

## Izpit pri predmetu Odločitveni sistemi (30.1.2012)

### 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 (30 točk): Datoteke v oblakih

Izbiramo brezplačno storitev za shranjevanje podatkov v oblaku (cloud). Spodnja tabela prikazuje nekaj takih storitev in njihovih karakteristik.

	Prostor (brezplačno)	Tip datotek	Enostavnost uporabe	Kompatibilnost	Poslušanje glasbe
Dropbox	2Gb	Vse	70	Povsod	NE
Skydrive	25Gb	Vse	40	PC, MAC	NE
Amazon cloud	5Gb	Vse	40	Android, PC, MAC	DA
Google	9Gb	Vse	10	Povsod	DA
Apple iCloud	5Gb	Vse	100	PC, MAC, iOS	DA

- a) Napišite tabelo "koristnosti". V prvem stolpcu (prostor) naj bodo koristnosti linearne med 0 in 5 (0 dobi koristnost 0, 5 dobi koristnost 100). Vse višje vrednosti dobijo koristnost 100. Enostavnost uporabe je že podana kot koristnost, pri kompatibilnosti vsak sistem prinese 25 točk in poslušanje glasbe dobi 0 ali 100 točko. Izločite tudi nepotrebne kriterije in manjvredne variante.

	Prostor (brezplačno)	Enostavnost uporabe	Kompatibilnost	Poslušanje glasbe
Dropbox	40	70	100	0
Google	100	10	100	100
Apple iCloud	100	100	75	100

Skydrive in Amazon cloud sta manjvredni varianti.

- b) Izberite najboljšo varianto glede na leksikografsko metodo (vrstni red: prostor, glasba, enostavnost uporabe, kompatibilnost, tip datotek).  
Apple iCloud
- c) Uteži za kriterije so: prostor (100), tip (10), enostavnost(50), kompatibilnost(30) in glasba(20). Katera varianta je najboljša glede na linearno odločitveni model. Uteži normirajte!  
uteži:  
prostor =  $100/200 = 0.5$   
enostavnost = 0.25  
kompatibilnost = 0.15  
glasba = 0.1

Dropbox:  $0.5 \cdot 40 + 0.25 \cdot 70 + 0.15 \cdot 100 + 0.1 \cdot 0 = 52.5$

Google:  $0.5 \cdot 100 + 0.25 \cdot 10 + 0.15 \cdot 100 + 0.1 \cdot 100 = 77.5$

Apple:  $0.5 \cdot 100 + 0.25 \cdot 100 + 0.15 \cdot 75 + 0.1 \cdot 100 = 96.5$

Apple najboljši.

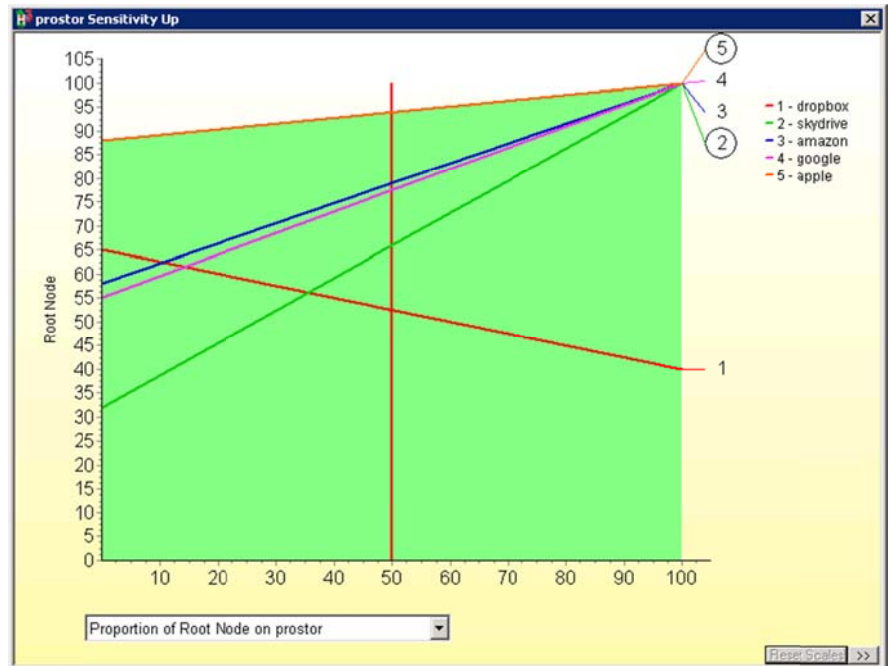
- d) Kaj bi se moralo spremeniti pri Dropbox-u, da bi postal najboljša varianta?

