

Data mining v Organonu

Ing. Jakub Gerlich
Mgr. Ludmila Dostálová, Ph.D.

Ostrava
23. únor 2015

Podpořeno projektem CZ.1.07/2.2.00/28.0209: Elektronické opory a e-learning pro obory výpočtového a konstrukčního charakteru



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Obsah

1. Data získaná ze systému Organon
2. Možnosti použití

Data získaná ze systému Organon

- Rozdíly oproti systémům eLogika a Barborka:
 - Otázky jsou pouze tvořivé
 - Je možné evidovat postup řešení
 - Nejsou zde připravené odpovědi
 - Jiná struktura databáze
- Nová evidovaná data:
 - Postup řešení
 - Výsledky jednotlivých částí řešení úkolu
 - Každá otázka má určitý **typ**, který je spojený s určitou typologií a udává, jaké dílčí úkoly (prvky typologie) musí student při řešení absolvovat (odstranění implikace, aplikace de morganova pravidla, zjednodušení, ...)

Možnosti použití

- Statistické metody, Shlukování, Rozhodovací stromy, Asociační pravidla
 - Výsledky kurzu, testů, otázek
 - Výsledky z typů, kategorií
 - Výsledky z dílčích částí řešení
- Analýza časových řad
 - Postup řešení

Shlukování

Support	Mean (odstranění spojek de morgan distributivita $p+-p=1$ $p*-p=0$)
0,08 (50 from 665)	(1 1 --- 1)
0,07 (46 from 665)	(0 0 --- ---)
0,06 (37 from 665)	(0 0 0 0)
0,03 (22 from 665)	(1 0.60 - 0.8 --- 1)
0,03 (21 from 665)	(1 0.60 - 0.8 --- 0.40 - 0.6)
0,03 (17 from 665)	(0.40 - 0.6 0.40 - 0.6 --- 0)
0,02 (12 from 665)	(1 0.80 - 1.0 1 1)
0,02 (12 from 665)	(1 0.60 - 0.8 0.40 - 0.6 1)
0,02 (11 from 665)	(0.60 - 0.8 0.60 - 0.8 --- 1)
0,02 (11 from 665)	(1 1 1 1)
0,02 (10 from 665)	(0.60 - 0.8 0.40 - 0.6 --- 0.40 - 0.6)
0,02 (10 from 665)	(1 0.40 - 0.6 --- 0.40 - 0.6)
0,02 (10 from 665)	(1 0.80 - 1.0 0.40 - 0.6 1)
0,02 (10 from 665)	(1 0.60 - 0.8 0.40 - 0.6 0.60 - 0.8)
0,01 (9 from 665)	(0.80 - 1.0 0.60 - 0.8 0.40 - 0.6 0.60 - 0.8)
0,01 (9 from 665)	(1 0.40 - 0.6 --- 0)
0,01 (9 from 665)	(1 0.80 - 1.0 1 0.60 - 0.8)

qROCK s threshold=1 - studenti musí mít stejné výsledky, aby byli ve stejném shluku

Shlukování

Support	Mean (odstranění spojek de morgan distributivita $p+-p=1$ $p^*-p=0$)
0,17 (64 from 369)	(0,04 0,02 0,00 0,00)
0,12 (45 from 369)	(0,86 0,30 0,19 0,02)
0,12 (43 from 369)	(0,91 0,61 0,48 0,57)
0,12 (43 from 369)	(0,93 0,72 0,51 0,94)
0,09 (34 from 369)	(0,97 0,78 0,01 0,94)
0,09 (32 from 369)	(0,95 0,80 0,92 0,81)
0,08 (31 from 369)	(0,74 0,44 0,01 0,60)
0,08 (30 from 369)	(0,99 0,94 0,96 0,99)
0,07 (27 from 369)	(0,95 0,73 0,92 0,55)
0,05 (20 from 369)	(0,90 0,57 0,95 0,06)

k-means s $k=10$ - je vytvořeno 10 shluků a studenti nemusí mít naprosto stejné výsledky

Rozhodovací stromy

Decision Tree			
count: 665	odstranění spojek		
	value: 0 count: 83 (12,48%)	de morgan	
		value: 0 count: 83 (12,48%)	p+-p=1 p*-p=0
			value: --- count: 46 (6,92%)
			value: 0 count: 37 (5,56%)
	value: 0.00 - 0.2 count: 3 (0,45%)	p+-p=1 p*-p=0	
		value: 0 count: 3 (0,45%)	
	value: 0.20 - 0.4 count: 12 (1,80%)	de morgan	
		value: 0 count: 2 (0,30%)	p+-p=1 p*-p=0
			value: 0 count: 2 (0,30%)
		value: 0.00 - 0.2 count: 2 (0,30%)	p+-p=1 p*-p=0
			value: 0 count: 1 (0,15%)
			value: 0.40 - 0.6 count: 1 (0,15%)
		value: 0.20 - 0.4 count: 8 (1,20%)	p+-p=1 p*-p=0
			value: 0 count: 7 (1,05%)

