Naive Bayes Algorithm (Classification)

* It is specifically used to solve the classification problem - Both binary and multiclass classification.

① Probability

② Bayes theorem.

Probability

Independent Events

Dependent Events

Independent Events

Rolling a Die {1, 2, 3, 4, 5, 6}

$Pr(1) = \frac{1}{6}$ (chances of getting 1)

$Pr(2) = \frac{1}{6}$ (chances of getting 2)

* above is an example of chance of an single event.

why it is called as Independent event?

* The probability of one event

... the probability of

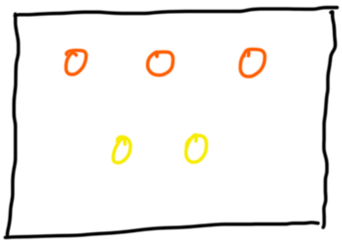
does not change the probability of the other event.

* If you consider the above example.

→ probability of occurring 1 when rolling a die does not change the probability of occurring 2.

Dependent Events

example, consider I have a bag of marbles



① what is the probability of removing an orange marble and then a yellow marble.

$$P(\text{orange}) = \frac{3}{5} \rightarrow 1^{st} \text{ Event} \left(\frac{\text{Total orange marbles}}{\text{Total marbles}}\right)$$

→ orange marble has been

 removed.

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

therefore, this dependent probability of this can be represented as

$$P(\text{orange and yellow}) = \underbrace{P(\text{orange})}_{1st\ event} * \underbrace{P\left(\frac{\text{yellow}}{\text{orange}}\right)}_{2nd\ event}$$

this is nothing but, Probability of orange has been already taken place.

this equation is called as conditional probability.

therefore, this can be evaluated as

$$= \frac{3}{5} \times \frac{1}{2}$$

$$= \frac{3}{10} \rightarrow P(\text{orange and yellow})$$

In a generic way this can be represented as

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

Probability of B given A (meaning P(A) has been already taken place)

Bayes Theorm

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

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

$$\boxed{P(A/B) = \frac{P(A)\ P(B/A)}{P(B)}}$$

$\Downarrow$

Bayes theorm.

$P(A/B)$ = 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.

Independent Features   Dependent features.

x_1	x_2	x_3	y
—	—	—	Yes

–	–	–	NO
–	–	–	Yes
–	–	–	NO

we know that Bayes theorem

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

this can be represented as,

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

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

Given our data set

Independent			Dependent
x_1	x_2	x_3	Y
–	–	–	Yes
–	–	–	NO
–	–	–	Yes
–	–	–	NO

the equation can be represented as,

$Pr(Yes/(x_1, x_2, x_3)) =$

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

$Pr(NO/(x_1, x_2, x_3)) =$

$P(NO) * P(x_1/NO) * P(x_2/NO) *$

$$\frac{P(x_3/NO)}{P(x_1) * P(x_2) * P(x_3)}$$

If you observe the above two equations the denominator is constant. So we will be removing the denominator.

Therefore

$$Pr(yes/(x_1, x_2, x_3)) = P(yes) * P(x_1/yes) * P(x_2/yes) * P(x_3/yes)$$

$$Pr(NO/(x_1, x_2, x_3)) = P(NO) * P(x_1/NO) * P(x_2/NO) * P(x_3/NO)$$

Lets consider the Below 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

Rain,

Temprature

	Yes	No	P (E / yes)	P (E / No)
Hot	2	2	2/9	2/5
mild	4	2	4/9	2/5
cold	3	1	3/9	1/5

Play (yes / no)		P (yes)	P (No)
yes	9	9/14	5/14
No	5		

Test (sunny, Hot) $\longrightarrow$ output ?

$$P(\text{yes} / \text{sunny}, \text{hot}) = P(\text{yes}) * P(\text{sunny}) * P(\text{hot})$$

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

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

$$= 0.031$$

$$P(\text{NO} / \text{sunny}, \text{hot}) = P(\text{NO}) * P(\text{sunny}/\text{no}) * P(\text{hot}/\text{no})$$

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

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

$$= 0.085$$

$$P(Yes/(sunny, hot)) = \frac{0.031}{(0.031 + 0.085)}$$

$$= 0.27$$

$$= 27\%$$

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

$$= 0.073$$

$$= 73\%$$

Now with the new test data

Outlook	Temprature	O/P
Sunny	Hot	73%

therefore, there are 73% chances of not (No) playing tennis and 27% chances of playing tennis.

⇓

test

The final outcome for this given [illegible] data will be No (0) — is not going to play tennis.