

A	B	Y
0	0	0
1	1	1
2	0	1
1	1	0

$$a) P(A=0 | Y=1) \stackrel{\text{MLE}}{=} \frac{0}{2} = 0$$

$$b) P(A=0 | Y=1) \stackrel{\text{LAPLACE}}{=} \frac{0+1}{2+3+2} = \frac{1}{7}$$

val
pos # val
pos
P+A P+B

A	B	Y
0	0	0
1	0	0
2	1	1
0	0	1
1	1	2
2	1	2

$$Y: \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} - 2 \text{ estimări}$$

$$\mathbb{P} \left\{ \begin{array}{l} I: A: \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} - 2 \text{ estimări} \\ MLE: \end{array} \right.$$

$$B: \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & 0 \end{pmatrix} - 1 \text{ estimare.}$$

$$MLE: \quad \text{II: } A: \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} - 2 \text{ estimări}$$

$$B: \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} - 1 \text{ estimare}$$

$$\text{III: } A: \begin{pmatrix} 0 & 1 & 2 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} - 2 \text{ estimări}$$

$$B: \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} - 1 \text{ estimare.}$$

b) II estimări minime

$$c) P(A=0, B=0)$$

$$i) Y_{MAP} = \underset{y \in \{0,1,2\}}{\operatorname{argmax}} P(Y=y | A=0, B=0)$$

$$F.B. \underset{y \in \{0,1,2\}}{\operatorname{argmax}} \frac{P(A=0, B=0 | Y=y) \cdot P(Y=y)}{P(A=0, B=0)} \xrightarrow{\text{constantă pozitivă în rap. cu y}}$$

$$= \underset{y \in \{0,1,2\}}{\operatorname{argmax}} P(A=0, B=0 | Y=y) \cdot P(Y=y)$$

$$\underset{\substack{\text{indep} \\ \text{cond}}}{\operatorname{argmax}} \underset{y \in \{0,1,2\}}{P(A=0 | Y=y) \cdot P(B=0 | Y=y) \cdot P(Y=y)}$$

$$P_0 = P(A=0 | Y=0) \cdot P(B=0 | Y=0) \cdot P(Y=0) \stackrel{\text{MLE}}{=} \frac{1}{2} \cdot 1 \cdot \frac{1}{3} = \frac{1}{6} = 0,1666$$

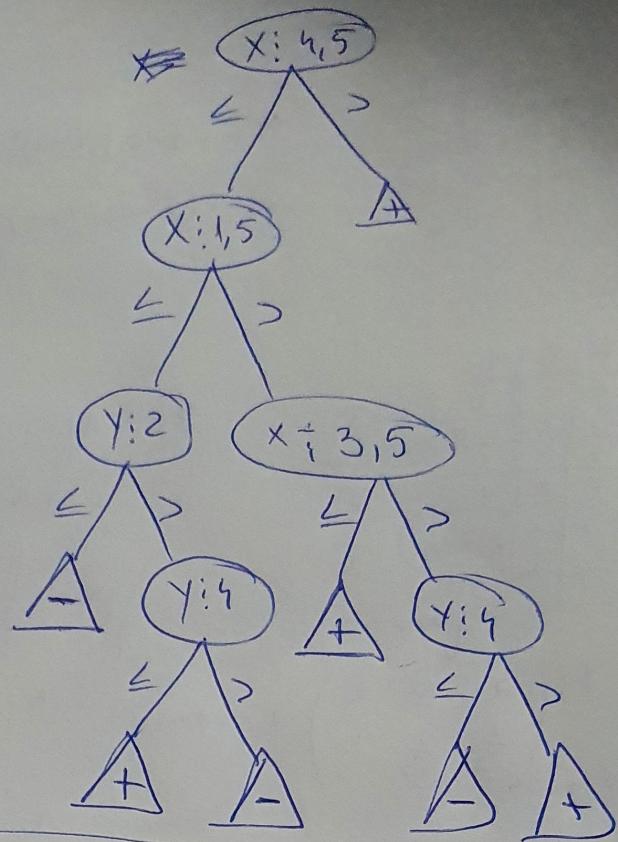
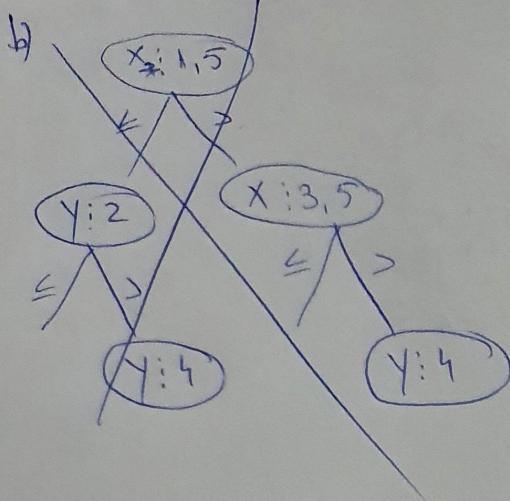
$$P_1 = P(A=0 | Y=1) \cdot P(B=0 | Y=1) \cdot P(Y=1) \stackrel{\text{MLE}}{=} \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{12} = 0,0833$$

