Machine Learning

Naive Bayes & Bayes Classifier

Marchelo Bragagnini
cesarbrma91@gmail.com
@MarchBragagnini

Universidad Católica
San Pablo

Centro de Investigación
e Innovación en
Ciencia Computación

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- | D+) = 0.1 (false negative)
- P(T+ | D+) = ___
- P(T+ | 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/

Machine Learning

Naive Bayes & Bayes Classifier

Marchelo Bragagnini
cesarbrma91@gmail.com
@MarchBragagnini

Universidad Católica
San Pablo

Centro de Investigación
e Innovación en
Ciencia Computación