

Izpit pri predmetu Odločitveni sistemi (4.9.2012)

Navodila:

Izpit se piše 90 minut. Literatura ni dovoljena. Vrednost naloge je označena pri vsaki nalogi. Prepisovanje od kolegov je prepovedano in bo kaznovano z diskvalifikacijo.

Naloga 1 (25 točk): Nakup novih oken

Odločamo se za zamenjavo oken v hiši, izbrali smo nekaj potencialnih proizvajalcev in v tabeli predstavili njihove ponudbe (resnična imena proizvajalcev so izpuščena).

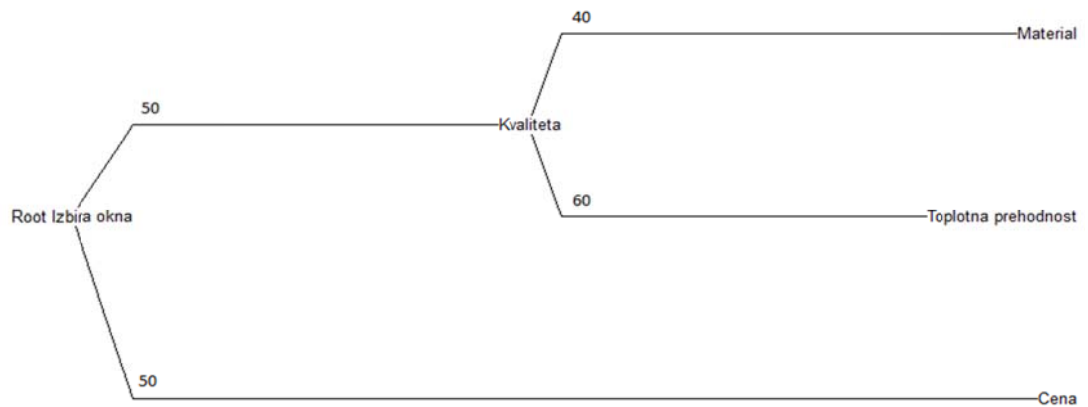
	Cena (EUR)	Kvaliteta materiala	Toplotna prehodnost (v W/m ² K)
Proizvajalec A	10.000,00 EUR	Dober	1.0
Proizvajalec B	9.500,00 EUR	Zelo dober	1.3
Proizvajalec C	10.500,00 EUR	Dober	1.2
Proizvajalec D	9.000,00 EUR	Povprečen	1.3

- a) Napišite tabelo "koristnosti". Odločili smo se, da dobi cena, ki je enaka ali manjša od 9.000,00 EUR koristnost 100, cene višje od 11.000,00 EUR dobijo 0, ostale so razporejene linearno. Podobno je pri toplotni prehodnosti, 1.0 ali manj je zelo dobro in dobi koristnost 100, več kot 1.5 je nesprejemljivo, vmes linearno. Za kvaliteto materiala smo določili fiksno lestvico in sicer: zelo dober dobi 100, dober 70 in povprečen 30.

	Cena (EUR)	Kvaliteta materiala	Toplotna prehodnost (v W/m ² K)
Proizvajalec A	50	70	100
Proizvajalec B	75	100	40
Proizvajalec C	25	70	60
Proizvajalec D	100	30	40

Tabela koristnosti.

- b) Izločite manjvredne variante.
Manjvredna varianta je od proizvajalca C, saj je po vseh kriterijih slabši od proizvajalca A.



- c) Kriterije smo uredili v hierarhično drevo. Uteži po posameznih vejah so napisane. Izberite najboljšo varianto glede na leksikografsko metodo.

Cena ima največjo globalno utež, zato je varianta od proizvajalca D po leksikografski metodi najboljša.

- d) Izračunajte normirane uteži za vse kriterije. Vse uteži naj se seštejejo v 1.

Cena: 0.5

Material: 0.2

Toplotna prehodnost: 0.3

- e) Izračunajte najboljšo varianto z uporabo linearne metode.

$$A: 0.5 \cdot 50 + 0.2 \cdot 70 + 0.3 \cdot 100 = 25 + 30 + 14 = 69$$

$$B: 0.5 \cdot 75 + 0.2 \cdot 100 + 0.3 \cdot 40 = 69.5$$

$$D: 0.5 \cdot 100 + 0.2 \cdot 30 + 0.3 \cdot 40 = 50 + 6 + 12 = 68$$

Najboljša varianta je B.

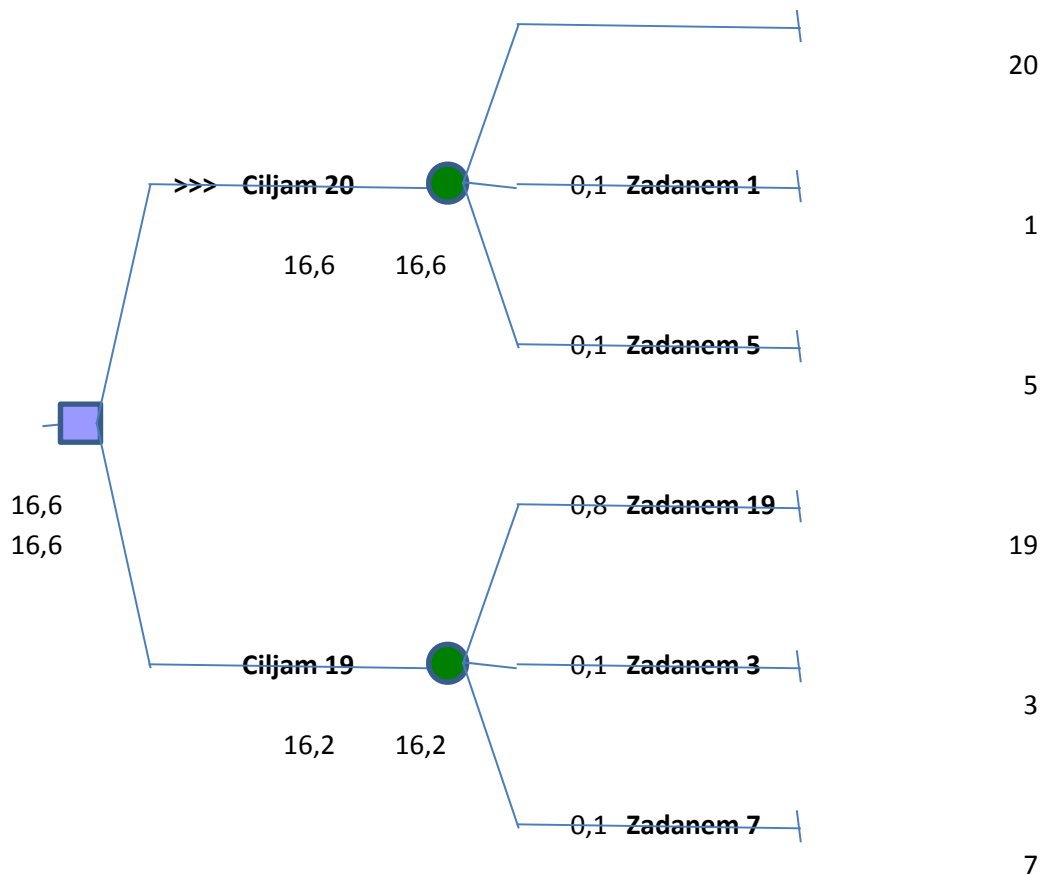
Naloga 2 (25 točk): Pikado

Pri pikadu večina ljudi cilja ali dvajsetico (20) ali devetnajstico (19).

Če ciljamo 20, lahko zgrešimo in zadanemo ali 1 ali 5. Naj bo verjetnost, da zgrešimo 20%. Kadar zgrešimo imata 1 in 5 enaki verjetnosti: obe 10%.

Če ciljamo 19 in zgrešimo, bomo zadeli ali 3 ali 7. Ker enako dobro ciljamo 19 in 20, je verjetnost, da zgrešimo, enaka.

- a) Narišite odločitveno drevo. Katero vrednost se nam bolj splača ciljati?



Bolj se splača ciljati 20, saj bomo v povprečju zadeli 16.6 točke.

- b) Naredite analizo občutljivosti glede na verjetnost zgrešitve. Pri kakšni verjetnosti zgrešitve bi se naša odločitev spremenila?

