GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Advanced Data Mining Techniques, PG_00048044		
Field of study	Informatics		
Date of commencement of studies	February 2023	Academic year of realisation of subject	2022/2023
Education level	second-cycle studies	Subject group	Optional 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 Software Engineering -> Faculty of Electronics, Telecommunications and Informatics		
Name and surname of lecturer (lecturers)	Subject supervisor	dr Paweł Weichbroth	
	Teachers	dr Paweł Weichbroth dr inż. Aleksandra Karpus dr inż. Agata Kołakowska	

Lesson types and methods of instruction

Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
Number of study hours	0.0	0.0	30.0	0.0	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 aim of the course is to introduce students to the exploration and visualization of large data sets using a scalable computing cluster using modern functional languages and statistical packages.

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K7_W05] Knows and understands, to an increased extent, methods of process and function support, specific to the field of study.	Student is able to explain and apply data mining algorithms.	[SW2] Assessment of knowledge contained in presentation
[K7_U07] can apply advanced methods of process and function support, specific to the field of study	Student is able to chose appropriate data mining methods and evaluate them.	[SU4] Assessment of ability to use methods and tools [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	Student is able to use his knowledge to solve real domain problems.	[SK5] Assessment of ability to solve problems that arise in practice
[K7_U04] can apply knowledge of programming methods and techniques as well as select and apply appropriate programming methods and tools in computer software development or programming devices or controllers using microprocessors or programmable elements or systems specific to the field of study, making assessment and critical analysis of the prepared software as well as a synthesis and creative interpretation of information presented with it	Student is able to analyze data and use different tools and methods for that purpose.	[SU4] Assessment of ability to use methods and tools [SU1] Assessment of task fulfilment

Subject contents	1. Basic programming in Python and libraries: pandas and numpy. 2. Data preparation and visualization. 3. Prediction and model evaluation. 4. Decision trees and random forest. 5. Cluster analysis. 6. Feature selection and extraction.
Prerequisites and co-requisites	
Assessment methods and criteria	

Subject passing criteria	Passing threshold	Percentage of the final grade
laboratory	51.0%	100.0%

Recommended reading	Basic literature	1. Allen Downey „Myśl w języku Python!: nauka programowania”, Helion, Gliwice, 2017. 2. Foster Provost, Tom Fawcett „Analiza danych w biznesie”, Helion, Gliwice, 2015. 3. Alberto Boschetti, Luca Massaron “Python: podstawy nauki o danych”, Helion, Gliwice, 2017.
	Supplementary literature	None
	eResources addresses	Adresy na platformie eNauczanie:
Example issues/ example questions/ tasks being completed		
Work placement	Not applicable	