Data Science & Machine Learning Lab

Introduction to the course

Flavio Giobergia

Course contents – theory

- Data science, Machine Learning algorithms & techniques
  - Data exploration and preparation
  - Classification
  - Regression
  - Clustering
  - Anomaly Detection
  - Association Rule extraction
  - Deep learning (notions)

Course contents – practice

- Python programming
  - NumPy
  - Pandas
  - Scikit-learn
  - PyTorch

Course contents – labs, case studies, RB

- 10 lab sessions
  - On open datasets
  - Covering theoretical & practical topics
  - Competitions
  - Exam simulations
- Company case studies
- Research Bites
  - Seminars on advanced DS&ML techniques, by PhD students

Lectures

📅 Mondays
- 🕒 16:00 → 17:30
- 📍 Room 2P

📅 Wednesdays
- 🕒 16:00 → 19:00
- 📍 Room 1B

Labs

- Starting Oct. 6
- 3 teams
  - Will be announced on the course website
  - (We're trying to handle overlaps with other courses! Send an email to claudio.savelli@polito.it)

📅 Mondays
🕙 10:00 → 13:00
📍 LAIB3B

📅 Tuesdays
🕗 8:30 → 11:30
📍 LAIB2B

📅 Thursdays
🕙 10:00 → 13:00
📍 5T

What's new?

- Topics:
  -  ~~Introduction to Python~~
  -  Removed redundancies between theory/practice
  -  Introduction to deep learning
  -  PyTorch basics
  -  Anomaly detection
- Exam
  - Updated partial scores (written: 22, project: 10)
  - New exercise type: implementation of a simple data science pipeline
  - Simplified exam rules (more on this in early December)

Material

All slides, labs, solutions will be uploaded on the course website
https://dbdmg.polito.it/dbdmg_web/2025/data-science-and-machine-learning-lab-2025-26/

Additional books:

- Tan, Steinbach, Karpathe, Kumar, *Introduction to data mining*, 2nd edition, Pearson, 2019
- Han, Kamber, Pei *Data mining: concepts and techniques*, 3rd edition, Morgan Kaufmann, 2011
- Kent D. Lee , *Python Programming Fundamentals*, Springer, 2015
- Jake VanderPlas, *Python Data Science Handbook: Essential Tools for Working with Data*, O'Reilly, 2016