

Specification for the book of courses

Study program		Communications and Information Technologies		
Module		Communications and Information Processing		
Type and level of studies		Master studies		
The name of the course		Big data analysis		
Lecturer (for lectures)		Milović M. Daniela, Đorđević T. Goran, Milić N. Dejan		
Lecturer/associate (for exercises)		Cvetković M. Aleksandra		
Lecturer/associate (for OFE)				
Number of ECTS		4	Course status (obligatory/elective)	Obligatory
Prerequisites				
Getting acquainted with the complete process of data acquisition, initial steps of understanding data, their visualization, hypothesis settings and analysis. The most important step in obtaining information from large data sets is the application of intelligent techniques for dimensional reduction of large datasets (data collected from different sensors) and application in practical applications (shape recognition).				
Course objectives				
Ability to extract and analyze large sets of data from different sources. Design and implementation of statistical models and neural networks in solving communication problems. Programming in statistical packages: R, Python.				
Course outcomes				
Ability to extract and analyze large sets of data from different sources. Design and implementation of statistical models and neural networks in solving communication problems. Programming in statistical packages: R, Python.				
Course outline				
Introduction to the analysis of large data sets - their characteristics, sources and challenges. Programming in statistical packages: R, Python. Probabilistic generative model, graphic models, neural networks. Methods for reducing the dimensionality of large data sets - PCA. Classification - Bayes Classifier, multidisciplinary classification. Regression - SVM regression. Mechanical learning algorithms and applications in robust prediction models.				
Theoretical teaching				
Auditory exercises, computer exercises.				
Practical teaching (exercises, OFE, study and research)				
Auditory exercises, computer exercises.				
Textbooks/references				
1 J. Izenman. Modern Multivariate Statistical Techniques. Springer 2008.				
2 Elements of Statistical Learning by Hastie, Tibshirani & Friedman				
3 Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman (2014.), Mining of Massive Datasets, Cambridge University Press				
4 Statistical Inference, Learning and Models in Big Data Franke et al, 2016				
5 Python Machine Learning, Sebastian Raschka, Packt Publishing, 2015				
Number of classes of active education per week during semester/trimester/year				
Lectures	Exercises	OFE	Study and research work	Other classes
2	1	0	0	0
Teaching methods				
Lectures, auditory exercises, homeworks, office hours.				
Grade (maximum number of points 100)				
Pre-exam duties		Points	Final exam	Points
Activity during lectures			Written exam	25
Exercises		30	Oral exam	25
Colloquia				
Projects		20		