

Data Mining and Matrices

00 – Organization

Rainer Gemulla, Pauli Miettinen

April 18, 2013

Lecture

- Website & news
 - ▶ <http://www.mpi-inf.mpg.de/departments/d5/teaching/ss13/dmm/>
 - ▶ Access from outside may require authentication
- When?
 - ▶ Lecture: Thursday, 10:15AM-11:45AM, R021, building E1.4 (MPI-INF)
 - ▶ Assignments: at home
- Who?
 - ▶ Rainer Gemulla
 - ▶ Pauli Miettinen
 - ▶ Contact: dmm13@mpi-inf.mpg.de
- For whom?
 - ▶ Master & grad school students
 - ▶ Computer Science / Computer and Communications Technology
 - ▶ 5 credit points
- Assumed knowledge (briefly refreshed in lecture)
 - ▶ Basic linear algebra
 - ▶ Basic background in data mining (e.g., IR&DM, ML)

Assignments

- 4 assignments in total
 - ▶ Analyze a dataset
 - ▶ Write a short essay
 - ▶ Implement an algorithm
 - ▶ ...
- Grading
 - ▶ Pass (you need at least 3 passes)
 - ▶ Excellent (you get 1 bonus point)
- Timeframe
 - ▶ 2 weeks time per assignment
 - ▶ Dates (tentative): Apr 25, May 16, Jun 20, Jul 11

Certificate

- Registration in HISPOS
 - ▶ Must!
- Pass ≥ 3 assignments
 - ▶ Must!
- Take the exam
 - ▶ Must!
 - ▶ At end of semester (probably Jul 29–Aug 9)
 - ▶ In English (or, upon prior request, in German)
 - ▶ Bonus points can be used to improve grade (-0.15 per bonus point)
 - ▶ We test *understanding*, not learning by heart

What to expect?

- From us
 - ▶ Lecture notes online day before lecture (hopefully!)
 - ▶ Feedback on your assignments
 - ▶ Discussion of issues and questions around the course
 - ★ Email us!
 - ★ Appointments on demand
 - ▶ Anything you'd like to add?

- From you
 - ▶ Be here (physically and mentally)
 - ▶ Be active (ask questions, answer questions)
 - ▶ Laptops for note-taking only

Literature

- David Skillicorn
Understanding Complex Datasets: Data Mining with Matrix Decompositions
Chapman and Hall, 2007

- Carl Meyer
Matrix Analysis and Applied Linear Algebra
Society for Industrial and Applied Mathematics, 2000
<http://www.matrixanalysis.com>

- More in lecture notes