Machine Learning

Naive Bayes & Bayes Classifier

Marchelo Bragagnini
cesarbrma91@gmail.com
@MarchBragagnini

Universidad Católica San Pablo

Centro de Investigación e Innovación en Ciencia Computación

Outline

1. Applications
2. Bayes' Theorem
3. Naive Bayes
4. Codes
5. Advantages & When to use Naive Bayes

Applications

1. Real time prediction
2. Multi class prediction
3. Text classification, spam filtering, sentiment analysis
4. Recommendation systems

Bayes’ Theorem

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

- Comments:
  - Bayes’ rule tells us how to ‘invert’ conditional probabilities, i.e. to find P(B|A) from P(A|B).

Example

Consider a routine screening test for a disease. Suppose the frequency of the disease in the population (base rate) is 0.5%. The test is highly accurate with a 5% false positive rate and a 10% false negative rate.

You take the test and it comes back positive. What is the probability that you have the disease?

Bayes' Theorem

Example

Events:

D+ = 'you have the disease'

D− = 'you do not have the disease

T+ = 'you tested positive'

T− = 'you tested negative'.

$$P(D+ \mid T+)$$

Bayes' Theorem

Example

Using :

- $P(D+) = 0.005$
- $P(D-) =$ _____
- $P(T- \mid D+) = 0.1$ (false negative)
- $P(T+ \mid D+) =$ ___
- $P(T+ \mid D-) =$ ____ (false positive)

Bayes' Theorem

Example

$$P(D+ \mid T+) = \frac{P(T+ \mid D+) \cdot P(D+)}{P(T+)}$$

$$P(D+ \mid T+) = \frac{P(T+ \mid D+) \cdot P(D+)}{P(T+ \mid D+) \cdot P(D+) + P(T+ \mid D-) \cdot P(D-)}$$

Naive Bayes

$$\hat{y} = \underset{k \in \{1, ..., m\}}{\operatorname{argmax}} \; p(C_k) \prod_{i=1}^{n} p(x_i \mid C_k)$$

Naive Assumption

$$p(C_k \mid x_1, x_2, \ldots, x_n) \propto p(C_k) \prod_{i=1}^{n} p(x_i \mid C_k)$$

Naive Bayes

Methodology

PARTE 1: Crear el modelo.

Para ello se necesitan **cuatro pasos**:

1. Calcular las probabilidades a priori de cada clase.
2. Para cada clase, realizar un recuento de los valores de atributos que toma cada ejemplo. Se debe distribuir cada clase por separado para mayor comodidad y eficiencia del algoritmo.
3. Aplicar la Corrección de Laplace, para que los valores "cero" no den problemas.
4. Normalizar para obtener un rango de valores [0,1].

PARTE 2:

1. Aplicar la fórmula de Naïve Bayes.

Naive Bayes

Example

Ejemplos	Atr. 1	Atr. 2	Atr. 3	Clase
x1	1	2	1	positiva
x2	2	2	2	positiva
x3	1	1	2	negativa
x4	2	1	2	negativa

For x5 = {1,1,1}, what is the class ?

[ ] http://naivebayes.blogspot.com/

Codes

With Scikit-Learn

1. Gaussian Naive Bayes

$$P(x_i \mid y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right)$$

2. Multinomial Naive Bayes

3. Bernoulli Naive Bayes

$$P(x_i \mid y) = P(i \mid y)x_i + (1 - P(i \mid y))(1 - x_i)$$

With Scikit-Learn

[ ] **1.9. Naive Bayes** http://scikit-learn.org/stable/modules/naive_bayes.html

Advantages & When to Use Naive Bayes

Advantages

- They are extremely fast for both training and prediction
- They provide straightforward probabilistic prediction
- They are often very easily interpretable
- They have very few (if any) tunable parameters

Practice

- When the naive assumptions actually match the data (very rare in practice)
- For very well-separated categories, when model complexity is less important
- For very high-dimensional data, when model complexity is less important

[ ] Python Data Science Handbook - Jake VanderPlas

Machine Learning

Naive Bayes & Bayes Classifier

Marchelo Bragagnini
cesarbrma91@gmail.com
@MarchBragagnini

Universidad Católica San Pablo

Centro de Investigación e Innovación en Ciencia Computación