

Data Mining II / Adv. Topics in Data Science

Presentation

Álvaro Figueira and Rita P. Ribeiro

2023/2024



The Course

- Classes
- Assessment
- Syllabus
- Main Resources and References

- Álvaro Figueira (arfiguei@fc.up.pt) and Rita P. Ribeiro (rpribeiro@fc.up.pt)
- **Classes:**
 - Lectures
 - Mondays, 11:00-12:30
 - Practical Lectures (Labs)
 - Mondays, 9h00-11h00
 - Mondays, 14h00-16h00
- **Webpage:**
 - On Moodle: [FCUP-CC4024_CC4061-2023/2024-META-2S](#)

- Evaluation Components

- Practical Assignment (groups 1-4 students): 8 points (40%)
- Final Exam: 12 points (60%)

- Important Notes

- Students need to attend, at least, 70% of the lectures and labs.
- Minimum grade of 30% in each component.
- Students who do not deliver the practical assignment are not eligible for the final test or supplementary exam.

- Frequent Pattern Mining
- Web Mining
- Information Retrieval
- Text Mining
- Introduction to Natural Language Processing
- Outlier Mining

Main Resources

- Books:
 - “Data Mining, The Textbook”, Charu C. Aggarwal, Springer, 2015
 - “Web Data Mining - Exploring Hyperlinks, Contents and Usage Data - Second Edition”, Bing Liu, Springer, 2011
 - “Data Mining with R: learning with case studies - Second Edition”, Luís Torgo, CRC Press, 2016
 - “Extração de Conhecimento de Dados - Data Mining, 3a edição”, João Gama et al., Edições Sílabo, 2017
 - “Introduction to Data Mining”, Pang-Ning Tan et al., Addison-Wesley, 2005.
- Software:
 - [R Project](#)
 - [RStudio](#) (Posit)

Main Resources (cont.)

- Additional References:
 - [An Introduction to R.](#)
W. N. Venables, D. M. Smith and the R Core Team, 2022.
 - [Bookdown](#)
(several free R books for Data Science)
 - [R Programming for Data Science](#)
Peng, R, 2022
 - [Quick-R](#)
 - [Advanced R](#)
Hadley Wickham
 - [RStudio Cheatsheets](#)

Additional Resources

- Video Lectures on Machine Learning
- KDD nuggets
- Data sets at UCI Repository
- Kaggle