CS422 : Data Mining
Homework - 2

MEGHA TATTI
A20427027.

Excercise 1

1.1 Tan, Chapter 3

Ex a]

2- (1) Gini index for the overall collection of training examples.

⇒ Gini index = $1 - \Sigma p(i|t)^2$.

$= 1 - \left[\left(\frac{10}{20}\right)^2 + \left(\frac{10}{20}\right)^2\right] = 1 - \left[\frac{1}{4} + \frac{1}{4}\right] = 1 - \frac{2}{4} = 0.5$

2- (2) Compute the Gini index for the customer ID attribute.

⇒ $1 - \left[\left(\frac{0}{1}\right)^2 + \left(\frac{1}{1}\right)^2\right] \Rightarrow 1 - 1 \Rightarrow 0$

2-(3) Compute the Gini index for the gender attribute.

⇒ for females,

Gini index = $1 - \left[\left(\frac{6}{10}\right)^2 + \left(\frac{4}{10}\right)^2\right]$

$= 1 - ((0.3)^2 + (0.2)^2) \Rightarrow$

$\Rightarrow 1 - ((0.6)^2 + (0.4)^2)$

$\Rightarrow 1 - (0.36 + 0.16)$

$\Rightarrow 0.48$

for males,

$$\text{Gini index} = 1 - \left(\left(\frac{4}{10}\right)^2 + \left(\frac{6}{10}\right)^2\right)$$

$$= 1 - 0.52$$

$$= 0.48.$$

Gini index for gender would be,

$$\left(\frac{10}{20}\right) * 0.48 + \left(\frac{10}{20}\right) * 0.48$$

$$\Rightarrow \underline{0.48}$$

2-(4) Compute the Gini index for the car type attribute using multiway split.

$\Rightarrow$ Family,

$$\text{Gini Index} = 1 - \left[\left(\frac{1}{4}\right)^2 + \left(\frac{3}{4}\right)^2\right]$$

$$\Rightarrow 1 - \left((0.25)^2 + (0.75)^2\right)$$

$$\Rightarrow 0.375$$

Luxury,

$$\text{G.I} = 1 - \left(\left(\frac{1}{8}\right)^2 + \left(\frac{7}{8}\right)^2\right)$$

$$\Rightarrow 0.219$$

Sports,

$$G.I = 1 - \left(\left(\frac{8}{8}\right)^2 + \left(\frac{0}{8}\right)^2\right)$$

$$\Rightarrow 1 - 1 = 0.$$

Gini index of car attribute is,

$$\left(\frac{4}{20}\right) * 0.375 + \left(\frac{8}{20}\right) * 0.219 + \frac{8}{20} * 0$$

$$\Rightarrow 0.163.$$

2-(5) Compute Gini Index of shirt size attribute using multiway split.

$\Rightarrow$ Small,

$$G.I = 1 - \left(\left(\frac{3}{5}\right)^2 + \left(\frac{2}{5}\right)^2\right)$$

$$\Rightarrow 1 - \left((0.6)^2 + (0.4)^2\right)$$

$$\Rightarrow 0.48$$

Medium,

$$G.I = 1 - \left(\left(\frac{3}{7}\right)^2 + \left(\frac{4}{7}\right)^2\right)$$

$$\Rightarrow 1 - \left((0.43)^2 + (0.57)^2\right)$$

$$\Rightarrow 0.49$$

Large,

$$GI = 1 - \left[\left(\frac{2}{4}\right)^2 + \left(\frac{2}{4}\right)^2\right]$$

$$\Rightarrow 1 - \left((0.5)^2 + (0.5)^2\right)$$

$$= 0.5.$$

Extra Large,

$$GI = 1 - \left[\left(\frac{2}{4}\right)^2 + \left(\frac{2}{4}\right)^2\right]$$

$$= 1 - \left((0.5)^2 + (0.5)^2\right)$$

$$= 0.5.$$

Gini index for shirt attribute is given by,

$$\left(\frac{5}{20}\right) * 0.48 + \left(\frac{7}{20}\right) * 0.49 + \left(\frac{4}{20}\right) * 0.5 + \left(\frac{4}{20}\right) * 0.5$$

$$\Rightarrow 0.4915.$$

2-(b) which attribute is better, Gender, Car or shirt?

⇒ Among the three attributes, car type is better as it has the lowest Gini index of all i.e, 0.163 whereas the other 2 have 0.48 (Gender) and ~~0.419~~ 0.49 (Shirt size).

2-(7) Why customer ID should not be used as the attribute test condition even though it has the lowest Gini.

⇒ Because each attribute has unique customer id. And Hence the attribute has no predictive power.

