

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
Informatika v sodobni družbi, univerzitetni študijski program prve stopnje	-	Drugi ali tretji	Četrty ali šesti
Informatics in Contemporary Society, first cycle Academic Study programme	-	Second or third	Fourth or sixth

Vrsta predmeta / Course type

Izbirni / Elective

Univerzitetna koda predmeta / University course code:

1-ISD-UN-IP-OZP-2019-05-13

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

Nosilec predmeta / Lecturer:

Jeziki / Languages:

Predavanja / Lectures: Slovenski, angleški / Slovene, English

Vaje / Tutorial: Slovenski, angleški / Slovene, English

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

Študent/študentka mora pred pristopom k izpitu pripraviti in zagovarjati raziskovalno nalogo in projekt.

Prerequisites:

Prior to the exam, the student has to prepare and present project work.

Vsebina:

- uvod: metode odkrivanja znanja, proces odkrivanja znanja, naloge podatkovnega rudarjenja, aplikacije podatkovnega rudarjenja, uporaba odkritega znanja pri inteligentnih, odločitvenih in ekspertnih sistemih, arhitektura sistemov podatkovnega rudarjenja

Content (Syllabus outline):

- introduction: methods of knowledge discovery process in databases (KDD), data mining tasks, applications of data mining (DM), decision-making and expert systems
- basics of machine learning: presentation of data, supervised and unsupervised learning, symbolic and

<ul style="list-style-type: none"> • osnove strojnega učenja: predstavitev podatkov, nadzorovano in nenadzorovano učenje, simbolične in ne-simbolične metode strojnega učenja • predobdelava: spoznavanje domene, priprava transformacija in čiščenje podatkov, transformacija prostora • tehnike podatkovnega rudarjenja: predikcijske metode (odločitvena drevesa (algoritem C4.5), k-najbližji sosed, Bayesova klasifikacija, genetski algoritmi), analiza gruč • obdelava rezultatov: evalvacija rezultatov (učna, testna množica vzorcev, metode validacije in križna validacije), interpretacija znanja in uporaba znanja 	<p>non-symbolic machine learning methods</p> <ul style="list-style-type: none"> • preprocessing: learning domain, the preparation of data, data cleansing, data transformation techniques • data mining techniques: decision trees (C4.5 algorithm), k-nearest neighbor, Bayesian classification, cluster analysis • evaluation of the results: evaluation of results (methods of validation and cross-validation), interpretation skills and use of knowledge
---	---

Temeljni literatura in viri / Readings:

<ul style="list-style-type: none"> • BRAMER, MAX (2007) <i>Principles of data mining. Vol. 180. London. Springer.</i> • FAYYAD, USAMA M., PIATETSKY-SHAPIRO, GREGORY in SMYTH, PADHRAIC (1996) From data mining to knowledge discovery in databases. <i>AI Magazine</i>, 17 (3), str. 37-54. • HAN, JIAWEI in KAMBER, MICHELINE (2006) <i>Data Mining: Concepts and Techniques, 2nd ed. The Morgan Kaufmann Series in Data Management Systems. GRAY, JIM, Series Editor. Morgan Kaufmann Publishers.</i> • HOFMANN MARKUS in KLINKENBERG RALF (ur.) (2013) <i>RapidMiner: Data Mining Use Cases and Business Analytics Applications. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. Chapman and Hall.</i> • Spletni viri, forumi skupnosti RapidMiner. • LIU, BING (2011) <i>Web Data, Exploring Hyperlinks, Contents, and Usage Data Second Edition. Springer.</i>

Cilji in kompetence:

<p><i>Učna enota prispeva k razvoju naslednjih splošnih in predmetno- specifičnih kompetenc:</i></p> <ul style="list-style-type: none"> • usposobljenost za samostojno in avtonomno uporabo, nadzor in vzdrževanje informacijsko komunikacijske tehnologije v organizaciji • obvladanje raziskovalnih metod, postopkov in procesov • razvoj (samo)kritične presoje • razvoj veščin in spretnosti pri uporabi znanja na področju družbenih ved s

Objectives and competences:

<p><i>The instructional unit contributes to the development of the following general and subject-specific competences:</i></p> <ul style="list-style-type: none"> • competence for independent and autonomous use, monitoring and maintenance of information communication technology in an institution • competence in research methods, procedures and processes • development of (self)critical judgement • development of abilities and skills for the
--

pomočjo reševanja teoretičnih ali empiričnih problemov

- sposobnost pridobivanja, selekcije, ocenjevanja in umeščanja novih informacij in zmožnost interpretacije v kontekstu družboslovja
- sposobnost zapisati problem v obliki algoritma in pretvorba algoritma v računalniški program z uporabo sodobnih programskih orodij,
- razumevanje in uporaba računalniških sistemov in arhitektur

use of knowledge in the field of social sciences with the aid of solving theoretic or empirical problems

- the ability to acquire, select, evaluate and insert new information and the ability to interpret within the context of social sciences
- ability to write down a problem in the form of an algorithm and the conversion of the algorithm into a computer programme with the use of modern software tools
- understanding and use of computer systems and architectures

Predvideni študijski rezultati:

Znanje in razumevanje:

Sposobnost študenta/študentke bo:

- poznavanje procesa odkrivanja zakonitosti v podatkih in njegove arhitekture
- poznavanje teoretičnih osnov strojnega učenja in pomembnejših podatkovno-rudarnih tehnik
- usposobljenost za zahtevnejšo obdelavo rezultatov
- uporaba nekaj najaktualnejših programskih orodij za podatkovno rudarjenje

Intended learning outcomes:

Knowledge and understanding:

Students will:

- understand the KDD process
- understand the common used machine learning techniques
- be able to solve complex problems with KDD
- use the latest software tools for data mining

Metode poučevanja in učenja:

- *predavanja* z aktivno udeležbo študentov (razlaga, diskusija, vprašanja, primeri, reševanje problemov)
- *vaje*: delo na vodenih primerih nalog odkrivanja zakonitosti v podatkih
- *vaje v računalniški učilnici*, pri katerih bodo študentje spoznali nekaj najaktualnejših programskih orodij za podatkovno rudarjenje (WEKA, RapidMiner)
- *projekt*, ki ga bodo študentje pripravili v manjših skupinah; vključeval bo konkreten problem, katerega bodo študentje v celoti rešili z metodami, spoznanimi na predavanjih in vajah

Learning and teaching methods:

- lectures with active participation of students (explanation, discussion, questions, examples, problem solving)
- explanations of the use cases in data mining
- lab work in computer lab, students will learn about latest software tools for data mining (weka, rapidminer)
- final project; students will prepare a project in small groups; it will include a concrete problem addressed by the student with methods and tools
- colloquia;

- kolokviji; z njimi bodo študentje stimulirani, da sproti študirajo obravnavano snov

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"> • laboratorijsko delo in aplikacija znanja na realni primer iz okolja 	50	<ul style="list-style-type: none"> • application of kdd on real world scenario in practice (laboratory seminar work)
<ul style="list-style-type: none"> • pisni/ ustni izpit 	50	<ul style="list-style-type: none"> • written/ oral exam