Naïve Bayes Classification

Naive Bayes

a supervised machine learning algorithm, which is used for classification tasks, it is one of the simplest supervised learning algorithms.

Naive Bayes assumption

It takes a strong assumption that all features are independent of each other. The effect of one feature on a given class or dataset is independent of the value on other features.

Naive Bayes Performance

is the fast, accurate and reliable algorithm. Naive Bayes classifiers have high accuracy and speed on large datasets.

Types:

- Gaussian Naive Bayes.
- Multinomial Naive Bayes.
- Bernoulli Naive Bayes

How does Naive Bayes work:

Example No	Color	Type	Origin	Stolen ?
1	**Red**	Sports	**Domestic**	YES
2	**Red**	Sports	**Domestic**	**NO**
3	**Red**	Sports	**Domestic**	YES
4	Yellow	Sports	**Domestic**	**NO**
5	Yellow	Sports	Imported	YES
6	Yellow	**SUV**	Imported	**NO**
7	Yellow	**SUV**	Imported	YES
8	Yellow	**SUV**	**Domestic**	**NO**
9	**Red**	**SUV**	Imported	**NO**
10	**Red**	Sports	Imported	YES

Q 1: New Instance = (Red , SUV , Domestic) yes or no ?

- **Step One :**

  P(YES) = 5/10

  P(NO) = 5/10

- **Step Two**

  Frequency Tables

Color	P(yes)	P(no)
Red	**3/5**	**2/5**
Yellow	**2/5**	**3/5**

Type	P(yes)	P(no)
Sports	**4/5**	**2/5**
SUV	**1/5**	**3/5**

Origin	P(yes)	P(no)
Domestic	**2/5**	**3/5**
Imported	**3/5**	**2/5**

- **Step three**

(Red , SUV , Domestic)

P(Yes | New instance) **(x1,x2,x3)**

5/10 * 3/5 * 1/5 * 2/5 = = 0.024

P(No | New instance) **(x1,x2,x3)**

5/10 * 2/5 * 3/5 * 3/5 = 0.072.

The (Red, SUV , Domestic) car is not stolen!

Sources:
https://scikit-learn.org/stable/modules/naive_bayes.html
https://www.datacamp.com/tutorial/naive-bayes-scikit-learn