

Naive Bayes Algorithm (Classification)

① Probability [Independent And Dependent Events]

② Baye's Theorem

③ Naive Baye's Math Intuition.

① Probability

Independent Events

Rolling a Dice $\{1, 2, 3, 4, 5, 6\}$

$$Pr(1) = \frac{1}{6} \quad Pr(2) = \frac{1}{6} \quad Pr(3) = \frac{1}{6}$$

$= \quad =$

Dependent Events

① What is the probability of first removing
a orange marble and then a yellow marble?

0	0	0
0	0	

$$\rightarrow P(O) = \frac{3}{5} \rightarrow 1^{\text{st}} \text{ Event}$$

0	0
0	0

$$\rightarrow P(Y) = \frac{2}{4} \rightarrow 2^{\text{nd}} \text{ Event}$$

$$\boxed{P(Y|O) = \frac{2}{4}} \Rightarrow \text{Conditional Probability}$$

$$Pr(O \text{ and } Y) = P(O) * \boxed{P(Y|O)} \Rightarrow \text{conditional probability}$$

$$= \frac{3}{5} * \frac{2}{4} = \boxed{\frac{3}{10}}$$

↓

$$\Pr(A \text{ and } B) = \Pr(A) * \Pr(B/A)$$

Bayes Theorem

$$\Pr(A \text{ and } B) = \Pr(B \text{ and } A)$$

$$\Pr(A) * \Pr(B/A) = \Pr(B) * \Pr(A/B)$$

$$\Pr(A/B) = \frac{\Pr(A) * \Pr(B/A)}{\Pr(B)}$$

Bayes Theorem .

$\Pr(A|B)$ = Probability of Event A given B has occurred

$\Pr(A)$ = Probability of Event A

$\Pr(B)$ = Probability of Event B

$\Pr(B/A)$ = Probability of Event B given A has occurred.

DATASET

x_1	x_2	x_3	$0/p \Rightarrow Y$
-	-	-	Yes
-	-	-	No
-	-	-	Yes
-	-	-	No
-	-	-	Yes

$$\Pr(Y/(x_1, x_2, x_3)) = \frac{\Pr(Y) * \Pr(x_1, x_2, x_3/Y)}{\Pr(x_1, x_2, x_3)}$$

$$\Pr(A/B) = \frac{\Pr(A) * \Pr(B/A)}{\Pr(B)}$$

$$\Pr(Y/(x_1, x_2, x_3)) = \frac{\Pr(y) * \Pr(x_1, x_2, x_3/y)}{\Pr(x_1, x_2, x_3)}$$

$$= \frac{\Pr(y) * \Pr(x_1/y) * \Pr(x_2/y) * \Pr(x_3/y)}{\Pr(x_1) * \Pr(x_2) * \Pr(x_3)}$$

<u>DATASET</u>			\downarrow Pred.
x_1	x_2	x_3	0/P \Rightarrow
-	-	-	Yes
-	-	-	No
-	-	-	Yes
-	-	-	No
-	-	-	Yes

$$\Pr(Y=Yes/(x_1, x_2, x_3)) = \frac{\Pr(Yes) * \Pr(x_1/Yes) * \Pr(x_2/Yes) * \Pr(x_3/Yes)}{\Pr(x_1) * \Pr(x_2) * \Pr(x_3)} = 0.60$$

Remove ~~$\Pr(x_1) * \Pr(x_2) * \Pr(x_3)$~~ \Downarrow
 = $\underline{}$ $\underline{1}$

$$\Pr(Y=No/(x_1, x_2, x_3)) = \frac{\Pr(No) * \Pr(x_1/No) * \Pr(x_2/No) * \Pr(x_3/No)}{\Pr(x_1) * \Pr(x_2) * \Pr(x_3)} = 0.40$$

