
Elektrotehnički fakultet u Beogradu
Katedra za računarsku tehniku i informatiku

Predmet: Pronalaženje skrivenog znanja (MS1PSZ)

Nastavnik: Prof. dr Veljko Milutinović

Asistent: Bojan Furlan

Ispitni rok: Septembar 2012.

Kandidat: _____

Broj Indeksa: _____ *E-mail:* _____

Ispit traje 1,5 sat. Dozvoljeno je korišćenje literature.

Zadatak 1 _____/10 *Usmeni* _____/20

Zadatak 2 _____/10 *Projekat* _____/60

Ukupno na ispitu: _____/20 *Ukupno na projektu:* _____/60

Ukupno: _____/100

Ocena: _____ (_____)

Napomena: Ukoliko u zadatku nešto nije dovoljno precizno definisano, student treba da uvede razumnu pretpostavku, da je uokviri (da bi se lakše prepoznala prilikom ocenjivanja) i da nastavi da izgrađuje preostali deo svog odgovora na temeljima uvedene pretpostavke. Ocenjivanje unutar potpitanja je po sistemu "sve ili ništa", odnosno nema parcijalnih poena. Kod pitanja koja imaju ponuđene odgovore treba **samo zaokružiti** jedan odgovor. Na ostala pitanja odgovarati **čitko, kratko i precizno**.

1. (10 poena) Data Mining

- a) (5) Dati skup podataka prikazuje vrednost temperature u zavisnosti od zadatih parametara. Sve numeričke vrednosti su normalizovane.

	Izgled	Vlažnost	Vetrovito	Temperatura
1.	sunčano	0.6	NE	1
2.	oblačno	0.7	NE	0.9
3.	kišovito	0.2	DA	0.3
4.	oblačno	0.3	DA	0.1
5.	kišovito	0.8	NE	0.8
6.	sunčano	0.4	DA	0.8
7.	kišovito	0.8	DA	0.5

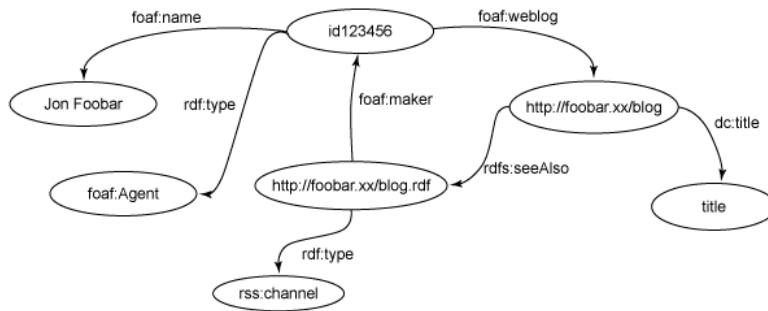
Za sledeće podatke izvršiti procenu vrednosti atributa *Temperatura* pomoću algoritma *k-Najbližih Suseda* (*k-Nearest Neighbors*). Uzeti da je $k=3$, kao meru distance koristiti euklidsko rastojanje za numeričke, a funkciju različitosti za kategoričke attribute.

a.)	kišovito	0.6	NE
-----	----------	-----	----

- b) (5) Koristeći C4.5 algoritam izgraditi stablo odlučivanja nad skupom podataka datim pod (a) koji u zavisnosti od zadatih parametara, predviđa da li će vreme biti vetrovito ili ne. Kontinualne vrednosti klasifikovati u 3 jednake kategorije (visoka, srednja i niska). Generisati stablo visine 1 (jedno grananje), a za svaki list dati meru njegove pouzdanosti (*confidence*) i zastupljenosti (*support*).

Rešenje:

2. (10 poena) Semantic Web



Na slici je prikazana struktura grafa jednog RDF dokumenta koji opisuje podatke vezane za urednike (*contributors*) jednog blog sajta.

- a) (3) Koji rezultat vraća sledeći SPARQL upit nad datim podacima?

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?url
WHERE {
    ?contributor foaf:name "Jon Fooobar" .
    ?contributor foaf:weblog ?url .
}
```

- b) (7) Napisati SPARQL upit koji za date podatke pronalazi ime osobe i naslov njenog blog-a, pri čemu je ta osoba tipa Agent i čija je URL adresa blog-a (weblog-a) <http://foobar.xx/blog> . (Nije potrebno navoditi tačan URL imenskog prostora svakog prefiksa)

Rešenje: