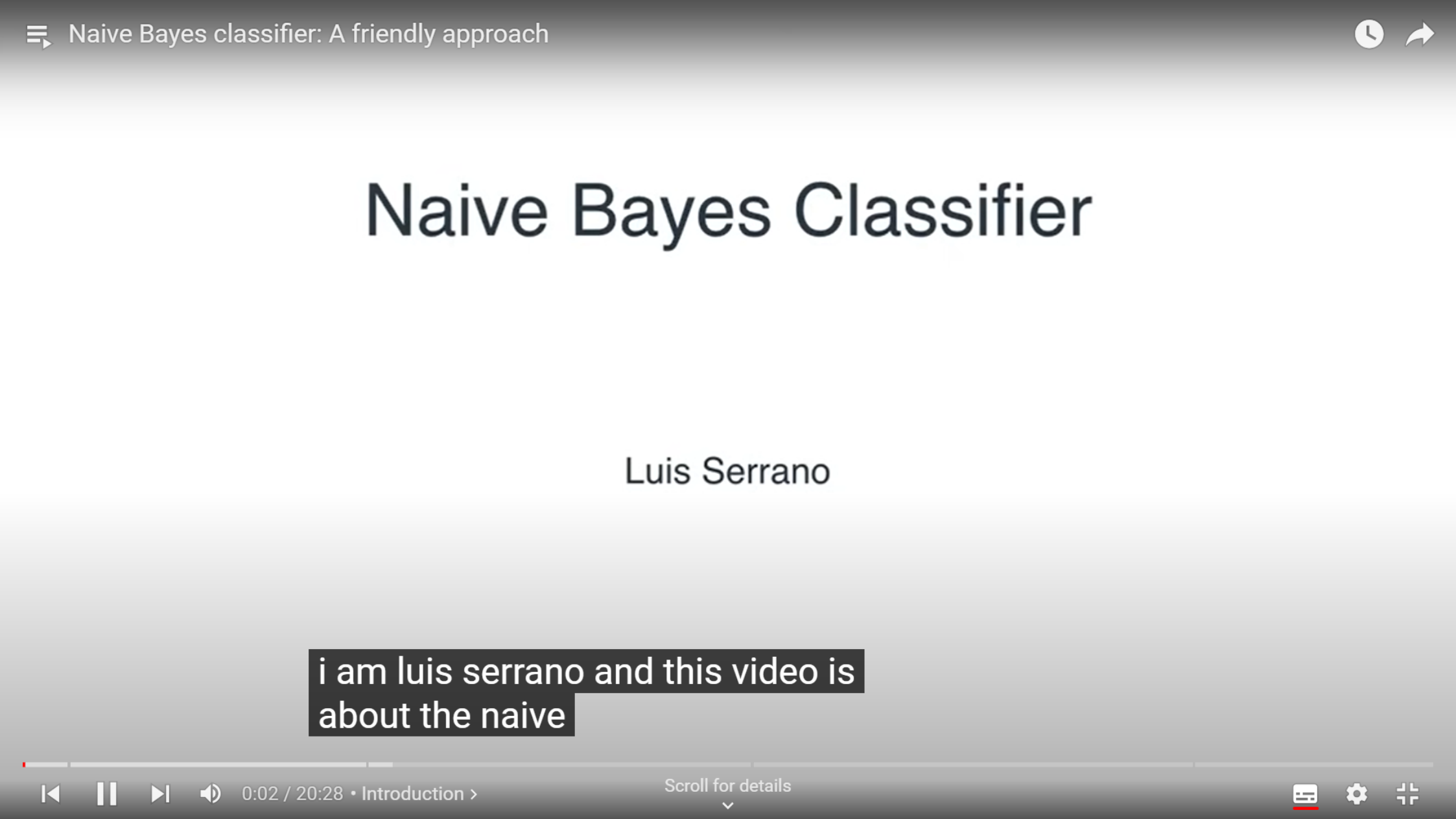
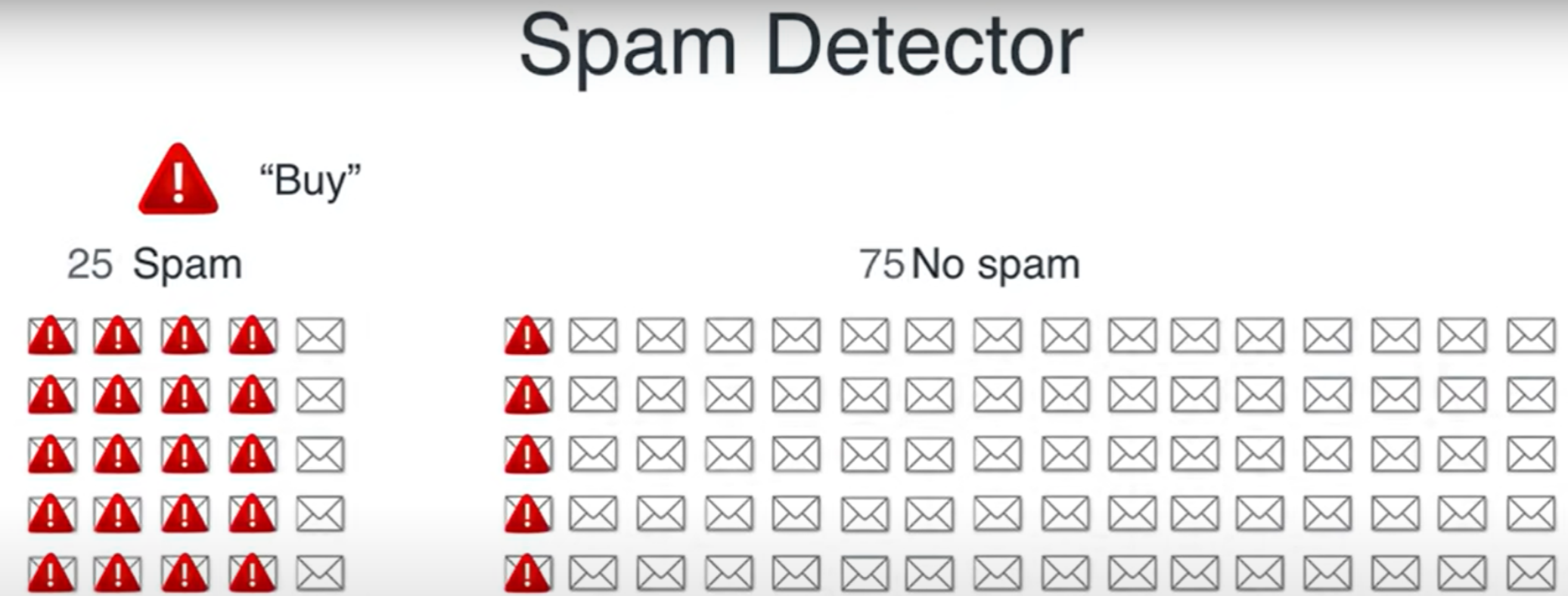
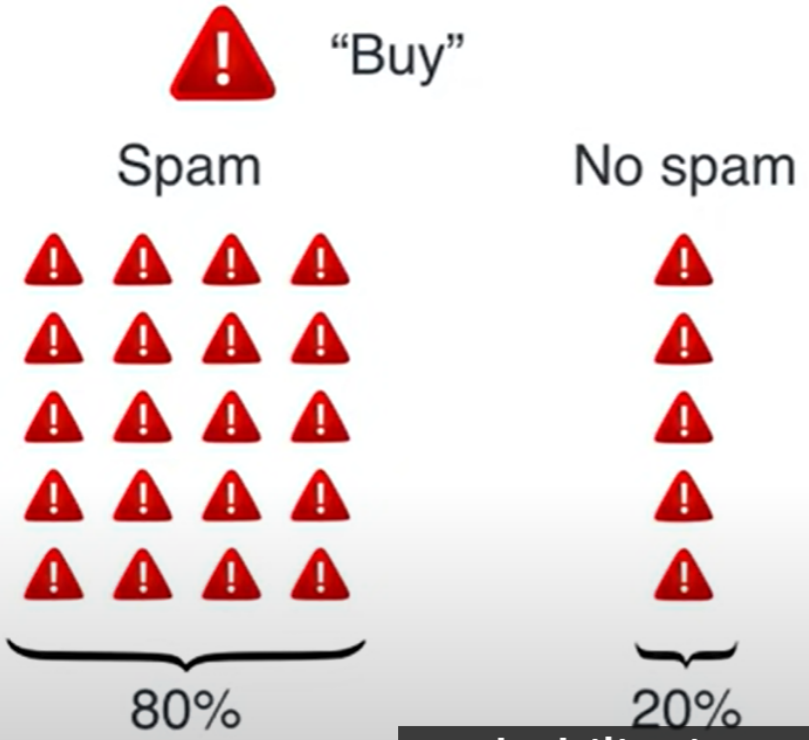

