4. **(Algoritmul Bayes Naiv: calculul ratei medii a erorii – exemplificare)**

Considerăm o problemă de clasificare binară în care fiecare exemplu de antrenament are două atribute binare $X_1, X_2 \in \{T, F\}$ şi eticheta/clasa $Y \in \{T, F\}$.

Presupunem că $P(Y = T) = 0.5$, iar $P(X_1 = T|Y = T) = 0.8$, $P(X_1 = F|Y = F) = 0.7$, $P(X_2 = T|Y = T) = 0.5$ şi $P(X_2 = F|Y = F) = 0.9$. (Se poate observa că atributul X_1 furnizează/constituie un indiciu întrucâtva mai puternic decât atributul X_2 în ce priveşte determinarea clasei unei instanţe oarecare.)

	$Y = T$	$Y = F$
$X_1 = T$	0.8	0.3
$X_1 = F$	0.2	0.7

	$Y = T$	$Y = F$
$X_2 = T$	0.5	0.1
$X_2 = F$	0.5	0.9

În cele ce urmează vom presupune că X_1 şi X_2 sunt independente în raport cu Y.

a. Calculaţi probabilităţile $P(X_1 = F|Y = T)$, $P(X_1 = T|Y = F)$, $P(X_2 = F|Y = T)$ şi $P(X_2 = T|Y = F)$ şi puneţi rezultatele în tabelele de mai sus. Asociaţi răspunsului dumneavoastră o justificare generală, sub forma unei formule din teoria probabilităţilor:

$$P(\neg A|B) = \ldots, \text{ unde } A \text{ şi } B \text{ sunt evenimente aleatoare oarecare.}$$

Demonstraţi formula respectivă.

b. Scrieţi regula de decizie a algoritmului Bayes Naiv pentru $X_1 = x_1$ şi $X_2 = x_2$.

c. Calculaţi rata medie a erorii produse de algoritmul Bayes Naiv, atunci când se folosesc ambele atribute, X_1 şi X_2. (Veţi da în prealabil definiţia ratei medii a erorii.)

***Sugestie:* Este convenabil să centralizaţi calculele într-un tabel de forma următoare:**

P(X₂|Y).

X_1	X_2	Y	$P(X_1, X_2, Y) = \ldots$ P	$Y_{NB}(X_1, X_2)$	$P(x_1, x_2, x_3, Y) = P \cdot P(x_3\|Y)$	$Y_{NB}(x_1, x_2, x_3)$
F	F	F	$0.7 \cdot 0.9 \cdot 0.5 = 0.315$	F	$0.315 \cdot 0.9 = 0.2835$	F
F	F	T	$0.2 \cdot 0.5 \cdot 0.5 = \boxed{0.05}$	F	$0.05 \cdot 0.5 = 0.025$	F
F	T	F	$0.7 \cdot 0.1 \cdot 0.5 = \boxed{0.035}$	T	$0.035 \cdot 0.1 = 0.0035$	T
F	T	T	$0.2 \cdot 0.5 \cdot 0.5 = 0.05$	T	$0.05 \cdot 0.5 = 0.025$	T
T	F	F	$0.3 \cdot 0.9 \cdot 0.5 = \boxed{0.135}$	T	$0.135 \cdot 0.9 = 0.1215$	F
T	F	T	$0.8 \cdot 0.5 \cdot 0.5 = 0.2$	T	$0.2 \cdot 0.5 = 0.1$	F
T	T	F	$0.3 \cdot 0.1 \cdot 0.5 = \boxed{0.015}$	T	$0.015 \cdot 0.1 = 0.0015$	T
T	T	T	$0.8 \cdot 0.5 \cdot 0.5 = 0.2$	T	$0.2 \cdot 0.5 = 0.1$	T

$x_3 \neq x_2$

de fapt, în final nu contează la calculul ratei (la d)

d. Să presupunem acum că se crează un nou atribut, X_3, care este o copie exactă a lui X_2. Aşadar, pentru fiecare exemplu de antrenament, atributele X_2 şi X_3 au aceeaşi valoare, $X_2 = X_3$. Răspundeţi la următoarele întrebări:

– Sunt X_2 şi X_3 independente condiţional în raport cu Y?

– Cât este rata medie a erorii pentru Bayes Naiv acum? (Atenţie! Distribuţia „adevărată“ a datelor nu s-a modificat.)

