

**UČNI NAČRT PREDMETA / COURSE SYLLABUS**

**Predmet:** Odkrivanje zakonitosti 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, visokošolski strokovni študijski program prve stopnje	-	Drugi	Četrtri
Informatics in Contemporary Society, first cycle Professional Study Programme	-	Second	Fourth

**Vrsta predmeta / Course type**

Izbirni / Elective

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

1-ISD-VS-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:**

The student is obliged to prepare and defend his/her research project before the admission to the examination.

**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.
- *Osnove strojnega učenja:* predstavitev podatkov, nadzorovano in nenadzorovano učenje, simbolične in

**Content (Syllabus outline):**

- Introduction: methods of knowledge discovery process in databases (KDD), data mining (DM) application, decision-making and expert systems, data mining systems architecture.
- The basics of machine learning: the presentation of data, supervised and unsupervised learning, symbolic and non-symbolic machine learning methods.

ne-simbolične metode strojnega učenja.

- *Predobdelava*: spoznavanje domene, priprava transformacija in čiščenje podatkov.
- Tehnike podatkovnega rudarjenja: Predikcijske metode: odločitvena drevesa (algoritem C4.5), k-najbližji sosed, Bayesova klasifikacija.
- *Obdelava rezultatov*: Evalvacija rezultatov (učna, testna množica vzorcev, metode validacije in križna validacija), interpretacija znanja in uporaba znanja.

- Learning domain, the preparation of data, data cleansing and transformation.
- The techniques of data mining: decision trees (C4.5 algorithm), k-nearest neighbor, bayesian classification.
- Evaluaton of the 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.(1997): *Strojno učenje*, Založba FERi, Ljubljana.
- U. M., Fayyad, G., Piatetsky-Shapiro, P., Smyth, P., and R., Uthurusamy, (eds.) (1996): *Advances in Knowledge Discovery and Data Mining*, AAAI Press / The MIT Press.
- Han, J. and Kamber, M. (2001): *Data Mining: Concepts and Techniques*, Morgan Kaufmann.
- Piatetsky-Shapiro, G. and Frawley, W.J. (1991): *Knowledge Discovery in Databases*, AAAI Press / The MIT Press.

### Cilji in kompetence:

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

- 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 pomočjo reševanja teoretičnih ali empiričnih problemov
- sposobnost pridobivanja, selekcije, ocenjevanja in umeščanja novih informacij in zmožnost interpretacije v kontekstu družboslovja

### Objectives and competences:

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

- 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 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

- sposobnost zapisati problem v obliki algoritma in pretvorba algoritma v računalniški program z uporabo sodobnih programskih orodij,
- razumevanje in uporaba računalniških sistemov in arhitektur

- 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:

*Študent/študentka:*

- se seznanijo s procesom odkrivanja zakonitosti v podatkih in njegovo arhitekturo
- se seznanijo s teoretičnimi osnovami strojnega učenja in pomembnejšimi podatkovno rudarnimi tehnikami
- se nauči uporabljati nekaj najaktualnejših programskih orodij za podatkovno rudarjenje.

### **Intended learning outcomes:**

Knowledge and understanding:

Student will gain:

- understanding of the DDR process
- understanding of the common used machine learning techniques
- use of 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 teh vajah bodo študentje spoznali nekaj najaktualnejših programskih orodij za podatkovno rudarjenje (WEKA, Rapid-I)
- *projekt*, ki ga bodo študentje pripravili v manjših skupinah. Vključeval bo konkreten problem, ki ga bodo morali študentje v celoti rešiti z metodami, spoznanimi na predavanjih in vajah.
- *kolokviji*: z njimi bodo študentje stimulirani, da sproti študirajo snov, ki bo obravnavana 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

<b>Načini ocenjevanja:</b>	Delež (v %) / Weight (in %)	<b>Assessment:</b>
Način (pisni izpit, ustno izpraševanje, naloge, projekt): <ul style="list-style-type: none"> <li>• pisni izpit</li> <li>• raziskovalna naloga</li> </ul>	50 50	Type (examination, oral, coursework, project): <ul style="list-style-type: none"> <li>• written exam</li> <li>• research paper</li> </ul>