
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 2014.
Datum: 29.08.2014.

Kandidat: _____

Broj Indeksa: _____

Ispit traje 1,5 sat.

Dozvoljeno je korišćenje literature (jedan A4 rukom pisan papir) i kalkulatora.

| | | | |
|--------------------------|-----------|----------------------------|-----------|
| <i>Zadatak 1</i> | _____ /15 | <i>Usmeni</i> | _____ /20 |
| <i>Zadatak 2</i> | _____ /15 | <i>Projekat</i> | _____ /50 |
| <i>Ukupno na ispitu:</i> | _____ /30 | <i>Ukupno na projektu:</i> | _____ /50 |

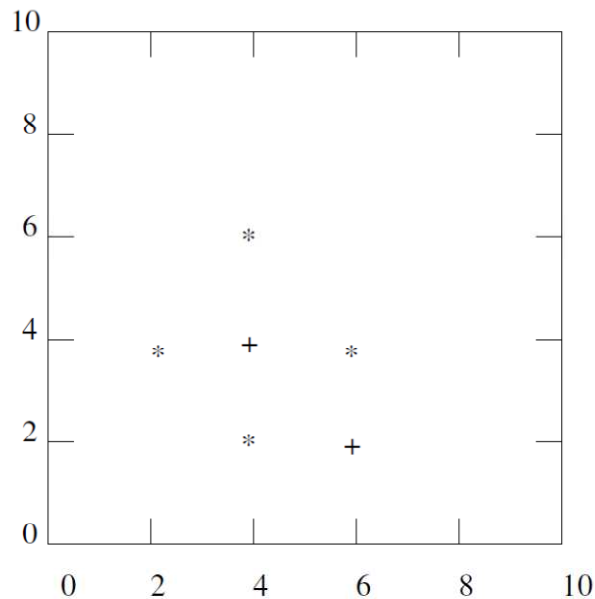
Ukupno: _____ /100

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. (15 poena) Data Mining

Na slici je predstavljen skup tačaka sa dva ishoda, gde * označava negativan, a + pozitivan ishod.



- a) (10) Predstaviti tačke sa slike tabelarno, a zatim koristeći C4.5 algoritam izgraditi stablo odlučivanja nad datim skupom podataka (za $x, y \in \{2, 4, 6\}$). U slučaju iste entropije, dati prednost atributu x .
- b) (5) Ukoliko je stvarni ishod, tj. verifikacioni skup opisan pomoću sledeće formule:

$$f(x_i, y_i) = \begin{cases} +, & \text{if } x_i = y_i \\ *, & \text{otherwise} \end{cases}$$

za generisano stablo odlučivanja pod a) odrediti tačnost (*Accuracy*) i preciznost (*Precision*):

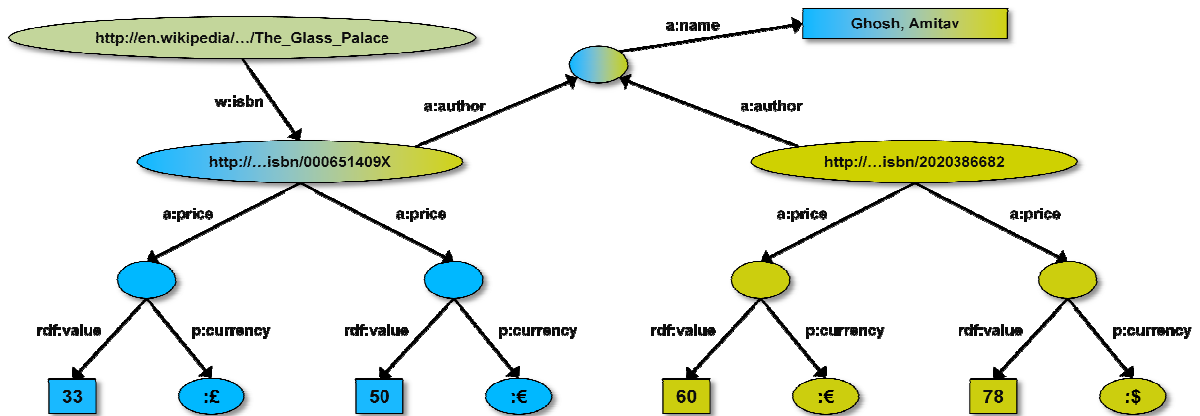
$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn}, \quad Precision = \frac{tp}{tp+fp}$$

t – true, f – false, p – positive, n – negative

Gde npr. tp označava tačno, a fp netačno prediktovanu pozitivnu vrednost.

Rešenje:

2. (15 poena) Semantic Web



Na slici je prikazana struktura grafa jednog RDF dokumenta koji opisuje podatke raspoloživih knjiga unutar jedne knjižare.

- (10) Napisati SPARQL upit koji za date podatke izlistava ISBN brojeve (`isbn`) i cenu (`price`) za knjige čije su cene date u valuti EUR (`:€`) i ukoliko postoji, dohvata Wikipedia link za datu knjigu.
- (5) Koji rezultat vraća ovaj SPARQL upit nad podacima sa slike?

Rešenje: