GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	TEXT MINING METHODS, PG_00067054
Field of study	Economic Analytics

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 Subject group related to scientific research in the field of study
Mode of study	Part-time studies (on-line)	Mode of delivery	at the university
Year of study	2	Language of instruction	Polish
Semester of study	4	ECTS credits	3.0
Learning profile	general academic profile	Assessment form	exam

Conducting unit	Department of Statistics and Econometrics -> Faculty of Management and Economics

Name and surname of lecturer (lecturers)	Subject supervisor	dr hab. Michał Pietrzak
	Teachers	

Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	8.0	0.0	16.0	0.0	0.0	24
	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	24	4.0	47.0	75

Subject objectives	Assesses the situation of economic phenomena, creating scenarios based on Text Mining research

Learning outcomes	Course outcome	Subject outcome	Method of verification
	[K7_W06] Knows and understands the principles of evaluating the reliability of utilized data, applying in-depth specialized knowledge in the field of economic analysis	analyzes economic phenomena on the basis of textual sources, selecting appropriate methods and techniques	[SW1] Assessment of factual knowledge
	[K7_U04] Prepares and delivers convincing presentations of the results of specialized analyses, providing in-depth interpretations during debates and meetings with diverse audiences.	critically evaluate the results of analyzes by presenting persuasive presentations along with their in-depth interpretation	[SU5] Assessment of ability to present the results of task

Subject contents	Introduction to Text Mining. Economic phenomena analyzed on the basis of textual sources - characteristics, empirical examples Text mining in business context. The process of obtaining text data by researchers, institutions, enterprises, official statistics services and their publication in databases The problem of accessibility and distribution of text data due to various criteria Scope and types of text data research, classification criteria, examples of research Methods and techniques used in Text Mining research Text categorization Topic Modelling Using Web Crawling in Text Mining Research Fundementals of Machine Learning Algorithms for Text Mining Document Summarization methods Sentiment Analysis Name Entity Recognition (NER) technique The problem of presenting text data and the results of Text Mining analysis The problem of assessing the situation of economic phenomena and creating scenarios based on Text Mining research
Prerequisites and co-requisites	

Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	Exam	60.0%	70.0%
	Test	60.0%	30.0%

Recommended reading	Basic literature	Jo, T. (2019). Text mining. Cham, Switzerland: Springer International Publishing Hofmann, M., & Chisholm, A. (Eds.). (2016). Text mining and visualization: Case studies using open-source tools. CRC Press Miner, G. (2012). Practical text mining and statistical analysis for non-structured text data applications. Academic Press
	Supplementary literature	Silge, J., & Robinson, D. (2016). tidytext: Text mining and analysis using tidy data principles in R. Journal of Open Source Software, 1(3), 37 Silge, J., & Robinson, D. (2017). Text mining with R: A tidy approach. " O'Reilly Media, Inc." Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. (2014). Automated data collection with R: A practical guide to web scraping and text mining. John Wiley & Sons Aydin, O. (2018). R Web Scraping Quick Start Guide: Techniques and tools to crawl and scrape data from websites. Packt Publishing Ltd.
	eResources addresses	Adresy na platformie eNauczanie:
Example issues/ example questions/ tasks being completed		
Work placement	Not applicable	

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