

Course title: Theory and Practice of Data-Mining	Neptun code: GEIAL417-a
Course coordinator: Dr. László József Kovács, PhD, dr. habil., professor	
type of lesson and number of lessons: lecture (2)	
method of evaluation: colloquium	
curriculum location of the subject: (autumn/spring semester): autumn and spring	
pre-study conditions (<i>if any</i>): -	
The task and purpose of the subject:	
The purpose of the subject is to present the goals and methods of data mining. Students learn the statistical and algebraic foundations of analysis. After a general overview of the basic data analysis frameworks, the most important clustering and classification methods are presented. Students learn the basics of neural network-based architecture..	
Course description:	
Operational model of data mining; Overview of main data mining methods, Overview of statistical methods: correlation, statistical tests; Algebraic elements, operations with matrices; Concept of classification, classification techniques: Bayes, decision tree, SVM classifier, Random Forest techniques, approximation theory, NN-MLP architecture, operational model and parameters, Efficiency indicators, Regularization techniques, k-fold validation.	
Required literature:	
<ol style="list-style-type: none"> 1. Dr. Kovács László: Data Analysis and Data Mining, moodle course (moodle.iit.uni-miskolc.hu) 2. J., Reis,M:Housley: Fundamentals of Data Engineering: Plan and Build Robust Data Systems , O'Reilly, 2022 3. D Larose, C. Larose: Data Mining and Predictive Analytics, Wiley, 2015 	
Recommended literature:	
<ol style="list-style-type: none"> 1. M Dunham: Data Mining: Introductory and Advanced Topics, 2020 2. Han J, Kamber M: DataMining, Concepts and Techniques, Panem , 2004 	