$$P_2 = P(A=0 | Y=2) \cdot P(B=0 | Y=2) \cdot P(Y=2) \stackrel{\text{MLE}}{=} 0 \cdot 0 \cdot \frac{1}{3} = 0$$

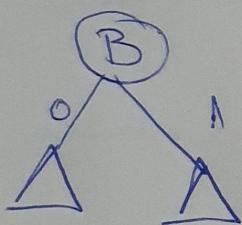
$$P_0 > P_1 > P_2 \Rightarrow Y_{NB} = 0. \quad iii) 0,1666$$

①

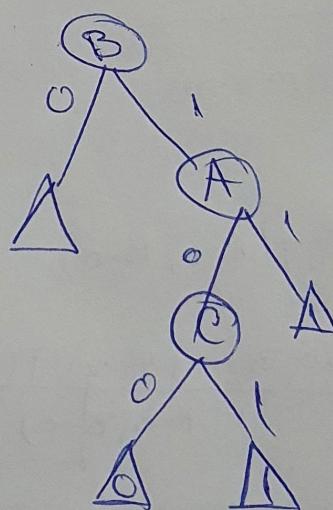
9) a) + freeze



10. a)



b)

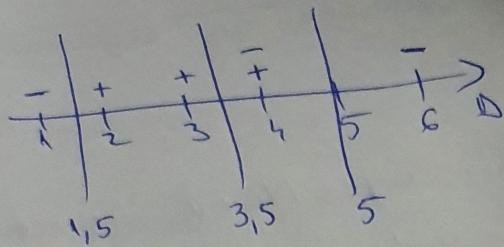


11. a)

c	y	0	-	+	+	-	-	c
2	0	1	2	3	4	5	6	
2	1	1	1	2	3	4	5	
1	1	1	1	3,5				
5	0	0	0	5,5				
6	0	0	0					
5	0	0	0					

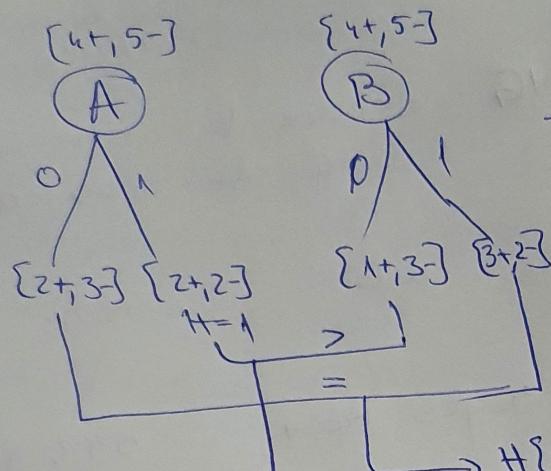
3 prague pt c

0	4
1	0
2	1
3	1
4	1
5	0
6	6



3 praguri.

