



Data mining concepts and techniques

Introduction to the PhD course

Flavio Giobergia, Danilo Giordano

Course content

- Introduction to **Data Science**
- **Data preparation** for Data Mining
- Association **rules**, itemset mining
- **Classification** techniques, supervised learning
- **Cluster** analysis, unsupervised learning
- Seminar about current trends in Data Science
- **Practice** with a data mining tool and **Exam**



Course schedule

- Mon, Feb 3, 2025 14:00-18:00 4M
 - Introduction to the course, Datamining pipeline and Preprocessing
- Mon, Feb 10, 2025 14:00-18:00 4M
 - Association Rules: Introduction, Algorithms and Performance metrics
 - Classification: Introduction
- Mon, Feb 17, 2025 14:00-18:00 4M
 - Classification: Algorithms, Validation, Performance metrics
 - Clustering: introduction
- Mon, Feb 24, 2025 14:00-18:00 (?)
 - Clustering: Algorithms, Performance metrics
 - Deep learning introduction
- Fri, Feb 28, 2025 13:00-17:00 LABINF
 - Hands on the data with RapidMiner/Python, and Exam

Please visit the web page for updates:

https://dbdmg.polito.it/dbdmg_web/index.php/2021/11/17/data-mining-concepts-and-algorithms/

- Practice with a machine learning tool
 - Exploiting data mining **algorithms**
 - To analyze a **real dataset**
 - **Fill a report** with your results and send it via **email**
- Oral part about (**optional**) – required to pass with Merit
 - The applied algorithms
 - The data mining results
 - Data science solutions seen during the course
- The exam is performed soon after the last lecture
 - A dataset will be assigned to each of you
 - You will exploit different data mining algorithms to analyze the dataset
 - Fill the exam report with your results
 - Take the oral part

➤ Reference books

- Tan, Steinbach, Kumar, *Introduction to data mining*, Pearson, 2006
- Han, Kamber, *Data mining: concepts and techniques*, Morgan Kaufmann, 2006