

## # Naive Bayes:

Naive Bayes's methods are a set of supervised learning algorithms based on applying Bayes's theorem with the naive assumption of conditional independence between every pair of features given the value of class variables.

Bayes's theorem state the following relationship  
Given class variables  $y$  and dependent  
Feature vector  $x_1$  through  $x_n$

$$P(y|x_1, \dots, x_n) = \frac{P(y) \cdot (P(x_1, \dots, x_n|y))}{P(x_1, \dots, x_n)}$$

$$= \frac{P(y) \prod_{i=1}^n (x_i|y)}{P(x_1, \dots, x_n)}$$

$$\text{Pr} \left[ P(y|x_1, \dots, x_n) \propto P(y) \prod_{i=1}^n (x_i|y) \right]$$



Outlook	Temp	Humidity	Windy	Play
Sunny	hot	high	False	No
Sunny	hot	high	True	No
Overcast	hot	High	False	Yes
Rainy	mild	High	False	Yes
Rainy	cool	Normal	False	Yes
Rainy	cool	Normal	True	No
Overcast	cool	Normal	True	Yes
Sunny	mild	high	False	No
Sunny	cool	Normal	False	Yes
Rainy	mild	Normal	False	Yes
Sunny	mild	Normal	True	Yes
Overcast	mild	High	True	Yes
Overcast	hot	Normal	False	Yes
Rainy	mild	High	True	No

### Outlook

### play

	Yes	No	P(y)	P(N)		Yes	P(y)	P(N)
Hot Sunny	2	3	$\frac{2}{9}$	$\frac{3}{5}$		9	$\frac{9}{14}$	
overcast	4	0	$\frac{4}{9}$	$\frac{0}{5}$		5	$\frac{5}{14}$	
Rainy	3	2	$\frac{3}{9}$	$\frac{2}{5}$		14		
	9	5						

### temp:

	Yes	No	P(y)	P(N)
Hot	2	2	$\frac{2}{9}$	$\frac{2}{5}$
mild	4	2	$\frac{4}{9}$	$\frac{2}{5}$
Cool	3	1	$\frac{3}{9}$	$\frac{1}{5}$
	9	5		

### Humidity

	Yes	No	P(y)	P(N)
High	3	4	$\frac{3}{9}$	$\frac{4}{5}$
Normal	6	1	$\frac{6}{9}$	$\frac{1}{5}$
	9	5		

### Windy

	Yes	No	P(y)	P(N)
True	3	3	$\frac{3}{9}$	$\frac{3}{5}$
False	6	2	$\frac{6}{9}$	$\frac{2}{5}$
	9	5		



According to Naïve Bayes Theorem

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Suppose  $I$  is  
if Today's Outlook is rainy, temp is mild,  
Humidity is High and Windy is false  
then tennis playing or not.

Let's calculate Today (rainy, mild, High, false)

$$\begin{aligned} P(\text{Yes} | \text{today}) &= P(\text{rainy} | \text{Yes}) \times P(\text{mild} | \text{Yes}) \times P(\text{High} | \text{Yes}) \times P(\text{false} | \text{Yes}) \\ &\quad \times P(\text{Yes}) \\ &= \frac{3}{9} \times \frac{4}{9} \times \frac{3}{9} \times \frac{6}{9} \times \frac{9}{14} \end{aligned}$$

$$P(\text{Yes} | \text{today}) = 0.0212$$

$$\begin{aligned} P(\text{No} | \text{today}) &= P(\text{rainy} | \text{No}) \times P(\text{mild} | \text{No}) \times P(\text{High} | \text{No}) \times P(\text{false} | \text{No}) \times P(\text{No}) \\ &= \frac{8}{9} \times \frac{2}{5} \times \frac{2}{5} \times \frac{4}{5} \times \frac{5}{14} \end{aligned}$$

$$P(\text{No} | \text{today}) = 0.0183$$

$$P(\text{Yes}) = \frac{0.0212}{0.0212 + 0.0183}$$

$$P(\text{Yes}) = 0.537$$

$$P(\text{No}) = 1 - P(\text{Yes})$$

$$= 1 - 0.537$$

$$P(\text{No}) = 0.463$$

In this scenario the probability of ~~Yes~~ Yes is greater than No.

$$P(\text{Yes}) > P(\text{No})$$

$$0.537 > 0.463$$

In this particular scenario, Outlook is rainy, temp. is mild, Humidity is high and Windy is false.

$\therefore$  The output is Yes: (means All conditions are satisfied so we can play tennis.)