Ex 3

3-(1) Entropy =,

$$E(S) = \sum_{i=1}^{c} -P_i \log_2 P_i .$$

$$= -\left[\left(\frac{4}{9}\right)\log_2\left(\frac{4}{9}\right) + \left(\frac{5}{9}\right)\log_2\left(\frac{5}{9}\right)\right]$$

$$= -((-0.519) + (-0.471))$$

$$= 0.991$$

3-(2) For a_1, T = (+ve → 3, -ve → 1) | F = (+ve → 1, -ve → 4)

For a_2, T = (+ve → 2, -ve → 3) | F = (+ve → 2, -ve = 2)

Entropy for a_1

$$\Rightarrow \frac{4}{9}\left[-\left(\left(\frac{3}{4}\right)\log_2\left(\frac{3}{4}\right) + \left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right)\right)\right] + \frac{5}{9}\left[-\left[\left(\frac{1}{5}\right)\log_2\left(\frac{1}{5}\right) + \left(\frac{4}{5}\right)\log_2\left(\frac{4}{5}\right)\right]\right.$$

⇒ Entropy for a_1 (T)

$$\Rightarrow -\left[\left(\frac{3}{4}\right)\log_2\left(\frac{3}{4}\right) + \left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right)\right] \Rightarrow \boxed{0.81128}$$

Entropy for $a_1(F)$

$$E = -\left[\left(\frac{1}{5}\right)\log_2\left(\frac{1}{5}\right) + \left(\frac{4}{5}\right)\log_2\left(\frac{4}{5}\right)\right]$$

$$\Rightarrow \boxed{0.721.}$$

Entropy for $a_2(T)$

$$E = -\left[\left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right) + \left(\frac{3}{5}\right)\log_2\left(\frac{3}{5}\right)\right]$$

$$= \boxed{0.970}$$

Entropy for $a_2(F)$

$$E = -\left[\left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right) + \left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right)\right]$$

$$\Rightarrow \boxed{1}$$

Information gain

For $a_1 \Rightarrow (0.991) - \left[\left(\frac{4}{9}\right) * (0.81128) + \left(\frac{5}{9}\right) * (0.721)\right]$

$\Rightarrow 0.229$

For $a_2 \Rightarrow (0.991) - \left[\left(\frac{5}{9}\right) * (0.970) + \left(\frac{4}{9}\right) * 1\right]$

$= 0.007.$

3-(3)

a_3	class label
1.0	+
6.0	+
5.0	−
4.0	+
7.0	−
3.0	−
8.0	−
7.0	+
5.0	−

Sorted a_3	Class label	Split point	Entropy	Info gain
1.0	+	2.0	0.848	0.142
3.0	−	3.5	0.988	0.0024
4.0	+	4.5	0.918	0.072
5.0	−	5.5	0.983	0.0072
5.0	−			
6.0	+	6.5	0.972	0.018
7.0	+	7.5	0.888	0.102
7.0	−			

The best split for a_3 occurs at split point equals to 2.

~~Split 1~~ Split point 2.0.

$\leq(E) = -\left[\left(\frac{1}{1}\right)\log_2\left(\frac{1}{1}\right) + 0\log_2 0\right]$

$= 0.$

$>(E) = -\left[\left(\frac{3}{8}\right)\log_2\left(\frac{3}{8}\right) + \left(\frac{5}{8}\right)\log_2\left(\frac{5}{8}\right)\right]$

$\Rightarrow -((-0.530) + (-0.423))$

$= 0.954.$

Weighted average $= \left[\left(\frac{1}{9}\right) * 0\right] + \left(\frac{8}{9}\right) * 0.954\Big]$

$= 0.848.$

Information gain $\Rightarrow 0.991 - 0.848 = 0.143$

Split point 3.5.

$\leq(E) = -\left[\left(\frac{1}{2}\right)\log_2\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\log_2\left(\frac{1}{2}\right)\right]$

$= -[(-0.5) + (-0.5)) = 1.$

$>(E) = -\left[\left(\frac{3}{7}\right)\log_2\left(\frac{3}{7}\right) + \left(\frac{4}{7}\right)\log_2\left(\frac{4}{7}\right)\right)$

$= -[(-0.523) + (-0.461))]$

$= 0.985.$

Weighted avg $= \left[\left(\frac{2}{9}\right) * 1 + \left(\frac{7}{9}\right) * 0.985\right] = 0.988.$

Information gain $= 0.0024$ //

Split point 4.5.

$\leq(E) - \left[\left(\frac{2}{3}\right)\log_2\left(\frac{2}{3}\right) + \left(\frac{1}{3}\right)\log_2\left(\frac{1}{3}\right)\right]$

$= 0.918$

$>(E) = -\left[\left(\frac{2}{6}\right)\log_2\left(\frac{2}{6}\right) + \left(\frac{4}{6}\right)\log_2\left(\frac{4}{6}\right)\right]$

$= 0.918.$

Weighted avg $= \left[\left(\frac{3}{9}\right) * (0.918) + \left(\frac{6}{9}\right) * (0.918)\right]$

$= 0.918.$

Information gain $\Rightarrow 0.991 - 0.918 \Rightarrow \underline{0.072}$

Split point 5.5

$\leq(E) = -\left[\left(\frac{2}{5}\right)\log_2\left(\frac{2}{5}\right) + \frac{3}{5}\log_2\left(\frac{3}{5}\right)\right)$

