UČNI NAČRT PREDMETA / COURSE SYLLABUS	
Predmet:	Programiranje za podatkovne znanosti
Course title:	Data Science Programming

Študijski program in stopnja Study programme and level	Študijska smer Study field	Letnik Academic year	Semester Semester
Podatkovne znanosti, magistrski študijski program druge stopnje	-	Prvi	Prvi
The second cycle masters study programme Data Sciences	-	First	First

Vrsta predmeta / Course type	Obvezni / Obligatory
Univerzitetna koda predmeta / University course code:	2-PZ-MAG-PPZ-2020-06-30

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
35	-	45	-	-	130	7

Nosilec predmeta / Lecturer:	Izr. prof. dr. Borut Lužar

Jeziki / Languages:	**Predavanja / Lectures:**	Slovenski / Slovenian, Angleški / English
	Vaje / Tutorial:	Slovenski / Slovenian, Angleški / English

Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:	**Prerequisites:**
Specifičnih pogojev za vključitev v delo ni. Priporočeno je osnovno znanje vsaj enega izmed standardnih programskih jezikov. Pogoj za pristop k izpitu je priprava in zagovor projektne naloge.	There are no specific requirements for this subject. It is recommended for a student to be familiar with at least one of standard programming languages. To attend the exam, a student has to prepare and present a project assignment.

Vsebina:	**Content (Syllabus outline):**

- Uvod v Python (osnove, podatkovne strukture, funkcije, delo z datotekami, primeri uporabe)
- Uvod v IPython (osnove, zgodovina ukazov, interakcija z OS, razvijalska orodja, napredne funkcije)
- Osnove dela z NumPy (ndarray, univerzalne funkcije, procesiranje podatkov, delo z datotekami, linearna algebra)
- Uvod v Pandas (podatkovne strukture, nekaj glavnih funkcionalnosti, statistika s Pandas, manjkajoči podatki)
- Upravljanje podatkov (branje in pisanje, binarni podatkovni formati, interakcija s spletnimi API-ji, interakcija s podatkovnimi bazami)
- Priprava podatkov (kombiniranje in združevanje baz podatkov, preoblikovanje in pivotiranje, transformacije, priprava nizov)
- Risanje in vizualizacija (matplotlib, funkcije v Pandas, primeri)
- Uvod v strojno učenje s Pythonom (delo s scikit-learn)

- Introduction to Python (basics, data structures, functions, working with files, examples of usage)
- Introduction to IPython (basics, command history, interacting with OS, development tools, advanced features)
- NumPy basics (ndarray, universal functions, data processing, working with files, linear algebra)
- Introduction to Pandas (data structures, some main functionalities, statistics with Pandas, missing data)
- Data management (reading and writing, binary data formats, interaction with web APIs, interaction with databases)
- Data preparation (combining and merging datasets, reshaping and pivoting, transformations, string management)
- Plotting and visualization (matplotlib, functions in Pandas, examples)
- Introduction to machine learning with Python (working with scikit-learn)

Temeljni literatura in viri / Readings:

- W. McKinney: *Python for Data Analysis*, 2nd edition, O'Reilly Media, Inc., 2017.
- S. Raschka, M. Vahid: *Python Machine Learning*, 3rd edition, Packt Publishing, 2019.
- B. Lužar: Prosojnice iz predavanj in vaj pri predmetu Programiranje za podatkovne znanosti, Moodle, FIŠ.

Cilji in kompetence:

Splošne kompetence:

- Sposobnost analitičnega in algoritmičnega razmišljanja.
- Sposobnost uporabe različnih programskih rešitev za analizo podatkov.
- Obvladovanje sodobnih visoko zmogljivih orodij in specifične programske opreme za obdelavo podatkov.

Predmetno-specifične kompetence:

- Sposobnost razvoja programskih rešitev za analizo specifičnih podatkov.
- Sposobnost pridobivanja in priprave podatkov iz različnih virov ter njihove shranjevanje v ustrezno podatkovno strukturo.

