GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Mathematic fundamentals of data science, PG_00062725		
Field of study	Technologies for Industry 5.0		
Date of commencement of studies	October 2024	Academic year of realisation of subject	2025/2026
Education level	first-cycle studies	Subject group	Obligatory subject group 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	5.0
Learning profile	general academic profile	Assessment form	exam
Conducting unit			
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	30.0	15.0	0.0	0.0	0.0	45

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	45	5.0	75.0	125

Subject objectives

The aim of the course is to familiarize students with the mathematical foundations and methods used in data engineering.

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K6_K01] is aware of the need to constantly update and enrich knowledge and practical skills, and improve professional, personal and social competences	The student is aware of the need to constantly update knowledge in the rapidly developing field of data engineering. The student is engaged in independent searches for new information, participation in courses and trainings, and following the latest trends and achievements in the field of data analysis and machine learning.	[SK2] Assessment of progress of work
[K6_W01] demonstrates knowledge and understanding of mathematics, physics, chemistry and IT tools at the level necessary to formulate and solve typical engineering and technological problems	The student demonstrates knowledge of basic mathematical methods used in data analysis and exploration, such as multiple regression, correlation analysis, and normalization and standardization methods.	[SW1] Assessment of factual knowledge
[K6_U01] applies knowledge of mathematics, physics, chemistry, IT tools and other engineering disciplines to solve theoretical, engineering and technological problems	The student is able to use computer science and programming tools to implement algorithms such as the nearest neighbor algorithm, decision trees, neural networks, k-means algorithm and a priori algorithm.	[SU1] Assessment of task fulfilment

<table>
<tr><td>Subject contents</td><td>1. Basic tasks used in data mining: problem identification and definition; data preparation: data collection, cleaning, processing; preliminary data analysis.
2. Data processing: normalization; standardization; discretization of continuous variables.
3. Qualitative and numerical variables: types of variables; conversion of qualitative variables into numerical ones.
4. Multivariate relations: covariance analysis; use of correlation matrices.
5. Statistical inference: statistical hypotheses; statistical tests; confidence intervals.
6. Multiple regression: linear models; parameter estimation; model evaluation and interpretation
7. Nearest neighbor algorithm: classification methods, selection of the number of neighbors.
8. Decision functions and classification trees: construction of decision trees; pruning and validation
9. Comparison of different algorithms: algorithm evaluation metrics; cross-validation.
10. Use of neural networks: basics of neural networks; sigmoid activation function; backpropagation rules.
11. Hierarchical clustering methods: agglomerative and partitioning algorithms; distance measures.
12. K-means algorithm: data clustering; choosing the number of clusters.
13. A priori algorithm and association rules: detecting patterns in data, generating association rules.</td></tr>
<tr><td>Prerequisites and co-requisites</td><td>Knowledge of basic mathematics (linear algebra, calculus)

Knowledge of basic mathematics (linear algebra, calculus).</td></tr>
<tr><td>Assessment methods and criteria</td><td><table>
<tr><th>Subject passing criteria</th><th>Passing threshold</th><th>Percentage of the final grade</th></tr>
<tr><td>exam</td><td>50.0%</td><td>50.0%</td></tr>
<tr><td>course test</td><td>50.0%</td><td>50.0%</td></tr>
</table></td></tr>
<tr><td>Recommended reading</td><td><table>
<tr><td>Basic literature</td><td>[1] Daniel T. Larose, Okrywanie wiedzy z danych, PWN 2006

[2] Daniel T. Larose, Metody i modele eksploracji danych, PWN 2016</td></tr>
<tr><td>Supplementary literature</td><td>[3] Tadeusz Morzy, Eksploracja danych. Metody i algorytmy, PWN 2013</td></tr>
<tr><td>eResources addresses</td><td>Adresy na platformie eNauczanie:</td></tr>
</table></td></tr>
<tr><td>Example issues/ example questions/ tasks being completed</td><td></td></tr>
<tr><td>Work placement</td><td>Not applicable</td></tr>
</table>

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