



Subject card

Subject name and code	Data Warehousing, PG_00044140						
Field of study	Mathematics						
Date of commencement of studies	October 2023	Academic year of realisation of subject			2024/2025		
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	2	Language of instruction			Polish		
Semester of study	3	ECTS credits			4.0		
Learning profile	general academic profile	Assessment form			assessment		
Conducting unit	Division of Differential Equations and Applications of Mathematics -> Institute of Applied Mathematics -> Faculty of Applied Physics and Mathematics						
Name and surname of lecturer (lecturers)	Subject supervisor	dr hab. Paweł Pilarczyk					
	Teachers	dr hab. Paweł Pilarczyk mgr inż. Michał Krzemiński					
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	30.0	0.0	30.0	0.0	0.0	60
	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	60		5.0		35.0	100
Subject objectives	Theoretical and practical introduction to data mining and theoretical foundations of data warehousing.						

Learning outcomes	Course outcome	Subject outcome	Method of verification
	[K7_U13] Understands the mathematical foundations of the analysis of algorithms and computational processes, can construct algorithms with good numerical properties, used to solve typical and unusual mathematical problems.	The student knows mathematical foundations of the data mining algorithms that he or she is familiar with, including machine learning, and is capable of applying these algorithms to specific problems. The student knows mathematical foundations of data warehousing.	[SU1] Assessment of task fulfilment [SU4] Assessment of ability to use methods and tools
	[K7_W08] Knows advanced computation techniques, supporting the work of a mathematician and understand their limitations.	The student knows mathematical methods and software tools used in data mining.	[SW1] Assessment of factual knowledge
	[K7_W10] Knows the numerical methods used to find approximate solutions to mathematical problems (e.g. differential equations) posed by applied fields (e.g. industrial technologies, management, etc.).	The student can choose statistical methods and machine learning algorithms, as well as software environment appropriate for modeling real world problems.	[SW3] Assessment of knowledge contained in written work and projects
	[K7_U08] Knows probability distributions and their properties; is able to use them in practical issues, is familiar with the basics of statistics (estimation issues and hypothesis testing) and the basics of statistical data processing.	The student can conduct basic statistical analysis of numerical quantitative and qualitative data.	[SU2] Assessment of ability to analyse information [SU3] Assessment of ability to use knowledge gained from the subject
[K7_K02] Can precisely formulate questions to deepen own understanding of a given topic or find missing elements of reasoning, understands the need to clearly present selected achievements of higher mathematics to laymen.	The student can plan a data mining project, and then discuss its outcome.	[SK4] Assessment of communication skills, including language correctness	
Subject contents	Lecture: introduction to data mining and knowledge discovery in data. Data preprocessing and exploratory data analysis, Cross-Industry Standard Process for Data Mining (CRISP-DM). Statistical data analysis and machine learning. Methods for classification and data clustering, discovering association rules. Data warehousing, multidimensional modeling, OLAP. Laboratory: practical data mining and data exploration using Python and R.		
Prerequisites and co-requisites	Basic ability to write programs in R and in Python. Familiarity with basic statistical methods.		
Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	Group assignments and projects (laboratory)	60.0%	50.0%
	Quizzes (lecture)	60.0%	50.0%
Recommended reading	Basic literature Jacek Rumiński. Wprowadzenie do hurtowni i eksploracji danych. Gdańsk, Wydawnictwo Politechniki Gdańskiej, 2015. Daniel T. Larose. Data Mining. Metody i modele eksploracji danych. Warszawa, Wydawnictwo Naukowe PWN, 2012. The original English edition: Daniel T. Larose. Data Mining Methods and Models. Wiley-IEEE Press; 1st Ed., 2006. Andres Fortino. Data Mining and Predictive Analytics for Business Decisions. A Case Study Approach. Mercury Learning & Information, 2023.		

	Supplementary literature	Daniel T. Larose, Chantal D. Larose. Discovering Knowledge in Data. An Introduction to Data Mining, 2nd Ed., 2014. Jiawei Han, Micheline Kamber, Jian Pei. Data Mining. Concepts and Techniques. 3rd Ed. Elsevier, 2011.
	eResources addresses	Adresy na platformie eNauczenie: Hurtownie danych 2024 - Moodle ID: 37137 https://enauczenie.pg.edu.pl/moodle/course/view.php?id=37137
Example issues/ example questions/ tasks being completed	<p>Methods for data preprocessing.</p> <p>What is the difference between a data warehouse and an operational database?</p> <p>Methods for supervised data classification.</p> <p>Data clustering using the DBSCAN method.</p> <p>Kohonen networks and their relation to neural networks.</p>	
Work placement	Not applicable	

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