Probability Course

Lecture 5

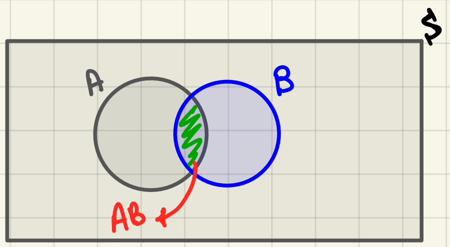
Created by : Israa Abdelghany
LinkedIn : www.linkedin.com/in/israa-abdelghany-4872b0222
GitHub : https://github.com/IsraaAbdelghany9

Session 5

we will discuss:-

- Conditional probability
- Baye's Rule
- Naïve Bayes classifier

venn diagram



$$P(A B) = P(A \cap B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B}$$

$$P(B|A) = \frac{P(AB)}{P(A)}$$

→ if A, B are independent (special case)

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

$$P(A B) = P(A)\,P(B)$$

$$P(AB) = P(B|A)\,P(A) = P(A|B)\,P(B)$$

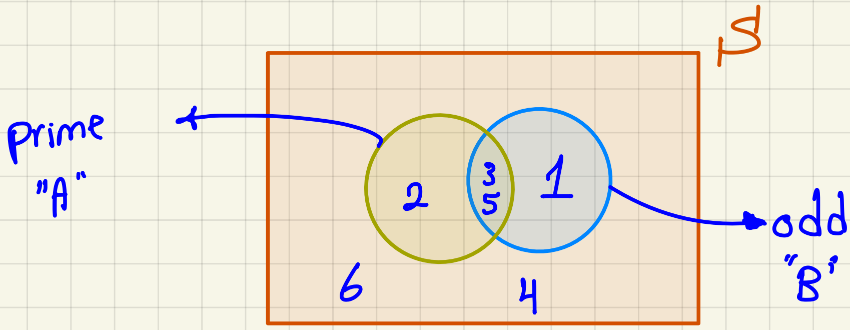
$$\underbrace{P(A|B)}_{\text{posterior}} = \frac{\overbrace{P(B|A)}^{\text{likelihood}}\ \overbrace{P(A)}^{\text{prior}}}{\underbrace{P(B)}_{\text{evidence}}}$$

prior — قبل ما اعرف إن ال B حصلت

Bayes' Rule

posterior — بعدما عرفت إن الـ B حصلت احتمال A|B

likelihood



P(odd and prime) $P(AB) = P(A,B) = P(A \cap B) = \frac{2}{6}$

p(prime given odd) $P(A|B) = \frac{2}{3}$

$2 \to$ I have 2 options in A

$3 \to$ I have 3 Numbers belong to B

$$P(B|A) = \frac{2}{3}$$

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{2/6}{3/6} = \frac{2}{3}$$

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{2/6}{3/6} = \frac{2}{3}$$

$$\underbrace{P_{X|Y}(x|y)}_{\text{Conditional Probability}} = \underbrace{P_{X,Y}(x,y)}_{\text{Joint Probability}} \Big/ \underbrace{P_Y(y)}_{\text{Marginal Probability}}$$

$$\underbrace{f_{X|Y}(x|Y)}_{\text{Conditional PDF}} = \underbrace{f_{X,Y}(x,y)}_{\text{Joint PDF}} \Big/ \underbrace{f_Y(y)}_{\text{Marginal PDF}}$$

$\Rightarrow$ also Can be written in terms of CDF!

$\int$ PDF

2 gaussian examples

Multiplication Rule $\equiv$ Chain Rule

$$P(ABCD) = P(A)P(B/A)P(C/AB)P(D/ABC)$$

→ Drawing cards without replacement..

$$P(A^1 A^2 A^3 A^4) = P(A^1)P(A^2/A^1)P(A^3/A^1A^2)P(A^4/A^1A^2A^3)$$

$$= \frac{4}{52} * \frac{3}{51} * \frac{2}{50} * \frac{1}{49}$$

$= \checkmark$

Birthday Paradox → for fun

monty hall problem

Bayes' Rule _ and total probability theorem

Ex:-

→ 1% of population suffer certain allergy.

test 98% accurate in detecting allergy.

