probability

Uncertainty

Data professionals

0 → → → - - - -

75%

Prefiction

AI

probability

* Type of probability!

1- Objective

2- ~~Subjective~~

~~Based on personal feeling.~~

based on statistics, experiments and mathematical measurements.

1- Random Experiment:

→ Statistical experiment:

→ process whose outcomes cannot be predicted with certainty.

2- Outcome:

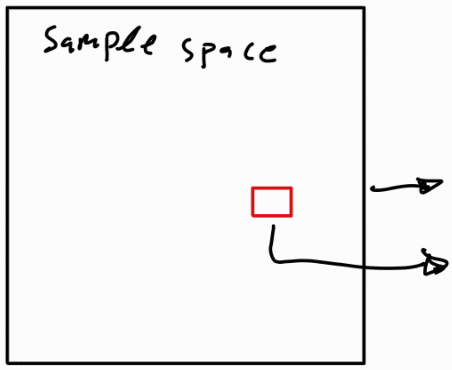
→ result of random experiment

$\{1, 2, 3, 4, 5, 6\}$

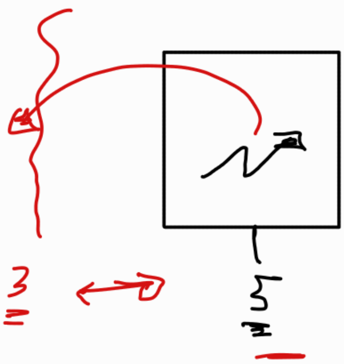
S

* Sample space

3- Event:



4- probability of event:



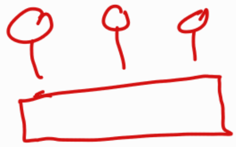
$$\frac{\text{number of } E}{\text{number of outcomes}}$$

$$0 \le \text{Prob of } E \le 1$$

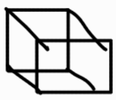
- Toss Two coin:

{HH TT HT TH}

{ [MMM MFM MFF
FFF FMM FFM
MMF FMF }



Ex



$S = \{1,2,3,4,5,6\}$

E1: odd number: $E2 = \{1,3,5\}$

$\boxed{P(E1)}$

$3 \to E$

$6 \to S$

E2: a number greater than 5.

$\{6\}$ $P(E2) = \frac{1}{6}$

E3: at most 4: $\{1,2,3,4\}$

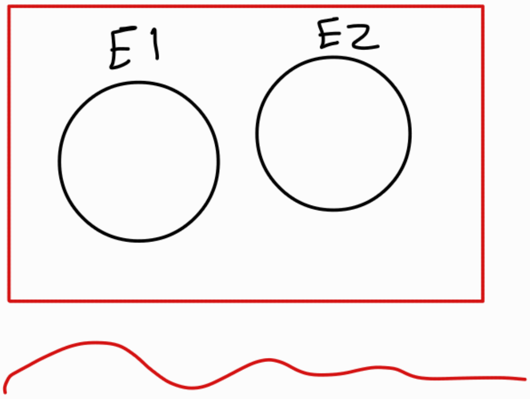
$P(E3) = \frac{4}{6} = 66.7\%$

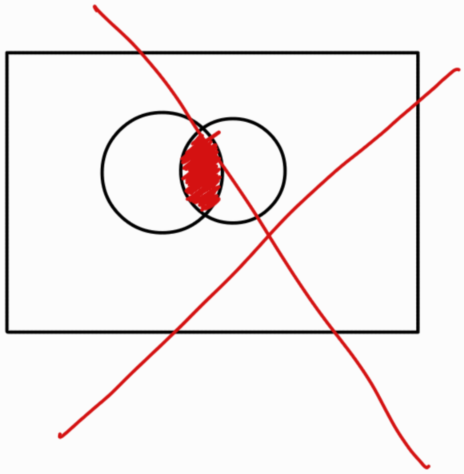
Event | Three Type of Events:

1- Mutually Exclusive
2- Independent Event
3- Dependent Event

1- Mutually Exclusive:

→ Two Event are mutually exclusive if they cannot occur at the same time





2- Independent Events:

Two Event are independent if the occurrence of one event does not change probability of the other event

3- Dependent Events:

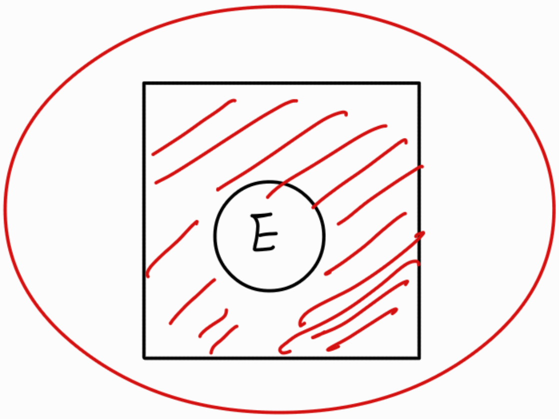
*** Three basic Rules:**

1- Complement Rule
2- Addition Rule
3- Multiplication Rule

1- Complement Rule:

$P(E) = 40\%$

$P'(E) = 60\%$



100%

1

$P(E) + P(E') = 100\%$

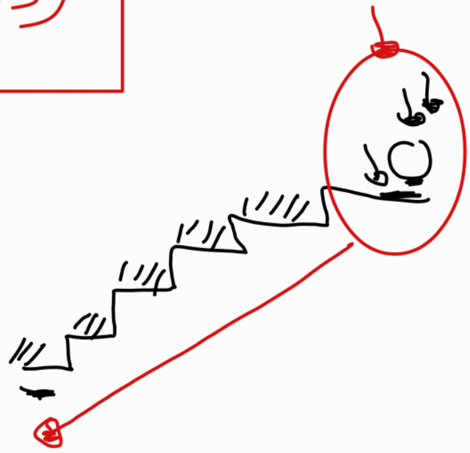
$P(E') = 1 - P(E)$

2- Addition Rule:

→ mutually Exclusive events

If A, B events ME

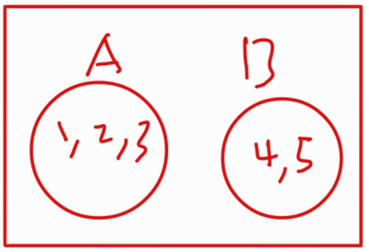
$$\boxed{P(A \text{ or } B) = P(A) + P(B)}$$



$S: \{1, 2, 3, 4, 5, 6\}$

$A: \{1, 2, 3\}$

$B: \{4, 5\}$



$P(A) = \frac{3}{6}$

$P(B) = \frac{2}{6}$

$$P(A \text{ or } B) = P(A) + P(B)$$

$$= \frac{3}{6} + \frac{2}{6} = \frac{5}{6}$$

$$P(A \text{ or } B) = P(A \cup B) = \underbrace{P(A) + P(B)}_{ME}$$

Independent Event

ME

$- P(A) \text{ or } P(B) = P(A) \cup P(B)$

$= P(A + B)$

$P(A) \text{ and } P(B) = P(A \cap B) = P(A \times B)$

Independent Event

"

" " " " "

" " " " " "

dependent Event

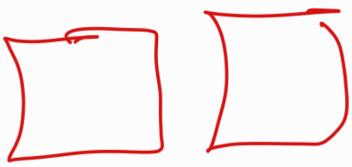
$P(A \text{ or } B) =$

$P(A + B) - P(A \text{ and } B)$

$P(A \text{ and } B) = P(A) \times P(B)$

$A \cup B$

Independent/ME



and

A B

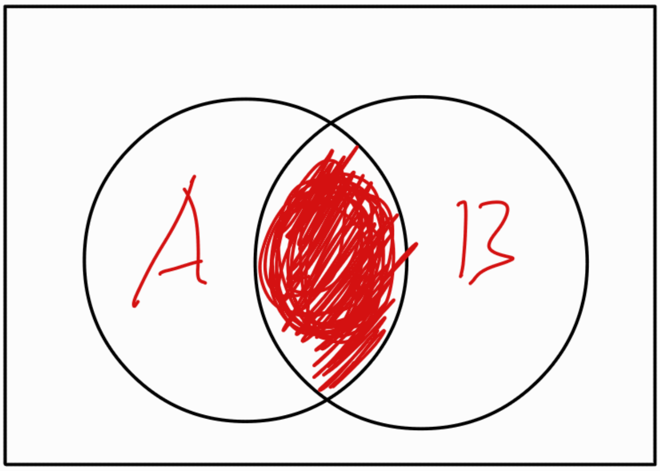
(1 2) (3 4) $A \cup B = \{1, 2, 3, 4\}$

$\cup$ ← or ← جمع ← اتحاد *

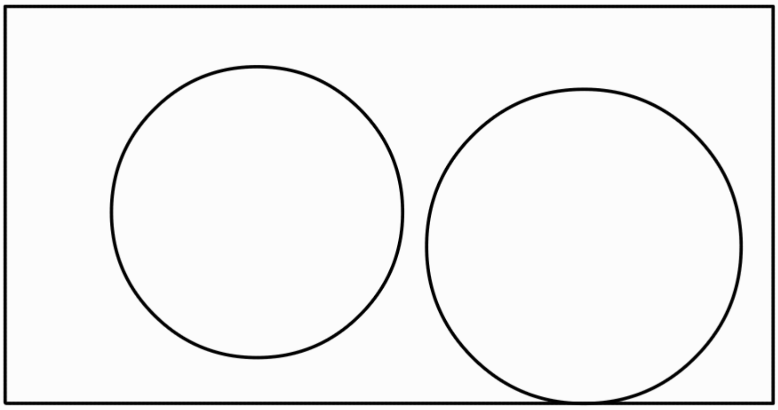
$A \cap B =$ Zero

$\cap$ ← and ← ضرب ← تقاطع *

الأشياء المشتركة



$A \cap B$



Set : (1, 2, 3) (1, 4, 5)

$\{1, 1, 2, 3, 4, 5\} - \{1\}$

$\{1, 2, 3, 4, 5\}$

* الاحتمالات توضع داخل Set ← خطأ دخل المجموعات

عدم التكرار ...

Type of E
- ME
  - Addition
  - Mult
- Ind
  - Add
  - mult
- Dep
  - Add
  - mult

$P(E1) = 30\%$

$P(E1') = 70\%$

$\cup$ → or → $+$

$\cap$ → and → $\times$

Ex: A bag consist of

8 red balls
7 blue balls
4 - Yellow balls
6 - green balls

S: 25

8 يوجد

→ what is the probability of:

1 - A red balls:

E: 8

$$P(E) = \boxed{\frac{8}{25}} = 0.32 = 32\%$$

2 - A blue balls: $\boxed{\frac{7}{25}}$

3. Blue ball in the first $\boxed{\text{and}}$ then green ball on the second with replacement

$$P(B \text{ and } g) = P(g) \times P(b) =$$

$$\frac{6}{25} \times \frac{7}{25} = \frac{42}{625} = 6.72\%$$

Or $+$ $\cup$

and $\times$ $\cap$

4- a yellow ball in the first

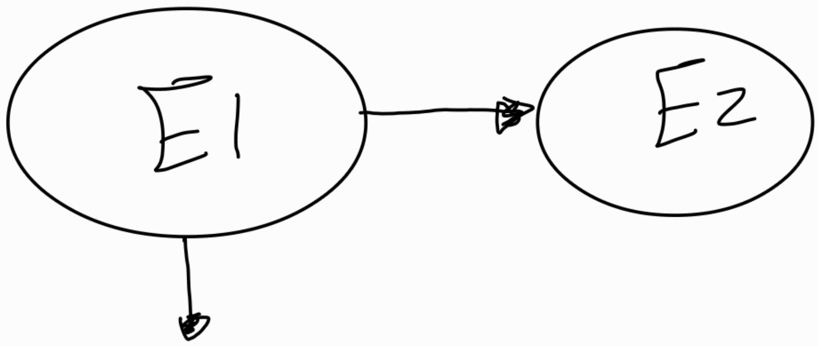
<u>and</u> then red ball on

the second

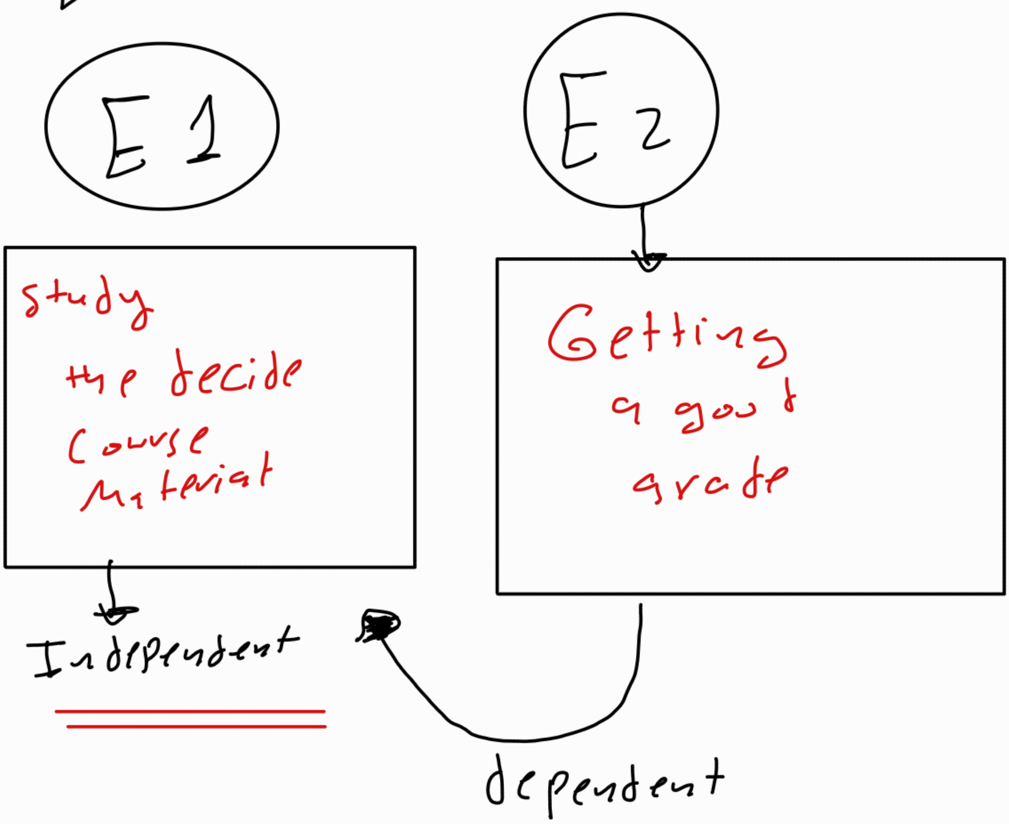
without replacement?

$$P(Y \text{ and } R) = \frac{4}{25} \cdot \frac{8}{24} = 5.3\%$$

Conditional probability



Ex 1:



- C.P: probability of an event occurring given that another event has already occurred.

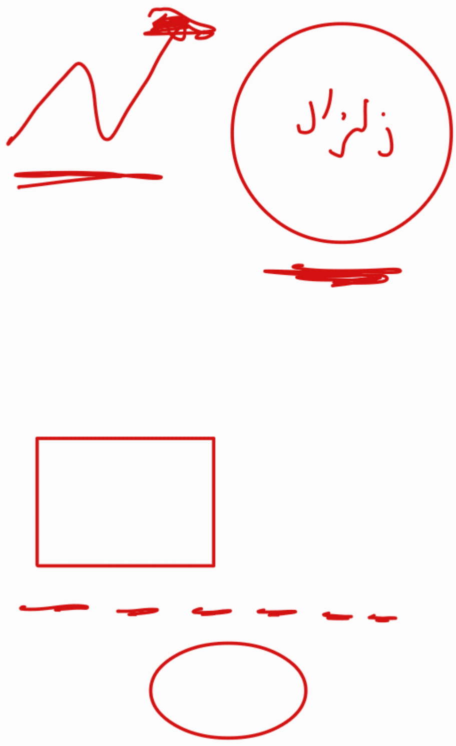
Event ← حدث يحدث الآن ولكن احتمالية حدوثه متأثرة بحدث قد انتهى

$$P(E_1 \mid E_2)$$

$$P(E1 \mid E2)$$

E1 → لم يحدث

E2: حدث وانتهى



→ A, B: Two Events with the Same Space of R.E

→ $\underline{A}$: $\boxed{\text{لم يحدث بعد}}$ → Prediction

→ $\underline{B}$: حدث وانتهى — الماضي / معطى / معلومة زيادة

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

$$= \frac{P(A \text{ and } B)}{P(B)}$$

$\underline{A}, \underline{B}$ Independent:

$$P(A \mid \underline{B}) = \frac{P(A \text{ and } B)}{P(B)}$$

$$\frac{P(A) \cdot \cancel{P(B)}}{\cancel{P(B)}} = P(A)$$

Ex: Toss a die and consider the events

$S: \{1, 2, 3, 4, 5, 6\}$ $\qquad F: \{3, 4\}$

$A: \{1, 2\}$ $\qquad B = \{2, 3, 4\}$

$C = \{2, 3, 5\}$ $\qquad D = \{2, 3, 5, 6\}$

$A, B, C, D \rightarrow$ Events

- $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{P(2)}{} = \dfrac{\frac{1}{6}}{\frac{3}{6}}$

$\frac{1}{3}$

- $P(C \mid D) = \dfrac{P(C \cap D)}{P(D)} = \dfrac{P(3)}{} = \dfrac{\frac{3}{6}}{\frac{4}{6}}$

$\frac{3}{4}$

- $P(A \mid D) = \dfrac{P(A \cap D)}{P(D)} = \dfrac{\frac{1}{6}}{\frac{4}{6}}$

$\frac{1}{4}$

$- P(A \mid f) = P(A)$

$\frac{2}{6}$

Bayes's Theorem

ارتفاع درجة الحرارة → نتيجة

إصابة بكورونا → سبب

$P($ارتفاع $\mid$ covid19$)$ ... →

$\boxed{P(\text{covid19} \mid \text{ارتفاع})}$ → $P($سبب $\mid$ نتيجة$)$

نار / سبب

دخان / نتيجة

$P($دخان $\mid$ نار$)$ → $P($نار $\mid$ دخان$)$ → مجهول

الموت — موت

السم — سبب

$P(\text{الموت} \mid \text{سم}) \rightarrow \uparrow \downarrow$

$P(\text{سم} \mid \text{موت}) \rightarrow$

Prediction

Classification

Bayes's Theorm

→ way to update the probability of event based on new information about the event.

Ali

$P(Ali)$

$P(Ali \mid \text{صحت})$

$$P(A \mid B) = \frac{P(B \mid A) * P(A)}{P(B)}$$

A	B
السم	الموت

$\rightarrow$

$$P(\underbrace{\text{السم}}_{\text{المطلوب}} \mid \underbrace{\text{موت}}_{\substack{\text{Evidence} \\ \text{Information}}}) = \frac{P(\text{الموت} \mid \text{السم}) * P(\text{السم})}{P(B)}$$

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

$P(A|B)$ — Posterior → بعد التحديث

$P(A)$ → Prior قبل التحديث

$\underline{P(A)}$ قبل

$P(B|A)$: → Likelihood → دالة الطبيعة

$P(B)$ → Evidence → التحديث

Ex:

A: Rain   B: Cloudy

$$P(\text{Rain} \mid \text{Cloudy}) = \frac{P(\text{cloudy} \mid \text{Rain}) * P(\text{Rain})}{P(\text{cloudy})}$$

$P(\text{Rain}) = 10\%$

$P(\text{Cloudy}) = 40\%$

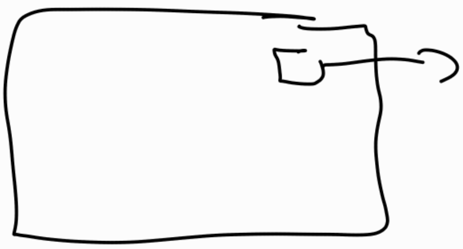
$P(\text{Cloudy} \mid \text{Rain}) = 50\%$

$$= \frac{0.5 * 0.1}{0.4}$$

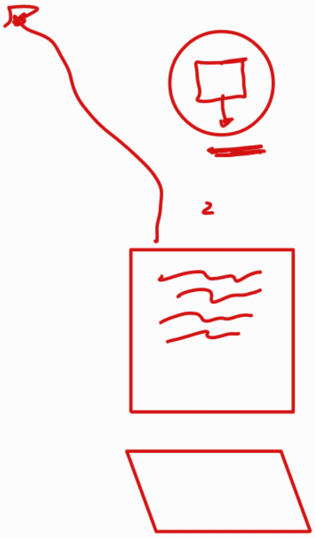
$$= 0.125 = 12.5\%$$

$$P(A) : \frac{E}{S}$$

$S \rightarrow$ الكلي



The Expanded Version of Bayes's Theorm



"predict the accuracy of a test"

$$P(\underset{\text{السم}}{A}|\underset{\text{الموت}}{B}) = \frac{P(B|A) * P(A)}{P(B|A) * P(A) + P(B|\text{not } A) * P(\text{not } A)}$$

