MGT-529

Data science and machine learning II

Thurm Boris, Vlachos Michail

Cursus	Sem.	Type
Managmt, dur et tech	MA3	Obl.

Language of teaching	English
Credits	3
Withdrawal	Unauthorized
Session	Winter
Semester	Fall
Exam	During the semester
Workload	90h
Weeks	14
Hours	**3 weekly**
Lecture	2 weekly
Exercises	1 weekly
Number of positions	**40**

It is not allowed to withdraw from this subject after the registration deadline.

Remark

Only given in 2022-23

Summary

This class discusses advanced data science and machine learning (ML) topics: Recommender Systems, Graph Analytics, and Deep Learning, Big Data, Data Clouds, APIs, Clustering. The course uses the Wolfram Language. Outcome: coding exercises in ML using real-data and a big ML project deliverable.

Content

- Big Data/Cloud, API
- Clustering
- Association rules
- Recommender system
- Graph Analytics
- Deep Learning
- Applications in Wolfram Language on Text, Audio, Images Analysis

Keywords

Data science, Machine learning, Algorithms, Big Data, Clustering, Recommender Systems, Deep Learning, Wolfram Language

Learning Prerequisites

Required courses

Data Science and Machine Learning I (MGT-492)

Important concepts to start the course

- Fundamental Probability and Statistics concepts
- Fundamental ML topics: cost function and optimization, gradient descent, K-fold cross-validation, overfitting, model calibration, confusion matrix, curse of dimensionality
- Basic ML methods: Regression (linear regression, ridge regression), Classification (logistic regression, k-NN classification, decision trees, random forests), Dimensionality reduction (PCA, ISOMAP, t-SNE)
- Basic understanding of Neural Networks and Text Analytics (text representation, sentiment classification, similarity search)
- Basic programming skills in Wolfram Language

Learning Outcomes

By the end of the course, the student must be able to:

- Choose an appropriate Machine Learning method for a given task
- Design and Conduct a data science project
- Investigate data, data types, and problems with the data
- Implement ML algorithms in Wolfram Language

Transversal skills

- Plan and carry out activities in a way which makes optimal use of available time and other resources.
- Demonstrate the capacity for critical thinking
- Access and evaluate appropriate sources of information.
- Collect data.
- Use a work methodology appropriate to the task.

Teaching methods

- Lectures
- Lab sessions: coding exercices
- Data Science projects

Expected student activities

The students are expected to:

- attend lectures and lab sessions;
- work on the weekly theory and coding exercises;
- complete assignments (graded);
- conduct data science projects making use of the theory learned during lectures and code developed during lab sessions (graded)

Assessment methods

- Coding assignments: 50%
- Project: 50%

Supervision

Office hours	No
Assistants	Yes
Forum	No
Others	Slack channel

Resources

Virtual desktop infrastructure (VDI)

No

Notes/Handbook

Slides will be made available on the course Moodle page. Notebooks will be made available in a GitHub repository.
Book: **Introduction to Machine Learning, Etienne Bernard (2022)**
https://www.amazon.com/Introduction-Machine-Learning-Etienne-Bernard-ebook/dp/B09PSRXFZ9/

Moodle Link

- https://go.epfl.ch/MGT-529