Dovolj je, da odkrijemo mejo pri kateri se odločitev spremeni. To se zgodi, ko sta vrednosti odločitvenih vej enaki. Naj bo "p" verjetnost zgrešitve, potem je

$$20 * (1-p) + 1 * p/2 + 5 * p/2 = 19 * (1-p) + 3 * p/2 + 7 * p/2$$

$$1-p = 2p$$

$$p = 1/3$$

Če je verjetnost, da zgrešimo večja kot 1/3, se nam bolj splača ciljati 19.

Naloga 3 (40 točk): Priporočila

V tabeli so ocene uporabnikov za nekaj izdelkov.

	Izdelek 1	Izdelek 2	Izdelek 3	Izdelek 4
Uporabnik X	5	3	4	?
Uporabnik 2	3	3	3	3
Uporabnik 3	4	1	3	4
Uporabnik 4	4	2	5	5
Uporabnik 5	1	4	2	1

- a) Izračunajte povprečno oceno za posamezne uporabnike. Kateri uporabnik daje v povprečju najvišjo oceno?
 $\text{avg}(U_X) = 4$; $\text{avg}(U_2) = 3$; $\text{avg}(U_3) = 3$; $\text{avg}(U_4) = 4$; $\text{avg}(U_5) = 2$. Najvišjo oceno daje uporabnik X in 4.
- b) Naj bo izračunana povprečna ocena kar napoved priporočilnega sistema za posameznega uporabnika. Kakšno MAE napako bi tak sistem naredil pri uporabniku 5?
 Tak sistem bi za uporabnika 5 vedno napovedal 2. MAE je v tem primeru:
 $\text{MAE} = (1+2+0+1)/4 = 1$.
- c) Ocenite podobnosti med produktom 4 in ostalimi produkti. Uporabite popravljeno cosinusno mero za računanje podobnosti. Formula je (a in b sta dva različna produkta, U je množica vseh uporabnikov):

$$\text{sim}(\vec{a}, \vec{b}) = \frac{\sum_{u \in U} (r_{u,a} - \bar{r}_u)(r_{u,b} - \bar{r}_u)}{\sqrt{\sum_{u \in U} (r_{u,a} - \bar{r}_u)^2} \sqrt{\sum_{u \in U} (r_{u,b} - \bar{r}_u)^2}}$$

$$\text{sim}(I_1, I_4) = ((3-3)*(3-3)+(4-3)*(4-3)+(4-4)*(5-4)+(1-2)*(1-2)) / (\text{sqrt}(2)*\text{sqrt}(3)) = 2/\text{sqrt}(6)$$

$$\text{sim}(I_2, I_4) = ((3-3)*(3-3)+(1-3)*(4-3)+(2-4)*(5-4)+(4-2)*(1-2)) / (\text{sqrt}(12)*\text{sqrt}(3)) = -6/6 = -1$$

$$\text{sim}(I_3, I_4) = ((3-3)*(3-3)+(3-3)*(4-3)+(5-4)*(5-4)+(2-2)*(1-2)) / (\text{sqrt}(1)*\text{sqrt}(3)) = 1/\text{sqrt}(3)$$

- d) Na podlagi "item-based collaborative filtering" tehnike napovedajte kakšno oceno bi uporabnik X dal izdelku 4. Za mero podobnosti uporabite podobnosti izračunane v prejšnji točki. Upoštevajte vse produkte z nenegativno podobnostjo.

$$\text{pred}(U_X, I_4) = (5 * 2/\text{sqrt}(6) + 4 * 1/\text{sqrt}(3)) / (2/\text{sqrt}(6) + 1/\text{sqrt}(3)) = 4.58$$

- e) Z uporabo tehnike "Slope one" napovedajte kakšno oceno bi dal uporabnik X izdelku 4?

$$\text{dev}(4,1) = (0+0+1+0)/4 = 1/4$$

$$\text{dev}(4,2) = (0+3+3-3)/4 = 3/4$$

$$\text{dev}(4,3) = 0$$

$$\text{pred}(U_X, 4) = ((5+1/4)*4 + (3+3/4)*4 + (4+0)*4) / 12 = 4.33$$

Naloga 4 (10 točk): Hladni zagon

- a) Kakšen problem označujemo z izrazom »hladni zagon« (angl. cold start)?
- b) Kako so na hladni zagon občutljive »collaborative« in kako »content-based« tehnike?