Dropbox mora dobiti 43 točk. Različne spremembe možne (ampak praktično vse).

- e) Desno je slika občutljivosti glede na kriterij "prostor". Kakšna bi morala biti normirana utež (med 0 in 1), da bi Dropbox postal najboljša varianta? Ocenite le približno.

Nikoli.

- f) Kaj bi se zgodilo s premico 5 (Apple), če bi ponujal le še 3Gb brezplačno? Kje bi se začela in kje končala? Začela isto, končala pa pri 60.

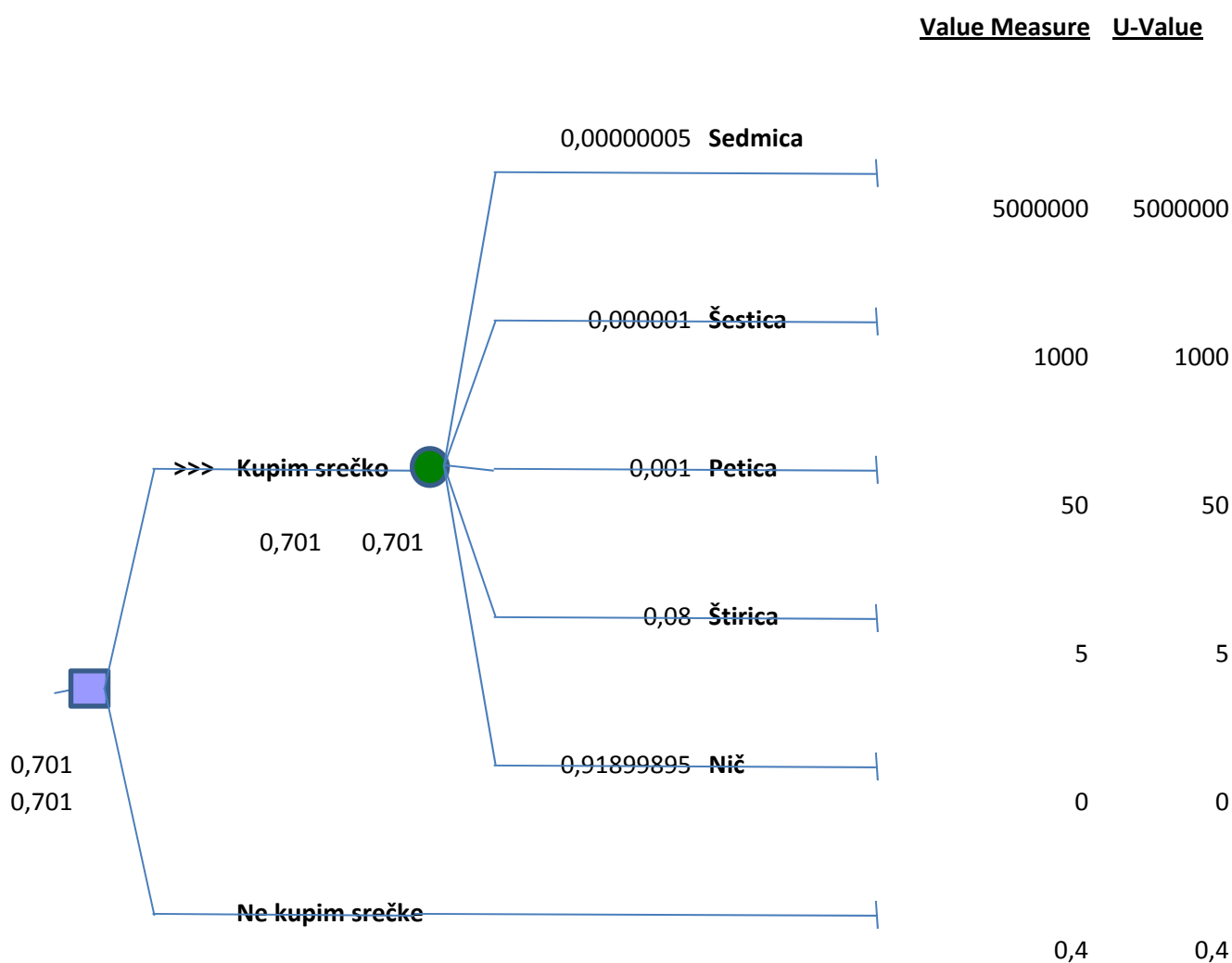


## Naloga 2 (15 točk): Loto

Na srečko loto, ki stane 0.40 EUR, ob nakupu napišemo 7 števil (med 1 in 39). Stroj potem naključno iz bobna izžreba 7 števil in mi, glede na ujemanje naših izbranih števil in naključno izžrebanih, dobimo ustrezno nagrado. Recimo, da za sedmico (vsa števila se ujemajo) dobimo 5mio EUR nagrade, za šestico 1000 EUR, za petico 50 EUR in štirico 5 EUR.

Naj bo verjetnost za zadetek sedmice na lotu  $p_7 = 5 \cdot 10^{-8}$ . Verjetnost za zadetek šestice je  $p_6 = 1 \cdot 10^{-6}$ , za petico je 0.001 in za štirico je 0.08. To so približne verjetnosti.

a) Narišite odločitveno drevo.



b) Izračunajte pričakovano nagrado. Ali se nam splača kupiti srečko ali ne?