Recall

97% accurate in detecting no allergy

→ if a person took the test, and tested +ve, what is the probability that this person actually has allergy?

A : has allergy

A\` : does not have allergy

احتمال شخص عشوائي عنده allergy

$P(A) = 0.01$

$P(A`) = 0.99$

B : test +ve

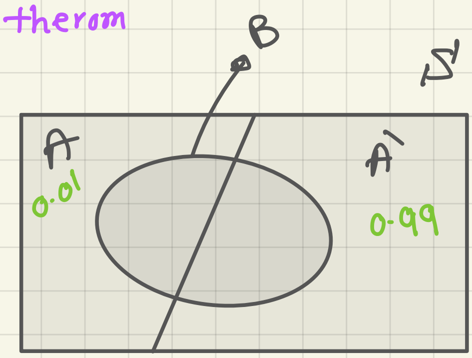
B\` : test -ve

$$P(A|B) = \frac{\overbrace{P(A)}^{\text{prior}}\ \overbrace{P(B|A)}^{\text{Liklihood}}}{\underbrace{P(B)}_{\text{evidence}}}$$

has allergy

tested +ve

I'll get it by total propability therom



$$P(A|B) = \frac{0.01 * 0.98}{0.0395}$$

$= 0.248$

belief update

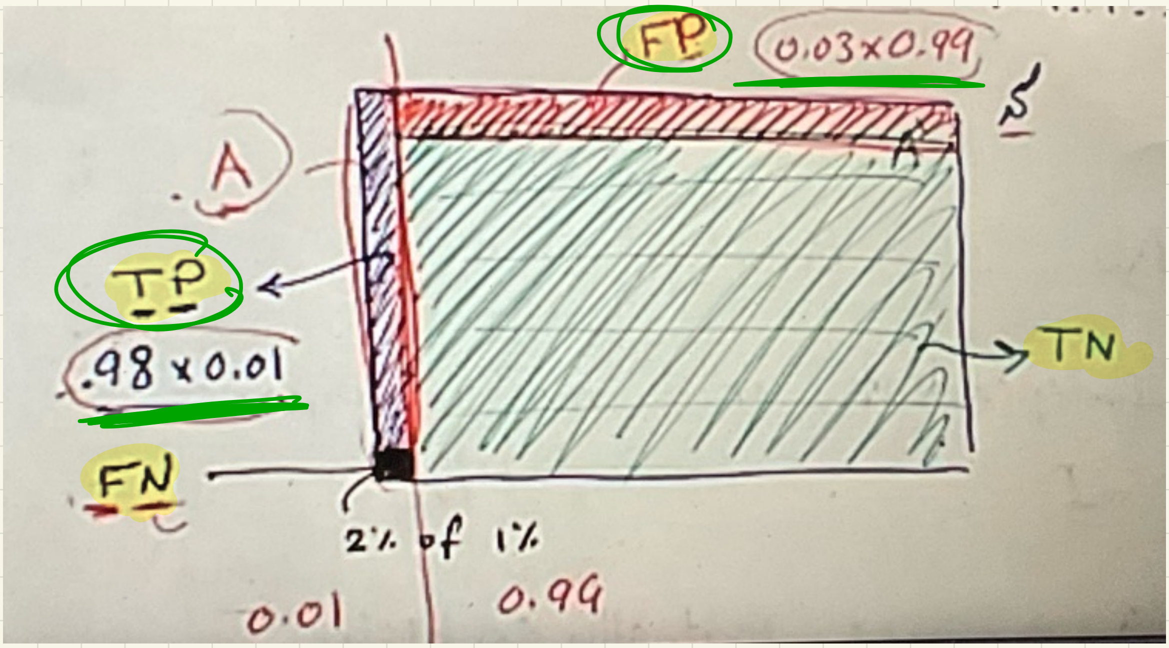
precision

$P(B) = P(A \cap B) + P(A` \cap B)$

$= P(B/A)\,P(A) + P(B/A`)\,P(A`)$

$= 0.98 * 0.01 + 0.03 * 0.99$

$= 0.0395$



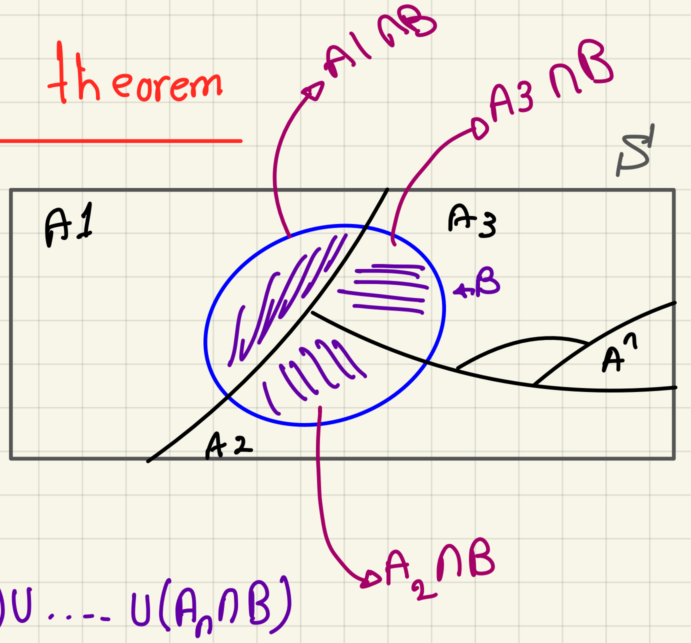
proof of the answer that FB ≅ 3 TP

total probability theorem

if $S = A_1 \cup A_2 \ldots\ldots \cup A_n$

A_i are disjoint events

"$i = 1 \longrightarrow n$"



$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)$$

$$P(B) = \sum_{i=1}^{n} P(A_i\, B)$$

Confusion Matrix

predictions

ground truth	P	N
✓	TP	FN
X	FP	TN

(Binary classifiers)

استعر التطبيقات

- Precision ↓ $= \dfrac{TP}{\text{all positive}} = \dfrac{TP}{TP + FP}$ as the example

- Recall ↑ $= \dfrac{TP}{\text{real } r} = \dfrac{TP}{TP + FN}$ (Blue rectangle) people actually has allergy & did test

- F1 - score = harmonic mean of precision & recall $= \dfrac{2}{\frac{1}{pre.} + \frac{1}{recall}}$

أصغر من أصغر واحد فيهم
زى حساب محصلة
مقاومات على
التوالى

- Accuracy ↑ $= \dfrac{\text{Correct classifications}}{\text{all classifications}} = \dfrac{TP + TN}{TP + TN + FP + FN}$

