

Sveučilište u Rijeci
ODJEL ZA INFORMATIKU
Radmile Matejčić 2, Rijeka
Akademska 2014/2015. godina

OTKRIVANJE ZNANJA U PODACIMA

Studij: Diplomski studij informatike (IKS, PI)
Web stranica predmeta: http://www.inf.uniri.hr/index.php?option=com_classes&Itemid=40&task=displayclass&id=2295
ECTS bodovi: 6
Nastavno opterećenje: 2 + 2

Nositeljica predmeta:

prof. dr.sc. Maja Matetić
e-mail: majam@inf.uniri.hr
web stranica: <http://www.uniri.hr>
Ured: Radmile Matejčić 2, kabinet O-407
Vrijeme konzultacija: po dogovoru e-mailom, nakon nastave, srijedom u 10:00

OTKRIVANJE ZNANJA U PODACIMA

Razvijanje općih i specifičnih kompetencija (znanja i vještina)

Predmet uvodi osnovne koncepte, zadatke, postupke i tehnike u području otkrivanja znanja u podacima dubinskom analizom podataka (data mining). Naglasak je na različitim problemima otkrivanja znanja u podacima i njihovim rješenjima. Studenti će u okviru predmeta upoznati proces otkrivanja znanja u podacima i probleme koji se pri tome javljaju, naučiti će različite tehnike dubinske analize podataka te će ih primijeniti u rješavanju problema otkrivanja znanja u podacima uporabom alata i sustava. Studenti će upoznati i primjere aplikacija za otkrivanje znanja u podacima.

Korespondentnost i korelativnost programa

Program kolegija je u korelaciji sa programima kolegija: Inteligentni sustavi 1 i 2. Navedeni kolegiji osiguravaju potrebno predznanje za kolegij.

Okvirni sadržaj predmeta

Uvod. Pretprocesiranje podataka. Klasifikacija i predviđanje. Klasifikacija indukcijom stabla odlučivanja. Bayesova klasifikacija. Klasifikacija "širenjem unatrag". Klasifikacija temeljena na pravilima. kNN klasifikator. Vrednovanje točnosti klasifikatora ili predviđanja. Fuzija postupaka—povećanje točnosti. Postupci grupiranja. Postupci dijeljenja. Hijerarhijsko grupiranje. konceptualno grupiranje. Postupci temeljeni na gustoći. Vrednovanje klastera. Otkrivanje čestih uzoraka, asocijacija i korelacija. Algoritam apriori. Otkrivanje asocijacija temeljenih na ograničenjima. Presentacija i demonstracija projekta.

Oblici provođenja nastave i način provjere znanja

Predavanja, laboratorijske vježbe. Pohađanje nastave, kolokviji, seminarski rad.

Popis literature potrebne za studij i polaganje ispita

1. Han, J. and Kamber, M., Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann, 2006.
2. Witten, I. H. and Frank, E., Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2011.

Popis literature koja se preporučuje kao dopunska

1. R.S. Michalski, I. Bratko, and M. Kubat, Machine Learning and Data Mining, John Wiley and Sons Ltd, 1998.
2. Mitchell, T.M., Machine Learning, McGraw-Hill, 1997.

Način praćenja kvalitete i uspješnosti izvedbe predmeta

Kroz ustrojeni sustav osiguranja kvalitete nastave Odjela za informatiku.

Preduvjeti za upis predmeta

Nema

R. BR.	OČEKIVANI ISHODI
1.	Objasniti što je dubinska analiza podataka (otkrivanje znanja u podacima) i kako se može primijeniti u rješavanju stvarnih problema
2.	Prepoznati je li tehnika dubinske analize podataka prihvatljiva za dani problem.
3.	Prepoznati je li tehnika dubinske analize podataka prihvatljiva za dani problem.
4.	Proći korake procesa otkrivanja znanja i napisati izvještaj o rezultatima izvođenja postupka otkrivanja znanja u podacima.
5.	Analizirati izlaz dobiven od aplikacije za dubinsku analizu podataka i vrednovati naučene rezultate.
6.	Prepoznati nekoliko strategija dubinske analize podataka i odlučiti u kojem slučaju je koja od njih najbolji izbor.
7.	Objasniti na koji način nekoliko različitih tehnika dubinske analize grade modele sa ciljem rješavanja problema.
8.	Opisati tipove problema koji se mogu rješavati kombinacijom ekspertnih sustava, pristupom rješavanja problema i strategijom dubinske analize podataka.
9.	Primijeniti programsku podršku koju su upoznali u okviru predmeta za rješavanje stvarnih problema.

