Bayes' Theorem

Example : Telecom Customers

- A telecom firm has many customers. Each customer either talks for the duration of more than 100 minutes or less than 100 minutes. The firm has launched a plan for the customers who talk more specially to optimize the amount spent by them on bills.
- Call Centre staff had been instructed to call some customers. In that operation, some customers bought the new plan and others didn’t.
- In this case each customer is a record, and the response of interest, Y = {Bought ,Not Bought}, has two classes: C1 = Bought and C2 = Not Bought.

Conditional Probabilities

- A conditional probability of event A given event B [denoted by P(A|B)] represents the chances of event A occurring only under the scenario that event B occurs.
- In the response example, we may be interested in P(bought| Talk Time >=100, gender=Male), also P(bought| Talk Time >=100, gender=Female), as we have gender as additional feature of the customers

BAYES FORMULA

- The Bayes theorem gives us the following formula to compute the probability that the record belongs to class Ci:

$$P(C_i|X_1,\ldots,X_p) = \frac{P(X_1,\ldots,X_p|C_i)P(C_i)}{P(X_1,\ldots,X_p|C_1)P(C_1)+\cdots+P(X_1,\ldots,X_p|C_m)P(C_m)}.$$

Where

Ci : classes of interest

$X_1, X_2, \ldots X_p$: Variables which co-exist with Classes of interest

Example

Talks for more than 100 min? (TT >= 100)	Gender	Response
y	male	not bought
n	male	not bought
n	female	not bought
n	female	not bought
n	male	not bought
n	male	not bought
y	male	bought
y	female	bought
n	female	bought
y	female	bought

Bayes' Formula Calculations

$P(Buy|Male, TT \geq 100)$

$$= \frac{P(Male, TT\geq 100\ |Buy)\ P(Buy)}{P(Male, TT\geq 100\ |Buy)\ P(Buy)+P(Male, TT\geq 100\ |\ Not\ Buy)\ P(\ Not\ Buy)}$$

$$= \frac{P(Male|Buy)P(TT \geq 100|Buy)\ P(Buy)}{P(Male|Buy)P(TT \geq 100|Buy)P(Buy)+P(Male|\ Not\ Buy)P(TT \geq 100|Not\ Buy)P(Not\ Buy)}$$

$$= \frac{\frac{1}{4}\times\frac{3}{4}\times\frac{4}{10}}{\frac{1}{4}\times\frac{3}{4}\times\frac{4}{10} + \frac{4}{6}\times\frac{1}{6}\times\frac{6}{10}}$$

$= 0.529$

(TT >= 100)	Gender	Response
y	male	not bought
n	male	not bought
n	female	not bought
n	female	not bought
n	male	not bought
n	male	not bought
y	male	bought
y	female	bought
n	female	bought
y	female	bought

Bayes Probabilities

- For the conditional probability of bought behaviors given (TT >= 100) = y, gender = male , the numerator is a multiplication of the proportion of (TT >= 100) = y instances among the bought customers, times the proportion of gender = male instances among the bought customers, times the proportion of bought customers: (3/4)(1/4)(4/10) = 0.075.
- To get the actual probabilities, we must also compute the numerator for the conditional probability of not bought given (TT >= 100) = y, gender = male : (1/6)(4/6)(6/10) = 0.067.
- The denominator is then the sum of these two conditional probabilities (0.075 + 0.067 = 0.14).

Bayes Probabilities

- The conditional probability of bought behaviors given (TT >= 100) = y, gender = male is therefore 0.075/0.14 = 0.53.
- Similarly,
  - P(bought | (TT >= 100) = y, gender = female) = 0.87,
  - P(bought | (TT >= 100) = n, gender = male) = 0.07,
  - P(bought | (TT >= 100) = n, gender = female) = 0.31.

Output of Bernoulli Naïve Bayes

	TT_gt_100_y	Gender_male
0	False	False
1	False	True
2	True	True
3	True	False

P(Buy) P(Won't Buy)

```
[0.31034483, 0.68965517]
[0.06976744, 0.93023256]
[0.52941176, 0.47058824]
[0.87096774, 0.12903226]
```