



Subject card

Subject name and code	Statistics II, PG_00021508						
Field of study	Mathematics						
Date of commencement of studies	October 2024	Academic year of realisation of subject			2024/2025		
Education level	second-cycle studies	Subject group					
Mode of study	Full-time studies	Mode of delivery			at the university		
Year of study	1	Language of instruction			Polish		
Semester of study	2	ECTS credits			5.0		
Learning profile	general academic profile	Assessment form			exam		
Conducting unit	Department of Nonlinear Analysis and Statistics -> Faculty of Applied Physics and Mathematics						
Name and surname of lecturer (lecturers)	Subject supervisor		dr hab. Karol Dziejduł				
	Teachers		dr hab. Karol Dziejduł				
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	30.0	0.0	15.0	0.0	15.0	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		60.0	125
Subject objectives	The lecture shows the role of the Vapnik Chervonenkis dimension for the correct estimation of regression functions using decision trees, i.e. what makes us have the consistency theorem. Formulas of information criteria are derived: AKAIKE criterion, Fisher criterion. We show the role of the Rao-Cramer inequality in the context of the limits of knowledge, i.e. the role of uncertainty. The laboratory has a completely applied nature. GLM models, regression models using decision trees, neural networks. Random forest concept.						
Learning outcomes	Course outcome		Subject outcome		Method of verification		
Subject contents	1. Regression estimator, Cantelli Gliwienko's theorems and regression consistency theorems 2. Matrix factorization 3. Combinatorics according to Vapnik and Czervonenkis. Vapnik Chervonenkis class generation 4. Theorem Rao-Blackwell the Hodges-Lehman theorem. 5. Kulback Leibler's information, AKAIKE information criteria. 6. Cramer Rao's inequality. 7. The role of information 8. Laboratory - GLM models, regression models using decision trees, neural networks. Random forest concept						
Prerequisites and co-requisites	Mathematical Statistics and Statistics with SAS, Probabilistics						
Assessment methods and criteria	Subject passing criteria		Passing threshold		Percentage of the final grade		
	Half the exercises half an oral examination		60.0%		100.0%		

Recommended reading	Basic literature	J. Bartoszewicz Wykłady ze Statystyki matematycznej PWN Warszawa 1989 Sadanori Konishi, Genshiro Kitagawa: "Information Criteria and Statistical Modeling" Springer Series in Statistics 2008 Leo Breiman, Jerome Friedman, Charles J. Stone, R.A. Olshen Classification and Regression Trees Taylor & Francis, 1984
	Supplementary literature	R. Zieliński Siedem wykładów wprowadzających do statystyki matematycznej PWN Warszawa 1990 E.L. Lehmann, G. Casella, Theory of Point Estimation Springer Texts in Statistics 2nd Edition 1998 19
	eResources addresses	Podstawowe https://drive.pg.edu.pl/s/7XJfAXz1lbdCVW - The entire lecture can be found in https://drive.pg.edu.pl/s/7XJfAXz1lbdCVW files lecture2.pdf and statisticslecture.pdf Uzupełniająca Adresy na platformie eNauczanie:
Example issues/ example questions/ tasks being completed	1. AKAIKE 2. VC dimension, examples, complexity measures (theorem 1.6) 3. estimator for decision trees. Its design. 4. Cantella Gliwienko's theorem and generalizations from the perspective of the VC dimension. 5. PCA or matrix decomposition theorems (applications of PCA) 6. State space, statistical space, example	
Work placement	Not applicable	

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