Naïve Bayes Algorithm

Sunday, April 9, 2023 8:11 PM

Classification

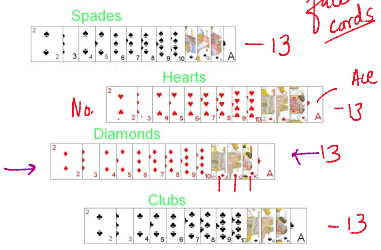
Linear — continous
Logistic — categorical / classification

Decision — continous
Random — classification

Naive Bayes → classification

Probability : Favourable outcome / Total outcome

$P(H) = \frac{1}{2} = 0.5$



you have a diamond find a prob that it is a queen

Conditional Probability : Some Cond^n

$$P(\text{queen} \mid \text{diamond}) = \frac{P(\text{diamond} \mid \text{queen}) \times P(\text{queen})}{P(\text{diamond})}$$

queen ↑ Prob that you need to find

diamond / given

$P(A/B)$ → what Prob of A given B

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

$$P(\text{queen} \mid \text{diamond}) \Rightarrow \frac{P(\text{diamond} \mid \text{queen}) \times P(\text{queen})}{P(\text{diamond})}$$

$P(\text{queen} \mid \text{diamond}) \Rightarrow \frac{P(\text{diamond} \mid \text{queen}) \times P(\text{queen})}{P(\text{diamond})}$

Prob | given

$= \frac{\frac{1}{4} \times \frac{4}{52}}{\frac{13}{52}} = \Rightarrow 1/13$

Bayes Theoreom

$\rightarrow P(D \mid Q) \leftrightarrow P(Q \mid D)$

Rain — Chirag football — Chirag football — Rain

3 | 10 — 10/10

$P(\text{Screen} \mid \text{IPHONE})$ — give — 70%.

$P(\text{Ipnow} \mid \text{Screen})$ — given — 0.001

Day	x_1 Outlook	x_2 Temperature	x_3 Humidity	x_4 Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes — y
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No — N
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

x_1 – Outlook
x_2 – Temp
x_3 – Humidity
x_4 – Wind

$P(y_i \mid x_1 x_2 x_3 x_4) = P(x_1 x_2 x_3 x_4 \mid y_{es}) \times P(y_{es})$

$$P(y_{es} \mid x_1 x_2 x_3 x_4) = \frac{P(x_1 x_2 x_3 x_4 \mid yes) \times P(yes)}{P(x_1 x_2 x_3 x_4)}$$

$$P(No \mid x_1 x_2 x_3 x_4) = \frac{P(x_1 x_2 x_3 x_4 \mid No) \times P(No)}{P(x_1 x_2 x_3 x_4)}$$

$$P(Slap \mid x_1 x_2 x_3 x_4) = \frac{P(x_1 x_2 x_3 x_4 \mid Slap) \times P(Slap)}{P(x_1 x_2 x_3 x_4)}$$

Slap > yes > No

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

$P(y \mid x_1 x_2 x_3 x_4) = P(x_1 x_2 x_3 x_4$

$\rightarrow P(x_1 x_2$

$P(x_1 \mid yes) \times$

2

Outlook

Outlook { Sunny, Overcast, Rain

Hu

$\rightarrow$ Test

$\rightarrow$ Sunn

$$\begin{cases} \frac{5}{7} < \frac{10}{17} \end{cases}$$

Demo
sum

$\text{yes}) \times P(\text{yes})$

$x_4)$

$\times P(x_2|\text{yes}) \times P(x_3|\text{yes}) \times P(x_4|\text{yes})$

	yes	No
→	2/9	3/5
	4/9	0
	3/9	2/5

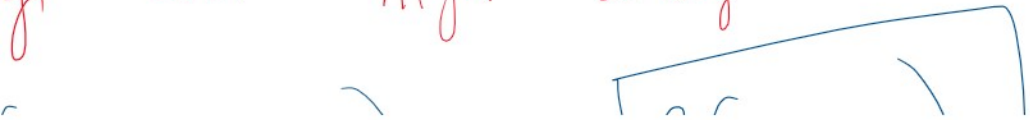
Temp	yes	No
hot	2/9	2/5
Mild	4/9	2/5
cool	3/9	1/5

midy	y	N
high	3/9	4/5
Normal	6/9	1/5

Windy	y	N
Strong	3/9	3/5
weak	6/9	2/5

↓ ↓

y cool high Strong



→ $P(\text{yes} \mid s, c, h, s) = P($

$= P$

$P(\text{NO})$

$P(schs|yes) \times P(yes)$

$P(S|yes) \times P(c|yes) \times P(h|yes) \times P(s|yes) \times P(yes)$

(S: Sunny)

$2/9 \times 3/9 \times 3/9 \times 3/9 \times 9/14$

$= 1/189 = 0.0053$

$P(S|No) \times P(C|No) \times P(h|No) \times P(s|No) \times P(No)$

$3/5 \times 1/5 \times 4/5 \times 3/5 \times 5/14$

$= 0.206$

>>> P(yes)

Naive

	Document ID	Keywords in the document	Class h
Training Set	1	Love Happy Joy Joy Happy	Yes
	2	Happy Love Kick Joy Happy	Yes
	3	Love Move Joy Good	Yes
	4	Love Happy Joy Love Pain	Yes
	5	Joy Love Pain Kick Pain	No
	6	Pain Pain Love kick	No

Spam

Non

Test

Love Pain Joy Love Kick

$P(\text{Love}|\text{yes}) = 5/19$

$P(\text{Pain}|\text{yes}) = 1/19$

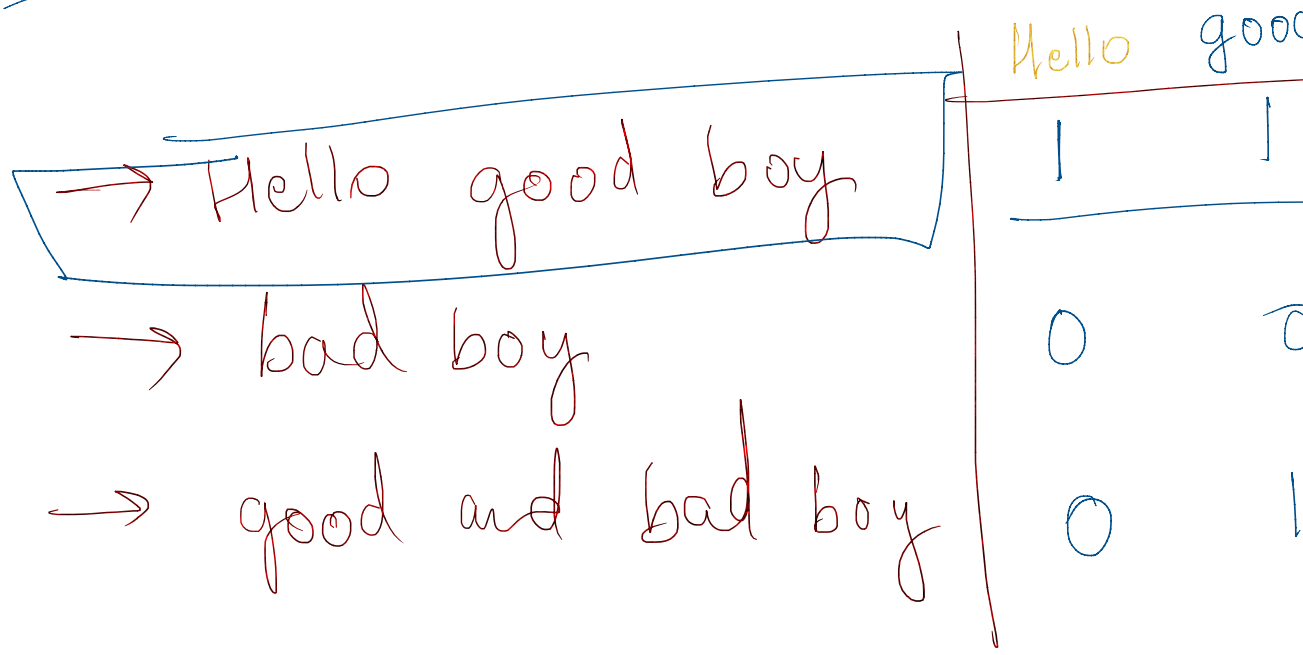
$P(\text{Lov}$

How many words are there in yes

e/NO) = 2/9

	Love	Pain	Joy	Love
yes →	5/19	1/19	5/19	5/19
No →	2/9	4/9	1/a	2/9

BOW



Kick

$\frac{1}{19} \times P(yes) \times \frac{4}{6} = 0.0000067$

$\frac{2}{9} \times P(No) \times \frac{2}{6} = 0.00018$

d	boy	bad	and
	1	0	0

0 1 1 0

1 1 1