3. ~~H~~ ~~H~~



→ B este nodul mai bun

(cu entropia media mai mică ↔ cu & mai mare)

$$H\{a-, b+\} = H\{a+, b-\}$$

$$0 \leq H\{a+, b-\} \text{ pt Bernoulli} \leq 1$$

$$H\{a+, b-\} < 1 \Leftrightarrow a \neq b$$

$$\rightarrow \# H\{1+, 3-\} < H\{2+, 2-\}$$

$$2. H\{2+, 3-\} \quad H\{4+, 5-\}.$$

$$H\{2+, 3-\} = H\{2-, 3+\}$$

$$H\{4+, 5-\} = H\{4-, 5+\}$$

$$H\{2+, 3-\} = \frac{\min(2, 3)}{2+3} = \frac{2}{5} \quad \left\{ \begin{array}{l} \frac{9}{5} \\ \Rightarrow \frac{2}{5} \boxtimes \frac{4}{9} \end{array} \right\} \Rightarrow H\{2+, 3-\} < H\{4+, 5-\}$$

$$H\{4+, 5-\} = \frac{\min(4, 5)}{4+5} = \frac{4}{9} \quad \left\{ \begin{array}{l} \frac{5}{9} \\ \frac{18}{45} \boxtimes \frac{20}{45} \end{array} \right\}$$

3. 1. ~~c)~~ Setul de date este consistent \Rightarrow eroarea la $\Delta_3 = 0$

b) $\text{err}_1 = \frac{1}{2} \cdot 0 \quad (0, 1, 1) - \text{e corect} \quad \left(\frac{0}{2} = 0 \right)$

$\text{err}_2 = \frac{1}{2} \quad (1, 1, 1) - \text{diferă} \Rightarrow \frac{1}{2} \text{ una dintre ele este clasificată greșit } (\frac{1}{2}) \quad \text{⊗}$

8. Este ales ID3 deoarece are cea mai mică eroare la validare

b) UNDERFITTING : Bayes Naïv deosebește arătăre de la antrenare foarte mare: 54%.

~~classificare~~ dacă luăm ca minim 10%. ID3 este Gini nu are UNDERFITTING

~~OVERFITTING~~

OVERFITTING : eroare la antrenare relativ mică față de eroarea la validare.

ID3 cu Gini

5%

40%

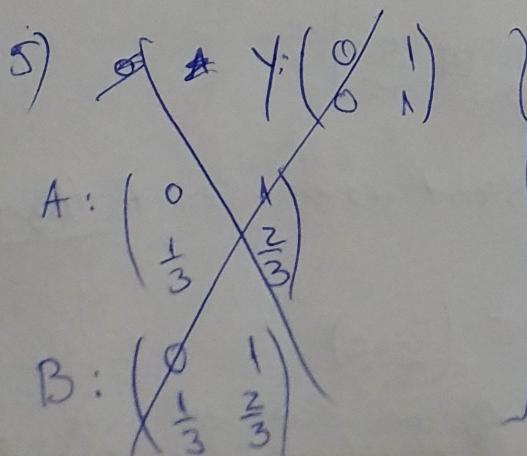
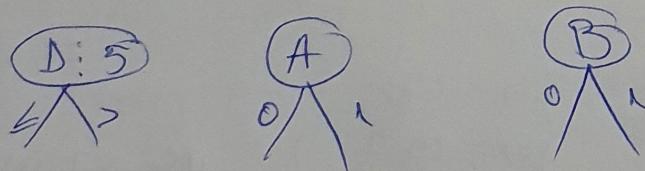
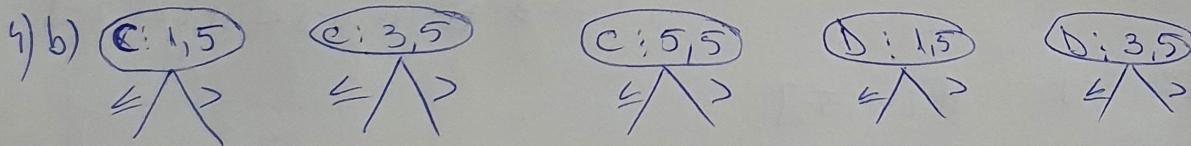
~~AVEM~~
OVERFITTING

