

Poslovna inteligenca

1. izpitni rok

25. januar 2013

Priimek in ime (tiskano): _____

Vpisna številka: _____

Naloga	1	2	3	4	5	Vsota
Vrednost	10	5	10	5	8	38
Točk						

- [10] 1. Naš avto potrebuje nove zimske gume. Ob pregledu rezultatov nemškega testa za gume ustrezne velikosti smo izbor zožili na 5 variant, ki so dobile oceno *dobra guma*.

Končno oceno gum na različnih podlagah želimo izračunati tako, da ima suha podlaga 40% utež, ostale podlage pa imajo med seboj enake uteži. Oceno zaokrožimo navzgor na 1 decimalno mesto.

guma	cestišče				cena
	suho	mokro	sneg	led	
C - W.C. TS850	2,4	2,3	1,9	2,0	107,60 EUR
M - A. A4	2,2	2,4	2,0	2,1	111,30 EUR
D - SP W.S. 4D	2,5	2,2	2,3	2,4	98,60 EUR
G - U.G. 8	2,5	2,1	2,3	2,3	90,70 EUR
N - WR D3	2,5	2,5	2,3	2,4	92,90 EUR

Skladno z danimi funkcijami koristnosti pretvorimo oceno na cestiščih in najugodnejšo ceno, ki smo jo našli na spletu, v razrede.

Cestišče	razred	Cena	razred
ocena < 2,3	odl	cena < 95 EUR	odl
2,3 <= ocena < 2,6	dobr	95 EUR <= cena < 100 EUR	dobr
2,6 <= ocena < 2,7	spr	100 EUR <= cena < 105 EUR	spr
2,9 <= ocena	nespr	105 EUR <= cena	nespr

Oba kriterija združimo v končno oceno po pravilih iz programa DEXi v naslednji tabeli:

	Cestišča	Cena	Guma
	56%	44%	
1	nespr	*	nespr
2	*	nespr	nespr
3	spr	>=spr	spr
4	dobr	spr:dobr	dobr
5	>=dobr	spr	dobr
6	>=dobr	odl	odl
7	odl	>=dobr	odl

Katero gumo boste izbrali in zakaj?

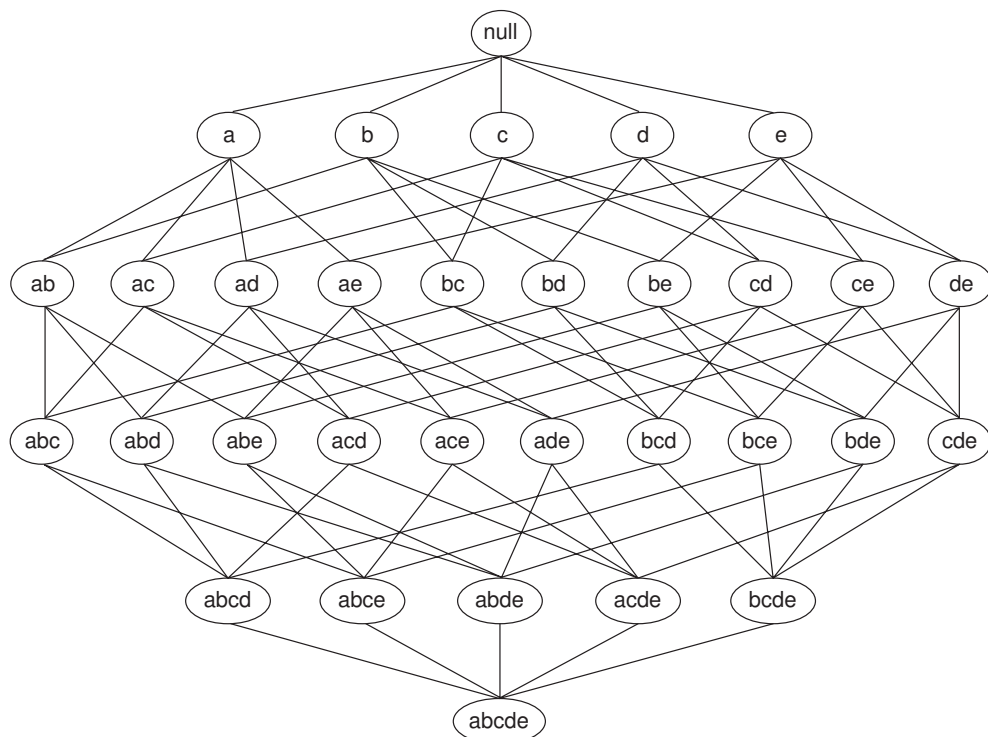
Stran je prazna, da lahko nanjo rešujete nalogo.