Naive Bayes Classifier
Luis Serrano
i am luis serrano and this video is
about the naive



so there's five so let's forget about
all the others and just

Spam Detector



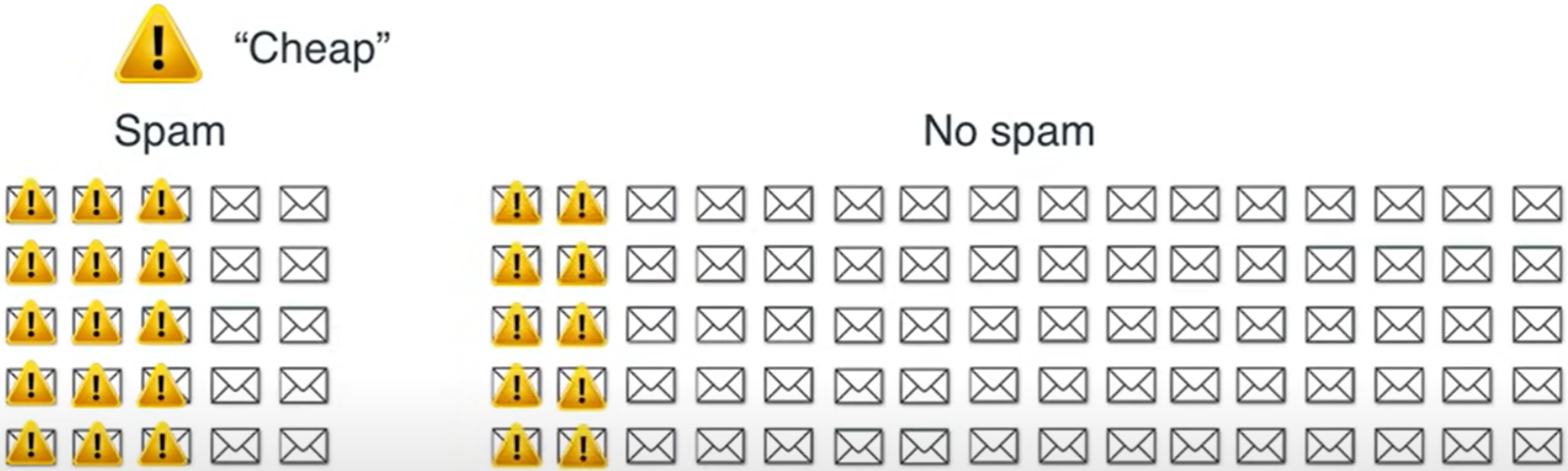
Quiz: If an e-mail contains the word "buy", what is the probability that it is spam?

- ○ 40%
- ○ 60%
- ● 80%
- ○ 100%

Solution:
80%

probability is gonna be 80 percent that it's spam if it contains

Spam Detector

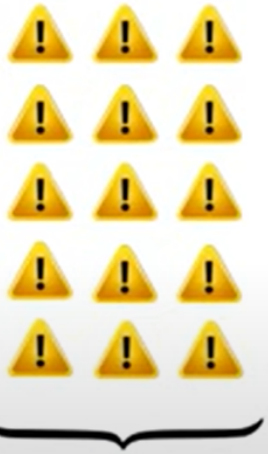


be in 15 of them and from the non-spam
ten of them I

Bayes Theorem

"Cheap"

Spam



60%

No spam

40%

Quiz: If an e-mail contains the word "cheap", what is the probability that it is spam?

- ○ 40%
- ● 60%
- ○ 80%
- ○ 100%

Solution:
60%

split and therefore the solution is 60%
so

Spam Detector

"Buy" and "Cheap"

Spam

No spam

the words buy and cheap so that's some good data and

Spam Detector

"Buy" and "Cheap" ⟶ 100% ?

Spam

100%

No spam

0%

Quiz: If an e-mail contains the words "buy" and "cheap", what is the probability that it is spam?

- ○ 40%
- ○ 60%
- ○ 80%
- ● 100%

Solution:
100%

right well maybe you're being skeptical like

Spam Detector

100 e-mails

5 “Buy”

10 “Cheap”

would contain the words buy and cheap so let's think 5 out of 100 is

Spam Detector

100 e-mails

5 “Buy”

10 “Cheap”

5% “Buy”

10% “Cheap”

0.5% “Buy” and “Cheap”

Spam Detector

Spam

25 e-mails
20 “Buy”
15 Cheap

4/5
3/5
→ 12/25

Spam Detector

Spam

25 e-mails
20 “Buy”
15 Cheap

4/5
3/5 → 12/25 × 25 = 12 “Buy” and “Cheap”

Spam Detector

No spam

75 e-mails

5 “Buy” 1/15

10 “Cheap” 2/15 → 2/225 × 75 = 2/3 “Buy” and “Cheap”

Spam Detector

"Buy" and "Cheap"

Spam

No spam


12


2/3

94.737%

Quiz: If an e-mail contains the words "buy" and "cheap", what is the probability that it is spam?

$$\frac{12}{12 + 2/3} = \frac{36}{38}$$

$$= 94.737\%$$

Naive Bayes

	Spam		No spam	
Total	25		75	
Buy	20	4/5	5	1/15
Cheap	15	3/5	10	2/15
Buy & Cheap	12	12/25	2/3	2/225

Naive Bayes

	Spam		No spam	
Total	25		75	
Buy	20	4/5	5	1/15
Cheap	15	3/5	10	2/15
Buy & Cheap	12	12/25	2/3	2/225

$$\frac{12}{12 + 2/3} = \frac{36}{38} = 94.737\%$$

Bayes Theorem

S: Spam
H: Ham (not spam)
B: ‘Buy’

$$P(S \mid B) = \frac{P(B \mid S)\,P(S)}{P(B \mid S)\,P(S) + P(B \mid H)\,P(H)}$$

Naive Bayes

$$P(\text{“Buy”} \ \& \ \text{“Cheap”}) = P(\text{“Buy”})\ P(\text{“Cheap”})$$

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

Naive

Naive Bayes

S: Spam
H: Ham (not spam)
B: 'Buy'
C: 'Cheap'

$$P(S \mid B \cap C) = \frac{P(B \mid S)\,P(C \mid S)\,P(S)}{P(B \mid S)\,P(C \mid S)\,P(S) + P(B \mid H)\,P(C \mid H)\,P(H)}$$

$$P(\text{spam if "Buy" \& "Cheap"}) = \frac{\frac{20}{25}\,\frac{15}{25}\,\frac{25}{100}}{\frac{20}{25}\,\frac{15}{25}\,\frac{25}{100} + \frac{5}{75}\,\frac{10}{75}}$$