Naive Baye's Algorithm (Classification)

↳
- Binary Classification
- Multiclass Classification

① Probability

② Baye's Theorem

Independent Events

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

$Pr(1) = \dfrac{1}{6}$ $Pr(2) = \dfrac{1}{6}$ $Pr(3) = \dfrac{1}{6}$.

Dependent Events

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

```
0 0 0
0 0
```

↳ $P(0) = \dfrac{3}{5}$ → 1st Event → Orange marble has been Removed

```
0 0
0 0
```
→ $P(y) = \dfrac{2}{4} = \dfrac{1}{2}$ ⇒ 2nd Event ⇒ Removed the Yellow Marble

$$P(0 \text{ and } y) = \underset{\Downarrow 1st}{P(0)} * \boxed{\underset{\Downarrow IInd}{P(y/0)}} \rightarrow \text{Conditional Probability}$$

$$= 3/5 * \dfrac{1}{2} = \boxed{\dfrac{3}{10}}$$

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

Probability of B, given A is already occurred

Baye's Theorem

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

$$P(A) * P(B/A) = P(B) * P(A/B)$$

$$\boxed{P(A/B) = \frac{P(A) * P(B/A)}{P(B)}} \Rightarrow \text{Baye's Theorem.}$$

$P(A/B)$ = Probability of Event A given B has occored

$P(A)$ = Probability of Event A

$P(B)$ = Probability of Event B

$P(B/A)$ = Probability of Event B given A has Occured.

$$\boxed{P(A/B) = \frac{P(A) * P(B/A)}{P(B)}} \Rightarrow \text{Baye's Theorem.}$$

I/P features

↓ Dependent

x_1	x_2	x_3	y
—	—	—	Yes
—	—	—	No
—	—	—	Yes
—	—	—	No

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

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

Expanding it :-

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

I/p features

↓ Dependent

λ_1	x_2	x_3	y
—	—	—	Yes
—	—	—	No
—	—	—	Yes
—	—	—	No

$\left[\boxed{0.6\ 0}, 0.40\right]$
⇓
Yes

New tst data

$$Pr(Yes/(x_1, x_2, x_3)) = \frac{P(Yes) * P(x_1/Yes) * P(x_2/Yes) * P(x_3/Yes)}{\cancel{P(x_1) * Pr(x_2) * Pr(x_3)} \Rightarrow Constant} = \boxed{0.60}$$
⇓
Yes

$$Pr(No/(x_1, x_2, x_3)) = \frac{Pr(No) * Pr(x_1/No) * Pr(x_2/No) * Pr(x_3/No)}{\cancel{P(x_1) * Pr(x_2) * Pr(x_3)} \Rightarrow Constant} = 0.40$$

Reject it
Bcoz it does not affect the
outcome of classification

Lets Solve this Problem

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Outlook

	Yes	No	$P(E/Yes)$	$P(E/No)$
Sunny	2	3	$2/9$	$3/5$
Overcast	4	0	$4/9$	$0/5$
Rain	3	2	$3/9$	$2/5$

$\rightarrow$ Test (Sunny, Hot) $\rightarrow$ O/P   PLAY (Y/N)

Temperature

	Yes	No	$P(E/Yes)$	$P(E/No)$
Hot	2	2	$2/9$	$2/5$
Mild	4	2	$4/9$	$4/5$
Cool	3	1	$3/9$	$3/5$

		$P(Yes)$	$P(No)$
Yes	9	$9/14$	$5/14$
No	5		

$$P(Yes/Sunny, Hot) = \frac{P(Yes) * Pr(Sunny/Yes) * Pr(Hot/Yes)}{\cancel{Pr(Sunny) * Pr(Hot)}}$$

$$= \frac{9}{14}_{7} * \frac{2^1}{9} * \frac{2}{9}$$

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

$$P(No/(Sunny, Hot)) = P(No) * Pr(Sunny/No) * Pr(Hot/No)$$

$$= \frac{5}{14}_{7} * \frac{3}{5} * \frac{2^1}{5}$$

$$= \frac{3}{35} = 0.085$$

$$Pr(Yes/(Sunny, Hot)) = \frac{0.031}{(0.031 + 0.085)} = 0.27 = 27\%$$

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

Outlook	Temperature	O/p
New
Test $\Rightarrow$ | Sunny | Hot | 73% $\Rightarrow$ They will not play Tennis

27% $\Rightarrow$ They will play Tennis.

$\Downarrow$

0 $\Rightarrow$ Person is not $\overset{going\ to}{\underset{\wedge}{playing}}$ Tennis