

CSG 230 Fall '08 Quiz 2 Solution

1. DB

(abc)
dc(abe)ea
b(df)b
(cg)b(ac)

min. sup = 2

frequent item:

a ✓	e x
b ✓	f x
c ✓	g x
d ✓	

DB with frequent items

(abc)
dc(ab)a
bdb
cb(ac)

RESULT

<a> |
 |
<c> |
<d> |

<a>-proj DB

(-bc)
(-b)a
(-c)

frequent items

-b ✓ |
-c ✓ |

<(ab)>-proj DB

-c
a

<(ac)>-proj DB

∅ |

<(ab)> |
<(ac)> |

-proj DB

(-c)
a
db
(ac)

frequent items

a ✓ |

<ba> |

<c>-proj DB

(ab)a
b(ac)

frequent items

a ✓ |
b ✓ |

<(ca)>-proj DB

(-b)a
-c

<(cb)>-proj DB

a
(ac)

frequent items

a ✓ |

<ca> |
<cb> |
<cba> |

<d>-proj DB

c(ab)a
b

frequent items

b ✓ |

<db> |

2. (a)

Age	#buy	#not buy
<30	3	3
[30,40]	2	1
>40	3	0

$$P(c | \text{Age} < 30) = \frac{3}{6} \cdot \frac{1}{2} + \frac{3}{6} \cdot \frac{1}{2} = 1$$

$$P(c | \text{Age} \in [30,40]) = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{3} = \frac{2}{3} \cdot 0.58 + \frac{1}{3} \cdot 1.58 = 0.91$$

$$P(c | \text{Age} > 40) = 0$$

$$\text{So } P(c | \text{Age}) = \frac{6}{12} \cdot 1 + \frac{3}{12} \cdot 0.91 = 0.732$$

credit	#buy	#not buy
good	6	2
bad	2	2

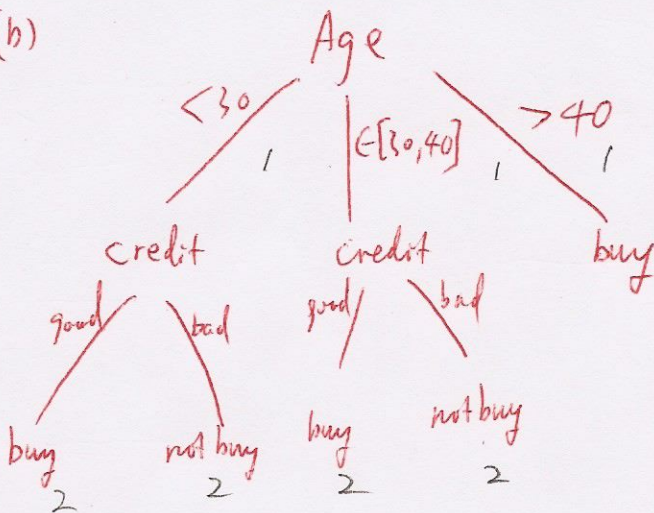
$$P(c | \text{credit} = \text{good}) = \frac{6}{8} \cdot \frac{1}{3} + \frac{2}{8} \cdot \frac{1}{4}$$

$$= \frac{3}{4} (2 - 1.58) + \frac{1}{4} \cdot 2 = 0.82$$

$$P(c | \text{credit} = \text{bad}) = 1$$

$$\text{So } P(c | \text{credit}) = \frac{8}{12} \cdot 0.82 + \frac{4}{12} \cdot 1 = 0.882$$

(b)



