

Izpit pri predmetu Odločitveni sistemi (13.2.2013)

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 (30 točk): Izbira SSD diska

V tabeli smo zbrali podatke o SSD diskih (vrednosti so približne):

	Kapaciteta[Gb]	Cena [EUR]	Hitrost pisanja [Mb/s]	Hitrost branja [Mb/s]
Samsung 840	128	140	550	400
Intel 330	180	130	500	450
Plextor M5	128	150	550	350
Crucial M4	128	100	400	200

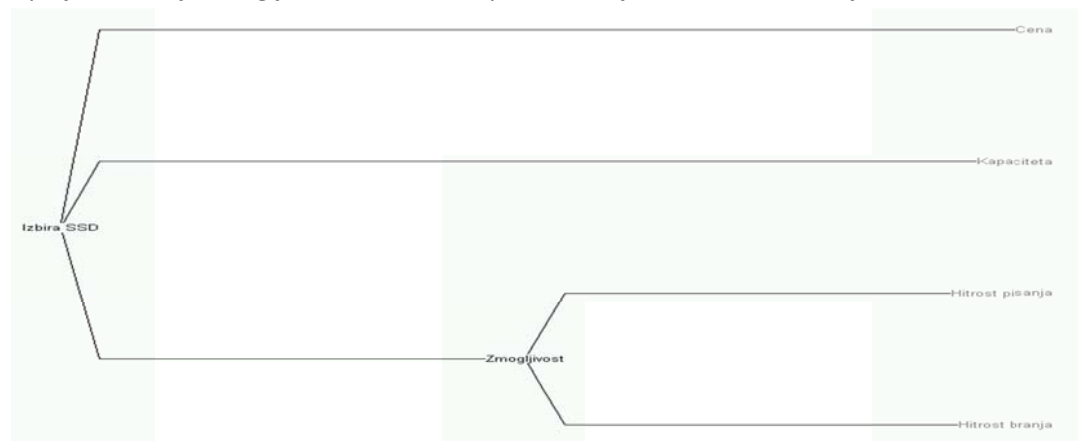
- a) (5 točk) Izločite manjvredne variante.

Plextor M5 (slabša od Samsung 840)

- b) (5 točk) Izračunajte koristnosti posameznih vrednosti. Povsod uporabite relativno skalo.

	Kapaciteta[Gb]	Cena [EUR]	Hitrost pisanja [Mb/s]	Hitrost branja [Mb/s]
Samsung 840	0	0	100	80
Intel 330	100	25	67	100
Crucial M4	0	100	0	0

- c) (5 točk) Narišite hierarhičen model kriterijev, kjer hitrost pisanja in hitrost branja združite v izpeljan kriterij "zmogljivost". Cena in kapaciteta naj ostaneta ločeni veji.



- d) (5 točk) Izberite najboljšo varianto, če imajo kapaciteta, cena in zmogljivost enako utež.

Kriterija "hitrost pisanja" in "hitrost branja" sta enako pomembna.

Samsung: $100 \cdot \frac{1}{6} + 80 \cdot \frac{1}{6} = 30$

Intel: $100 \cdot \frac{1}{3} + 25 \cdot \frac{1}{3} + 67 \cdot \frac{1}{6} + 100 \cdot \frac{1}{6} = 69.5$

Crucial: $100 \cdot \frac{1}{3} = 33$

e) (10 točk) Na sliki je prikazana analiza občutljivosti za ceno.

a. Kaj lahko ugotovite glede variant 1 in 3?

