



Thuật toán

❑ Ví dụ: Car theft (trộm cắp xe)

- Attribute gồm Color , Type , Origin và subject là Stolen? (có thể bị đánh cắp hay không?)

- Data set

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes



Thuật toán

						số mẫu yes theo thuộc tính/ tổng số mẫu yes		số mẫu no của thuộc tính/tổng mẫu no			
Frequency Table		Stolen?		Likelihood Table		Stolen?		Predictor prior Probability			
		Yes	No			Yes	No	tổng mẫu của thuộc tính/tổng số mẫu			
Color	Red	3	2	Color	Red	0.6	0.4	$P(x) = P(\text{Red})$		0.5	
	Yellow	2	3		Yellow	0.4	0.6	$P(x) = P(\text{Yellow})$		0.5	
Type	Sports	4	2	Type	Sports	0.8	0.4	$P(x) = P(\text{Sports})$		0.6	
	SUV	1	3		SUV	0.2	0.6	$P(x) = P(\text{SUV})$		0.4	
Origin	Domestic	2	3	origin	Domestic	0.4	0.6	$P(x) = P(\text{Domestic})$		0.5	
	Imported	3	2		Imported	0.6	0.4	$P(x) = P(\text{Imported})$		0.5	
				Class prior Probability $P(c)$		0.5	0.5				
						$P(\text{Yes})$	$P(\text{No})$				
				Yêu cầu: Red, SUV, Domestic => Stolen?							
				$P(x c) = P(\text{Red} \text{Yes})$		0.6	0.4	$P(x c) = P(\text{Red} \text{No})$			
				$P(x c) = P(\text{SUV} \text{Yes})$		0.2	0.6	$P(x c) = P(\text{SUV} \text{No})$			
				$P(x c) = P(\text{Domestic} \text{Yes})$		0.4	0.6	$P(x c) = P(\text{Domestic} \text{No})$			
		$P(c x) = P(x c) * P(c)/P(x)$		$P(c x) = P(\text{Yes} \text{Red})$		0.6	0.4				
				$P(c x) = P(\text{Yes} \text{SUV})$		0.25	0.75				
				$P(c x) = P(\text{Yes} \text{Domestic})$		0.4	0.6				
		$P(\text{Yes} X) = P(\text{Red} \text{Yes}) * P(\text{SUV} \text{Yes}) * P(\text{Domestic} \text{Yes}) * P(\text{Yes})$		Red, SUV, Domestic(Yes)		0.024					
				Red, SUV, Domestic(No)			0.072	$P(\text{No} X) = P(\text{Red} \text{No}) * P(\text{SUV} \text{No}) * P(\text{Domestic} \text{No}) * P(\text{No})$			
				Vậy kết luận là với Red, SUV, Domestic thì kết quả là NO							



Thuật toán

□ Hãy thực hiện yêu cầu bên dưới bằng cách tính Naïve Bayes

Tid	Refund	Marital Status	Taxable Income	Evade
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Với Refund = No, Marital Status = Married, và Income = 120K thì Evade=?



Thuật toán

□ Hãy thực hiện yêu cầu bên dưới bằng cách tính Naïve Bayes

	Outlook	Temperature	Humidity	Wind	Play Golf
Day					
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Outlook	Temp	Humidity	Windy	Play
Rain	Cool	High	Strong	?



Thuật toán

□ Hãy thực hiện yêu cầu bên dưới bằng cách tính Naïve Bayes

Name	Give Birth	Can Fly	Live in Water	Have Legs	Class
human	yes	no	no	yes	mammals
python	no	no	no	no	non-mammals
salmon	no	no	yes	no	non-mammals
whale	yes	no	yes	no	mammals
frog	no	no	sometimes	yes	non-mammals
komodo	no	no	no	yes	non-mammals
bat	yes	yes	no	yes	mammals
pigeon	no	yes	no	yes	non-mammals
cat	yes	no	no	yes	mammals
leopard shark	yes	no	yes	no	non-mammals
turtle	no	no	sometimes	yes	non-mammals
penguin	no	no	sometimes	yes	non-mammals
porcupine	yes	no	no	yes	mammals
eel	no	no	yes	no	non-mammals
salamander	no	no	sometimes	yes	non-mammals
gila monster	no	no	no	yes	non-mammals
platypus	no	no	no	yes	mammals
owl	no	yes	no	yes	non-mammals
dolphin	yes	no	yes	no	mammals
eagle	no	yes	no	yes	non-mammals

Give Birth	Can Fly	Live in Water	Have Legs	Class
yes	no	yes	no	?