Rozhodovací stromy

		value: 0.40 - 0.6 count: 36 (5,41%) value: 0.60 - 0.8 count: 27 (4,06%) value: 1 count: 39 (5,86%)
	value: 0.80 - 1.0 count: 58 (8,72%)	p+-p=1 p*-p=0 value: 0.40 - 0.6 count: 8 (1,20%) value: 0.60 - 0.8 count: 19 (2,86%) value: 1 count: 31 (4,66%)
	value: 1 count: 72 (10,83%)	p+-p=1 p*-p=0 value: --- count: 6 (0,90%) value: 0 count: 2 (0,30%) value: 0.40 - 0.6 count: 1 (0,15%) value: 0.60 - 0.8 count: 2 (0,30%) value: 1 count: 61 (9,17%)

Asociační pravidla

Antecedent	Consequent	Confidence	Support
(p+-p=1 p*-p=0: 1)	(odstranění spojek: 1)	0,74	0,21 (141 from 665)
(odstranění spojek: 1)	(p+-p=1 p*-p=0: 1)	0,42	0,21 (141 from 665)
(de morgan: 0.60 - 0.8)	(odstranění spojek: 1)	0,59	0,17 (116 from 665)
(odstranění spojek: 1)	(de morgan: 0.60 - 0.8)	0,34	0,17 (116 from 665)
(odstranění spojek: 0)	(de morgan: 0)	1,00	0,12 (83 from 665)
(de morgan: 0)	(odstranění spojek: 0)	0,80	0,12 (83 from 665)
(de morgan: 1)	(odstranění spojek: 1)	1,00	0,11 (72 from 665)
(p+-p=1 p*-p=0: 0.40 - 0.6)	(odstranění spojek: 1)	0,54	0,11 (72 from 665)
(odstranění spojek: 1)	(p+-p=1 p*-p=0: 0.40 - 0.6)	0,21	0,11 (72 from 665)
(odstranění spojek: 1)	(de morgan: 1)	0,21	0,11 (72 from 665)
(p+-p=1 p*-p=0: 1)	(de morgan: 0.60 - 0.8)	0,34	0,10 (65 from 665)
(de morgan: 0.60 - 0.8)	(p+-p=1 p*-p=0: 1)	0,33	0,10 (65 from 665)
(p+-p=1 p*-p=0: 1)(de morgan: 1)	(odstranění spojek: 1)	1,00	0,09 (61 from 665)
(de morgan: 1)	(p+-p=1 p*-p=0: 1)	0,85	0,09 (61 from 665)
(de morgan: 1)	(odstranění spojek: 1)(p+-p=1 p*-p=0: 1)	0,85	0,09 (61 from 665)
(odstranění spojek: 1)(de morgan: 1)	(p+-p=1 p*-p=0: 1)	0,85	0,09 (61 from 665)
(odstranění spojek: 1)(p+-p=1 p*-p=0: 1)	(de morgan: 1)	0,43	0,09 (61 from 665)
(p+-p=1 p*-p=0: 1)	(de morgan: 1)	0,32	0,09 (61 from 665)
(p+-p=1 p*-p=0: 1)	(odstranění spojek: 1)(de morgan: 1)	0,32	0,09 (61 from 665)
(odstranění spojek: 1)	(p+-p=1 p*-p=0: 1)(de morgan: 1)	0,18	0,09 (61 from 665)
(p+-p=1 p*-p=0: 0.40 - 0.6)	(de morgan: 0.60 - 0.8)	0,45	0,09 (60 from 665)

Asociační pravidla

Antecedent	Consequent	Confidence	Support
(odstranění spojek: 1)	$(p*0=0 \ p*1=p \ p+0=p \ p+1=1: \text{---})$	0,61	0,31 (207 from 665)
(odstranění spojek: 1)	$(p+-p=1 \ p*-p=0: 1)$	0,42	0,21 (141 from 665)
(odstranění spojek: 1)	$(p+p=p \ p*p=p: 1)$	0,40	0,21 (137 from 665)
(odstranění spojek: 1)	$(p+q=p+(-p*q) \ p*q=p*(-p+q): \text{---})$	0,38	0,20 (130 from 665)
(odstranění spojek: 1)	$(p+q=p+(-p*q) \ p*q=p*(-p+q): \text{---})(p*0=0 \ p*1=p \ p+0=p \ p+1=1: \text{---})$	0,38	0,20 (130 from 665)
(odstranění spojek: 1)	$(p=(p+q)*(p+-q) \ p=(p*q)+(p*-q): \text{---})$	0,37	0,19 (127 from 665)
(odstranění spojek: 1)	$(p+q=p+(-p*q) \ p*q=p*(-p+q): \text{---})(p=(p+q)*(p+-q) \ p=(p*q)+(p*-q): \text{---})$	0,37	0,19 (127 from 665)
(de morgan: 0)	$(p+p=p \ p*p=p: 0)$	1,00	0,16 (104 from 665)
(odstranění spojek: 0)	$(p+p=p \ p*p=p: 0)$	1,00	0,12 (83 from 665)
(odstranění spojek: 0) (de morgan: 0)	$(p+p=p \ p*p=p: 0)$	1,00	0,12 (83 from 665)
(odstranění spojek: 1)	$(p=(p+q)*(p+-q) \ p=(p*q)+(p*-q): 1)$	0,24	0,12 (80 from 665)
(odstranění spojek: 1)	$(p*0=0 \ p*1=p \ p+0=p \ p+1=1: 1)$	0,23	0,12 (77 from 665)

Děkuji za pozornost