– Explicaţi ce se întâmplă cu algoritmul Bayes Naiv.

a) $P(X_1 = F | Y = T) = 1 - P(X_1 = T | Y = T) = 1 - 0.8 = 0.2$
$P(X_1 = T | Y = F) = 1 - P(X_1 = F | Y = F) = 1 - 0.7 = 0.3$
$P(X_2 = F | Y = T) = 1 - P(X_2 = T | Y = T) = 1 - 0.5 = 0.5$
$P(X_2 = T | Y = F) = 1 - P(X_2 = F | Y = F) = 1 - 0.9 = 0.1$,
~~toate~~ deoarece $P(\neg A | B) = P(\overline{A} | B) = 1 - P(A|B)$.

Demonstrăm că $P(\neg A|B) = 1 - P(A|B) \Leftrightarrow$
$\Leftrightarrow P(\neg A|B) + P(A|B) = 1 \Leftrightarrow \frac{P(B \cap \neg A)}{P(B)} + \frac{P(B \cap A)}{P(B)} = 1 \Leftrightarrow$
$\Leftrightarrow P(B \cap A) + P(B \cap \neg A) = P(B)$. (1)

Cum $A, \bar{A}$ - disjuncte $\Rightarrow A \cap B$, $\neg A \cap B$ - disjuncte $\Rightarrow$
$\Rightarrow P(A \cap B) + P(\neg A \cap B) = P((A \cap B) \cup (\neg A \cap B)) \overset{\text{distrib.}}{=}$

propr. de aditivit. numărabilă.

$\overset{\text{distrib.}}{=} P(B \cap (A \cup \bar{A})) = P(B \cap \Omega) = P(B)$ (2)

Din (1) şi (2) $\Rightarrow \boxed{P(\neg A|B) = 1 - P(A|B)}$.

b) $\hat{Y}_{NB}(x_1, x_2) = \underset{y \in \{T, F\}}{\operatorname{argmax}} P(Y = y | X_1 = x_1, X_2 = x_2) \overset{\text{T.B.}}{=} \underset{y \in \{T,F\}}{\operatorname{argmax}}$

$= \underset{y \in \{T,F\}}{\operatorname{argmax}} \frac{P(X_1 = x_1, X_2 = x_2 | Y = y) \cdot P(Y = y)}{\underbrace{P(X_1 = x_1, X_2 = x_2)}_{\text{constant}}} \overset{\text{N.B.}}{=}$

$= \boxed{\underset{y \in \{T,F\}}{\operatorname{argmax}} P(X_1 = x_1 | Y = y) \cdot P(X_2 = x_2 | Y = y) \cdot P(Y = y)}$

c) Rata medie a erorii: $P(\hat{Y}_{NB} \neq Y)$

Scriem calculele direct în tabel, folosind formula:
$P(x_1, x_2, Y) = P(x_1, x_2 | Y) \cdot P(Y) \overset{\text{ind.}}{=} P(x_1 | Y) \cdot P(x_2 | Y) \cdot P(Y)$.

Cum X_1 şi X_2 sunt independente faţă de Y (ipoteză) $\Rightarrow$ distr. lui Bayes Naiv este ~~egală~~ identică cu cea reală $\Rightarrow$ putem calcula rata medie a erorilor pe tabelul dat $\Rightarrow$

$\Rightarrow P(\hat{Y}_{NB} \neq Y) = \sum_{x_1} \sum_{x_2} P(\hat{Y}_{NB}(x_1, x_2) \neq Y | X_1 = x_1, X_2 = x_2) =$
$= 0.05 + 0.035 + 0.135 + 0.015 = \boxed{0.235}$

d) Evident, X_2 şi X_3 nu sunt independente în raport cu Y.
Algoritmul Bayes Naiv va avea o rată medie a erorilor mai mare decât înainte, tocmai pt. că s-a presupus independenţa lui X_2 şi X_3 în raport cu Y (sursă de eroare). Vom calcula noua distribuţie în acelaşi tabel, prelungind tabelul.

$P(x_1, x_2, x_3 | Y) \overset{\text{N.B.}}{=} P(x_1, x_2 | Y) \cdot P(x_3 | Y) \overset{x_3 = x_2}{=}$
$= P(x_1, x_2 | Y) \cdot P(x_2 | Y)$. Continuarea pe ciornă...

Răspuns: R.m.e $= 0.3$.

Foarte bine!