

# Maschinelles Lernen: Symbolische Ansätze



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

Wintersemester 2009/2010

8. Übungsblatt für den 15.12.2009

## Aufgabe 1: RISE

Gegeben sei folgende Beispielmenge:

Outlook	Temperature	Humidity	Wind	PlayTennis
Overcast	Hot	High	False	Yes
Rainy	Cool	Normal	True	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	High	False	No
Overcast	Mild	High	True	Yes
Sunny	Cool	Normal	False	Yes
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	True	No
Sunny	Hot	High	False	No
Sunny	Mild	Normal	True	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Hot	High	True	No
Rainy	Mild	High	False	Yes
Overcast	Hot	Normal	False	Yes

- a) Wenden Sie den Algorithmus RISE (Foliensatz Instance-based Learning, Folie 42) auf den obigen Datensatz an. Berechnen Sie nur die erste Iteration und diese nur für die negativen Beispiele.

Bedenken Sie, dass die Accuracy berechnet wird, indem man für jede Regel alle Beispiele durchgeht und jeweils das nächste Beispiel nimmt, das nicht schon ohnehin abgedeckt wird. Dies ist notwendig, da sonst alle Beispiele immer mit den Regeln klassifiziert werden würden, die aus ihnen erstellt wurden und sich folglich die Accuracy nie verbessern könnte.

Benutzen Sie zur Berechnung der Distanz der Attribute die *Value Difference Metric* und nehmen Sie hierbei an, dass  $k = 1$  gilt. Als Vereinfachung müssen Sie die *VDM* nicht normieren.

Zur Bestimmung des Abstandes eines Beispiels zu einer Regel verwenden Sie die euklidische Distanz wobei Sie das Ziehen der Wurzel weglassen können.

Müssen Sie 2 Regeln zusammenfassen, so ist es Voraussetzung, dass diese die gleiche Klasse vorhersagen.

## Aufgabe 2: KD-Trees

- a) Bauen Sie einen KD-Tree aus der folgenden 2D Punktmenge auf und zeichnen Sie sowohl den Baum als auch die grafische Lösung im 2D-Raum:  
{4, 8}, {7, 4}, {5, 10}, {1, 3}, {2, 2}, {9, 1}, {10, 7}, {8, 10}, {6, 6}.
- b) Wenden Sie 1-NN für die folgenden beiden Queries {7, 9} und {1, 1} auf den Baum an und geben Sie die genaue Traversierung des Baumes an.