GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	, PG_00066700		
Field of study	Technical Physics		
Date of commencement of studies	February 2025	Academic year of realisation of subject	2024/2025
Education level	second-cycle studies	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	2.0
Learning profile	general academic profile	Assessment form	assessment
Conducting unit	Institute of Physics and Applied Computer Science -> Faculty of Applied Physics and Mathematics		
Name and surname of lecturer (lecturers)	Subject supervisor	dr inż. Michał Piłat	
	Teachers	dr inż. Michał Piłat	

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	2.0	18.0	50

Subject objectives

The aim of the course is to familiarize students with the basics of descriptive statistics with particular emphasis on techniques used in data science. The course will present the mathematical foundations of statistics and methods of implementing them in the Python programming language within built-in libraries such as Numpy, Scipy, Sympy, Pandas and Matplotlib. The effects of education are to be knowledgeable of basic statistical concepts, the ability to use statistical tools for analytical purposes and methods of implementing them in programming systems.

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K7_U06] can apply obtained knowledge of physics to exact sciences, natural and technical sciences	Student can perform the analysis of data describing, for example, medical results, population of a chosen region.	[SU4] Assessment of ability to use methods and tools
[K7_U05] can plan and conduct theoretical calculations, experimental research and computer simulations, critically analyze their results, draw conclusions and form reasoned opinions	Student can describe a given set of data using statistical methods and perform an analysis of data by using special libraries implementedin Python.	[SU3] Assessment of ability to use knowledge gained from the subject
[K7_W03] has knowledge of current development paths and discoveries in the scope of physics and related fields of science and technology	Student has knowledge on basic concepts of statistics and programming technics in Python.	[SW1] Assessment of factual knowledge

Subject contents

Exploratory data analysis

- Estimates of location
- Estimates of variability
- Exploring the data distribution
- Exploring binary and categorical data
- Correlation
- Exploring two or more variables

Data and sampling distribution

- Differences between sample and population
- Types of discrete distibutions
- Types of continuous distributions
- Confidence intervals

Statistical experiments and significance testing

- Hypothesis tests
- Statistical significance
- Errors of I and II kind
- Chi-square test

Regression and prediction

- Linear regression
- Polynomial regression
- Multiple linear regression
- Regression diagnostics

Classification

- Naive Bayes
- Discriminant analysis
- Logistic regression
- Evaluating Classification Models

Statistical machine learning

- K-nearest neighbours
- Tree models

Unsupervised Learning

- Principal Components Analysis
- K-means clustering

Prerequisites and co-requisites

Basics of programming in Python

Basics of calculus and linear algebra.

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Project	50.0%	50.0%
Test	50.0%	50.0%

Recommended reading

Basic literature	Peter Bruce, Andrew Bruce, Peter Gedeck, "Practical statistics for Data Scientists. 50+ Essential Concepts Using R and Python ", O'Reilly, Boston 2020 Robert Johansson "Numerical Python. Scientifing Computing and Data Science Applications with Numpy, SciPy i Matplotlib" Apress, 2019
Supplementary literature	Aurelien Geron "Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow", O'Reilly, Boston, 2019
eResources addresses	Podstawowe https://enauczanie.pg.edu.pl/moodle/course/view.php?id=44865 - Course on eNauczanie Adresy na platformie eNauczanie: Statystyka praktyczna w data science - Moodle ID: 44865 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=44865

Example issues/ example questions/ tasks being completed	Describe 3 discrete and continuous probability distributions. Describe k-neighbours method Describe chi-square test. Analyse a given set of data.
Work placement	Not applicable

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