Naive Baye's Algorithm (Classification)

↳ Probability and Baye's Theorem

Independent Event

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

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

Dependent Event

What is the probability of removing a white marble and then a yellow marble?

↳ $P(w) = \frac{3}{5}$ → 1st Event

$P(y/w) = \frac{2}{4} = \frac{1}{2}$ → 2nd Event

↳ Conditional probability

$$P(w \text{ and } y) = P(w) * P(y/w) \Rightarrow \text{Independent Event}$$

↳ Conditional probability

Baye's Theorem

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

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

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

Dataset

Independent feature.- → O/P

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad - - - - \quad x_n \qquad y$

$$P(y/x_1, x_2, x_3 \cdots x_n) = \frac{P(y) * P(x_1, x_2, x_3, x_4 \cdots x_n/y)}{P(x_1, x_2, x_3 \cdots x_n)}$$

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

x_1	x_2	x_3	x_4	y
–	–	–	–	Yes
–	–	–	–	No
–	–	–	–	Yes

$$P(Yes/x_1, x_2, x_3, x_4) = \frac{P(Yes) * P(x_1/Yes) * P(x_2/Yes) * P(x_3/Yes) * P(x_4/Yes)}{\cancel{P(x_1) * P(x_2) * P(x_3) * P(x_4)}}$$

Constant ← $P(x_1) * P(x_2) * P(x_3) * P(x_4)$ ✗

$$P(No/x_1, x_2, x_3, x_4) = \frac{P(No) * P(x_1/No) * P(x_2/No) * P(x_3/No) * P(x_4/No)}{\cancel{P(x_1) * P(x_2) * P(x_3) * P(x_4)}}$$

Constant ← $P(x_1) * P(x_2) * P(x_3) * P(x_4)$ ✗

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

Temprature

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

→ Test (Sunny, Hot) → O/P

PLAY (Y)

Yes 9

No 5

P(Yes) $\boxed{9/14}$ P(No) $\boxed{5/14}$

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

Constant ← P(Sunny) * P(Hot)

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

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

$$P(No/(Sunny, Hot)) = \frac{P(No) * P(Sunny/No) * P(Hot/No)}{P(Sunny) * P(Hot)}$$

X ← P(Sunny) * P(Hot)

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

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

$$P(yes/(sunny, hot)) = \frac{0.031}{0.031 + 0.085} = 0.27 = 27\%$$

$$P(No/(sunny, hot)) = \frac{0.085}{0.031 + 0.085} = 0.73 = 73\%$$

Test data (Sunny, hot) = 73% They will not play Tennis

$\Rightarrow$ Person is not going to play Tennis