GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Advanced Data Preparation in Machine Learning, PG_00064487		
Field of study	Informatics		
Date of commencement of studies	February 2027	Academic year of realisation of subject	2026/2027
Education level	second-cycle studies	Subject group	Optional subject group Specialty subject group Subject group related to scientific research in the field of study
Mode of study	Full-time studies	Mode of delivery	at the university
Year of study	1	Language of instruction	Polish
Semester of study	1	ECTS credits	2.0
Learning profile	general academic profile	Assessment form	assessment
Conducting unit	Department of Computer Architecture -> Faculty of Electronics Telecommunications and Informatics -> Faculties of Gdańsk University of Technology		
Name and surname of lecturer (lecturers)	Subject supervisor	dr inż. Jan Cychnerski	
	Teachers	dr inż. Jan Cychnerski	

Lesson types

Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
Number of study hours	15.0	0.0	0.0	15.0	0.0	30

E-learning hours included: 0.0

Learning activity and number of study hours

Learning activity	Participation in didactic classes included in study plan	Participation in consultation hours	Self-study	SUM
Number of study hours	30	4.0	16.0	50

Subject objectives

The main goal of the subject is to present practical methods of solving problems using artificial intelligence techniques: dataset construction, architecture choice, artificial intelligence algorithms training, model selection and testing.

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K7_U12] is able, to an increased extent, to analyze the operation of components and systems related to the field of study, as well as to measure their parameters and study their technical characteristics, and to plan and carry out experiments related to the field of study, including computer simulations, interpret the obtained results and draw conclusions	The student is able to select machine learning methods in the context of the requirements of the problem being solved.	[SU1] Assessment of task fulfilment
[K7_U08] while identifying and formulating engineering tasks specifications and solving these tasks, can: - apply analytical, simulation and experimental methods, - notice their systemic and non-technical aspects, - make a preliminary economic assessment of suggested solutions and engineering work	The student is able to conduct training and perform tests of the operation of a neural network or other artificial intelligence algorithm.	[SU1] Assessment of task fulfilment
[K7_K02] is ready to provide critical evaluation of received content and to acknowledge the importance of knowledge in solving cognitive and practical problems	The student is able to assess the difficulty of problems solved by machine learning methods and search for information helpful in solving them.	[SK5] Assessment of ability to solve problems that arise in practice
[K7_W04] knows and understands, to an increased extent, the principles, methods and techniques of programming and the principles of computer software development or programming devices or controllers using microprocessors or other elements or programmable devices specific to the field of study, and organization of work of systems using computers or such devices	The student has the advanced knowledge necessary to design systems based on artificial intelligence. The student understands the principles of operation and ways of using commonly used libraries and environments providing self-learning algorithms. The student knows the practical limitations and the best ways to apply methods and systems of artificial intelligence. The student understands the reasons of these limitations and their practical effects.	[SW1] Assessment of factual knowledge

Subject contents

Course content – lecture

1. General information on the practical aspects of training artificial intelligence algorithms
2. Preparation of training, validation, and test datasets
3. Preprocessing, normalization, and augmentation of training data
4. Selection of machine learning methods in the context of the requirements of the problem being solved
5. Methods of testing and evaluating the effectiveness and efficiency of artificial intelligence algorithms

Course content – project

1. Defining requirements for raw data collection
2. Raw data collection, analysis, and quality control
3. Defining requirements for the data annotation process
4. Data annotation process, analysis, and annotation quality control
5. Training and advanced validation of machine learning models

Prerequisites and co-requisites

Basic knowledge of artificial intelligence area, basic knowledge of Python programming language

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
project	50.0%	50.0%
written test	50.0%	50.0%

Recommended reading

Basic literature	1. James, Gareth, et al. An introduction to statistical learning. Vol. 112. New York: springer, 2013. 2. Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning, http://www.deeplearningbook.org/ 3. Scikit-learn Tutorials, http://scikit-learn.org/stable/tutorial/index.html
Supplementary literature	1. Andrew Ng, Machine Learning Yearning, http://www.mlyearning.org/
eResources addresses	

Example issues/ example questions/ tasks being completed

Sample issues:
- Training and testing datasets preparation for classifier training
- Performing training and testing of a neural network or other artificial intelligence algorithm
- Description of methods and measures for quality, efficiency and performance of machine learning algorithms

Practical activites within the subject	Not applicable

Document generated electronically. Does not require a seal or signature.