Pričakovana nagrada  $\sim 0,7$  EUR. Ker je cena 0.4 EUR, se nam splača kupiti.

c) Kako bi morala biti določena glavna nagrada, da se nam še splača kupiti srečko? Vse ostale nagrade naj ostanejo enake.

Ni pomembno, kakšna je nagrada, saj že samo nagrada za štirico ( $0.8 \cdot 5$ ) odtehta strošek srečke. Vedno se splača kupiti srečko.

### Naloga 3 (15 točk): Nakupi

Iz baze spletne trgovine smo dobili naslednje podatke o nakupih. Znak X pomeni, da je kupec produkt kupil.

	prenosnik	miška	miza	stol
kupec 1	X	X		X
kupec 2			X	X
kupec 3	X	X		
kupec 4		X	X	X
kupec 5	X	X		
kupec 6			X	X
kupec 7	X			X

- a) z Jaccardovo mero podobnosti izračunajte podobnost med prenosnikom in stolom.

$$\text{sim}(\text{stol}, \text{prenosnik}) = 2/7$$

- b) Poiščite najbolj podobna produkta (glede na Jaccarda).

$$\text{sim}(\text{stol}, \text{miška}) = 2/7$$

$$\text{sim}(\text{stol}, \text{miza}) = 3/5$$

$$\text{sim}(\text{miza}, \text{prenosnik}) = 0$$

$$\text{sim}(\text{miza}, \text{miška}) = 1/6$$

$$\text{sim}(\text{miška}, \text{prenosnik}) = 3/5$$

Najbolj podobna sta dva para: stol & miza ter miška in prenosnik.

- c) Napišite eno povezovalno pravilo, ki ima zaupanje 1 in podporo vsaj 0.4.

IF miza THEN stol.