AKTIVNOSTI I OCJENJIVANJE STUDENATA

VRSTA AKTIVNOSTI	ECTS	ISHODI UČENJA	SPECIFIČNA AKTIVNOST	METODA PROCJENJIVANJA	BODOVI MAX.
Pohađanje nastave	1	1-12	Prisutnost studenata	Popisivanje (evidencija)	0
Praktični rad - vježbe	1	1-12	Dubinska analiza uporabom odgovarajućih alata	Rješavanje zadataka i predaja rješenja na Mudri	20 (prag 15)
Istraživanje / Projekt	1	1-12	Istraživanje uporabom odgovarajućih alata	Prezentacija i obrana istraživanja	30 (prag 10)
Seminarski rad	1	1-12	Izrada i prezentacija seminarskog rada	Ocjena rada i izlaganja	25 (prag 10)
Kolokviji	1	1-12	Kolokvij (pismeni)	Kolokvij: 20	25 (prag 10)
Završni ispit (neobavezan)	1	1-12	Pismeni ispit	Opcionalno (za višu ocjenu): 30 bodova, ovisno o stupnju točnosti i potpunosti	
UKUPNO	6				100

Na završni ispit student izlazi samo ako skuplja bodove za višu ocjenu. Završni ispit nije obavezan.

Varijanta sa završnim ispitom: Ako student izlazi na završni ispit koji vrijedi 30 bodova, bodovi koje je skupio tijekom nastave se iz skale od 100 pretvaraju u bodove u skali od 70.

Varijanta bez završnog ispita: Ukoliko od 100 bodova student ima bar 50 bodova, dobiva ocjenu dovoljan (2), D (50-69.9 =>2, 70-79.9=>3, 80-89.9=>4, 90-100=>5). Studenti koji od 100 bodova steknu od 40 do 49.9 bodova izlaze na popravni ispit.

Primjeri ocjenjivanja:

Bodovi (skala od 100)		Ocjena
98	A	5
72	C	3
63	D	2
58	E	2
55	E	2
89	B	4
71	C	3
49	Fx	1 popravni
30	F	1 pad

OTKRIVANJE ZNANJA U PODACIMA,
PREDAVANJA I VJEŽBE, 2014/2015 IZVEDBENI PLAN NASTAVE

Nositelj kolegija: Izv. prof. dr. sc. Maja Matetić

Tjedan	Predavanja / Vježbe	Prostorija	Tema
1.	3.10. (P)		Predstavljanje, detaljno upoznavanje studenata s bodovima i obvezama. Uvod u dubinsku analizu podataka. Dubinska analiza podataka i strojno učenje.
	3.10. (V)		
2.	10.10. (P)		Ulaz: Koncepti, instance i atributi. Izlaz: Predstavljanje otkrivenog znanja Uvod u alat za dubinsku analizu podataka - WEKA.
	10.10. (V)		
3.	17.10. (P)		Preprocesiranje podataka: čišćenje, integracija, redukcija, transformacija, diskretizacija podataka. Generiranje hijerarhije koncepata. Izvedbe u alatu WEKA.
	17.10. (V)		
4.	24.10. (P)		Algoritmi – osnovni postupci dubinske analize podataka. Izvedbe u alatu WEKA.
	24.10. (V)		
5.	31.10. (P)		Vrednovanje rezultata dubinske analize podataka.
	31.10. (V)		
6.	7.11. (P)		Klasifikacija i predviđanje. Klasifikacija indukcijom stabla odlučivanja. Bayesova klasifikacija. Klasifikacija "širenjem unatrag".
	7.11. (V)		
7.	14.11. (P)		Klasifikacija temeljena na pravilima. kNN klasifikator. Vrednovanje točnosti klasifikatora ili predviđanja. Fuzija postupaka—povećanje točnosti.
	14.11. (V)		
8.	21.11. (P)		1. kolokvij
	21.11. (V)		
9.	28.11. (P)		Postupci grupiranja. Postupci dijeljenja. Hijerarhijsko grupiranje. konceptualno grupiranje. Postupci temeljeni na gustoći.
	28.11. (V)		
10.	5.12. (P)		Vrednovanje klastera. Otkrivanje čestih uzoraka, asocijacija i korelacija. Algoritam apriori. Otkrivanje asocijacija temeljenih na ograničenjima.
	5.12. (V)		
11.	12.12. (P)		Primjene: dubinske analize podataka: učenje iz skladišta podataka, učenje iz tijekova podataka, dubinska analiza teksta, dubinska analiza mrežnih stranica
	12.12. (V)		
12.	19.12. (P)		2. kolokvij

	19.12. (V)		
13.	9.01. (P)		Prezentacije projekta i seminarskog rada
	9.01. (V)		
14.	16.01. (P)		Prezentacije projekta i seminarskog rada
	16.01. (V)		
15.	23.01. (P)		Prezentacije projekta i seminarskog rada
	23.01. (P)		

Ispitni rokovi

Prvi i drugi redoviti rok:

03.02.2015.

17.02.2015.

Izvanredni rok:

17.03.2015.

Treći rok:

15.09.2015.