~~$\Pr(x_1) * \Pr(x_2) * \Pr(x_3)$~~

lets Sowc This Problem

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

Outlook

Sunny

Yes

No

$$P(E|Y_{11}) \quad P(E|N_0)$$

$$2/9 \quad 3/5$$

Overcast

2

0

$$4/9 \quad 0/5$$

Rain

3

$$3/9 \quad 4/5$$

Temperature

$$(Sunny, Hot) = 0/1$$

Yes/No

$$Yes \quad No \quad P(E|Y_{11}) \quad P(E|N_0)$$

$$Hot \quad 2 \quad 2 \quad 2/9 \quad 4/5 \quad Yes \quad 9 \quad P(Y) = 9/14$$

$$Mild \quad 4 \quad 2 \quad 4/9 \quad 2/5 \quad No \quad 5 \quad P(Y|No) = 5/14$$

$$Cool \quad 3 \quad 1 \quad 3/9 \quad 1/5$$

$$Pr(Y_{11} | (Sunny, Hot)) = \frac{Pr(Yes) * Pr(Sunny|Yes) * Pr(Hot|Yes)}{Pr(Sunny) * Pr(Hot)}$$

$$= 9/14 * 2/9 * 4/5$$

$$= \frac{2}{63} = \boxed{0.031}$$

$$Pr(N_0 | (Sunny, Hot)) = \frac{Pr(No) * Pr(Sunny|No) * Pr(Hot|No)}{Pr(Sunny) * Pr(Hot)}$$

Constant

$$= 5/14 * 3/5 * 2/5$$

$$= \frac{0.085}{0.031}$$

Finally

$$\Pr(\text{Yes} | (\text{Sunny}, \text{hot})) = \frac{0.031}{0.031 + 0.085} = 0.27 = 27\%$$

$$\Pr(\text{No} | (\text{Sunny}, \text{hot})) = \frac{0.085}{0.031 + 0.085} = 0.73 = 73\%$$

Now DATA $\left[\begin{matrix} \text{Sunny}, \text{Hot} \end{matrix} \right] \Rightarrow \boxed{73\%} \Rightarrow \text{No} \Rightarrow \text{O}$
 \downarrow
 $\boxed{\text{Person will NOT play}}$

Variants of Naive Bayes

① Bernoulli Naive Bayes

② Multinomial Naive Bayes

③ Gaussian Naive Bayes

① Bernoulli Naive Bayes

Whenever your features are following a Bernoulli Distribution, then we use Bernoulli Naive Bayes

Bernoulli $\rightarrow 0, 1$

Dataset

f_1	f_2	f_3	O/P	$P(Success) = 1 = P$
Yes	Pass	Male	Yes	$P(Fail) = 0 = 1 - P$
No	Fail	Female	No	
Yes	Pass	Male	Yes	
No	Pass	Female	No.	
Yes	Pass	Female	No.	

② Multinomial Naive Bayes $\Rightarrow IP = Text$

Dataset : Sentiment Analysis

IP

Review Message O/P

The product is really good Positive

The product is bad Negative

↓

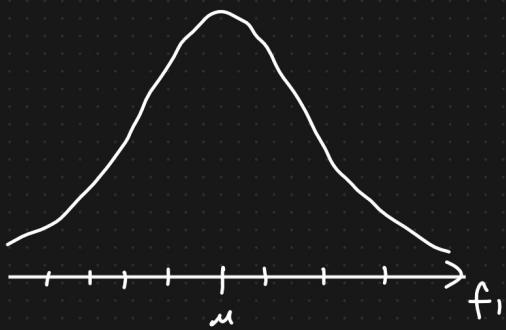
Numerical Values \Rightarrow Natural Language Processing

- ① BOW
- ② Tf-IDF
- ③ Word2Vec

③ Gaussian Naive Bayes

If the features are following Gaussian Distribution then we use Gaussian Naive Bayes Algorithm to solve classification problem

[IRIS] Dataset



[continuous feature]

	Age	Height	Weight	Yes/No
25	170	78		
28	160	75		
32	150			
34	140			