



Subject card

Subject name and code	Data mining, PG_00045365						
Field of study	Data Engineering						
Date of commencement of studies	October 2021	Academic year of realisation of subject				2022/2023	
Education level	first-cycle studies	Subject group				Obligatory subject group in the field of study 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				English	
Semester of study	4	ECTS credits				3.0	
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 inż. Wojciech Waloszek				
	Teachers		dr inż. Wojciech Waloszek				
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	15.0	0.0	0.0	30
	E-learning hours included: 0.0						
Data Mining 2022/2023 - Moodle ID: 29556 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=29556							
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		5.0		40.0	75
Subject objectives	The goal of the course is to familiarize students with methods of data mining and to present a basic wireframe for data mining process. The main task of data mining, like prediction, classification and segmentation, are discussed and algorithms used for the tasks are presented. CRISP-DM is shown as an example of a data mining process.						
Learning outcomes	Course outcome	Subject outcome			Method of verification		
	[K6_U03] analyses problems and creates appropriate models, data structures and algorithms (including heuristic and numerical ones), assesses their computational complexity, estimates errors of the received solutions	A student designs the data mining process, runs it, and evaluates the results			[SU1] Assessment of task fulfilment [SU2] Assessment of ability to analyse information		
	[K6_W08] Knows the models and structure of the data mining process and their multidimensional analysis and can assess the results of such analyses	A Student enlists basic tasks of data mining process. A Student demonstrates data mining models. A Student assesses created data mining models.			[SW1] Assessment of factual knowledge		
[K6_U06] Independently solves complex engineering tasks using literature, materials and devices, prepares extensive documentation of the developed solution using appropriate description techniques.	Student enlists and presents algorithms and data structures used in the data mining process.			[SU3] Assessment of ability to use knowledge gained from the subject			
Subject contents	1. Data Mining Basics 2. Data Mining Process and Its Role within an Organisation 3. Representations of Data and Knowledge 4. Review of Basic Classifiers 5. Decision Trees 6. Classification Rules 7. Association Rules 8. Clustering 9. Examples of Numerical Prediction in Data Mining 10. Sources of Bias and Errors in the Data Mining Process 11. Engineering Input and Output 12. Other Data Mining Techniques						
Prerequisites and co-requisites	No requirements						
Assessment methods and criteria	Subject passing criteria		Passing threshold		Percentage of the final grade		
	Practical exercise		50.0%		50.0%		
	written test		50.0%		50.0%		

Recommended reading	Basic literature	Basic literature: Ian H. Witten, Eibe Frank: "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann 2005. David J. Hand, Padhraic Smyth, Heikki Mannila: "Principles of Data Mining", The MIT Press 2001. Daniel T. Larose: "Metody i modele eksploracji danych", PWN 2008 (In Polish).
	Supplementary literature	No recommendations
	eResources addresses	
Example issues/ example questions/ tasks being completed	Sample issues: Exploratory Data Analysis. Generating decision trees. Segmentation of data sets.	
Work placement	Not applicable	