مسموم

بطريقة

OFF

33%

70%

$$= \frac{P(\text{الموت} | \text{السم}) * P(\text{السم})}{P(\text{الموت} | \text{السم}) * P(\text{السم}) + P(\text{الموت} | \text{غير السم}) * P(\text{غير السم})}$$

100%

السم

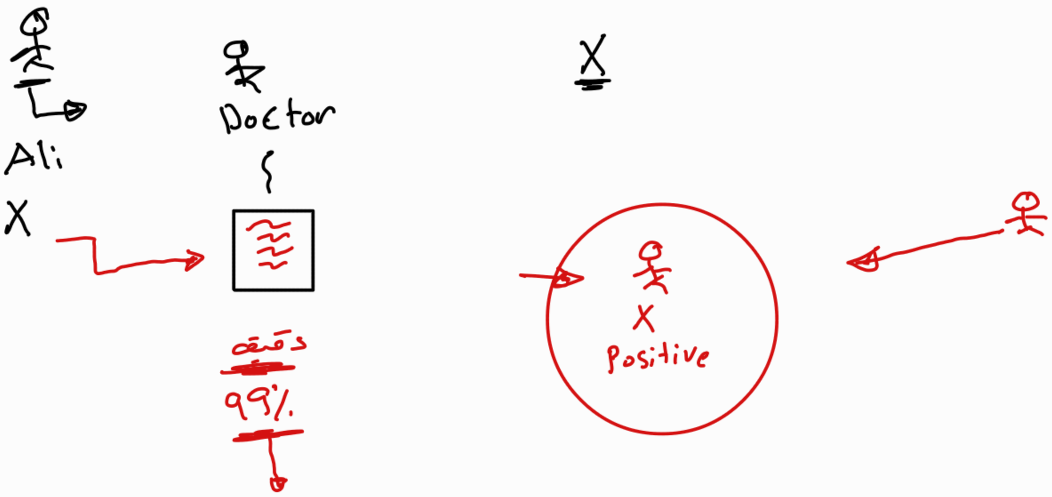
$P(\text{السم}) = 30$

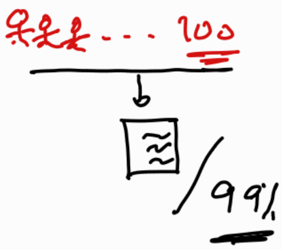
$P(\text{not السم}) = 70\%$

دخان

نار

$$P(\text{نار} \mid \text{دخان}) = \frac{P(\text{دخان} \mid \text{نار}) * P(\text{نار})}{P(\text{دخان} \mid \text{نار}) * P(\text{نار}) + P(\text{دخان} \mid \text{غير نار}) * P(\text{غير نار})}$$





99%

If 100 people diagnosed Sick

99 TP

1 FP

If 100 people Health

99 TN

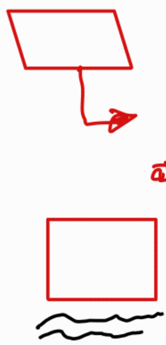
1 FN

TP	FP
FN	TN

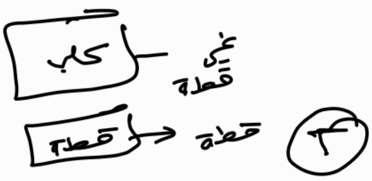
$$\frac{TP + TN}{TP + TN + FP + FN}$$

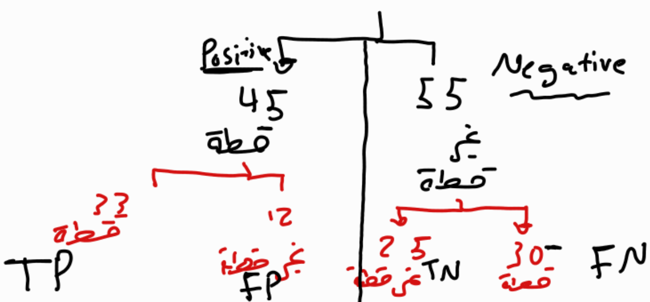
$$= \frac{99 + 99}{99 + 99 + 1 + 1} = \frac{198}{200} = 99\%$$

