

§ GDAŃSK UNIVERSITY § OF TECHNOLOGY

Subject card

Subject name and code	Business Data Processing, PG_00048279							
Field of study	Informatics							
Date of commencement of studies	February 2023		Academic year of realisation of subject		2023/2024			
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	2		ECTS credits		4.0			
Learning profile	general academic profile		Assessment form		exam			
Conducting unit	Department of Software Engineering -> Faculty of Electronics, Telecommunications and Informatics							
Name and surname	Subject supervisor	dr inż. Aleksandra Karpus						
of lecturer (lecturers)	Teachers		dr inż. Aleksandra Karpus					
			dr inż. Michał Wróbel					
			dr inż. Wojciech Waloszek					
			dr inż. Grzegorz Gołaszewski					
			or Inz. Teresa ∠awadzka					
			dr Paweł Weichbroth					
		dr hab. inż. Agnieszka Landowska						
Lesson types and methods	Lesson type	Lecture	Tutorial	Laboratory	Projec	t	Seminar	SUM
of instruction	Number of study hours	15.0	0.0	30.0	0.0		0.0	45
	E-learning hours inclu	uded: 0.0		-				
Learning activity and number of study hours	Learning activity	Participation i classes incluc plan	n didactic led in study	Participation in consultation hours		Self-study		SUM
	Number of study hours	45		8.0		47.0		100
Subject objectives	The aim of the course is to introduce modern methods of data processing, taking into account various goals of data processing, and various characteristics of stored data.							

Learning outcomes	Course outcome	Subject outcome	Method of verification			
	[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 can do the data analysis. He applies different tools and techniques for that purpose.	[SU1] Assessment of task fulfilment			
	[K7_W42] Knows and understands, to an increased extent, the principles and trends in the analysis and design of local and distributed IT systems and the basics of computer modeling and computerization of complex cognitive and decision-making processes.	Student knows how to build data mining systems.	[SW1] Assessment of factual knowledge			
	[K7_W04] Knows and understands, to an advanced extent, the principles, methods and techniques of programming and the principles of computer software development or programming devices or controllers using microprocessors or programmable elements or systems specific to the field of study, and organisation of systems using computers or such devices	Student knows different data analysis techniques and tools.	[SW1] Assessment of factual knowledge			
	[K7_W41] Knows and understands, to an increased extent, the standards, production methods, life cycle and development trends of software as well as information systems and applications.	Student knows CRISP methodology.	[SW1] Assessment of factual knowledge			
	[K7_U07] can apply advanced methods of process and function support, specific to the field of study	Student can choose data mining models and evaluate them.	[SU1] Assessment of task fulfilment [SU4] Assessment of ability to use methods and tools			
Subject contents	1. Data quality aspects.					
	2. Emotion recognition in Informatics.					
3. R language in data mining.						
	5. Recommender systems.					
	6. Deep Learning.					
Prerequisites and co-requisites Basic knowledge about relational databases.						
	Basic knowledge about methods and algorithms of data mining.					

Assessment methods	Subject passing criteria	Passing threshold	Percentage of the final grade			
and criteria	Tasks during the second workshop	50.0%	12.0%			
	Tasks during the first workshop	50.0%	12.0%			
	Tasks during the fifth workshop	50.0%	12.0%			
	Tasks during the forth workshop	50.0%	12.0%			
	Egzam	50.0%	28.0%			
	Tasks during the third workshop	50.0%	12.0%			
	Tasks during the sixth workshop	50.0%	12.0%			
Recommended reading	Basic literature	A. Maydanchik, Data Quality Assessment, Technics Publication, 2007				
		D. McGilvray, Executing Data Quality Projects, Morgan Kaufman, 2008				
		Webster J.J.: Tokenization as the initial phase in NLP, 15th conference on Computational linguistics, COLING, vol. 4, Association for Computational Linguistics Stroudsburg, 1992, s.1106-1110 Xu J., Croft W.B.: Corpus-Based Stemming Using Cooccurrence of Word Variants, ACM Transactions on Information Systems, Vol. 16, f 1, 1998. s. 61-81				
		Rajaraman A., Ullman J.D.: Mining of Massive Datasets, Cambridge University Press, New York 2012				
		Ramos J.: Using TF-IDF to Determine Word Relevance in Docu Queries, In Proceedings of the First instructional Conference or Machine Learning iCML-03, 3-8 December 2003, Piscataway, U				
		D. Mendrala, M. Szeliga: SQL 2008. Usługi biznesowe. Analiza i eksploracja danych. Helion 2009.				
		Avril Coghlan, A Little Book of R For Time Series, Release 0.2, 2016 https://media.readthedocs.org/pdf/a-little-book-of-r-for-time-series/la alittle-book-of-r-for-time-series.pdf Robert Nau, Principles and risks of forecasting, Fuqua School of Business, Duke University, September 2014, https://people.duke.ed ~rnau/Principles_and_risks_of_forecasting Robert_Nau.pdf				
		Vito Ricci, R functions for time series analysis by R.0.5 26/11/04, https:// cran.rproject.org/doc/contrib/Ricci-refcard-ts.pdf				
	Supplementary literature	T.C. Redman, Data Driven: Profiting from Your Most Important Business Asset, Harvard Business Review Press, 2008				
		Ingesoll G.S., Morton T.S., Farris A. organize and manipulate it, Manning	L.: Taming Text How to find, g, Shelter Island, 2013			
		Walter Zucchini, Oleg Nenadic, Time http://www.statoek.wiso.unigoettinge sommer03/ts_r_intro.pdf	e Series Analysis with R - Part I, en.de/veranstaltungen/zeitreihen/			
	eResources addresses	Adresy na platformie eNauczanie: Przetwarzanie Danych w Biznesie 2 https://enauczanie.pg.edu.pl/moodle	2023/24 - Moodle ID: 32689 e/course/view.php?id=32689			
Example issues/ example questions/ tasks being completed	1. Analyze the sample time series2. Asses quality of data.3. Data mining using R language.					
Work placement	Not applicable					