Naïve Bayes classifier

$$P(A|B) = \frac{P(A)\,P(B|A)}{P(B)}$$

Binary classifier

$X \longrightarrow$ male or female

$Y \longrightarrow$ cat or not cat

$P(\text{male} \mid \text{features}) = ?$ a

$P(\text{female} \mid \text{features}) = ?$ b

$a > b \Rightarrow$ male

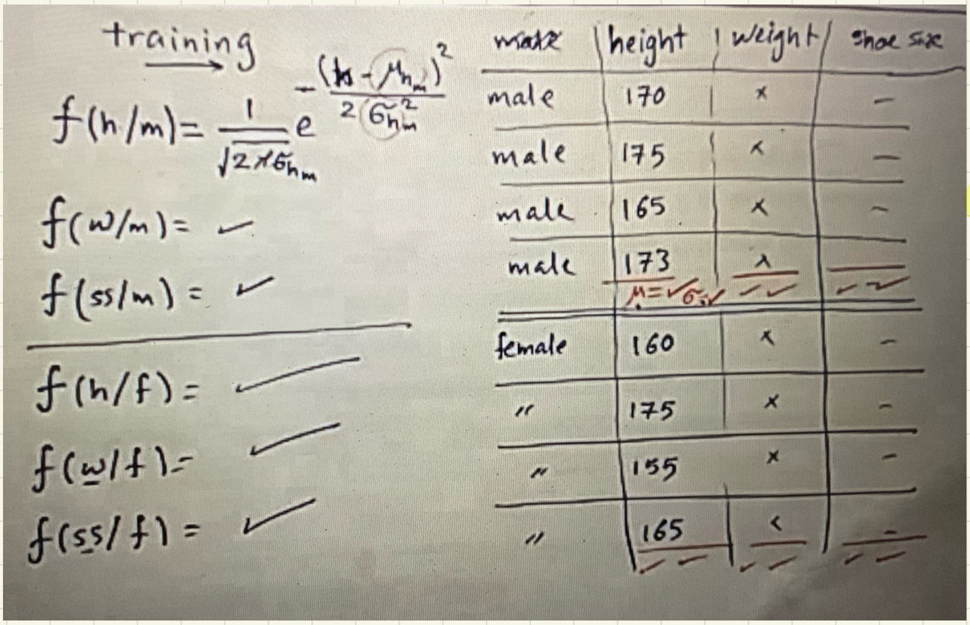
$b > a \Rightarrow$ female

$p(\text{cat} \mid \text{features}) = 0.35$

$p(\text{cat} \mid \text{features}) = 0.65$

$\Rightarrow$ not cat

training in Naive Bayes Classifier



training

$$f(h/m) = \frac{1}{\sqrt{2\pi\sigma_{h_m}}} e^{-\frac{(h-\mu_{h_m})^2}{2\sigma_{h_m}^2}}$$

$f(w/m) = \checkmark$

$f(ss/m) = \checkmark$

$f(h/f) = \checkmark$

$f(w/f) = \checkmark$

$f(ss/f) = \checkmark$

~~male~~	height	weight	shoe size
male	170	x	-
male	175	x	-
male	165	x	-
male	173	x	-
	$\mu = \checkmark$ $\sigma = \checkmark$	✓ ✓	✓ ✓
female	160	x	-
"	175	x	-
"	155	x	-
"	165	x	-
	✓ ✓	✓ ✓	✓ ✓

Wikipedia example

Binary classifiers & Naïve Bayes classifier

$$p(male \mid features) = \frac{p(male)\ p(features \mid male)}{p(features)}$$

$$p(female \mid features) = \frac{P(female)\ P(features \mid female)}{p(features)}$$

⟹ Bayes' Rule we spent alot of time calculating the P(B) [denominator]
But in Naïve I don't have to because I am compairing probability and both have same denominator.

Naïve assumption:- ⟹ but assume independence! ②

$$p(\vec{features} \mid \overset{class}{male})$$

$$= p(\underbrace{weight, height, shoe\ size}_{\text{Joint probability}} \mid class)$$

③ assume independence

⟹ features are dependent (not independent) ⟹ fact ①

③ assuming independence:-

$$P(\vec{features} \mid male) = P(height \mid male) \cdot p(w \mid male) \cdot p(shoesize \mid male)$$

$$P(\vec{features} \mid female) = P(h \mid female) \cdot p(weight \mid F) \cdot p(s.s \mid F)$$

$$\frac{1}{\sqrt{2\pi}\ \sigma_{h/f}}\, e^{-\frac{(n-\mu_{h/f})^2}{2\sigma^2_{h/f}}}$$