

COURSE SYLLABUS

AI Systems in Production, 7.5 credits

AI-system i produktion, 7.5 högskolepoäng

Course Code:	T2AIPN	Education Cycle:	Second-cycle level
Confirmed:	Sep 01, 2025	Disciplinary domain:	Technology
Valid From:	Jan 18, 2027	Subject group:	Computer Technology
		Specialised in:	A1F Second cycle, has second-cycle course/s as entry requirements
		Main field of study:	Computer Science

Intended Learning Outcomes (ILO)

On completion of the course the student shall:

Knowledge and understanding

- demonstrate comprehension of the challenges and complexities of managing AI/ML systems in production environments.
- display knowledge of scalable architectures, pipelines, deployment, and monitoring strategies for AI/ML-enabled systems.
- demonstrate comprehension of the best practices for data versioning, configuration management, and continuous deployment and performance optimization in AI/ML workflows.
- demonstrate comprehension of the concepts of reproducibility, traceability, and explainability in AI/ML systems.

Skills and abilities

- demonstrate the ability to design, implement, and automate production-ready AI/ML pipelines to deploy AI models and services in production environments.
- demonstrate skills of ensuring quality of the entire AI/ML pipeline with test automation and other quality assurance techniques.
- demonstrate the ability to set up monitoring systems to track drifts (model/data/concept), system performance, and errors in real-time.

Judgement and approach

- demonstrate the ability to assess the trade-offs between system scalability, maintainability, and operational complexity.

Content

The course explores the challenges and solutions involved in deploying, managing, and maintaining AI/ML-enabled systems in production environments. It covers architectural design, deployment pipelines, data versioning, configuration management, and monitoring strategies to ensure scalability, reliability, and performance of AI systems. Students will gain hands-on experience using appropriate languages and tools, such as Python, Git, MLFlow, Prometheus, and Databricks, to develop, deploy, and monitor AI systems in real-world scenarios.

The course includes the following elements:

- Architecture and Design for AI-Enabled Systems in Production
- Deployment Pipelines and Model Lifecycle Management
- Data Versioning and Configuration Management

- Quality Attributes and Operational Challenges
- Monitoring, Logging, and Error Management

Type of instruction

The teaching mainly consists of lectures, assignments, and workshops.

Language of instruction is English.

Entry requirements

Passed courses at least 90 credits within the major subject computer engineering, computer science, informatics, information systems or information technology, including a minimum of 15 credits in mathematics and at least 30 credits in programming/software development, or alternatively passed courses at least 150 credits from the programme Computer Science and Engineering, and taken Developing AI-enabled Systems, 7.5 credits and Data Science, 7.5 credits.

Examination and grades

The course is graded 5, 4, 3 or U.

Registration of examination:

Name of the Test	Value	Grading
Examination ¹	3.5 credits	5/4/3/U
Assignment	4 credits	G/U

¹Determines the final grade of the course, which is issued only when all course units have been passed.

Course literature

Please note that changes may be made to the reading list up until eight weeks before the start of the course.

Kästner, C. (2025). Machine learning in production: from models to products. MIT Press, <https://mlip-cmu.github.io/book/>.

Huyen, C. (2024). AI Engineering: Building Applications with Foundation Models. O'Reilly Media, Incorporated.

Bass, L., Lu, Q., Weber, I., & Zhu, L. (2025). Engineering AI systems: architecture and DevOps essentials. Addison-Wesley Professional.

Chip, H. (2022). Designing machine learning systems: An iterative process for production-ready applications.