

**UČNI NAČRT PREDMETA / COURSE SYLLABUS**

**Predmet:** Odkrivanje znanja v podatkih  
**Course title:** Knowledge Discovery from Data

Študijski program in stopnja Study programme and level	Študijska smer Study field	Letnik Academic year	Semester Semester
Računalništvo in spletne tehnologije, visokošolski strokovni študijski program prve stopnje	-	Drugi	Četrta
Computer Science and Web Technologies, first cycle Professional Study Programme	-	Second	Fourth

**Vrsta predmeta / Course type**

Obvezni / Obligatory

**Univerzitetna koda predmeta / University course code:**

2-RST-VS-OZP-2020-05-14

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
30	-	45	-	-	105	6

**Nosilec predmeta / Lecturer:** izr. prof. dr. Biljana Mileva Boshkoska

**Jeziki / Languages:**

**Predavanja / Lectures:** Slovenski / Slovenian, Angleški / English

**Vaje / Tutorial:** Slovenski / Slovenian, Angleški / English

**Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:**

Pogoj za vključitev je vpis v 2. letnik študija. Študent/študentka mora pred pristopom k izpitu pripraviti in zagovarjati seminarsko nalogo.

**Prerequisites:**

Prerequisite for inclusion is enrolment into the second year of the study. Prior to the exam, the student has to prepare and defend seminar work.

**Vsebina:**

- Uvod v področje tehnologije znanj.
- Proces odkrivanja znanja v podatkih.
- Različni modeli odkrivanja znanja v podatkih (raziskovalni, industrijski, hibridni).
- Evolucija sistemov za odkrivanje znanja v podatkih.
- Podatkovna skladišča, podatkovna kocka in analize tipa OLAP (online analytical processing).

**Content (Syllabus outline):**

- Introduction to Knowledge Technologies.
- Description of the Knowledge Discovery in Databases (KDD) process.
- Different models of KDD (academic, industrial, hybrid).
- Evolution of KDD systems.
- Data warehouses, data cube and OLAP (online analytical processing) analysis.

- Preplet odkrivanja znanja v podatkih in poslovnih procesov v organizaciji (business intelligence).
- Strojno učenje: problem učenja iz podatkov, podatkovno rudarjenje, odločitvena drevesa, Bayesov klasifikator, analiza grozdenja, asociacijska pravila, najbližji sosed).
- Študije realnih primerov odkrivanja znanja iz podatkov / spletnih virov / besedil.
- Pregled orodij za podatkovno rudarjenje.

- KDD and business processes (business intelligence).
- Machine learning: learning from data, data mining, decision tree induction, Naive Bayes classifier, cluster analysis, association rules, k-NN,).
- KDD Case studies on real world scenarios.
- Overview of KDD tools.

### Temeljni literatura in viri / Readings:

- Bramer, M. (2007). *Principles of data mining. Vol. 180*. London: Springer.
- Fayyad, U. M., Piatetsky-Shapiro, G. & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), str. 37–54.
- Han, J. & Kamber, M. (2006). *Data Mining: Concepts and Techniques* (2nd ed.). The Morgan Kaufmann Series in Data Management Systems. Gray, Jim, Series Editor. Morgan Kaufmann Publishers.
- Hofmann, M. & Klinkenberg, R. (ur.) (2013). *RapidMiner: Data Mining Use Cases and Business Analytics Applications*. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. Chapman and Hall.
- Spletni viri, forumi skupnosti RapidMiner.
- Liu, B. (2011). *Web Data, Exploring Hyperlinks, Contents, and Usage Data* (2nd ed.). Springer.
- Rogel-Salazar, J. (2020). *Advanced Data Science and Analytics with Python*. Chapman and Hall/CRC.