$= 0.970$

$>(E) = -\left[\left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right) + \left(\frac{2}{4}\right)\log_2\left(\frac{2}{4}\right)\right]$

$= 1.$

Weighted avg $= \left[\left(\frac{5}{9}\right) * 0.970\right] + \left[\left(\frac{4}{9}\right) * 1\right]$

$= 0.983$

Information gain $= 0.991 - 0.983 = \underline{0.0071}.$

Split point 6.5

$\leq(E) = -\left[\left(\frac{3}{6}\right)\log_2\left(\frac{3}{6}\right) + \left(\frac{3}{6}\right)\log_2\left(\frac{3}{6}\right)\right] = 1.$

$>(E) = -\left[\left(\frac{1}{3}\right)\log_2\left(\frac{1}{3}\right) + \left(\frac{2}{3}\right)\log_2\left(\frac{2}{3}\right)\right] = 0.918.$

weighted avg $\Rightarrow \left[\left(\frac{6}{9}\right) * 1\right] + \left[\left(\frac{3}{9}\right) * 0.918\right] = 0.972.$

Information gain = $0.991 - 0.972 = 0.018$

Split point 7.5

$\leq(E) = -\left[\left(\frac{4}{8}\right)\log_2\left(\frac{4}{8}\right) + \left(\frac{4}{8}\right)\log_2\left(\frac{4}{8}\right)\right] = 1$

$>(E) = -\left[\left(\frac{0}{1}\right)\log_2\left(\frac{0}{1}\right) + \left(\frac{1}{1}\right)\log_2\left(\frac{1}{1}\right)\right] = 0.$

weighted avg $= \left[\left(\frac{8}{9}\right) * 1\right] + \left[\left(\frac{1}{9}\right) * 0\right] = 0.888$

Information gain = $0.991 - 0.888 = 0.102.$

3-(4) According to information gain, a_1 produces the best split as it has higher gain

---

3-(5)

For a_1 (T)

$$\Rightarrow 1 - \left[\frac{3}{4}, \frac{1}{4}\right] = 1 - \frac{3}{4} = 0.25$$

a_1(F)

$$\Rightarrow 1 - \left[\frac{1}{5}, \frac{4}{5}\right] = 1 - \frac{4}{5} = 0.2$$

Avg error rate $\Rightarrow \frac{4}{9} * 0.25 + \frac{5}{9} * 0.2$

$\Rightarrow 0.222$

For a_2 (T)

$$\Rightarrow 1 - \left[\frac{2}{5}, \frac{3}{5}\right] = 1 - \frac{3}{5} = 0.4$$

a_2(F)

$$\Rightarrow 1 - \left[\frac{2}{4}, \frac{2}{4}\right] = 1 - \frac{2}{4} = 0.5$$

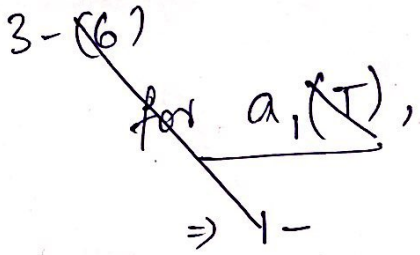
Avg error rate $= \frac{5}{9} * 0.4 + \frac{4}{9} * 0.5$

$\Rightarrow 0.222 + 0.222$.

$= 0.444$

According to classification error rate, the best split is a_1 due to low classification error rate. The classification error depicts the accuracy of the sample set, The higher the classification error the more error the sample set contains.

3-(b)



3-(b)

for a_1 ⇒

$$\text{Gini Index} = \frac{4}{9}\left[1-\left[\left(\frac{3}{4}\right)^2+\left(\frac{1}{4}\right)^2\right]\right]+\frac{5}{9}\left[1-\left[\left(\frac{1}{5}\right)^2+\left(\frac{4}{5}\right)^2\right]\right]$$

$$= 0.3444.$$

for a_2 ⇒

$$\text{Gini Index} = \frac{4}{9}\left[1-\left[\left(\frac{2}{5}\right)^2+\left(\frac{3}{5}\right)^2\right]\right]+\frac{4}{9}\left[1-\left[\left(\frac{2}{4}\right)^2+\left(\frac{2}{4}\right)^2\right]\right]$$

$$= 0.4889.$$

Since gini index of a_1 is small, it has the better split.

Ex 5

5 (a). The contingency tables after splitting on attributes A and B are :-

	A=T	A=F
+	4	0
−	3	3

	B=T	B=F
+	3	1
−	1	5

overall entropy before splitting,

$$E_{original} = -(0.4 \log 0.4 + 0.6 \log 0.6)$$

$$= 0.97.$$

The information gain after splitting on A.

$$E_A(T) = -\left(\left(\frac{4}{7}\right)\log_2\left(\frac{4}{7}\right) + \left(\frac{3}{7}\right)\log_2\left(\frac{3}{7}\right)\right)$$

$$= 0.985.$$

$$E_A(F) = -\left(\left(\frac{3}{3}\right)\log_2\left(\frac{3}{3}\right) + \left(\frac{0}{3}\right)\log_2\left(\frac{0}{3}\right)\right)$$

$$= 0.$$

Info gain Δ, $\Rightarrow E_{orig} - (7/10)\, E_A(T) - (3/10)\, E_A(F)$

$\Rightarrow 0.2813$

The information gain after splitting on B is

$$E_B(T) = -\left(\left(\frac{3}{4}\right)\log_2\left(\frac{3}{4}\right) + \left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right)\right) = 0.811$$

$$E_B(F) = -\left(\left(\frac{1}{6}\right)\log_2\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\log_2\left(\frac{5}{6}\right)\right) = 0.650.$$

$$\Delta = E_{orig} - (4/10)\, E_B(T) - \left(\frac{6}{10}\right) E_B(F) = 0.2565.$$

Therefore, attribute A will be chosen to split the node.

5 (b)

The overall gini before splitting is

$$G_{orig} = 1 - 0.4^2 - 0.6^2 = 0.48$$

The gain in gini after splitting on A,

$$G_A(T) = 1 - \left(\frac{4}{7}\right)^2 - \left(\frac{3}{7}\right)^2 = 0.489$$

$$G_A(F) = 1 - \left(\frac{3}{3}\right)^2 - \left(\frac{0}{3}\right)^2 = 0.$$

$$\Delta = 0.48 - \left(\frac{7}{10}\right) * 0.489 - \left(\frac{3}{10}\right) * 0$$

$$= 0.1371$$

The gain in gini after splitting on B,

$$G_B(T) = 1 - \left(\frac{1}{4}\right)^2 - \left(\frac{3}{4}\right)^2 = 0.375$$

$$G_B(F) = 1 - \left(\frac{1}{6}\right)^2 - \left(\frac{5}{6}\right)^2 = 0.2778$$

$$\Delta = 0.48 - \left(\frac{4}{10}\right) * 0.375 - \left(\frac{6}{10}\right) * 0.2778$$

$$= 0.1633$$

Therfore attribute B will be chose to split the node.

5-(c) Yes, even though these measures have similar range and monotonous behavior, their respective gains, which are scaled differences of the measures, do not necessarily behave in the same way, as in results from part (a) & (b)

1.2 Tan, Chapter 4.

Question 18.

18-(a) Consider 100 total values, from this,
50 → training set & the rest are used for testing

The error rate is the percent of incorrect classifications. Now for this case, for the records of 50% positive and 50% negative, the decision tree is predicting that all the records are positive.

So,

	Predicted	
	negative	positive
Negative(AV)	0	25
Positive(AV)	0	25

$$\text{error rate} = \frac{(\text{False+ve} + \text{False-ve})}{(\text{T+ve}) + FP + TN + fN} \Rightarrow \frac{25+0}{25+0+0+25}$$

$$\Rightarrow 0.5 = 50\%$$

18-(2) positive class with probability 0.8
negative class with probability 0.2

predicted.

Actual		positive PV	Negative PV
	positive AV	20	5
	negative AV	20	5

Error rate = (FP + FN) / (TP + TN + FP + FN)

$\Rightarrow \frac{20+5}{20+5+20+5} \Rightarrow \frac{25}{50} = 0.5$

= 50%.

18-(3) From the set of 50 records,

(33.33) $\frac{2}{3} * 50$ → belongs to → +ve class.

(16.66) $\frac{1}{3} * 50$ → belongs to → -ve class

	positive PV	Negative PV
positive AV	33.33	0
Negative AV	16.66	0

Error rate $\Rightarrow \frac{16.66 + 0}{33.33 + 0 + 16.66 + 0} \Rightarrow \frac{16.66}{49.99}$

$\Rightarrow 0.333 \Rightarrow 33\%$.

18-(4) From 50 records,

Actual +ve → 33.33
Actual -ve → 16.66.

predicted +ve => 22.22.
predicted -ve => 5.55.

	positive PV	Negative PV
positive AV	22.22	11.11
Negative AV	11.11	5.55

Error rate => $\frac{11.11 + 11.11}{22.22 + 5.55 + 11.11 + 11.11}$

=> $\frac{22.22}{49.99}$ => 0.44 = 44%

1.3 Zaki, Chapter 8

Ex 1).

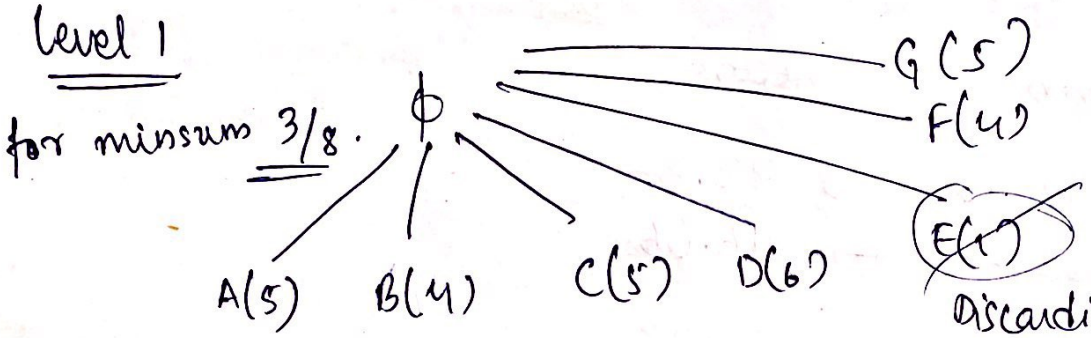
(a)

tid	itemset
t_1	ABCD
t_2	ACDF
t_3	ACDEG
t_4	ABDF
t_5	BCG
t_6	DFG
t_7	ABG
t_8	CDFG.

	A	B	C	D	E	F	G
t_1	1	1	1	1			
t_2	1		1	1		1	
t_3	1		1	1	1		1
t_4	1	1		1		1	
t_5		1	1				1
t_6				1		1	1
t_7	1	1					1
t_8			1	1		1	1

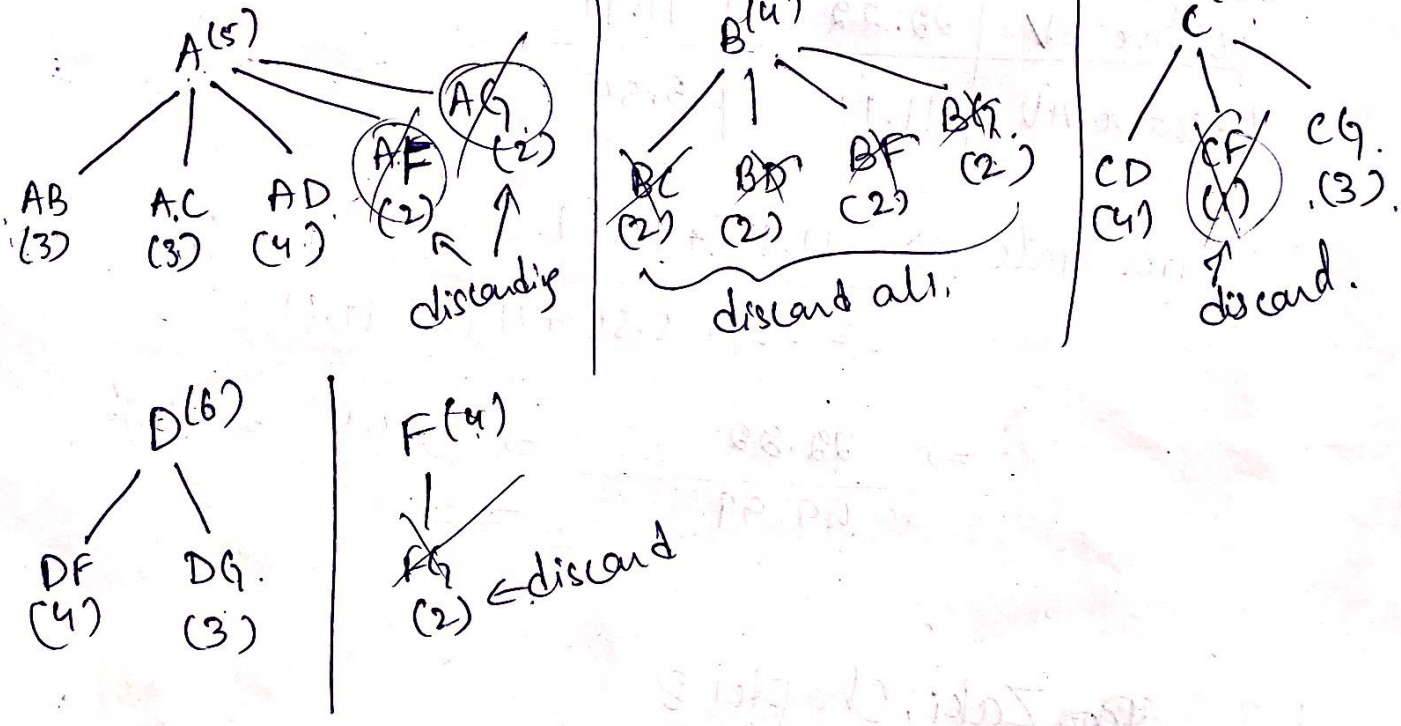
level 1

for minsum 3/8.