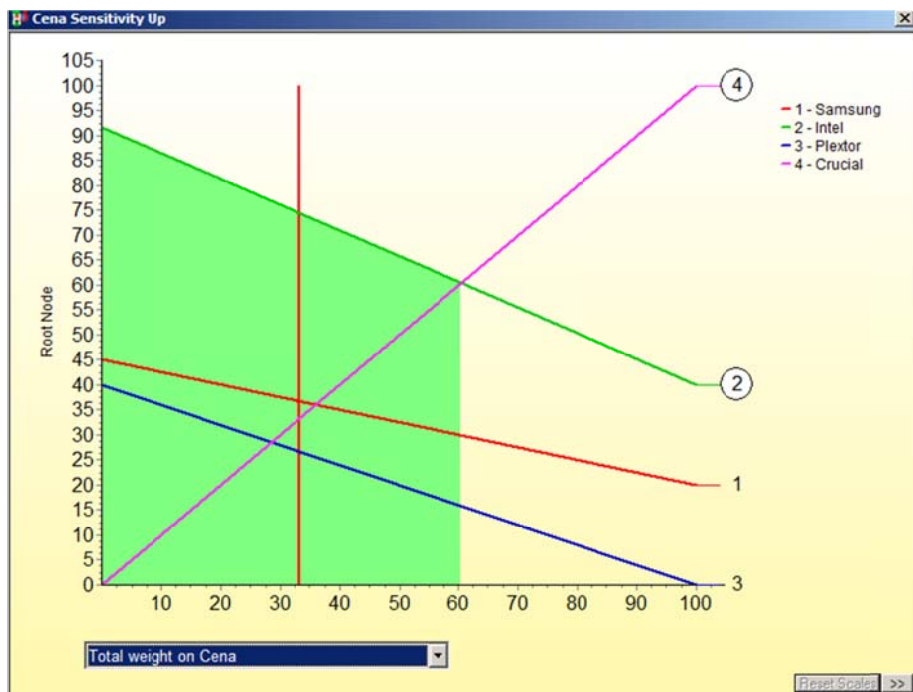
Varianta 1 bo boljša od variante 3 ne glede na to, kakšno utež bomo nastavili za ceno.

b. Pri kateri uteži se spremeni najboljša varianta?

0.6

c. Ali je model občutljiv? Zakaj?

Model ni občutljiv, saj je potrebna velika sprememba uteži, da se spremeni najboljša varianta.

