

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-2016-10-01

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
30	-	15	-	30	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 o data mining (DM), decision-making and expert systems
- basics of machine learning: presentation of data, supervised and unsupervised learning, symbolic and

• 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	non-symbolic machine learning methods • 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:

• RapidMiner : data mining use cases and business analytics applications, (Chapman & Hall/CRC data mining and knowledge discovery series), Markus Hofmann, ur., Ralf Klinkenberg, ur., Boca Raton, CRC Press, cop. 2014. • Kononenko, I.: <i>Strojno učenje</i>, Založba FEFRI, Ljubljana. • U. M., Fayyad, G., Piatetsky-Shapiro, P., Smyth, P., and R., Uthurusamy, (eds.) (1996): <i>Advances in Knowledge Discovery and Data Mining</i>, AAAI Press / The MIT Press. • Han, J. and Kamber, M. (2001): <i>Data Mining: Concepts and Techniques</i>, Morgan Kaufmann. • Piatetsky-Shapiro, G. and Frawley, W.J. (1991): <i>Knowledge Discovery in Databases</i>, AAAI Press / The MIT Press.
--

Cilji in kompetence:

<i>Učna enota prispeva k razvoju naslednjih splošnih in predmetno- specifičnih kompetenc:</i> • obvladanje raziskovalnih metod, postopkov in procesov • razvoj kritične in samokritične presoje • sposobnost pridobivanja, selekcije, ocenjevanja in umeščanja novih spoznanj in zmožnost njihove interpretacije • sposobnost načrtovanja informacijsko-komunikacijske tehnologije in sistemov • sposobnost upravljanja informacijsko-komunikacijske

Objectives and competences:

<i>The instructional unit contributes to the development of the following general and subject-specific competences:</i> • knowledge of research methods, procedures and processes in the field of social sciences • development of critical and self critical evaluation • ability to design information and communication technologies and systems • ability to manage information and communications technologies and systems • development of the skills and expertise through solving theoretical
--

- tehnologije in sistemov
- razvoj veščin in spretnosti pri uporabi znanja s pomočjo reševanja teoretičnih ali empiričnih problemov

or empirical problems

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

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;

Načini ocenjevanja:	Delež (v %) / Weight (in %)	Assessment:
Način (pisni izpit, ustno izpraševanje, naloge, projekt): • pisni izpit • raziskovalna naloga	50 50	Type (examination, oral, coursework, project): • written exam • research paper