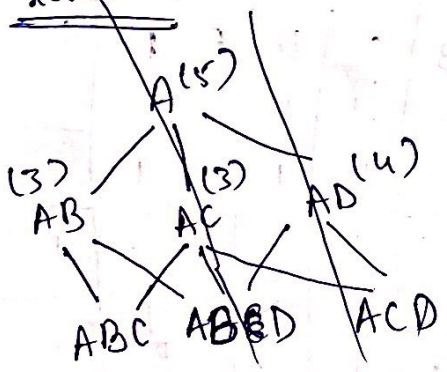


Discarding E as it has less than minsup value.

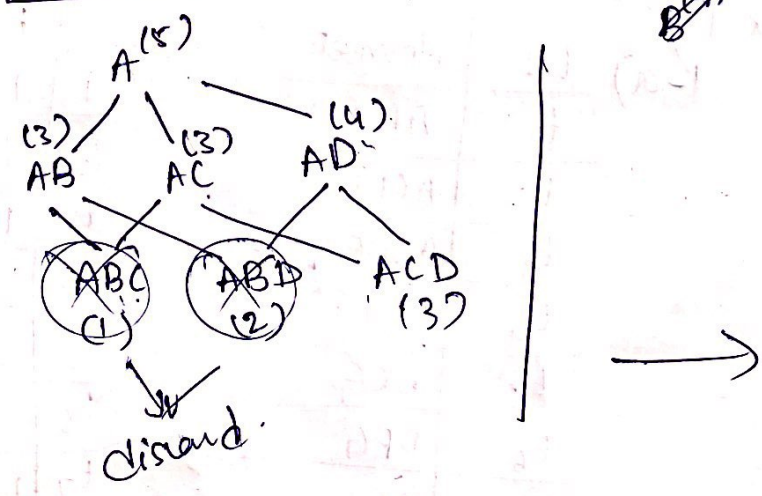
level 2

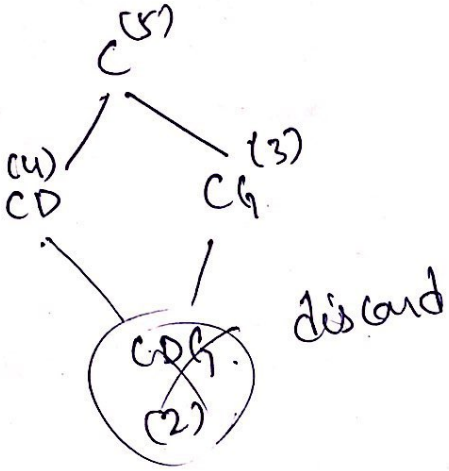


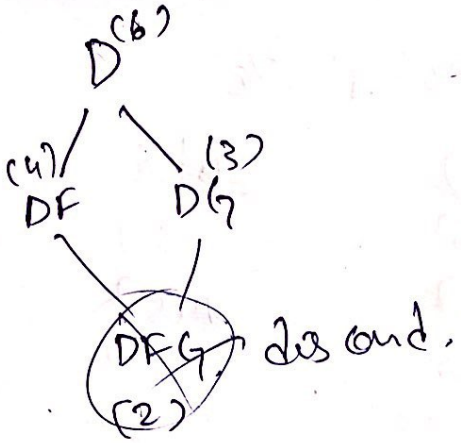
level - 3



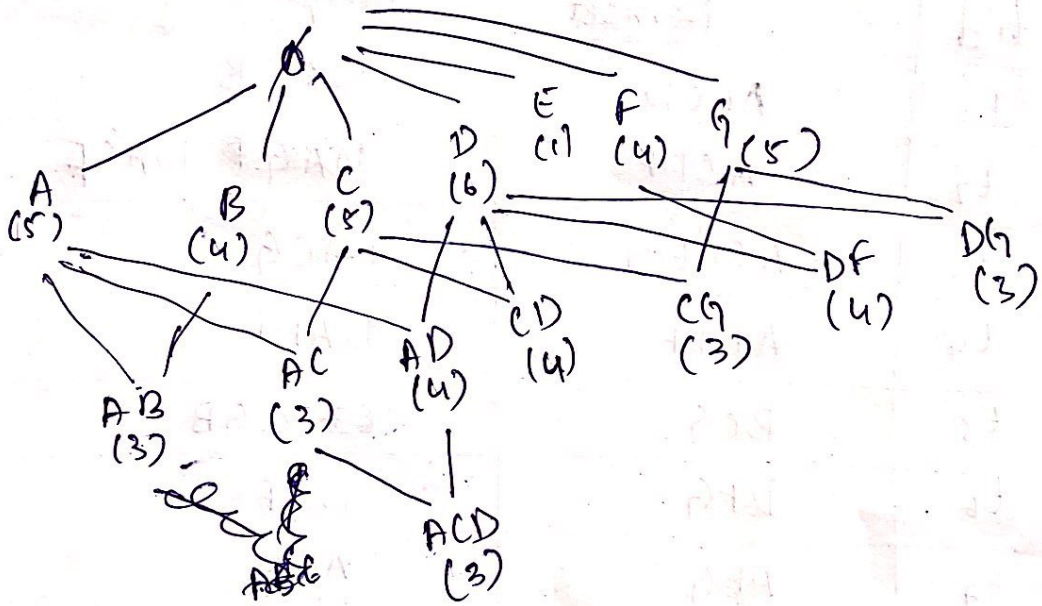
level - 3







overall would be

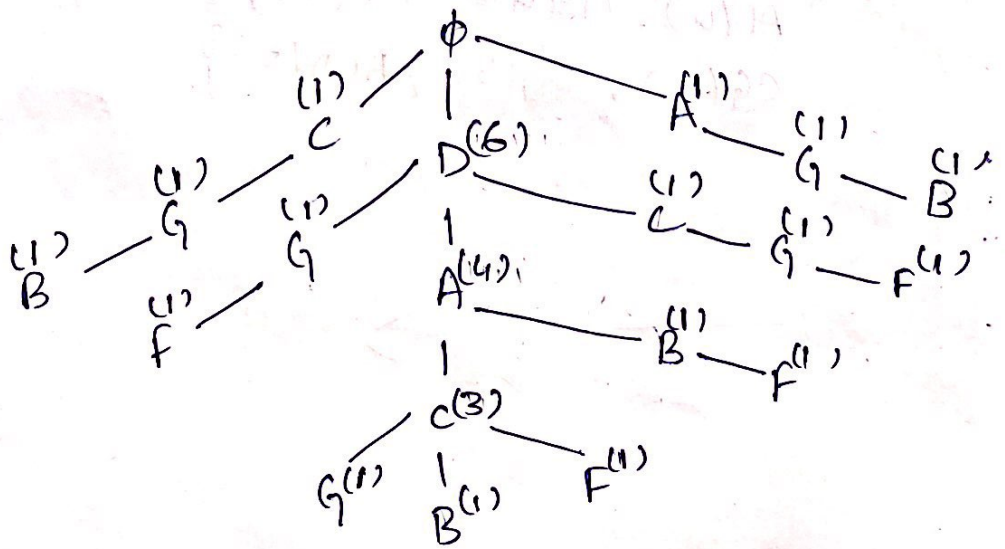


Frequent items = {D(6), A(5), C(5), G(5), B(4), F(4), AD(4), CD(4), DF(4), AB(3), AC(3), CG(3), DG(3), ACD(3)}.

1-(b) minsup = 2/8.

A(5)	sorted
B(4)	D(6)
C(5)	A(5)
D(6)	C(5)
E(1)x	G(5)
F(4)	B(4)
G(5)	F(4)

tid	itemset	frequently sorted
t_1	ABCD	DACB
t_2	ACDF	DACF
t_3	ACDEG	DACG.
t_4	ABDF	DABF.
t_5	BCG	CGB
t_6	DFG	DGF.
t_7	ABG	AGB.