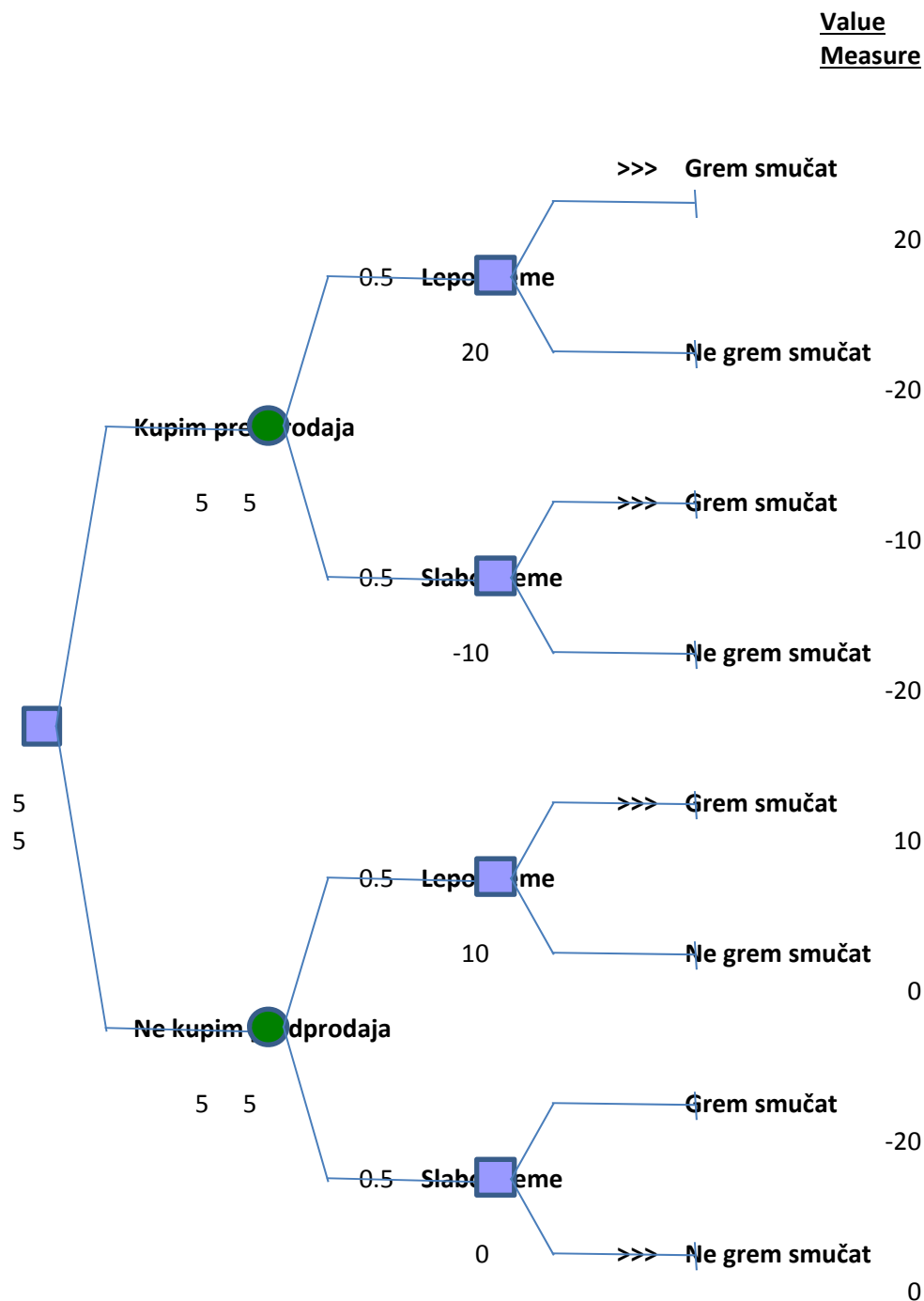


Naloga 2 (25 točk): smučanje

Jutri, po zadnjem izpitu, bi radi šli smučat na bližnje smučišče. Odločiti se moramo, ali naj karto kupimo danes v predprodaji ali ne. Problem je, ker ne vemo, kakšno vreme bo jutri. Karta v predprodaji stane 20EUR, na smučišču (na dan smučanja) pa 30 EUR.

Kako se bomo odločili, če želimo maksimizirati naš užitek? Vrednost užitka v lepem vremenu ocenimo na 40 EUR, če je grdo pa na le 10 EUR. Pozor: jutri se še vedno lahko odločimo, da ne gremo smučat.

- a) (10 točk) Narišite odločitveno drevo. Verjetnost za lepo vreme označite s spremenljivko P .



- b) (5 točk) Kakšna mora biti verjetnost za lepo vreme, da se splača kupiti karto v predprodaji? Kot označeno v drevesu, smo pri $P=0.5$ indiferentni glede nakupa v predprodaji. Pri $P>0.5$ se bolj splača kupiti v predprodaji.
- c) (10 točk) Predpostavimo, da je verjetnost za lepo vreme 70%. Koliko je vredna informacija o vremenu za naslednji dan, če nam jo posreduje zelo sposoben meteorolog, ki se nikoli ne zmoti?
Če je $P=70\%$, potem je vrednost drevesa: 11. Če bi poznali informacijo o vremenu za naslednji dan, potem:

Če je slabo vreme, se nam ne splača smučati, saj je vrednost smučanja negativna. Če bo lepo, se nam splača smučati in seveda kupiti karto v predprodaji; vrednost bo 20.

$$\text{Ocena} = 0,3 * 0 + 0,7 * 20 = 14.$$

Vrednost informacije o vremenu je $14-11=3$ EUR.

Naloga 3 (35 točk): Priporočanje filma

V spodnji tabeli so navedene ocene uporabnikov za nekaj filmov. Imamo dve možni oceni:

1 (film mu je bil všeč) in 0 (film mu ni všeč).