### Cilji in kompetence:

*Učna enota prispeva k razvoju naslednjih splošnih in predmetno-specifičnih kompetenc:*

#### *Splošne kompetence:*

- poznavanje in razumevanje tehnoloških procesov ter sposobnost za njihovo analizo, sintezo in reševanje njihovih posledic
- obvladanje raziskovalnih metod, postopkov in procesov na področju tehničnih ved
- razvoj kritične in samokritične presoje
- sposobnost fleksibilne uporabe znanja v praksi
- sposobnost za reševanje konkretnih tehničnih in analitičnih problemov z uporabo ustreznih metod in postopkov
- sposobnost povezovanja koherentno obvladanega temeljnega znanja,

### Objectives and competences:

*The instructional unit contributes to the development of the following general and subject-specific competences:*

#### *General competences:*

- knowledge and understanding of technical processes together with the ability of their analysis, synthesis and solutions of their consequences
- mastering research methods procedures and processes in the field of technical sciences
- development of critical and self-critical judgement
- ability to use the acquired knowledge in practice in a flexible manner
- ability to solve technical and analytical problems using appropriate methods and procedures

pridobljenega pri obveznih predmetih, ter njegova uporaba v praksi

- sposobnost pridobivanja, selekcije, ocenjevanja in umeščanja novih informacij in zmožnost interpretacije v ustreznem kontekstu

*Predmetno-specifične kompetence:*

- razvoj veščin in spretnosti pri uporabi znanja na področju tehnoloških ved s pomočjo reševanja teoretičnih in empiričnih problemov
- razvoj veščin za gradnjo modelov nad podatki.
- spoznavanje področij rudarjenja podatkov, strojnega učenja
- spoznavanje področij rudarjenja besedil, spleta in multimedijskih vsebin

- ability to link coherent fundamental knowledge acquired at compulsory subjects and its application in practice
- ability to find, select, evaluate and position the new information as well as appropriate, context-aware interpretation

*Subject-specific competences:*

- development of skills and abilities using the knowledge in the fields of technical sciences by solving theoretical and empirical problems
- development of skills necessary for the construction of data models
- acquire knowledge about data mining, machine learning
- acquire knowledge about text mining, the web and multimedia content

#### **Predvideni študijski rezultati:**

Znanje in razumevanje:

*Študent/študentka:*

- pozna pomen in vlogo odkrivanja znanja v podatkih
- razume pomen standardizacije KDD procesov
- pozna podroben opis evaluacije in aplikacije v CRISP-DM procesnem modelu
- pozna osnovne metode za zajem in pripravo podatkov
- pozna proces odkrivanja znanja v podatkih s posebnim poudarkom na izbranih metodah rudarjenja podatkov
- razume osnovne metode rudarjenja podatkov
- spozna področje rudarjenja podatkov, strojnega učenja
- spozna področje rudarjenja besedil, spleta in multimedijskih vsebin
- pozna delo z razpoložljivimi orodji za podatkovno rudarjenje
- pridobi veščine za samostojno uporabo KDD procesov v praksi

#### **Intended learning outcomes:**

Knowledge and understanding:

*The student:*

- knows and understands role of KDD
- understands the meaning of standardization of KDD processes
- provide detailed description of evaluation and deployment in CRISP-DM proces model
- understands KDD process and basic DM methods
- understands fields of data mining, machine learning
- understands methods for gathering data from text, internet resources, and media
- knows how to implement a KDD process from real world in KDD-DM tools
- gather skills for application of KDD process into practice

#### **Metode poučevanja in učenja:**

- *predavanja z aktivno udeležbo študentov (razlaga, diskusija, vprašanja, primeri, reševanje problemov)*

#### **Learning and teaching methods:**

- *lectures with active student participation (explanation, discussion, questions, examples problem solving)*

- *seminarske vaje* (refleksija izkušenj, projektno delo, timsko delo, metode kritičnega mišljenja, študije primerov)
- individualne in skupinske laboratorijske vaje (načrtovanje KDD procesov z RapidMiner OKPO orodjem, dodatna razlaga na primerih, obravnava specifičnih vprašanj)

- *seminars* (reflexion about experiences, project work, team work, method of critical thinking, case studies)
- individual and group *work in laboratory* (design of KDD processes with RapidMiner tool, additional explanation, treatment of specific questions)

Delež (v %) /

Weight (in %)