#### Naloga 4 (25 točk): Priporočaj članek

Imate "content-based" priporočilni sistem, ki gradi modele svojih uporabnikov z naivnim Bayesovim klasifikatorjem na podlagi predizbranih atributov (evro, sedmica, rokomet in BMW). Spodaj je podana tabela zbranih podatkov (ocen like/dislike in poleg za vsak atribut 0 = te besede ni v povzetku ali 1 = ta beseda je v povzetku) za tri uporabnike ter trije povzetki člankov.

uporabnik	evro	sedmica	rokomet	BMW	like/dislike
Valerija	0	0	1	0	like
Valerija	1	0	0	1	dislike
Tanja	1	1	0	1	like
Tanja	0	1	0	1	like
Tanja	1	0	1	0	like
Tanja	1	1	0	0	dislike
Tanja	0	0	1	1	dislike
Tanja	1	0	0	0	dislike
Aleksej	1	0	1	0	like
Aleksej	1	1	0	0	dislike
Aleksej	0	1	0	0	like

**povzetek 1:** »Sedmica je izžrebana. Dobitnik je zmagovalni listič vplačal v Mariboru. Izžrebane številke so 14, 16, 19, 20, 27, 31 in 38. Dodatna številka je 9. Dobitnik je bogatejši za 4.872.229,44 evra.«

**povzetek 2:** »Danska moška rokometna reprezentanca je osvojila zlato kolajno na letošnjem evropskem prvenstvu v Srbiji, potem ko je v finalni tekmi v Beogradu premagala domačo izbrano vrsto z 21:19.«

**povzetek 3:** »Fotografi so ujeli BMWjevo sedmico med zimskimi testiranj. Facelift verzija prihaja z očitno spremenjenimi odbijači in seveda hitom poletja, LED tehnologijo – v vsem dobro znanih, Angel Eyesih.«

- (a) Zapišite povzetke vseh treh člankov v atributni obliki, primerni za nadaljnje napovedovanje. Pri tem upoštevajte le smiselne korene besed (atributov).

Članek	evro	sedmica	rokomet	BMW
povzetek 1	1	1	0	0
povzetek 2	0	0	1	0
povzetek 3	0	1	0	1

- (b) Sestavite naivni Bayesov klasifikator za uporabnico Tanjo.

$$P(\text{like}) = P(\text{dislike}) = 0.5$$

$$P(\text{evro} | \text{like}) = 2/3$$

$$P(\text{sedmica} | \text{like}) = 2/3$$

$$P(\text{rokomet} | \text{like}) = 1/3$$

$$P(\text{BMW} | \text{like}) = 2/3$$

$$P(\text{evro} | \text{dislike}) = 2/3$$

$$P(\text{sedmica} | \text{dislike}) = 1/3$$

$$P(\text{rokomet} | \text{dislike}) = 1/3$$

$$P(\text{BMW} | \text{dislike}) = 1/3$$

- (c) Za vse tri članke na podlagi modela iz točke (b) in njihovih povzetkov povejte ali bi jih priporočilni sistem Tanji priporočil ali ne.

$$P(\text{like} | \text{povzetek1}) = P(\text{like}) * P(\text{evro} | \text{like}) * P(\text{sedmica} | \text{like}) * (1 - P(\text{rokomet} | \text{like})) * (1 - P(\text{BMW} | \text{like})) = 0.5 * 2/3 * 2/3 * 2/3 * 1/3$$

$$P(\text{dislike} | \text{povzetek1}) = 0.5 * 2/3 * 1/3 * 2/3 * 2/3$$

$$P(\text{like} | \text{povzetek2}) = 0.5 * 1/3 * 1/3 * 1/3 * 2/3$$

$$P(\text{dislike} | \text{povzetek2}) = 0.5 * 1/3 * 2/3 * 1/3 * 2/3$$

Ne predlaga povzetka 2!

$$P(\text{like} | \text{povzetek3}) = 0.5 * 1/3 * 2/3 * 2/3 * 2/3$$

$$P(\text{dislike} | \text{povzetek3}) = 0.5 * 1/3 * 1/3 * 2/3 * 1/3$$

Predlaga povzetek 3.

### Naloga 5 (10 točk)

Na kratko skicirajte kako bi implementirali razlago odločitev v priporočilnem sistemu iz prejšnje naloge (NB CB PS).

Na kratko opišite en primer hibridnega priporočilnega sistema.