[5] 2. Dani so transakcijski podatki v obliki nakupovalnih košaric:

ID	kupljeni izdelki
1	{a, b, d, e}
2	{b, c, d}
3	{a, b, d, e}
4	{a, c, d, e}
5	{b, c, d, e}
6	{b, d, e}
7	{c, d}
8	{a, b, c}
9	{a, d, e}
10	{b, d}

Algoritem *Apriori* gradi skupine izdelkov velikosti $k + 1$ z združevanjem pogostih skupin (frequent itemsets) velikosti k . Za algoritem *Apriori* s spodnjo mejo podpore (minimum support) 30% na skici označi skupine izdelkov s črko:

- N, če *Apriori* za to skupino izdelkov sploh ne računa podpore.
- F, če *Apriori* računa podporo za to skupino izdelkov in ugotovi, da je večja ali enaka minimalni.
- I, če *Apriori* računa podporo za to skupino izdelkov in ugotovi, da je manjša od minimalne.



$$\sigma(X) = |\{t_i | X \subseteq t_i, t_i \in T\}| \quad s(X \rightarrow Y) = \sigma(X \cup Y)/N \quad c(X \rightarrow Y) = \sigma(X \cup Y)/\sigma(X)$$

3. Dana je množica primerov $(a \dots e)$, ki so opisani s štirimi atributi $(x_1 \dots x_4)$.

	x_1	x_2	x_3	x_4
a	1	0	0	2
b	2	0	3	4
c	0	0	1	2
d	4	1	2	1
e	0	1	4	3

Primere bi želeli predstavili kot točke v Evklidski ravnini.

- [2] (a) Predlagajte kriterijsko funkcijo, ki jo naj optimiziramo pri taki predstavitvi (projekciji).
- [6] (b) Predlagajte primerno projekcijo in postavitev točk primerno utemeljite. Pri tem si pomagajte z matriko razdalj med primeri.
- [1] (c) Kako se imenuje algoritem, s katerim lahko nalogo te vrste rešite programsko?
- [1] (d) Kakšne so prednosti take predstavitve pred predstavitvijo dendrograma hierarhičnega razvrščanja?

4. Pri vpeljavi linearne in logistične regresije smo omenjali regularizacijo.

- [2] (a) Čemu služi regularizacija?
- [1] (b) Čemu je enak regularizacijski člen pri logistični in linearni regresiji?
- [2] (c) Stopnjo regularizacije določimo s parametrom λ . Dana je množica učnih podatkov. Na kakšen način izberemo ustrezno vrednost tega parametra?

5. Pet prijateljev smo povprašali, katere športe imajo najraje. Odgovore podaja spodnja tabela.

	tango	salsa	plezanje	smučanje	bordanje	plavanje	kolesarjenje
Agnieszka			♡		♡	♡	♡
Marko	♡		♡		♡	♡	♡
Nejc		♡		♡			♡
Sara			♡	♡	♡		♡
Urška		♡			♡		

- [4] (a) Agnieszka želi poskusiti nekaj novega. Pomagajte ji: razvrstite tango, salso in smučanje glede na njene pričakovane preference. Utemeljite izbran postopek.
- [4] (b) Bi z večjo gotovostjo priporočili smučanje Agnieszki ali Urški? Zakaj?

$$s_c(u, u') = \frac{\mathbf{r}_u \mathbf{r}_{u'}}{|\mathbf{r}_u| |\mathbf{r}_{u'}|} \quad d_e(u, u') = |\mathbf{r}_u - \mathbf{r}_{u'}| \quad s_e(u, u') = \frac{1}{1+d_e(u, u')} \quad s_J = \frac{\|\mathbf{r}_u \cap \mathbf{r}_{u'}\|}{\|\mathbf{r}_u \cup \mathbf{r}_{u'}\|}$$