	12 opic	Titanik	Gladiator	4 poroke in pogreb	Tolpe New Yorka	Ana Karenina
Miha	1	0	1	0	1	?
Maja	1	1	?	1	0	0
Mojca	1	?	1	0	?	1
Marko	1	1	1	1	1	1

- a) (20 točk) Na podlagi »user-based collaborative filtering« tehnike napovejte kakšno oceno bi Miha dal filmu Ana Karenina? Za mero podobnosti uporabite Pearsonov količnik. Upoštevajte vse uporabnike s pozitivnim količnikom.

$\text{sim}(\text{Miha}, \text{Maja}) < 0$ (ocenjujeta praktično nasprotno).

$$\text{sim}(\text{Miha}, \text{Mojca}) = ((1-3/5)*(1-3/4) + (1-3/5)*(1-3/4) + (0-3/5)*(0-3/4)) / (4/25 + 4/25 + 9/25) / (1/16 + 1/16 + 9/16) = 0.95$$

$\text{sim}(\text{Miha}, \text{Marko}) = 0$ (Marko ocenjuje konstantno).

$$\text{pred}(\text{Miha}, \text{Ana Karenina}) = \text{avg}(\text{Miha}) + (\text{ocena}(\text{Mojca}, \text{Ana Karenina}) - \text{avg}(\text{Mojca})) * \text{sim}(\text{Miha}, \text{Mojca}) / \text{sim}(\text{Miha}, \text{Mojca}) = 3/5 + (1-3/4) * 0.955 / 0.955 = \mathbf{0.85}$$

- b) (15 točk) Ali bi priporočilni sistem, ki izbira glede na vsebino in uporablja metodo naivni Bayes, Mihi priporočil film Ana Karenina? Filme opišemo z žanri:

	drama	akcija	romantika	zgodovinski
12 opic	DA	NE	NE	NE
Titanik	DA	NE	DA	DA
Gladiator	DA	DA	DA	DA
4 poroke in pogreb	NE	NE	DA	NE
Tolpe New Yorka	DA	DA	NE	DA
Ana Karenina	DA	NE	DA	DA

1. Naredimo model naivnega Bayesa:

$$P(\text{like}) = 3/5$$

$$P(\text{drama}=\text{DA} \mid \text{like}) = 1 \quad P(\text{drama}=\text{DA} \mid \text{dislike}) = 1/2$$

$$P(\text{akcija}=\text{DA} \mid \text{like}) = 2/3 \quad P(\text{akcija}=\text{DA} \mid \text{dislike}) = 0$$

$$P(\text{romantika}=\text{DA} \mid \text{like}) = 1/3 \quad P(\text{romantika}=\text{DA} \mid \text{dislike}) = 1$$

$$P(\text{zgodovinski}=\text{DA} \mid \text{like}) = 2/3 \quad P(\text{zgodovinski}=\text{DA} \mid \text{dislike}) = 1/2$$

2. Napovejmo za Ano Karenino:

$$\text{Like: } P(\text{like}) * P(\text{drama}=\text{DA} \mid \text{like}) * (1 - P(\text{akcija}=\text{DA} \mid \text{like})) * P(\text{romantika}=\text{DA} \mid \text{like}) *$$

$$P(\text{zgodovinski}=\text{DA} \mid \text{like}) = 3/5 * 1/3 * 1/3 * 2/3 = 2/45$$

$$\text{Dislike: } P(\text{dislike}) * P(\text{drama}=\text{DA} \mid \text{dislike}) * (1 - P(\text{akcija}=\text{DA} \mid \text{dislike})) * P(\text{romantika}=\text{DA} \mid \text{dislike}) *$$

$$P(\text{zgodovinski}=\text{DA} \mid \text{dislike}) = 2/5 * 1/2 * 1 * 1 * 1/2 = 1/10$$

Večja vrednost pride za dislike, zato filma Mihi ne bi priporočali.

Naloga 4 (10 točk): Kako bi združili metodi iz 3. naloge v hibridni priporočilni sistem?

Napišite, kateri način za združevanje se vam zdi najbolj primeren in na kratko opišite, kako bi ga implementirali.