

Joshua Mitchell
CS 7312 Assignment 3

1. The following database has 5 transactions. Let min sup = 60 percent (3/5) and min conf = 80 (4/5) percent.

$$\begin{bmatrix} \{m, o, n, k, e, y\} \\ \{d, o, n, k, e, y\} \\ \{m, a, k, e\} \\ \{m, u, c, k, y\} \\ \{c, o, k, i, e\} \end{bmatrix}$$

- (40 points). Run the Apriori algorithm and list C_i (candidate itemset) and L_i (frequent itemset) for each iteration i . For C_i , please show the result after applying self-joining and pruning.
- List all the strong association rules in the format of {item 1, item 2} \Rightarrow {item 3}.

$C_1 \Rightarrow L_1 \Rightarrow C_2 \Rightarrow L_2 \Rightarrow C_3 \Rightarrow L_3$:

set	sup		set	sup		set	sup		set	sup		set	sup
m	3		mo	1		mko	1		mke	2		oke	3
o	3		mk	3		mky	2		oky	2		oey	2
n	2		me	2		key	2		kmo	1			
k	5	\Rightarrow	my	2	\Rightarrow								
e	4		ok	3									
y	3		oe	3									
d	1		oy	2									
a	1		ke	4									
u	1		ky	3									
c	2		ey	2									
i	1												

Strong association rules:

- $o \Rightarrow k, e$: 60% sup, 100% conf
- $o, e \Rightarrow k$: 60% sup, 100% conf
- $o, k \Rightarrow e$: 60% sup, 100% conf

2. For a given dataset with $\min \text{sup} = 8$ (absolute support), the closed patterns are $\{a,b,c,d\}$ with support of 9, $\{a,b,c\}$ with support of 11, and $\{a,b,d\}$ with support of 13.

- (10 points). List all the max-patterns.
- (30 points). List all the frequent patterns together with their absolute support values.

Pattern	Max?	Frequent?	Absolute Support Value
abcd	Yes	Yes	9
abc	No	Yes	11
abd	No	Yes	13
bcd	No	Yes	9
acd	No	Yes	9
ab	No	Yes	13
ac	No	Yes	11
ad	No	Yes	13
bc	No	Yes	11
bd	No	Yes	13
cd	No	Yes	9
a	No	Yes	13
b	No	Yes	13
c	No	Yes	11
d	No	Yes	13