GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Data Warehousing, PG_00044140		
Field of study	Mathematics		
Date of commencement of studies	October 2024	Academic year of realisation of subject	2025/2026
Education level	second-cycle studies	Subject group	
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	Divison of Differential Equations and Applications of Mathematics -> Institute of Applied Mathematics -> Faculty of Applied Physics and Mathematics -> Wydziały Politechniki Gdańskiej		
Name and surname of lecturer (lecturers)	Subject supervisor	dr hab. Paweł Pilarczyk	
	Teachers	mgr inż. Michał Krzemiński dr hab. Paweł Pilarczyk	

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

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
Quizzes (lecture)	60.0%	50.0%
Group assignments and projects (laboratory)	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	
Example issues/ example questions/ tasks being completed	Methods for data preprocessing. What is the difference between a data warehouse and an operational database? Methods for supervised data classification. Data clustering using the DBSCAN method. Kohonen networks and their relation to neural networks.	
Work placement	Not applicable	

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