Naive Bayes

→ Naive Bayes is a classification algorithm based on Bayes theorem that assumes all features are independent, and uses probability to determine which category a data point belongs to.

① Independent Events

ⓐ Rolling a Dice

$\{1, 2, 3, 4, 5, 6\}$

→ $Pr(1) = \frac{1}{6}$ $P(2) = 1/6$

$P(3) = 1/6$

→ In this case one event is not impacting another event which is called as Independent Events.

ⓑ Tossing a coin

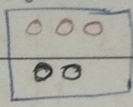
$P(H) = 1/2$ $P(T) = 1/2$

→ No matter how many times i toss this is my probability (ratio)

→ This doesn't impact another event → is called as Independent Event

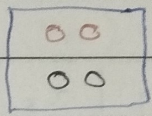
② Dependent Events

Eg : Bag of marbles

① ⟺  $P(R) = 3/5$

∴ Total I'm having 3 and 3 red marbles

② Remaining marbles (another example)

 ~~P(...)~~ $P(B) = \frac{2}{4}$

∴ Total no of marbles :- 4

∴ Total no of black marbles :- 2

↓

$P(B/R) = 2/4$

⇒ This is an Eg of Dependent Event where one event is impacting another event.

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

→ This term is called conditional probability

→ This is the formula for probability of Red and Black (Dependent Event)

Bayes theorem :- Derive

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

(always same) $P(A \text{ and } B) = P(B \text{ and } A)$ — Cometitive property

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

↳ (Probability of A multiply by Probability of B given A)
= (Probability of B multiply by Probability of A given B)

Baye's theorem formula

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

↳ This baye's theorem is extensively used in Naive Baye's

↳ We solve classification problem using this formula

Eg :-

x_1	x_2	x_3	y
-	-	-	-
-	-	-	-
-	-	-	-

(3 Independent & 1 Dependent feature)

↳ How this formula applies to this features or dataset

∴ B :- Dependent Feature (y)
A :- Independent Features (x_1, x_2, x_3)

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

↳ Using this formula we find the Probability of y

↳ This whole formula is trained using training datasets

Page No.
Date

Bayes theorem :- Implementation

$$\boxed{P(B|A) = \frac{P(B) * P(A|B)}{P(A)}} \longrightarrow \text{Bayes theorem}$$

Note: A :- Independent feature
B :- Dependent feature

Supervised Machine Learning ↙ classification

x_1, x_2, x_3, x_4 → Independent feature

X_1	X_2	X_3	X_4 } Input	o/p (y)
-	-	-	-	0 (No)
-	-	-	-	1 (Yes)

- y :- Dependent feature

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

$$= \frac{P(y) * P(x_1/y) * P(x_2/y) * P(x_3/y) * P(x_4/y)}{P(x_1, x_2, x_3, x_4 \dots x_n)} \longrightarrow \text{constant (ignore it)}$$

Equation for machine learning classification problem

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

→ constant (ignore it)

Problem statement

Page No.
Date

Problem statement

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hoot	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

* Final Feature :- Outlook

- How many Category outlook has : 3 (How many Yes or No is their in this category)

1) sunny 2) Overcast 3) Rain

①

	Yes ~~or~~	No	P(Y)	P(N)
Sunny	2	3	2/9	3/5
Overcast	4	0	4/9	0/5
Rain	3	2	3/9	2/5
	9	5 = 14		

~~How do we write it in equation~~

② Temperature

	Yes	No	P(Y)	P(N)
Hot	2	2	2/9	2/5
Mild	4	2	4/9	2/5
Cold	3	1	3/9	1/5
	9	5 = 14		

③ Play Tennis

Yes 9 P(Y) - 9/14

No 5 P(N) - 5/14 independent feature

New Data → Test (Sunny, Hot) → ?

$$\text{① } P(\underset{y}{Yes} \mid \underset{x_1}{Sunny}, \underset{x_2}{Hot}) = \frac{P(Yes) * P(Sunny \mid Yes) * P(Hot \mid Yes)}{P(Sunny) * P(Hot)}$$ - constant (ignore it)

$$\text{② } P(\underset{y}{No} \mid \underset{x_1}{Sunny}, \underset{x_2}{Hot}) = \frac{P(No) * P(Sunny \mid No) * P(Hot \mid No)}{P(Sunny) * P(Hot)}$$ - constant (ignore it)

* Equate

→ ① $P(Yes \mid Sunny, Hot) = \cancel{9}/\cancel{14}^{7} * \cancel{2}^{1}/9 * 2/9$

$= \frac{2}{7 \times 9} = \frac{2}{63} = 0.031$

→ ② $P(No \mid Sunny, Hot) = \cancel{5}/\cancel{14}^{7} * 3/5 * \cancel{2}^{1}/5$

$= \frac{3}{7 * 5} = \frac{3}{35} = 0.085$

Final :- Output into proper number

$P(Yes \mid Sunny, Hot) = \frac{0.031}{0.031 + 0.085} = 27\%$

$P(No \mid Sunny, Hot) = \frac{0.085}{0.031 + 0.085} = 73\%$

→ This indicates that : → Sunny, Hot → No → 0.73