t_8	CDFG	DCGF.

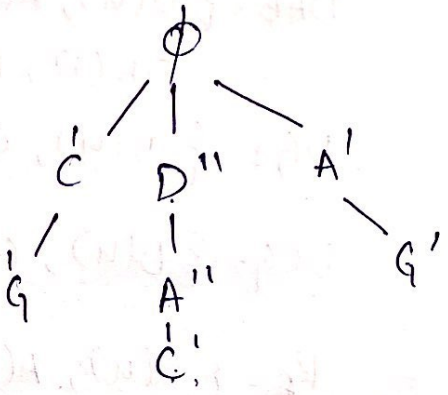


projection :-

(1) R_A => $D^{(4)}$

=> $\{AD(4)\}$

(2) R_B => DAC, DA, CG, AG.



DAC = { D(2), A(3), C(2), DA(2) }

~~DA~~

DA = { D(2), A(3), DA(2) }.

CG = { C(2), G(2), CG(1) }

AG = { A(3), G(2), AG(1) }.

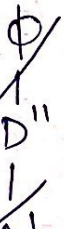
R_B = { D(2), A(3), C(2), DA(2), G(2) }

= { BD(2), BA(3), BC(2), BDA(2), BG(2) }.

(3) R_C => DA, D

DA = { D(2), A(1), DA(1) }

D = { CD(2) } ~~CA(1), CDA(1)~~



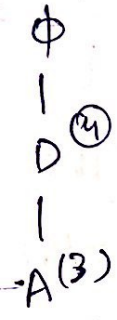
(4) R_D => None.

(5) R_E =>

(3) R_C => DA, D.

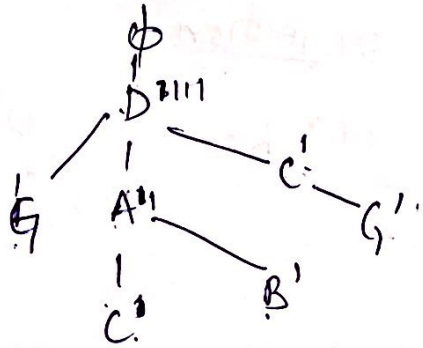
DA = { D(4), A(3), DA(3) }

= { CD(4), CA(3), CDA(3) }



(4) R_D = None.