Objectives and competences:

General competences:

- Ability of analytical and algorithmic thinking.
- Ability to use various software solutions for data analysis.
- Mastering cutting edge high performance tools and corresponding software for data processing.

Subject-specific competences:

- Ability for an independent development of software solutions for specific data analysis
- Ability to retrieve and prepare data from various sources and storing them in an appropriate data structure.

Predvideni študijski rezultati:	**Intended learning outcomes:**
Znanje in razumevanje: • Študentje bodo spoznali enega izmed najbolj razširjenih programskih jezikov za obdelavo podatkov • Podrobno bodo spoznali napredne knjižnice programskega jezika Python. • Razumeli bodo koncepta niti in vzporednega procesiranja ter ju bili sposobni uporabiti v praksi. Prenosljive spretnosti: • Pridobljeno znanje bo nadgradilo razumevanje razvoja algoritmov. • Podobno, bodo znanja podlaga za študij bolj specifičnih predmetov s področja strojnega učenja in statistike.	Knowledge and understanding: • Students will get acquainted with one of the most common programming languages used for data analysis. • Students will get acquainted with advanced libraries in Python programming language. • Students will understand the concepts of threads and parallel processing and will be able to use them in practice. Transferable skills: • Obtained knowledge will upgrade the understanding of algorithms development. • Similarly, the unit is a basis for study of more specific subjects from the fields of machine learning and statistics.

Metode poučevanja in učenja:	**Learning and teaching methods:**
• Predavanja z aktivno udeležbo študentov (razlaga, diskusija, vprašanja, primeri). • Vaje (reševanje nalog).	• Lectures with active students' participation (explanations, discussion, questions, examples). • Laboratory work (solving problems).

Načini ocenjevanja:	Delež (v %) / Weight (in %)	**Assessment:**
• pisni izpit • projektna naloga	60 % 40 %	• written exam • project assignment

Reference nosilca / Lecturer's references:

- L. Bezegová, B. Lužar, M. Mockovčiaková, R. Soták, R. Škrekovski, Star edge colorings of some classes of graphs, J. Graph Theory 81 (2016), 73-82.
- P. Gregor, B. Lužar, R. Soták, On incidence coloring conjecture in Cartesian products of graphs, Discrete Appl. Math. 213 (2016), 93-100.
- P. Gregor, B. Lužar, R. Soták, Note on incidence chromatic number of subquartic graphs, J. Combin. Optim. 34 (2017), 174-181.

- M. Bonamy, M. Knor, B. Lužar, A. Pinlou, R. Škrekovski, On the difference between the Szeged and the Wiener index, Appl. Math. Comput. 312 (2017), 202-213.
- B. Lužar, M. Petruševski, R. Škrekovski: On vertex-parity edge-colorings, J. Combin. Optim. 35 (2018), 373-388.
- V. Andova, B. Lidický, B. Lužar, R. Škrekovski: On facial unique-maximum (edge-) coloring, Discrete Appl. Math. 237 (2018), 26-32.
- B. Lužar, P. Ochem, A. Pinlou: On repetition thresholds of caterpillars and trees of bounded degree, Electron J. Combin. 25 (2018), #P1.61.
- B. Lužar, J. Przybyło, R. Soták: New bounds for locally irregular chromatic index of bipartite and subcubic graphs, J. Combin. Optim. 36(4) (2018), 1425-1438.
- B. Lužar, M. Mockovčiaková, R. Soták: Note on list star edge-coloring of subcubic graphs, J. Graph Theory 90(3) (2018), 304-310.
- F. Dross, B. Lužar, M. Maceková, R. Soták: Note on 3-choosability of planar graphs with maximum degree 4, Discrete Math. 342(11) (2019), 3123-3129.
- B. Lužar, M. Mockovčiaková, P. Ochem, A. Pinlou, R. Soták: On non-repetitive sequences of arithmetic progressions: the cases $k \in \{4,5,6,7,8\}$, Discrete Appl. Math. 279 (2020), 106-117.