GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	BIG DATA, PG_00067653		
Field of study	Economic Analytics		
Date of commencement of studies	October 2025	Academic year of realisation of subject	2026/2027
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	2	Language of instruction	Polish
Semester of study	3	ECTS credits	4.0
Learning profile	general academic profile	Assessment form	assessment
Conducting unit	Department Of Statistics And Econometrics -> Faculty Of Management And Economics -> Wydziały Politechniki Gdańskiej		
Name and surname of lecturer (lecturers)	Subject supervisor		
	Teachers		

Lesson types and methods of instruction

Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
Number of study hours	15.0	0.0	45.0	0.0	0.0	60

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	60	5.0	35.0	100

Subject objectives

Demonstrates in-depth comprehensive preparation for the analysis of large data sets

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K7_U01] creates innovative solutions for complex and unstructured processes, considering unpredictable environmental conditions through the synthesis of information from various sources.	Formulates innovative solutions to complex problems based on the analysis of large and unstructured datasets by integrating information from various sources and applying advanced machine learning algorithms under conditions of uncertainty.	[SU3] Assessment of ability to use knowledge gained from the subject
[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.	Knows advanced technologies to process large datasets, assessing their reliability and quality based on in-depth knowledge of economic analytics in order to prepare data for solving complex decision-making problems.	[SW1] Assessment of factual knowledge

Subject contents

Overview of Big Data. Types of Digital Data, Introduction to Big Data
Big data programming tools (e.g., Hadoop, MongoDB, Spark, etc.). Using Spark with R
Big data extraction and integration
Big data storage; Technologies for Handling Big Data Introduction to Hadoop HDFS (Hadoop Distributed File System)
Dig Deep to understand the fundamental of MapReduce and HBase
Hadoop MapReduce in R; Integrating Hadoop and R RHIPE; RHadoop
Data Analytics with R and Hadoop data preprocessing, visualising data
Big Data Analysis and Machine Learning supervised and unsupervised ML algorithms. Spark Machine Learning with R
Importing and exporting data from various DBs (RMySQL, RSQLite, RHive, RHBase). Using SparkSQL with R
Big Data Analytics with BigR Deep learning algorithms with R & H2O

Prerequisites and co-requisites

Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	Exam	60.0%	60.0%
	Test	60.0%	40.0%
Recommended reading	Basic literature	Hamstra, M., & Zaharia, M. (2013). Learning Spark: lightning-fast big data analytics. O'Reilly & Associates Densmore, J. (2021). Data pipelines pocket reference. O'Reilly Media Drabas, T., & Lee, D. (2017). Learning PySpark. Packt Publishing Ltd Haines, S. (2022). Modern Data Engineering with Apache Spark: A Hands-on Guide for Building Mission-critical Streaming Applications. Apress	
	Supplementary literature	Warren, J., & Marz, N. (2015). Big Data: Principles and best practices of scalable realtime data systems. Simon and Schuster Ilijason, R. (2020). Beginning Apache Spark Using Azure Databricks: Unleashing Large Cluster Analytics in the Cloud. Apress	
	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.