ÓBUDA UNIVERSITY
JOHN VON NEUMANN FACULTY OF INFORMATICS

Institute of Cyberphysical Systems			Semester 5. of the curriculum 2025-26-1			
Name of the subject:	Code of the subject:	Credits:	Weekly hours:			
				lec	sem	lab
Introduction to data science	NKXBA1EBNF		full-time	2	0	2
Responsible person for the subject: Masoumeh Vali			Classification: senior lecturer			
Subject lecturer(s):						
Prerequisites:	NKXAB1EBNF	Databases				
Way of the assessment:	mid-term grade					

Course description	
Goal:	The aim of the course is to provide a practical approach to the basic concepts and processes of data science. Through real-life application examples from practice, students will gain precise theoretical and practical hands-on knowledge by experiencing the material in depth. Machine learning algorithms form the backbone of the theoretical knowledge, while practical exercises provide a practical representation of the theory through the use of the Python language.
Course description:	Python basic, data cleaning, data visualization, data preprocessing, supervised, unsupervised learning, overfitting, underfitting, model validation, learning/validation/testing sets, cross-validation, Bias-Variance, least squares, Linear Regression, Gradient Method, Maximum-likelihood estimation, Logistic regression, Learning/validation/testing set, Cross-validation, Bias-Variance tradeoff, Precision-Recall, F1-score, ROC curve, SVM, Neural networks, Decision trees, Random forests, Boosting, Unsupervised learning, Clustering, K-means clustering, Reinforcement learning.

Lecture schedule	
Education week	Topic
1.	Python Basics
2.	Python Basics, Statistics in data science
3.	Import Data, Data Visualization
4.	Cleaning of Imported data, Data preprocessing
5.	Supervised learning - types of validation metrics
6.	Supervised learning - types of validation metrics
7.	Unsupervised learning, clustering, K-means
8.	Reinforcement learning
9.	Neural networks, Deep learning
10.	Deep learning, Regularization
11.	Convolutional Neural Networks, transfer Learning
12.	Practical examples of data mining problems
13.	Theoretical and Laboratory test
14.	Theoretical and Laboratory test (replacement)

Mid-term requirements	
Conditions for obtaining a mid-term grade/signature	During the semester, students will take final tests in week 13 based on the lecture and lab material. Student can obtain maximum 50-50 points for each test, the sum of which will be the final score (max. 100 points). A minimum of 26 points is required for both the final and final exams to pass the course.

Assessment schedule	
Education week	Topic
13.	Theoretical and Laboratory test

14.	Theoretical and Laboratory test (replacement)

Method used to calculate the *mid-term grade* (to be filled out only for subjects with mid-term grades)

The final semester mark is the average of the scores of the tests, for which a maximum of 100 points can be obtained. For a satisfactory mark, 52 points, 63 points for a medium, 74 points for a good and 85 points for an excellent mark are required.

Type of the replacement

Type of the replacement of written test/mid-term grade/signature	During the first week of the exam period both test can be replaced.

Type of the exam (to be filled out only for subjects with exams)

Calculation of the exam mark (to be filled only for subjects with exams)

Final grade calculation methods:

The final semester mark is the average of the scores of the tests, for which a maximum of 100 points can be obtained. For a satisfactory mark, 52 points, 63 points for a medium, 74 points for a good and 85 points for an excellent mark are required.

References

Obligatory:	Lecture slides of each week will be placed on https://elearning.uni-obuda.hu/
Recommended:	Sweigart, A. (2019). *Automate the boring stuff with Python: practical programming for total beginners*. No Starch Press. Han, J., Pei, J., & Tong, H. (2022). *Data mining: concepts and techniques*. Morgan kaufmann. Kane, F. (2017). *Hands-on data science and python machine learning*. Packt Publishing Ltd. Barry, P. (2016). *Head first Python: A brain-friendly guide*. " O'Reilly Media, Inc.". Lutz, M. (2013). *Learning python: Powerful object-oriented programming*. " O'Reilly Media, Inc.". Slatkin, B. (2019). *Effective Python*. Addison-Wesley Professional. Ramalho, L. (2022). *Fluent python*. " O'Reilly Media, Inc.". VanderPlas, J. (2016). *Python data science handbook: Essential tools for working with data*. " O'Reilly Media, Inc.". Grus, J. (2019). *Data science from scratch: first principles with python*. O'Reilly Media. Müller, A. C., & Guido, S. (2016). *Introduction to machine learning with Python: a guide for data scientists*. " O'Reilly Media, Inc.".
Other references:	Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press. Kelleher, J. D. (2019). *Deep learning*. MIT press.