4.

(Clasificare bayesiană: calculul ratei medii a erorilor pentru diverşi clasificatori bayesieni)

Fie funcţia $y = (A \wedge B) \vee \neg(B \vee C)$, unde A, B şi C sunt variabile aleatoare binare independente, fiecare dintre ele având posibilitatea să ia valoarea 0 cu probabilitate de 50%.

a. Câţi parametri trebuie să estimeze clasificatorul Bayes Naiv pentru a învăţa funcţia y? Enumeraţi aceşi parametri. *Atenţie*: $P(\neg x)$ nu va fi socotit ca parametru dacă $P(x)$ a fost deja estimat ca parametru.

b. Care este rata medie a erorii pentru clasificatorul Bayes Naiv la învăţarea conceptului y, presupunând că avem o infinitate de date de antrenament? *Convenţie*: În cazul în care, pentru o setare oarecare a variabilelor A, B şi C, cele două probabilităţi calculate de către algoritmul Bayes Naiv în vederea determinării valorii y_{NB} sunt egale, convenim că algoritmul va lua decizia $y_{NB} = 1$.

c. Câţi parametri trebuie să estimeze clasificatorul Bayes Corelat?

d. Care este rata medie a erorii pentru clasificatorul Bayes Corelat la învăţarea conceptului y, presupunând acelaşi lucru ca mai sus? *Atenţie*: Nu este nevoie să calculaţi efectiv această rată; este suficient să indicaţi valoarea ei şi să o justificaţi printr-un *raţionament calitativ.*

e. Considerăm un alt clasificator de tip Bayes, care presupune că A este independent de C, condiţionat de B şi y — în contrast cu clasificatorul Bayes Naiv, care presupune că variabilele A, B şi C sunt independente două câte două în raport cu y.

Arătaţi că acest clasificator Bayes va avea nevoie să estimeze mai puţini parametri decât clasificatorul Bayes Corelat la învăţarea conceptului y, şi totuşi va obţine aceeaşi rată medie a erorii (considerând că este valabilă aceeaşi presupoziţie în legătură cu datele de antrenament). Veţi calcula — fie lucrând efectiv, fie în baza unui raţionament concis dar riguros — rata medie a erorii pentru acest clasificator.

A	B	C	$A \wedge B$	$B \vee C$	$\neg(B \vee C)$	y
0	0	0	0	0	1	1
0	0	1	0	1	0	0
0	1	0	0	1	0	0
0	1	1	0	1	0	0
1	0	0	0	0	1	1
1	0	1	0	1	0	0
1	1	0	1	1	0	1
1	1	1	1	1	0	1

a) Bayes Naiv:

$y = \arg\max_{y \in Y} P(y \mid A=a, B=b, C=c) = \arg\max_{y \in Y} P(A=a, B=b, C=c \mid y=y) \cdot P(y=y)$

$= \arg\max_{y \in Y} P(A=a, B=b, C=c \mid y=y) \cdot P(y=y)$ $\quad P(A=a, B=b, C=c) \overset{BN}{=}$

$= \arg\max P(A=a \mid y=y) \cdot P(B=b \mid y=y) \cdot P(C=c \mid y=y) \cdot P(y=y)$

BN trebuie să estimeze:

$P(y=0) \Rightarrow$ află $P(y=1) = 1 - P(y=0)$

$P(A=0 \mid y=0) \Rightarrow$ află $P(A=1 \mid y=0)$ $\quad P(C=0 \mid y=0) \Rightarrow \ldots$

$P(A=0 \mid y=1) \Rightarrow$ află $P(A=1 \mid y=1)$ $\quad P(C=0 \mid y=1) \Rightarrow$

$P(B=0 \mid y=0) \Rightarrow$ — $P(B=1 \mid y=0)$

$P(B=0 \mid y=1) \Rightarrow P(B=1 \mid y=1)$

$\Rightarrow$ în total $\cancel{6}+1 = \cancel{7}$ parametri (pe cazul general: $2 \cdot n + 1$, unde n = nr atribute.

c) Bayes Corelat va avea de estimat: $2^{n+1} - 1 = 2^4 - 1 = 15$ $\Rightarrow$

$y = \arg\max_{y \in Val(y)} P(A=a, B=b, C=c \mid y=y) \cdot P(y=y) =$

$= \arg\max P(A=a, B=b, C=c, y=y)$

$P(y=0) \Rightarrow P(y=1) = 1 - P(y=0)$

Din fiecare grup de 2^3 perechi de valori posibile pt A, B şi C una din ele e dedusă pe baza celorlalte. $\Rightarrow (2^3 - 1)$ calcule $\cdot 2 \Rightarrow 2^4 - 2$

valori y

$2^4 - 2 + 1 \Rightarrow 2^4 - 1 = 15$.

b)