(5) R_F = DAC, DAB, DG, DCG.



DAC = { D(4), A(2), C(2), DA(2), DC(2), AC(1) }

DAB = { D(4), A(2), B(1), DA(2), DB(1), AB(1) }

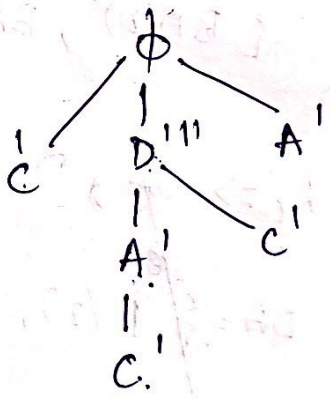
DG = { D(4), G(2), DG(2) }

DCG = { D(4), C(2), G(2), DC(2), CG(1), DG(2) }

R_F = { D(4), A(2), C(2), DA(2), DC(2), DG(2) }

⇒ { FD(4), FA(2), FC(2), FDA(2), FDC(2), FDG(2) }

(6) R_G = DAC, C, D, A, DC



DAC = { D(3), A(2), C(3), DA(1), DC(2), AC(1) }

DC = { D(3), C(3), DC(2) }

R_G = { D(3), A(2), C(3), DC(2) }

⇒ { GD(3), GA(2), GC(3), GDC(2) }

The frequent itemset ⇒ 25 items

{ D(6), A(5), C(5), G(5), B(4), F(4), AD(4), BD(2), BA(3), BC(2), BDA(2), BG(2), CD(4), CA(3), CDA(3), FD(4), FA(2), FC(2), FDA(2), FDC(2), FDG(2), GD(3), GA(2), GC(3), GDC(2) }

Q4)

tid	itemset
t_1	ACD
t_2	BCE
t_3	ABCE
t_4	BDE
t_5	ABCE
t_6	ABCD

Rules that one can generate from the set ABE

⇒ let $a = \{ABE\}$

calculating all the subsets of a.

$a = \{AB(3), AE(2), BE(4), A(4), B(5), E(4), ABE(2), \phi(6)\}$.

for each subset, generating a rule.

* $\{AB\} \longrightarrow \{E\}$

$$\text{confidence } (C) = \frac{supp(ABE)}{supp(AB)} \Rightarrow \frac{2}{3} = 0.66$$

* $\{AE\} \longrightarrow \{B\}$

$$C = \frac{supp(ABE)}{supp(AE)} \Rightarrow \frac{2}{2} = 1.$$

* $\{BE\} \longrightarrow \{A\}$

$$C = \frac{supp(ABE)}{supp(BE)} \Rightarrow \frac{2}{4} = \frac{1}{2} = 0.5.$$

* $\{ABE\} \longrightarrow \{\phi\}$

$$C = \frac{2}{2} = 1$$

* $\{A\} \rightarrow \{BE\}$

$$C = \frac{supp(ABE)}{supp(A)} = \frac{2}{4} = 0.5$$

* $\{B\} \rightarrow \{AE\}$

$$C = \frac{supp(ABE)}{supp(B)} = \frac{2}{5} = 0.4.$$

* $\{E\} \rightarrow \{AB\}$

$$C = \frac{supp(ABE)}{supp(E)} = \frac{2}{4} = 0.5$$

$\{B\} \rightarrow \{ABE\}$
$C = \frac{2}{6} = \frac{1}{3}$

Q 6)

6-a) Number of itemsets $\Rightarrow 2^k - 1 = 2^{11} - 1 = 2047$.

6-b) More than or equal to support of X. (option iv).

1.4 Multiclass classification

The multiclass.Rmd's confusion matrix created by 'R' is as follows :-

Actual value	predicted value: setosa	predicted value: versicolor	predicted value: virginica
setosa	8	0	0
Versicolor	0	10	2
virginica	0	1	9

→ for setosa, the binary confusion matrix is,

	setosa	versicolor + virginica
setosa	8	0.
versicolor + virginica	0	22

sensitivity => $\frac{TP}{TP+FN} = \frac{8}{8+0} = 1.0$

specificity => $\frac{TN}{TN+FP} = \frac{22}{22+0} = 1.0$

precision = $\frac{TP}{TP+FP} = \frac{8}{8+0} = 1.0$

→ for versicolor

AV / PV

	versicolor	setosa + virginica
versicolor	10.	2.
setosa + virginica	1	17.

sensitivity => $\frac{10}{10+2} = \frac{10}{12} = 0.83$

specificity => $\frac{17}{18} = 0.94$

precision => $\frac{10}{11} = 0.91$

→ for virginica,

AV \ PV	virginica	setosa + versicolor
virginica.	9	1
setosa + versicolor	2	18

sensitivity => $\frac{9}{10} = 0.90$

specificity => $\frac{18}{20} = 0.90.$

precision => $\frac{9}{11} = 0.82$

=