3. (a) $P(\text{buy}) = \frac{8}{12} = 0.67$

$$P(\text{age} < 30 | \text{buy}) = \frac{3}{8} = 0.375$$

$$P(\text{age} < 30 | \text{not buy}) = \frac{3}{4} = 0.75$$

$$P(\text{age} > 40 | \text{buy}) = \frac{3}{8} = 0.375$$

$$P(\text{age} > 40 | \text{not buy}) = \frac{0}{4} = 0$$

$$P(\text{credit} = \text{good} | \text{buy}) = \frac{6}{8} = 0.75$$

$$P(\text{credit} = \text{good} | \text{not buy}) = \frac{2}{4} = 0.5$$

note: $P(\text{not buy})$
 $P(\text{age} \in [30,40] | \text{buy})$
 $P(\text{age} \in [30,40] | \text{not buy})$
 $P(\text{credit} = \text{bad} | \text{buy})$
 $P(\text{credit} = \text{bad} | \text{not buy})$ } can be calculated. But if you compute them, also correct.

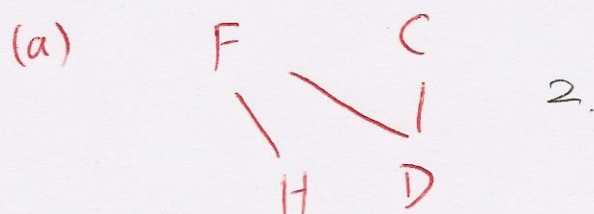
(b) $P(\text{age} < 30 | \text{buy}) \times P(\text{credit} = \text{good} | \text{buy}) \times P(\text{buy})$
 $= 0.375 \times 0.75 \times 0.67 = 0.194$

$$P(\text{age} < 30 | \text{not buy}) \times P(\text{credit} = \text{good} | \text{not buy}) \times P(\text{not buy})$$

$$= 0.75 \times 0.5 \times 0.33 = 0.124$$

So predict BUY!

4. $F = \text{fever}$, $H = \text{headache}$, $C = \text{cough}$, $D = \text{cold}$



(b) We know already $P(H|F) = 0.25$, $P(D|F \wedge C) = 0.2$
we need to find

$$P(F), P(C)$$

$$P(H|\sim F), P(D|\sim F \wedge C), P(D|F \wedge \sim C), P(D|\sim F \wedge \sim C)$$

$$(c) P(D|H) = \frac{P(D \wedge H)}{P(H)} \quad 2$$

$$P(D \wedge H) = P(D \wedge H \wedge F \wedge C) + P(D \wedge H \wedge \sim F \wedge C) \\ + P(D \wedge H \wedge F \wedge \sim C) + P(D \wedge H \wedge \sim F \wedge \sim C)$$

$$= P(D|F \wedge C) * P(H|F) * P(F) * P(C)$$

$$+ P(D|\sim F \wedge C) * P(H|\sim F) * P(\sim F) * P(C)$$

$$+ P(D|F \wedge \sim C) * P(H|F) * P(F) * P(\sim C)$$

$$+ P(D|\sim F \wedge \sim C) * P(H|\sim F) * P(\sim F) * P(\sim C)$$

$$P(H) = P(H|F) P(F) + P(H|\sim F) P(\sim F) \quad 2$$

note: $P(\sim C) = 1 - P(C)$

$$P(\sim F) = 1 - P(F)$$

$$5. (a) \text{ new CF}_1 = \left\langle \underset{2}{3}, \left(\underset{1}{29}, \underset{1}{3} \right), \underset{2}{286} \right\rangle$$

$$R(\text{new CF}_1) = \frac{1}{3} \sqrt{\frac{3 \times 286 - (841 + 9)}{2}} = \frac{1}{3} \sqrt{8} \doteq 0.94$$

$$(b) \text{ new CF}_2 = \left\langle \underset{2}{2}, \left(\underset{1}{19}, \underset{1}{11} \right), \underset{2}{282} \right\rangle$$

$$R(\text{new CF}_2) = \frac{1}{2} \sqrt{\frac{2 \times 282 - (271 + 121)}{2}} = \frac{1}{2} \sqrt{172}$$

(c)

$\langle 3, (29, 3), 286 \rangle$	2
$\langle 1, (10, 10), 200 \rangle$	2