierte Ähnlichkeitsschätzer

Gegeben sei eine multirepräsentierte Datenmenge  $DB$  mit den 3 Repräsentationen  $R_1, R_2, R_3$ . Mittels multirepräsentierter Ähnlichkeitsschätzung wurden die folgenden Similarity-Likelihood-Funktionen  $L_{sim}^i$  bestimmt.



- Diskutieren Sie welche Aussagen sich über den Nutzen jeder Repräsentation anhand der jeweiligen Kurve treffen lassen.
- Wie stark wirkt sich jede Repräsentation auf die globale Ähnlichkeitsabschätzung aus.