

。 GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Statistics II, PG_00021508							
Field of study	Mathematics							
Date of commencement of studies	October 2023		Academic year of realisation of subject		2023/2024			
Education level	second-cycle studies		Subject group		Optional 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	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 -> Wydziały Politechniki Gdańskiej							
Name and surname of lecturer (lecturers)	Subject supervisor		dr hab. Karol Dziedziul					
	Teachers		Paweł Wieczyński dr hab. Karol Dziedziul					
Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Projec	t	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				
	[K7_U10] In a selected field, can examine evidence, in which, if necessary, also can use tools from other branches of mathematics, can identify one's own interests and develop them; in particular, is able to establish contact with specialists in their field, e.g. understand their lectures intended for young mathematicians.	Student has the ability to analyze data and prove the correctness of reasoning based on theory.	[SU3] Assessment of ability to use knowledge gained from the subject				
	[K7_W05] Has enhanced knowledge of a selected branch of mathematics: knows most classical definitions and theorems and their proofs, Understands problems being examined, Knows relations between problems from particular field with other branches of mathematics, theoretical and applied	Understanding of statistical structures within decision trees. Is able to indicate the role of ELM at the theoretical level and in packages.	[SW1] Assessment of factual knowledge				
	[K7_W12] Knows well at least one symbolic computation software package and one statistical data processing package.	Use SAS, R or Python programs is developed	[SW3] Assessment of knowledge contained in written work and projects				
	[K7_U08] Knows probability distributions and their properties; is able to use them in practical issues, is familiar with the basics of statistics (estimation issues and hypothesis testing) and the basics of statistical data processing.	basically, concepts resulting from the basics are developed, i.e. parametric estimation and hypothesis verification	[SU4] Assessment of ability to use methods and tools				
Subject contents							
	1. Regression estimator, Cantelli Gliwienko's theorems and regression consistency theorems2. Matrix factorization3. Combinatorics according to Vapnik and Czervonenkis. Vapnik Chervonenkis class generation4. Theorem Rao-Blackwell the Hodges-Lehman theorem.5. Kulback Leibler's information, AKAIKE information criteria.6. Cramer Rao's inequality.7. The role of information8. 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 Statyst 1989	yki matematycznej PWN Warszawa				
	Sadanori Konishi, Genshiro Kitagawa: "Information Criteria and Statistical Modeling" Springer Series in Statistics2008						
		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 inStatistics 2nd Edition 199819	
	eResources addresses	Podstawowe https://drive.pg.edu.pl/s/7XJjfAXz1lbdCVW - The entire lecture can be found in https://drive.pg.edu.pl/s/7XJjfAXz1lbdCVW files lecture2.pdf and statisticsalecture.pdf Uzupełniające Adresy na platformie eNauczanie: Statystyka II 2024 - Moodle ID: 38036 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=38036	
Example issues/ example questions/ tasks being completed	1. AKAIKE2. VC dimension, examples, complexity measures (theorem 1.6)3. festimator 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.