COURSE DESCRIPTION CARD - SYLLABUS

Course name
Machine learning operations (MLOps)

Course

Field of study
Artificial Intelligence
Area of study (specialization)
-
Level of study
Second-cycle studies
Form of study
full-time

Year/semester
1/2
Profile of study
general academic
Course offered in
English
Requirements
elective

Number of hours

Lecture
15
Laboratory classes
15
Other (e.g. online)

Tutorials
Projects/seminars

Number of credit points

3

Lecturers

Responsible for the course/lecturer:
prof. Mikolaj Morzy, Ph.D., D.Sc.
email: mikolaj.morzy@put.poznan.pl
tel: 61 665-3420
Faculty of Computing and Telecommunications
Piotrowo 2, 60-965 Poznan

Responsible for the course/lecturer:

Prerequisites

Student partaking this course should have basic knowledge in the areas of machine learning and data processing. Student should be able to solve basic problems related to information systems' design and implementation (versioning, testing, code integration). Basic command of Python programming language is required. Student should be able to utilize external APIs and integrate external resources. Student should also understand the need to broaden one's competences and be able to effectively cooperate within a project team. Regarding social competences student should exhibit qualities such as

responsibility, endurance, intellectual curiosity, creativity, respect for other people, ability to work in a group.

Course objective

The main objective of the course is to present a rich set of tools required to apply machine learning solutions to information systems in practice. During lectures students learn techniques and tools for statistical model productization, in particular, tools allowing for a seamless integration of machine learning models with the existing computing infrastructure or ecosystem. During laboratories students experiment with the tools using various scenarios. The list of topics covered by the course includes:

- versioning of data and statistical models,
- workflow management tools for machine learning,
- tools for data annotation for machine learning models' training,
- monitoring of the learning process,
- practical aspects of machine learning models' productization,
- tools for managing of machine learning projects.

Course-related learning outcomes

Knowledge

Student has advanced and in-depth knowledge of machine learning and of practical aspects of productization of machine learning solutions [K2st_W1].

Student has knowledge of good practices related to the development and practical application of machine learning solutions in information systems, in particular, good practices related to the testing and verification of statistical models [K2st_W2].

Student has detailed knowledge of the collection, annotation, and versioning of data used for machine learning models' training [K2st_W3].

Student is aware of the currently available tools, libraries, and frameworks for development, integration, and maintenance of information systems utilizing machine learning solutions [K2st_W4].

Student understands the full life cycle of information systems utilizing machine learning solutions, is able to evaluate the correctness of the life cycle and identify non-trivial dependencies between subsequent steps of the life cycle (data collection and annotation, training of statistical models, validation and optimization of statistical models, monitoring of statistical models) [K2st_W5]

Student knows basic methods, techniques and tools used to design complex information systems and is able to apply these methods, techniques and tools to the specific case of designing systems utilizing machine learning solutions [K2st_W6].

Skills

Student knows how to use various APIs and documentation of complex information systems [K2st_U1].

During the design of an information system utilizing machine learning solutions the student is able to plan and carry out measurement experiments in accordance with scientific method and to interpret the results of conducted experiments [K2st_U3].

Student knows how to use visualization tools to monitor the process of machine learning model training [K2st_U4].

Student is able to use agile programming methodologies to manage a project containing statistical models at its core. Student understands the need to include domain expertise and knowledge in the process of designing machine learning based solutions [K2st_U5].

Student is able to construct information systems with machine learning components using the containerization technique which allows a seamless integration of the existing workflow with new tools and solutions [K2st_U6].

Student is able to evaluate the time required to complete development steps of an information system utilizing statistical models [K2st_U7].

Student understands the nature of the technological debt incurred by machine learning and is able to use available tools to mitigate the negative impacts that potential errors in the statistical model training process can project on the remaining part of an information system [K2st_U8].

During the design of an information system utilizing machine learning solutions the student is able to analyze available data in search for latent biases and correlations. The student is able to analyze the life cycle of data and spot threats to the integrity of the model training process [K2st_U9].

Student is able to design and develop a solution to a given economic, technological, or social problem using machine learning tools [K2st_U10].

Social competencies

Student understands the exceptional dynamics of the area of machine learning and is aware of the existence of multiple tools [K2st_K1].

Student understands the need for life-long learning in the area of machine learning tools due to a rapid cycle of tool exchange [K2st_K2].

Student can effectively communicate in a group and collaborate within the ramifications of agile programming techniques [K2st_K3].

Methods for verifying learning outcomes and assessment criteria

Learning outcomes presented above are verified as follows:

Knowledge obtained during lectures is verified on the basis of a report prepared by students either individually, or in pairs. The report documents a selected tool for machine learning. The tool is presented in a seminar form during the last lecture of the course.

Knowledge obtained during laboratories is verified during the semester using short programming exercises for individual completion by students. In addition, at the end of the course the students prepare a group project (2-4 persons). The project should present the entire life cycle of an information system with machine learning component.

Programme content

General ramifications of machine learning tools are presented during lectures. Laboratories present detailed case studies of tool usage at a given stage of the development of a machine learning-based system. Topics covered during the course include:

- data versioning
- manual and programming annotation of data, including text and image data
- synthetic data generation
- feature stores
- tools for the visualization of the machine learning processing
- tools for containerization of machine learning systems
- machine learning workflows
- universal platforms for machine learning

Teaching methods

Lecture: multimedia presentation, seminar with students' presentations, information retrieval

Laboratory: programming examples, small individual programming exercises

Project: practical problem solution, group work, design thinking, documentation

Bibliography

Basic

1. Géron, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, 2019.

2. Raschka, Sebastian, and Vahid Mirajalili. Python machine learning. No. 1. Packt Publishing,, 2019.

Additional

1. Pustejovsky, James, and Amber Stubbs. Natural Language Annotation for Machine Learning: A guide to corpus-building for applications. O'Reilly Media, Inc. 2021.

2. Zheng, Alice, and Amanda Casari. Feature engineering for machine learning: principles and techniques for data scientists. O'Reilly Media, Inc., 2018.

.

Breakdown of average student's workload

	Hours	ECTS
Total workload	75	3.0
Classes requiring direct contact with the teacher	30	1.5
Student's own work (literature studies, preparation for laboratory classes/tutorials, preparation for tests/exams, project preparation)[1]	45	1.5

[1] niepotrzebne skreślić lub dopisać inne czynności