Module Description, available in: EN

Machine Learning and Data in Operation

General Information

Number of ECTS Credits

3

Module code

TSM_MachLeData

Valid for academic year

2024-25

Last modification

2023-06-20

Coordinator of the module

Vanni Galli (SUPSI, vanni.galli@supsi.ch)

Explanations regarding the language definitions for each location:

- Instruction is given in the language defined below for each location/each time the module is held.
- Documentation is available in the languages defined below. Where documents are in several languages, the percentage distribution is shown (100% = all the documentation).
- The examination is available 100% in the languages shown for each location/each time it is held.

	Lausanne			Lugano	Zurich		
Instruction	**X** E 100%				**X** E 100%		
Documentation	**X** E 100%				**X** E 100%		
Examination	**X** E 100%				**X** E 100%		

Module Category

TSM Technical scientific module

Lessons

2 lecture periods and 1 tutorial period per week

Entry level competences

Prerequisites, previous knowledge

- Basic knowledge of machine learning, deep learning, data management and data engineering.
- Good command of an imperative programming language, basic knowledge of Python.
- Basic knowledge of probability, statistics, linear algebra (vectors, matrices).

Brief course description of module objectives and content

This module presents powerful techniques to manage the lifecycle of machine learning models, covering in particular baseline models, infrastructure (clusters, cloud, edge AI and resource management) and tooling (frameworks), model training and debugging, model evaluation and tuning, data management (sources, storage, versioning, privacy), systems testing (CI/CD) and explainability, deployment (batch, service, edge), monitoring (data drift) and continual learning. Emphasis is placed on practical tools, real use-case scenarios, and the relevant hardware and software platforms.

Additional topics such as business requirements and objectives, project management for ML, team structure, user experience as well as responsible use of ML systems, including sustainable AI, are also considered.

Aims, content, methods

Learning objectives and acquired competencies

- Recognising the complete lifecycle of machine learning projects, from data requirements to development, deployment, and monitoring.
- Demonstrating skills in maintaining ML code and data, version and integrate it, and define appropriate environments, with emphasis on practical applications such as data cleaning and preprocessing.
- Deploying ML models at scale, monitoring their performance and adapting models to changing requirements, with a focus on assessing and adjusting to data drift and shifts in data distribution.
- Analysing relevant tools and real use-case scenarios, such as real-time services management; critically analysing the implications and applications in practical scenarios.
- Selecting software and hardware platforms based on the requirements of different scenarios, demonstrating a thorough understanding of the needs and constraints of each.
- Extracting and integrating insights from guest lectures by industry professionals (subject to availability), demonstrating the ability to interpret expert knowledge from scientific literature and online resources, and applying it effectively to complement their hands-on experience.

Contents of module with emphasis on teaching content

- Brief recap of machine learning and deep learning.
- Introduction to the lifecycle of a Machine Learning project.
- Understanding data needs and requirements for ML projects (e.g. versioning, storage, processing, labeling, augmentation, simulation).
- ML Development: defining the environment, maintaining the ML code, integrating ML code (versioning, evaluation, baselines).
- ML Deployment: running models at scale (e.g. batch vs online, model compression, cloud / edge deployment), ensuring system availability, monitoring performance, adapting to changes (data distribution shifts, failures, metrics, logging, continual learning).
- Exploration of tools and real-world scenarios.
- Overview of relevant hardware and software platforms.
- Selection of advanced topics such as:
    - Trustworthy AI (incl. regulatory aspects).
    - Guest lecture(s) from industry professionals (subject to availability).
    - Project management and business perspective (e.g. job roles, teams).

Teaching and learning methods

Classroom teaching; programming exercises using MLOps tools and Python (among others); guest lectures from industry professionals (subject to availability).

Literature

- Chip Huyen , “Designing Machine Learning Systems: An Iterative Process for Production-Ready Application”, O-Reily, 2022

- Noah Gift & Alfredo Deza, “Practical MLOps - Operationalizing Machine Learning Models”, O’Reilly, 2021

- Scientific literature and articles as discussed during the lectures

Assessment

Certification requirements

Module uses certification requirements

Certification requirements for final examinations (conditions for attestation)

Group project development: conducted throughout the semester with the final presentation taking place during the last lecture of the module. An in-person presence for the presentation at the last lecture is mandatory.

The group project contributes to 1/3 of the final grade.

Basic principle for exams

As a rule, all the standard final exams for modules and also all resit exams are to be in written form

Standard final exam for a module and written resit exam

Kind of exam

written

Duration of exam

120 minutes

Permissible aids

Aids permitted as specified below:

Permissible electronic aids

No electronic aids permitted

Other permissible aids

No other aids permitted

Special case: Resit exam as oral exam

Kind of exam

oral

Duration of exam

30 minutes

Permissible aids

No aids permitted