

Data Mining

a.y. 2019-2020

Anna Monreale

Dino Pedreschi

Computer Science Department, University of Pisa

Data Mining

Data Mining (12 CFU) – Code: 420AA

- **DM I: Foundations of Data Mining** (first semester)
- **DM II: Advanced topics on Data Mining and case studies** (Second semester)
- For students in **Data Science & Business Informatics**

Data Mining (9 CFU) Code: 309AA

- For students in **Computer Science**

Data Mining: Fundamentals (6 CFU) Code: 676AA

- For students in **Informatica Umanistica**

Data Mining (first semester)

- **Instructors**

- Dino Pedreschi, pedre@di.unipi.it
- Anna Monreale, anna.monreale@unipi.it

- **Classes**

- Monday 14:00 - 16:00 Room C1
- Wednesday 16:00 - 18:00 Room A1
- Friday 11:00 - 13:00 Room C1

- **NO Class for Students from Data Science & Business Informatics and Informatica Umanistica after mid November (the last lesson will be around 22 Nov, date to be confirmed)**
- **After that date classes are dedicated to Computer Science Master Degree**

Data Mining

- **Web Page**

- <http://didawiki.cli.di.unipi.it/doku.php/dm/start>

- **Learning Material**

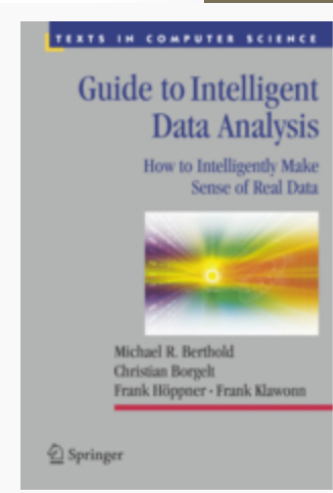
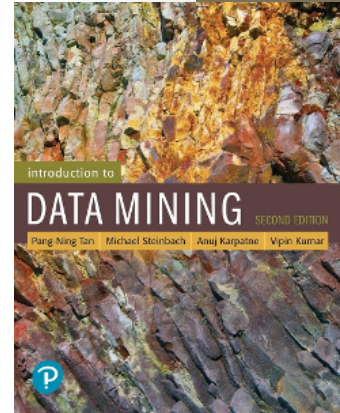
- **Textbook:** Pang-Ning Tan, Michael Steinbach, Vipin Kumar **Introduction to DATA MINING** Addison Wesley, ISBN-13: 978-0-13-312890-1, 2018

- **Textbook:** Berthold, M.R., Borgelt, C., Höppner, F., Klawonn, F. **GUIDE TO INTELLIGENT DATA ANALYSIS**. Springer Verlag, 1st Edition., 2010. ISBN 978-1-84882-259-7

- **Slides of the classes**

- **Data mining software**

- **KNIME:** The Konstanz Information Miner. Download page
- **Scikit-learn:** python library with tools for data mining and data analysis



Exam

- A **written exam**, with exercises and questions about methods and algorithms presented during the classes. It can be done:
 - Before Christmas holidays
 - During the exam sessions: Winter Session 2020, Summer Sessions 2020
- A **project**, consisting in exercises that require the use of data mining tools for analysis of data. Exercises include: *data understanding, market basket analysis, clustering analysis and classification*.
 - The project has to be performed by max 4 people.
- An **oral exam** that includes: (1) discussing the project report with a group presentation; (2) discussing topics presented during the classes, including the theory of the parts already covered by the written exam.
- **Oral exam is not mandatory if your written exam is passed by the mid-test**

Mailing list

- Send as soon as possible a mail to anna.monreale@unipi.it with the following data:
 - subject: "DM"
 - text:
 - name
 - surname
 - student-id
- You will be included in a mailing list of urgent news