Discussion 12/3/18

Monday, December 3, 2018 10:01 AM

	APP α_1	Rating α_2	Inc α_3	Bal α_4	OK
1	1	0	0	1	0
2	0	0	1	0	0
3	1	1	0	1	1
4	0	1	1	1	1
5	0	1	1	0	0
6	1	1	1	0	1
7	1	1	1	1	1
8	1	0	1	0	0
9	1	1	0	0	0

Use the GSCA Algorithm

$\alpha_1 \wedge \alpha_2 \wedge \ldots \alpha_n \supset ok$

Find which atom has the largest : $r = \frac{n_+}{n} = \frac{\text{positive instance } \supset ok}{\text{all instances}}$

$r_{app} = \frac{3}{6}$ $r_{rating} = \frac{4}{6}$ $r_{Inc} = \frac{3}{6}$ $r_{bal} = \frac{3}{4}$

BAL $\supset$ ok

Now we compare it w/ balance (BAL)

$r_{app} = 2/3$ $r_{rating} = \frac{3}{3}$ ← The ones that match w/ BAL

$r_{Inc} = \frac{2}{2}$

we take rating over inc b/c rating has a larger pool

RATING $\wedge$ BAL $\supset$ OK

$\vee$ (APP $\wedge$ Rating $\wedge$ Inc) $\supset$ ok → if we run the algorithm again

Now we take out cases 3, 4, 7

3 5, 7 ⇒ take out these rows in the HW

HW 5 #1

H/W 5 #1

	A	B	C	D	E
1	1	1	1	1	1
2	0	0	1	0	1
4	0	1	0	1	0
6	0	0	1	0	1
8	0	1	0	0	0
9	1	1	1	0	1

$r_A = \frac{2}{2}$ $r_B = \frac{2}{3}$

$r_C = \frac{4}{4}$ $r_D = \frac{1}{2}$

$C \to E$ covers all of it

#2

Apriori Algorithm

support $\{I_1\}$ = (1) number of transactions that contain I_1

(2) number of trans. containing I_1 / All trans

either definition works,
but we usually use (1)

$$\text{confidence}\{I_1 \to I_2\} = \frac{\text{support}\{I_1, I_2\}}{\text{support}\{I_1\}}$$

from H/W table

Itemset	Support
I_1	4 > 2
I_2	7 > 2
I_3	6 > 2
I_4	4 > 2
I_5	3 > 2

sup $\{I_1\}$ = 1

sup $\{I_3\}$ = 2

I_1 I_2 I_4
I_2 I_3
I_3 I_5

We have a threshold of 2

- If not greater, take off list

Item set	Support
I_1, I_2	2 ✓
I_1, I_3	2 ✓
I_1, I_4	2 ✓
I_1, I_5	0 ✗ ← eliminate
⋮	
I_4, I_5	

⋮	
I_4, I_5	

Then, we do another table w/ 3 items

Itemset	Sup
$I_1, I_2 I_3$	

The final answer should be like

$$\text{Confidence}\{I_1, I_2 \to I_3\} = \text{Sup}\{I_1, I_2, I_3\}$$