

Naïve Bayes Example

Author - Kanav Bansal

You can find me on **Linkedin** -

www.linkedin.com/in/kanavbansal

Or <http://www.thataiguy.com>

Consider this binary classification data set:

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

→ X_q Rain Hot Normal Weak ?

Given $X_q \rightarrow$ Outlook = 'Rain' & Temp = 'Hot' &
Humidity = 'Normal' & Wind = 'Weak'

Find 1. $P(\text{PlayTennis} = \text{Yes} | X_q) \propto P(\text{PlayTennis} = \text{Yes}) \times$

$$P(\text{Outlook} = \text{Rain} | \text{PlayTennis} = \text{Yes}) \times$$

$$P(\text{Temp} = \text{Hot} | \text{PlayTennis} = \text{Yes}) \times$$

$$P(\text{Humidity} = \text{Normal} | \text{PlayTennis} = \text{Yes}) \times$$

$$P(\text{Wind} = \text{Weak} | \text{PlayTennis} = \text{Yes})$$

2. $P(\text{PlayTennis} = \text{No} | X_q) \propto P(\text{PlayTennis} = \text{No}) \times$

$$P(\text{Outlook} = \text{Rain} | \text{PlayTennis} = \text{No}) \times$$

$$P(\text{Temp} = \text{Hot} | \text{PlayTennis} = \text{No}) \times$$

$$P(\text{Humidity} = \text{Normal} | \text{PlayTennis} = \text{No}) \times$$

$$P(\text{Wind} = \text{Weak} | \text{PlayTennis} = \text{No})$$

Frequency Tables →

1. Outlook vs PlayTennis

	Yes	No	
Sunny	2	3	5
Overcast	4	0	4

2. Temperature vs PlayTennis

	Yes	No	
Hot	2	2	4
Mild	4	2	6
Cool	3	1	4

Sunny	2	3	5
Overcast	4	0	4
Rainy	3	2	5
	9	5	

Hot	2	2	4
Mild	4	2	6
Cool	3	1	4
	9	5	

3. Humidity vs Play Tennis

	Yes	No	
High	3	4	7
Normal	6	1	7
	9	5	

4. Wind vs Play Tennis

	Yes	No	
Strong	3	3	6
Weak	6	2	8
	9	5	