**Načini ocenjevanja:**

**Assessment:**

Način (pisni izpit, ustno izpraševanje, naloge, projekt):	Delež (v %) / Weight (in %)	Type (examination, oral, coursework, project):
<ul style="list-style-type: none"> <li>• laboratorijsko delo in aplikacija znanja na realni primer iz okolja</li> </ul>	40	<ul style="list-style-type: none"> <li>• application of kdd on real world scenario in practice (laboratory seminar work)</li> </ul>
<ul style="list-style-type: none"> <li>• pisni izpit</li> </ul>	60	<ul style="list-style-type: none"> <li>• written exam</li> </ul>

**Reference nosilca / Lecturer's references:**

- STROJNIK, Lidija, STOPAR, Matej, ZLATIC, Emil, KOKALJ, Doris, NAGLIČ, Mateja, ŽENKO, Bernard, ŽNIDARŠIČ, Martin, BOHANEK, Marko, MILEVA BOSHKOSKA, Biljana, LUŠTREK, Mitja, GRADIŠEK, Anton, POTOČNIK, Doris, OGRINC, Nives. Authentication of key aroma compounds in apple using stable isotope approach. *Food chemistry*, ISSN 0308-8146. [Print ed.], 2019, vol. 277, str. 766-773, doi: 10.1016/j.foodchem.2018.10.140. [COBISS.SI-ID 31834663].
- BOŠKOSKI, Pavle, DEBENJAK, Andrej, MILEVA BOSHKOSKA, Biljana. Rayleigh copula for describing impedance data - with application to condition monitoring of proton exchange membrane fuel cells. *European journal of operational research*, ISSN 0377-2217. [Print ed.], 2018, vol. 266, no. 1, str. 269-277, doi: 10.1016/j.ejor.2017.08.058. [COBISS.SI-ID 30736167].
- GRAŠIČ, Valerij, KOS, Andrej, MILEVA BOSHKOSKA, Biljana. Classification of incoming calls for the capital city of Slovenia smart city 112 public safety system using open Internet of Things data. *International journal of distributed sensor networks*, ISSN 1550-1477. [Online ed.], 2018, vol. 14, no. 9, str. 1-12, ilustr. <https://journals.sagepub.com/doi/pdf/10.1177/1550147718801703>, doi: 10.1177/1550147718801703. [COBISS.SI-ID 2048569107].
- MILJKOVIĆ, Dragana, LAVRAČ, Nada, BOHANEK, Marko, MILEVA BOSHKOSKA, Biljana. Discovering dependencies between domains of redox potential and plant defence through triplet extraction and copulas. *International journal of intelligent engineering informatics*, ISSN 1758-8723, 2018, vol. 6, no. 1/2, str. 61-77. <http://www.inderscience.com/info/ingeneral/forthcoming.php?jcode=ijiei>, doi: 10.1504/IJIEI.2018.10012065. [COBISS.SI-ID 2048463379].
- MILEVA BOSHKOSKA, Biljana, LIU, Shaofeng, CHEN, Huilan. Towards a knowledge management framework for crossing knowledge boundaries in agricultural value chain. *Journal of decision systems*, ISSN 1246-0125, [in press] 2018, 15 str., doi: 10.1080/12460125.2018.1468173. [COBISS.SI-ID 31392807].
- ZHAO, Guoqing, LIU, Shaofeng, LOPEZ, Carmen, LU, Haiyan, ELGUETA, Sebastian, CHEN, Huilan, MILEVA BOSHKOSKA, Biljana. Blockchain technology in agri-food value chain management : a synthesis of applications, challenges and future research directions. *Computers in industry*, ISSN 0166-3615. [Print ed.], 2019, vol. 109, str. 83-99, doi: 10.1016/j.compind.2019.04.002. [COBISS.SI-ID 32345127].

- GRAŠIČ, Valerij, KOS, Andrej, MILEVA BOSHKOŠKA, Biljana. Classification of incoming calls for the capital city of Slovenia smart city 112 public safety system using open Internet of Things data. *International journal of distributed sensor networks*, ISSN 1550-1477. [Online ed.], 2018, vol. 14, no. 9, str. 1-12, doi: 10.1177/1550147718801703. [COBISS.SI-ID 2048569107].