# Selected parts of data mining – a course intro

**Jiří Kléma, Filip Železný**

Department of Computer Science,
Czech Technical University in Prague

IDA
Intelligent Data Analysis
RESEARCH GROUP

http://cw.felk.cvut.cz/wiki/courses/xp36vpd/start

# The course content

- Modern and recent data mining methods,
- modern and recent machine learning methods
  - a sister ML course planned, currently joined with this DM course,
- Related bachelor and master courses (not direct prerequisites)
  - Pattern recognition,
  - Symbolic/statistical machine learning,
  - Statistical data analysis,
- Resources
  - books: Rajaraman, A., Leskovec, J., Ullman, J. D.: Mining of Massive Datasets, Cambridge University Press, 2011.
  - research papers: see examples,
  - recent issues in relevant journals: Data Mining and Knowledge Discovery, Machine Learning,
  - recent tutorials, major ML/DM conferences: KDD, Discovery Science, ICML, ECML, NIPS/NeurIPS.

# Expectations

- Reading class
  - no "labs", presentations and discussions only,
  - much effort spent on home work,
- each participant will prepare and give two presentations
  - around 60min each + discussion,
  - the first one a broader review,
  - the second one can touch/extend your current research if DM/ML related,
- you are expected to regularly attend the class
  - you are supposed to read a review paper recommended for the topic before presentations of the other students,
  - and prepare a few questions for discussion.
- the course will be concluded with an exam
  - rather a formal one,
  - still, every student should review the presented topics before the exam.

# The main issues for the first class

- Introduce each other
  - name, supervisor, dissertation topic, research interests,
- The initial discussion about the first assignments
  - the topics should be settled at least two weeks in advance ...
  - ... and publicly agreed by the audience.