

t.STDM - Statistisches Data Mining

Person responsible for the course: Andreas Ruckstuhl, rkst

Credits: 4

Valid for: 2011/2012

Last saved: 16.08.2011 11:18

Learning objectives:

The students know the data mining process, know some of the methods from supervised and unsupervised statistical learning. They can assess the appropriateness of the methods for the available data and can apply the discussed methods practically using a data analysis / data mining software.

Course content:

Data mining process (CRISP), principal component analysis (PCA), dissimilarities, multidimensional scaling, hierarchical cluster analysis and partitioning (k-nearest neighbours), logistic classification, nearest-neighbour methods, decision trees, random forest, boosting, performance evaluation of classification methods.

R (a freely available language and environment for statistical computing and graphics) is used to apply the introduced methods.

Previous knowledge:

Statistisches Modellieren (StMo) or Datenanalyse und Prognose (DP)

Teaching method:

Type of lesson:	Number of lessons per week:
Lecture	14 x 2L
Tutorial/Practicum	14 x 2L
Group teaching	
Block instruction	
Seminar	

Assessment:

According to the table or as specified in writing by the lecture at the beginning of the semester!

Number	Type	Weighting
1	End of term exam	0.65
1	Exam during the semester	0.2 as bonus
1	Further assessments	0.15

Language of instruction:

Deutsch

Instruction material:

Skript

Comments:

