Naive Bayes Algorithm (Classification)

1. Probability [Independent And Dependent Events]
2. Bayes Theorem
3. Naive Bayes Math Intuition.

1. Probability

Independent Events

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

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

$=$ $=$

Dependent Events

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

① → $P(O) = 3/5$ → 1st Event

② → $P(Y) = \frac{2}{4}$ → 2nd Event

⇓

$\boxed{P(Y/O) = 2/4}$ ⇒ Conditional Probability

$Pr(O \text{ and } Y) = P(O) * \boxed{P(Y/O)}$ ⇒ Conditional Probability

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

⇓

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

Bayes Theorem

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

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

Bayes Theorem

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

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

P(A) = Probability of Event A

P(B) = Probability of Event B

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

DATASET

x_1	x_2	x_3	↓ Predict O/P ⇒ Y
–	–	–	Yes
–	–	–	No
–	–	–	Yes
–	–	–	No
–	–	–	Yes

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

(A = Y, B = (x_1, x_2, x_3))

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

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

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

DATASET

x_1	x_2	x_3	↓ Pred. O/P ⇒
–	–	–	Yes
–	–	–	No
–	–	–	Yes
–	–	–	No
–	–	–	Yes

$$Pr\left(y = Yes/(x_1, x_2, x_3)\right) = \frac{Pr(Yes) * Pr(x_1/Yes) * P(x_2/Yes) * Pr(x_3/Yes)}{\cancel{Pr(x_1) * Pr(x_2) * Pr(x_3)}} = 0.60 \Rightarrow 1$$

Remove

$$Pr\left(y = No/(x_1, x_2, x_3)\right) = \frac{Pr(No) * Pr(x_1/No) * Pr(x_2/No) * Pr(x_3/No)}{\cancel{Pr(x_1) * Pr(x_2) * Pr(3)}} = 0.40$$

Lets Solve 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

	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

(Sunny, Hot) = O/P

Temperature

	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

Yes/No

Yes	9	P(Y) = 9/14
No	5	Pr(No) = 5/14

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

$$= 9/14 * 2/9 * 2/9$$

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

$$Pr(No/(Sunny, Hot)) = \frac{Pr(No) * Pr(Sunny/No) * Pr(Hot/No)}{\text{Constant}}$$

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

$= 0.085$

Finally

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

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

NEW DATA [Sunny, Hot] $\Rightarrow$ 73% $\Rightarrow$ No $\Rightarrow$ ○

27% $\Rightarrow$ Yes

$\Downarrow$

Person will Not play