

Data Mining II

DMA - Data mining: advanced topics and case studies

DM2, 2018

Instructors



Mirco Nanni

KDD Lab, ISTI-CNR

mirco.nanni@isti.cnr.it



Dino Pedreschi

KDD Lab, Dip. Informatica

pedre@di.unipi.it

Topics

Advanced/alternative data mining methods

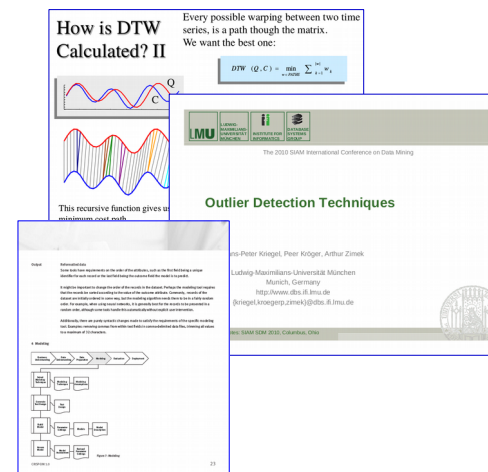
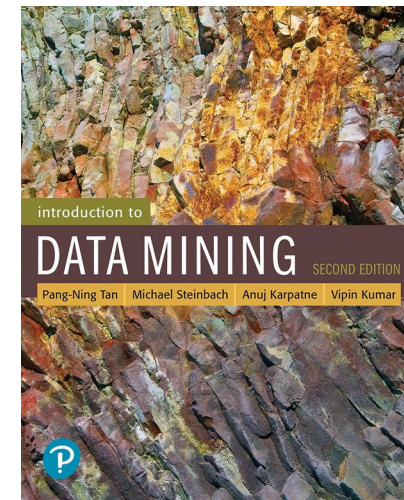
- Sequential patterns
- Time series
- Classification
 - Basic elements of KNN, SVM, Neural Networks, etc.
 - Validation methods
- Outlier/anomaly detection
- The data mining process at work (CRISP, etc.)
 - Sample case studies in the CRM domain

Environments

- **Knime**
 - Most classification and validation methods
 - Very few tools for time series
 - Some outlier detection tools
- **Python**
 - Most tasks are supported
- **Ad hoc external tools**
 - SPFM (sequential patterns)
 - ELKI (outlier detection)

Material

- **Some topics covered by textbook (Tan-Steinbach-Kumar)**
 - Most classification and validation methods
 - Sequential patterns
- **For the others, slides and ad hoc materials**
 - Time series
 - Outliers detection
 - DM / CRISP process



Classes

DM 2

Classes - Lezioni

| Day of week | Hour | Room |
|-------------|---------|------|
| Thursday | 14 - 16 | A1 |
| Friday | 16 - 18 | C1 |

Office hours - Ricevimento:

- Nanni : appointment by email, c/o ISTI-CNR

Exams

Same procedure as DM 1

- **Written exam**

- exercises and questions about all topics – sequential patterns, classification (advanced topics), times series, outliers detection, the data mining process
- Two (optional) mid-term exams will be given as replacement

- **Project(s)**

- Topics proposed during the classes
- A single comprehensive report to be sent 2 days before the oral exam
- Groups composed of up to 3 people
- To be performed any tool, including (but not necessarily) those presented during classes (Knime, Python, ...)

- **Oral exam**

- Short discussion of the project (group presentation, where possible)
- Questions on all topics presented during the classes

Web-site

- Same as DM 1: <http://didawiki.di.unipi.it/doku.php/dm/>

Data Mining A.A. 2017/18

DM 1: Foundations of Data Mining (6 CFU)

Instructors - Docenti:

- **Dino Pedreschi**
 - KDD Laboratory, Università di Pisa ed ISTI - CNR, Pisa
 - <http://www-kdd.isti.cnr.it>
 - ✉ [dino \[dot\] pedreschi \[at\] unipi \[dot\] it](mailto:dino@di.unipi.it)

Teaching assistant - Assistente:

- **Riccardo Guidotti**
 - KDD Laboratory, Università di Pisa and ISTI - CNR, Pisa
 - ✉ [guidotti \[at\] di \[dot\] unipi \[dot\] it](mailto:guidotti@di.unipi.it)

DM 2: Advanced topics on Data Mining and case studies (6 CFU)

Instructors:

- **Mirco Nanni, Dino Pedreschi**
 - KDD Laboratory, Università di Pisa and ISTI - CNR, Pisa
 - <http://www-kdd.isti.cnr.it>
 - ✉ [mirco \[dot\] nanni \[at\] isti \[dot\] cnr \[dot\] it](mailto:mirco@isti.cnr.it)
 - ✉ [dino \[dot\] pedreschi \[at\] unipi \[dot\] it](mailto:dino@di.unipi.it)