

。 GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Data collection and analysis, PG_00069090								
Field of study	Technical Physics, Materials Engineering, Mathematics, Nanotechnology, Nanotechnology								
Date of commencement of studies	October 2024		Academic year of realisation of subject			2025/2026			
Education level	second-cycle studies		Subject group			Optional 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			1.0			
Learning profile	general academic profile		Assessment form			assessment			
Conducting unit	Division of Theoretical Physics and Quantum Informaton -> Institute of Physics and Applied Computer Science -> Faculty of Applied Physics and Mathematics -> Wydziały Politechniki Gdańskiej						omputer		
Name and surname	Subject supervisor	ubject supervisor dr inż. Bartosz Reichel			el				
of lecturer (lecturers)	Teachers		dr inż. Bartosz Reichel						
Lesson types and methods	Lesson type	Lecture	Tutorial	Laboratory	Project	t	Seminar	SUM	
of instruction	Number of study hours	15.0	0.0	0.0	0.0		0.0	15	
	E-learning hours inclu	ided: 0.0		I					
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	15		2.0		8.0		25	
Subject objectives	 Ine subject aims to raminarize the audience with the elements of data collection (especially concerning physical - hardware) as well as with a very general introduction to the possibilities of their analysis 1. Collecting video data 2. Serial interfaces: UART, USB, Ethernet, 3. Mapped files, blocking. Advanced file access. 4. SQL/NoSQL databases, saving optimization 5. Web pages, collecting data on user behavior 6. Collecting data on user location, searching in GIS data 7. Stream processing (e.g. Apache Kafka) 8. MQTT as a data collection control protocol for IoT, MQTT brokers 9. Methods of management of analytical projects. Data analysis as a process. 10. Methods of preparing data for analysis in the context of data warehouses. 11. Classical algorithms: cluster analysis, feature detection, SURF, SIFT 12. Algorithms on the border of machine learning: Classification methods. Classification and regression trees Methods of assessing the effectiveness of models Machine learning methods, types of layers in neural networks 								
Learning outcomes	earning outcomes Course outcome		Subj	Subject outcome		Method of verification			
	[K7_K01] knows limitations of own knowledge, understands the need to learn and improve professional and personal competencies		Can evaluate the result obtained by himself (and the supporting tools).			USK2J Assessment of progress of work			
	[K7_W04] has enhan knowledge of mather numerical and simula applied in the descrip modelling of physical	i enhancedIs able to implement selected[SW1] Assessmermathematical,algorithms and/or use ready-madeknowledgel simulation methodsones (understands their operation,is able to parameterize themubysical phenomenaappropriately)		Assessment o	of factual				

Subject contents	The main goal of the course is to introduce students to fundamentals of data mining and data analysis. The course will provide examples of real-world problems in the fields of business, medicine and IT.							
	They will be presented to the processes of collecting and recording data encountered in practice-row data collection video, audio, data transferred via serial interfaces, data acquisition cards.							
	 Video data collection Serial interfaces: UART, USB, Ethernet Mapped files, blocking Database SQL / NoSQL, storage optimization Collecting data about web user behavior Collecting data about the user's location 							
	7. Streaming (eg Apache Kafka)							
	8. MQTT as a data collection control protocol for IoT, MQTT Brokers							
	Emphasis will be placed on how to solve real problems during analytical process. Scope of the course includes:							
	 Data analysis as a process. Data mining methodologies. Data preparation. Classification techniques: Naïve Bayes classifie, k-NN Regression methods Classification and regression trees. Optimalization. 							
	6. Assessing and comparing models performance							
	7. Machine learning methods, types of layers in neural networks.							
Prerequisites and co-requisites								
Assessment methods	Subject passing criteria	Passing threshold	Percentage of the final grade					
and criteria	Egzam	100.0%	100.0%					
Recommended reading	Basic literature	1. D. Hand, H. Mannila, P. Smyth, Principles of data mining, MIT Press, 2001 2. F. Provost, T. Fawcett, Data science for business, OReilly Media, 2013						
	Supplementary literature J.Axelson, Serial Port Complete: COM Ports, USB Virtual COM Ports, and Ports for Embedded Systems (Complete Guides series), 2007, Lakeview Research							
	eResources addresses							
Example issues/	Implementation of the circular buffer.							
example questions/ tasks being completed	Describe the process for data warehouse ETL?							
Work placement	Not applicable							

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