

2. Übung „Knowledge Discovery“

Sommersemester 2006

Infos zu den Rechnerübungen

- Für die Rechnerübungen ist ein leistungsfähiger Rechner erforderlich.
- Wir können Zugänge zu den Rechnern des Fachgebietes zur Verfügung stellen. Bitte melden Sie sich bei Andreas Hotho oder Robert Jäschke.
- Die Rechnerübungen finden in Gruppen zu zwei bis drei Personen statt.
- Melden Sie sich auf <https://gforge.cs.uni-kassel.de> an und registrieren Sie sich beim Projekt kddss2005.
- Der Informationsaustausch zur KDD-Praxisübung wird über die Mailinglisten des GForge Systems erfolgen.
- **Bonus:** einen 0.3 Notensprung (also z.B. von 1.3 auf 1.0)
- **Was heißt Erfolg?** Bis zum 13.07.2006 gibt jede Gruppe eine schriftliche Ausarbeitung von maximal 10 Seiten ab. Diese dokumentiert die Arbeit und die Ergebnisse der Gruppe, wobei der Schwerpunkt auf den durchgeführten Arbeiten liegt und erst sekundär auf den Ergebnissen.
- **Werkzeug:** Weka (<http://www.cs.waikato.ac.nz/ml/weka/>), alternative Tools, wie die Demoversion von kommerziellen Anbietern oder Tools wie R können auch eingesetzt werden.
- **Was ist zu tun?** Sie lösen die unten beschriebene Aufgabe!

Ein Teil der Übung am 13.7.2006 wird dazu genutzt, Ihre Erfahrungen mit Weka vorzustellen und zu diskutieren. Dabei sollen aufgefallene Merkmale und Probleme des Werkzeugs (Benutzerfreundlichkeit, ...) diskutiert werden und mit möglichen Alternativen verglichen werden. Außerdem erhält jede Gruppe die Möglichkeit kurz ihre Lösungen sowie den Lösungsweg präsentieren.

Aufgabe

Lösen Sie die Aufgabe des KDD Cup 98 und beschreiben sie ihr Vorgehen. <http://kdd.ics.uci.edu/databases/kddcup98/kddcup98.html>

Abstract

This is the data set used for The Second International Knowledge Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-98 The Fourth International Conference on Knowledge Discovery and Data Mining. The competition task is a regression problem where the goal is to estimate the return from a direct mailing in order to maximize donation profits.