TP	FP
TN	FN



صح
=
القطة
او غير القطة



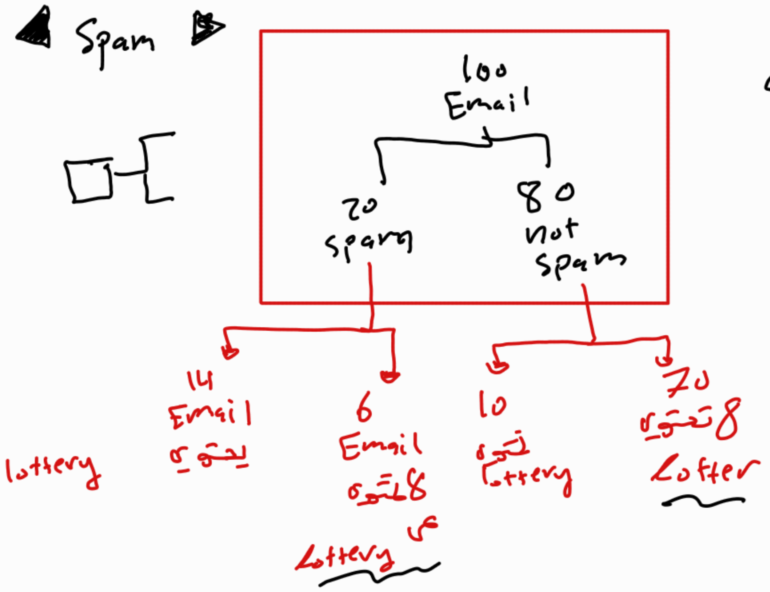


$$\frac{TP+TN}{TP+TN+fP+fN}$$

$$= \frac{0.33+0.25}{0.33+0.25+0.12+0.3} = \frac{0.58}{100} =$$

58%

Spam



classifer

lottery

* ما احتمالية ان يكون ال email الذي يحتوي ع lottery هو Spam

$$P(\text{Spam} \mid \text{lottery}) = \frac{P(\text{lottery} \mid \text{spam}) * P(\text{spam})}{P(\text{lottery} \mid \text{spam}) * P(\text{spam}) + P(\text{lottery} \mid \text{not spam}) * P(\text{not spam})}$$

lottery and spam = 14

$$P(\text{spam}) = \frac{20}{100} = 0.2$$

$$P(\text{not spam}) = 0.8$$

$$P(\text{lottery} \mid \text{spam}) = \frac{P(\text{lottery} \cap \text{Spam})}{P(\text{Spam})}$$

$$= \frac{\frac{14}{\cancel{100}}}{\frac{20}{\cancel{100}}} = \frac{14}{20} = 0.7$$

$$P(\text{lottery} \mid \text{not spam}) = \frac{P(\text{lottery} \cap \text{not spam})}{P(\text{not spam})}$$

$$\frac{\frac{10}{\cancel{100}}}{\frac{80}{\cancel{100}}} = \frac{10}{80} = 0.125$$

$$P(\text{spam} \mid \text{lottery}) = \frac{0.7 * 0.2}{0.7 * 0.2 + 0.125 * 0.8}$$

$$= 0.583 = \boxed{58.3}$$

احتمالية ان الايميل spam وهو يحتوي على lottery.

Prior

الاحتمال الأصلي

$$P(spam) = \frac{spam}{spam + not\ spam}$$

20%

Evendence

معلومة جديدة حول الاحتمال الأصلي وسوف تغيير الاحتمال اما بزيادة او بنقصان

Email contain Lottery

58%

Posterior

هو الاحتمال الجديد بعد معرفة الدليل او بعد التحديث

$$P(spam \mid lottery) = \frac{spam\ have\ lottery}{spam\ have\ lottery + not\ spam\ have\ lottery}$$

بمجرد ما أصبح هناك شرط أصبحنا نهتم فقط بالشرط نفسه في جميع حالته

lottery — spam / not spam

و فقط أهم بباقي الأمور (الاحتمالات).

