

Naive Bayesian: with the following training dataset, find the prediction.

Predict:

Outlook=sunny, Temperature=Cool, Humidity=High, Wind=True

===? Play golf or NOT?

outlook	temperature	humidity	wind	play
sunny	hot	high	false	no
sunny	hot	high	true	no
overcast	hot	high	false	yes
rainy	mild	high	false	yes
rainy	cold	normal	false	yes
rainy	cold	normal	true	no
overcast	cold	normal	true	yes
sunny	mild	high	false	no
sunny	cold	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
2 levels   14 rows				

Will find the answer according to following equation:

Likelihood function in Naïve Bayes is given as:

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

**Step 1:** create the sub tables based on the all exiting attributes in dataset.

1.	Outlook			
	Yes	No	P(Yes)	P(No)
Sunny	2	3	2/9	3/5
Overcast	4	0	4/9	0/5
Rainy	3	2	3/9	2/5
Total	9	5	100%	100%

2.	Temperature			
	Yes	No	P(Yes)	P(No)
Hot	2	2	2/9	2/5
Mild	4	2	4/9	2/5
Cool	3	1	3/9	1/5
Total	9	5	100%	100%

3.	Humidity			
	Yes	No	P(Yes)	P(No)
High	3	4	3/9	4/5
Normal	6	1	6/9	1/5
Total	9	5	100%	100%

4.	Wind			
	Yes	No	P(Yes)	P(No)
FALSE	6	2	6/9	2/5
TRUE	3	3	3/9	3/5
Total	9	5	100%	100%

**Step 2:** Find the probability of the “Yes” and “No “based on the training dataset.

$$P(\text{Yes})=9/14$$

$$P(\text{No})=5/14$$

**Step 3:** the result will be:

$$P(\text{Yes} | X_i) = \frac{P(\text{yes}) * P(\text{sunny} | \text{yes}) * p(\text{Cool} | \text{yes}) * P(\text{High} | \text{yes}) * P(\text{True} | \text{yes})}{p(\text{Suunny, cool, high, True})}$$

$P(\text{Yes} | X_i) = (9/14 * 2/9 * 3/9 * 3/9 * 3/9) / (5/14 * 4/14 * 7/14 * 6/14) = 0.2424$  This is the likelihood of YES- And the probability of YES is around 20%.

$$0.2424 / (0.2424 + 0.9421) = 0.2046 \sim 20\%$$

-----

$$P(\text{No} | X_i) = \frac{P(\text{No}) * P(\text{sunny} | \text{No}) * p(\text{Cool} | \text{No}) * P(\text{High} | \text{No}) * P(\text{True} | \text{No})}{p(\text{Suunny, cool, high, True})}$$

$P(\text{No} | X_i) = (5/14 * 3/5 * 1/5 * 4/5 * 3/5) / (5/14 * 4/14 * 7/14 * 6/14) = 0.9421$  This is the likelihood of NO- And the probability of NO is around 80%.

$$0.9421 / (0.2424 + 0.9421) = 0.7953 \sim 80\%$$

#### Step 4:

So, the result of the predict is “NO”. They are not allowed to play golf, when the Outlook=sunny, Temperature=Cool, Humidity=High, Wind=True.

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

$$0.9421 > 0.2424$$