Bayes Naïv

54%

55% ~~NU AVEM~~
OVERFITTING

c) ~~Pruning~~ La ID3 contracarăm OVERFITTINGUL prin pruning.



$$\begin{aligned}
 & \text{(ORO)} \rightarrow P(A=0|Y=0) \cdot P(B=0|Y=0) \cdot P(Y=0) \\
 & = 0. \quad \text{eroare} \\
 & P(A=0|Y=1) \cdot P(B=0|Y=1) \cdot P(Y=1) \\
 & = \frac{1}{3} \cdot \frac{1}{3} \cdot 1 = \frac{1}{9} \rightarrow Y_{NB} = 1 \\
 & \text{eror NB} = 1
 \end{aligned}$$

$$(0,0) : \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$$

$$(0,1) : \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$$

prima linie: de testare: $(0,0,0)$

$(0,0)$ va fi clasificat $1 \neq 0$ $\text{err}_1 = 1$

a doua linie: de testare $(0,0,1)$

$(0,0)$ va fi clasificat $0 \neq 1$ $\text{err}_2 = 1$

a treia linie: de testare: $(1,1,1)$

$(1,1)$ va fi clasificat $1 = 1$ $\text{err}_3 = 0$

a patra linie: de testare $(1,1,0)$

$(1,1)$ va fi clasificat $1 = 1$ $\text{err}_4 = 0$

$$\text{err}_{\text{average}} = \frac{1}{2}$$

Bayes Optimal: $\arg \max_{P(A=0,1)} P(A=1, B) P(Y|A, B)$

	A	B	Y
1	0	0	0
2	0	0	1
3	1	1	1
4	1	1	1

	A	B	Y
1	0	0	1
2	1	1	1
3	1	1	1
4	1	1	1

	A=0	B=0	B=1
1	$\frac{1}{3}$	0	
2	0	$\frac{2}{3}$	

	A	B	Y
1	0	0	0
2	0	0	0
3	1	0	0
4	1	0	0

	A=0	B=0	B=1
1	1	0	0
2	0	0	0

	A=0	B=0	B=1
1	0	0	0
2	0	0	0

$$1. P_0 = 0 \\ P_1^{\text{MALE}} = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9} \quad \left\{ \Rightarrow Y_{JB} = 1, \text{err}_1 = 1 \right.$$

$$2. q_0 = 1 - 1/2 = 1/2 \\ P_1 = 0 \cdot 1 = 0 \quad \left\{ \Rightarrow Y_{JB} = 0, \text{err}_2 = 1 \right.$$

	A	B	Y
1	0	0	0
2	1	0	0
3	0	1	0
4	1	1	0

	A=0	B=0	B=1
1	1	0	0
2	0	0	0

	A=0	B=0	B=1
1	0	0	0
2	0	0	0

$$3) P_0 = 0 \\ P_1 = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3} \quad \left. \begin{array}{l} \end{array} \right\} \Rightarrow Y_{JB} = 1 \\ err_3 = 0$$

$$4) \text{ similar zu 3)} \Rightarrow Y_{JB} = 1 \\ err_4 = 0$$

$$err_{JB} = \frac{1+1+0+0}{4} = \frac{2}{4} = \frac{1}{2}$$