Naive Bayes's

Prior: P(spam)

Evedence:
- lottery
- winning

Posterior:
P(spam | lottery)
P(spam | lottery & winning)

lottery — L
winning — w
spam → S

$$\frac{P(L \,\&\, w \mid S) * P(S)}{P(L \,\&\, w \mid S) * P(S) + P(\text{not } S) * P(L / w \mid \text{not } S)}$$

P(Lottery & winning | spam)

لهم علاقة ببعض

لكي نطبق قاعدة Naive
يجب ع الأدلة ان تكون
مستقلة .

lottery → Event
winning → Event
→ لازم مستقلة
Independent →

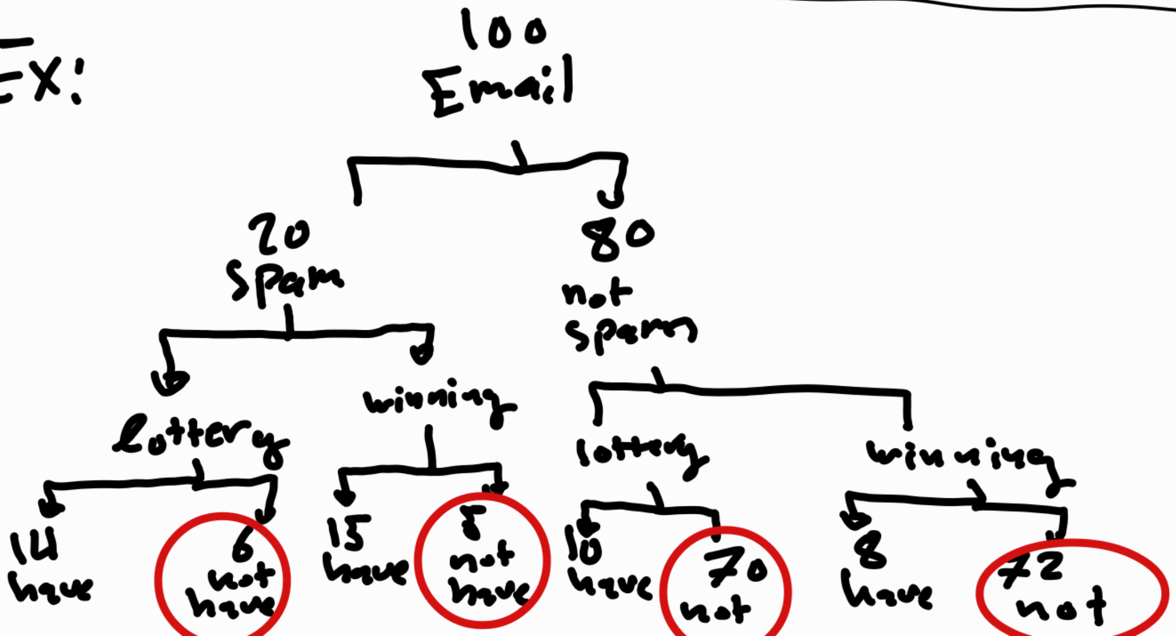
$P(\text{lottery} \ \& \ \text{winning} \mid \text{spam})$

$$P(\text{lottery} \mid \text{spam}) \cdot P(\text{winning} \mid \text{spam})$$

$P(\text{spam} \mid \text{lottery} \ \& \ \text{winning})$

$$\frac{P(\text{spam}) \cdot P(\text{lottery} \mid \text{spam}) \cdot P(\text{winning} \mid \text{spam})}{P(S) \cdot P(L \mid S) \cdot P(W \mid S) + P(\text{not } S) \cdot P(L \mid \text{not } S) \cdot P(W \mid \text{not } S)}$$

Ex:



$$\frac{0.2 \cdot \frac{14}{20} \cdot \frac{15}{20}}{0.2 \cdot \frac{14}{20} \cdot \frac{15}{20} + 0.8 \cdot \frac{10}{80} \cdot \frac{8}{80}}$$

$$= \boxed{0.913}$$

91%

إحتمالية ان يكون الايميل هو spam

علما بأنه يحتوي ع winning lottery