GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Business Data Processing, PG_00064475		
Field of study	Informatics		
Date of commencement of studies	February 2025	Academic year of realisation of subject	2024/2025
Education level	second-cycle studies	Subject group	Optional subject group Specialty 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	1	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. Aleksandra Karpus	
	Teachers	dr inż. Aleksandra Karpus dr inż. Michał Wróbel dr inż. Wojciech Waloszek dr inż. Grzegorz Gołaszewski dr inż. Teresa Zawadzka dr Paweł Weichbroth dr hab. inż. Agnieszka Landowska	

Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	0.0	0.0	30.0	0.0	0.0	30
	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	30	6.0	39.0	75

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_W101] is able to make an in-depth identification of key objects and phenomena related to the field of study, as well as theories that describe them and applicable analytical and design methods	The student has knowledge of broadly understood data analysis including time series analysis and social network analysis. The student understands the role of known methods in application of anomaly detection as well as in the process of items recommendation.	[SW1] Assessment of factual knowledge [SW3] Assessment of knowledge contained in written work and projects
[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_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.	[SU4] Assessment of ability to use methods and tools [SU1] Assessment of task fulfilment
[K7_W04] knows and understands, to an increased extent, the principles, methods and techniques of programming and the principles of computer software development or programming devices or controllers using microprocessors or other elements or programmable devices specific to the field of study, and organization of work of systems using computers or such devices	Student knows different data analysis techniques and tools.	[SW1] Assessment of factual knowledge

Subject contents

1. Data quality aspects.

2. Emotion recognition in Informatics.

3. R language in data mining.

4. Time series in data mining.

5. Recommender systems.

6. Deep Learning.

7. Social Network Analysis.

Prerequisites and co-requisites

Basic knowledge about relational databases.

Basic knowledge about methods and algorithms of data mining.

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Tasks during the forth workshop	50.0%	16.66%
Tasks during the third workshop	50.0%	16.66%
Tasks during the sixth workshop	50.0%	16.66%
Tasks during the second workshop	50.0%	16.66%
Tasks during the first workshop	50.0%	16.7%
Tasks during the fifth workshop	50.0%	16.66%

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, Nr 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 Document Queries, In Proceedings of the First instructional Conference on Machine Learning iCML-03, 3-8 December 2003, Piscataway, USA 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/latest/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.edu/~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.L.: Taming Text How to find, organize and manipulate it, Manning, Shelter Island, 2013 Walter Zucchini, Oleg Nenadic, Time Series Analysis with R - Part I, http://www.statoek.wiso.unigoettingen.de/veranstaltungen/zeitreihen/sommer03/ts_r_intro.pdf
	eResources addresses	Adresy na platformie eNauczanie: Przetwarzanie danych w biznesie - lato 2024/25 - Moodle ID: 43091 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=43091
Example issues/ example questions/ tasks being completed	1. Analyze the sample time series 2. Asses quality of data. 3. Data mining using R language.	
Work placement	Not applicable	