

# Maschinelles Lernen: Symbolische Ansätze



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

Wintersemester 2008/2009

2. Projektaufgabe für den 4.12.2008

---

## Evaluation von Regellernern

---

In dieser Aufgabe sollen unterschiedliche Evaluierungsmethoden unter Verwendung von WEKA eingesetzt und deren Ergebnisse diskutiert werden. Wenden Sie den Regellerner JRip auf die Datensätze `anneal.arff`, `credit-g.arff`, `diabetes.arff`, `tic-tac-toe.arff` und `vehicle.arff`, die Sie hier herunterladen können, an. Teilen Sie hierzu jeden Datensatz zunächst in 2 gleich große, stratifizierte Teile, einer Trainingsmenge und Validierungsmenge, auf.

- a) Trainieren Sie nun JRip auf jeder dieser Trainingsmengen (ggf. auf Teilen dieser Mengen, siehe Cross-Validation usw.) und evaluieren Sie die Genauigkeit (prozentualer Anteil der korrekt klassifizierten Beispiele) der resultierenden Klassifizierer jeweils mittels:

- 1x5 Cross-Validation
- 1x10 Cross-Validation
- 1x20 Cross-Validation
- Leave-One-Out
- bzw. auf der Trainingsmenge

Wie schätzen Sie die Qualität der erhaltenen Genauigkeitsabschätzungen ein?

*Anmerkung: In dieser Teilaufgabe sollen vorerst keine Veränderungen an weiterführenden Einstellungen, wie z.B. andere Random-Seeds, vorgenommen werden.*

- b) Wiederholen Sie Aufgabe a) mit dem Unterschied, daß Sie nun eine 10x10 Cross-Validation zur Evaluation verwenden sollen. Wenden Sie hierzu zehnmal eine 1x10 Cross-Validation mit 10 unterschiedlichen Random-Seeds an und mitteln die erzielten Genauigkeiten.

Vergleichen Sie die so erzielte Genauigkeitsabschätzung mit den Abschätzungen aus der Aufgabe a).

Führt ihrer Meinung nach eine geschickte Auswahl von Random-Seeds zu einer besseren Abschätzung?

- c) Bestimmen Sie die Genauigkeit auf der Validierungsmenge (d.h. verwenden Sie diese als Testmenge). Wie schätzen Sie nun unter Annahme, daß es sich bei der Validierungsmenge um einen realen Anwendungsfall handelt, die Abschätzungen der Evaluierungsmethoden aus den Aufgaben a) und b) ein?