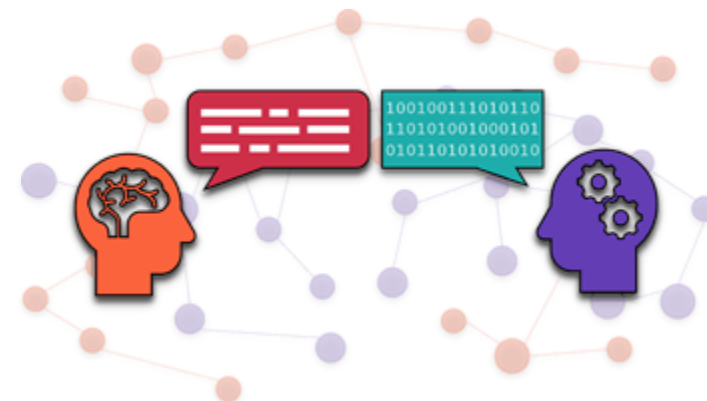


## :: U3 ::

# Procesamiento del Lenguaje Natural



**Anexo ::** Introducción a *Naive Bayes*

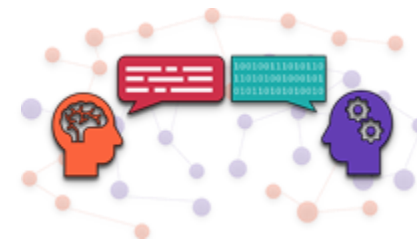
**Curso 2022-23**

# Tabla de contenidos



CIPFP Mislata  
Centre Integrat Públic  
Formació Professional Superior

1. ¿Qué es *Naive Bayes*?
2. Teorema de Bayes
3. Caso práctico
4. Modelo Naïve Bayes
5. Pros y contras de *Naive Bayes*



# 1. ¿Qué es *Naive Bayes*?



CIPFP Mislata  
Centre Integrat Públic  
Formació Professional Superior

- Los modelos de *Naive Bayes* son una clase especial de algoritmos de clasificación de Aprendizaje Automático
- Se basan en una técnica de clasificación estadística llamada “teorema de Bayes”
- Proporcionan una manera fácil de construir modelos con un comportamiento muy bueno debido a su simplicidad



## 2. Teorema de *Bayes*

$$P(A|R) = \frac{P(R|A)P(A)}{P(R)}$$

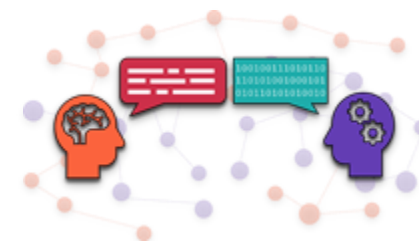
$P(A)$ : Probabilidad de A

$P(R|A)$ : Probabilidad de que se de R dado A

$P(R)$ : Probabilidad de R

$P(A|R)$ : Probabilidad posterior de que se de A dado R

- La fórmula de Bayes se basa en la noción estadística de “suceso”
- **A** y **R** son **sucesos**
  - cualquier posible resultado de un experimento aleatorio
  - o dicho de otro modo, a cualquier subconjunto de un espacio muestral, siendo el espacio muestral un conjunto de posible resultados





## 2. Teorema de *Bayes*

- **Ejemplo:** probabilidad de sacar un resultado concreto al tirar al aire una moneda



## 2. Teorema de *Bayes*

- **Ejemplo:** probabilidad de sacar un resultado concreto al tirar al aire una moneda
  - **A** = "sacar cara" → **P(A)** = 0,5
  - **B** = "sacar cruz" → **P(B)** = 0,5



## 2. Teorema de *Bayes*

- **Ejemplo:** probabilidad de sacar un resultado concreto al tirar al aire una moneda
  - **A** = "sacar cara" → **P(A)** = 0,5
  - **B** = "sacar cruz" → **P(B)** = 0,5
- Si tiramos la moneda 2 veces ... ¿Cuál es la probabilidad de sacar cara habiendo sacado antes cruz?



## 2. Teorema de *Bayes*

- **Ejemplo:** probabilidad de sacar un resultado concreto al tirar al aire una moneda
  - **A** = "sacar cara" →  **$P(A) = 0,5$**
  - **B** = "sacar cruz" →  **$P(B) = 0,5$**
- Si tiramos la moneda 2 veces ... ¿Cuál es la probabilidad de sacar cara habiendo sacado antes cruz?
  - **$P(A|B) = 0,5$**
  - **$P(B|A) = 0,5$**





## 2. Teorema de Bayes

- **Ejemplo:** probabilidad de sacar un resultado concreto al tirar al aire una moneda

- **A** = "sacar cara" → **P(A)** = 0,5

- **B** = "sacar cruz" → **P(B)** = 0,5

- Si tiramos la moneda 2 veces ... ¿Cuál es la probabilidad de sacar cara habiendo sacado antes cruz?

- **P(A|B)** = 0,5

- **P(B|A)** = 0,5



La probabilidad es la misma porque en este caso ambos sucesos son **independientes**



## 2. Teorema de Bayes

- **Ejemplo:** probabilidad de sacar un resultado concreto al tirar al aire una moneda

- **A** = "sacar cara" → **P(A)** = 0,5

- **B** = "sacar cruz" → **P(B)** = 0,5

- Si tiramos la moneda 2 veces ... ¿Cuál es la probabilidad de sacar cara habiendo sacado antes cruz?

- **P(A|B)** = 0,5



La probabilidad es la misma porque en este caso ambos sucesos son **independientes**

- **P(B|A)** = 0,5



A esta forma de representar la probabilidad de un suceso a partir de otro se llama **probabilidad condicionada**



## 2. Teorema de *Bayes*

- **Ejemplo:** en una estantería hay **60** novelas y **20** libros de poesía.  
Una persona **A** elige un libro al azar de la estantería y se lo lleva.  
A continuación otra persona **B** elige otro libro al azar.



## 2. Teorema de *Bayes*

- **Ejemplo:** en una estantería hay **60** novelas y **20** libros de poesía.  
Una persona **A** elige un libro al azar de la estantería y se lo lleva.  
A continuación otra persona **B** elige otro libro al azar.
- ¿Cuál es la probabilidad de que el libro seleccionado por **B** sea una novela?

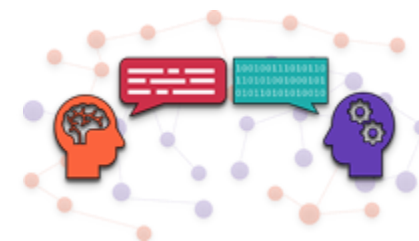


## 2. Teorema de *Bayes*

- **Ejemplo:** en una estantería hay **60** novelas y **20** libros de poesía.  
Una persona **A** elige un libro al azar de la estantería y se lo lleva.  
A continuación otra persona **B** elige otro libro al azar.
- ¿Cuál es la probabilidad de que el libro seleccionado por **B** sea una novela?

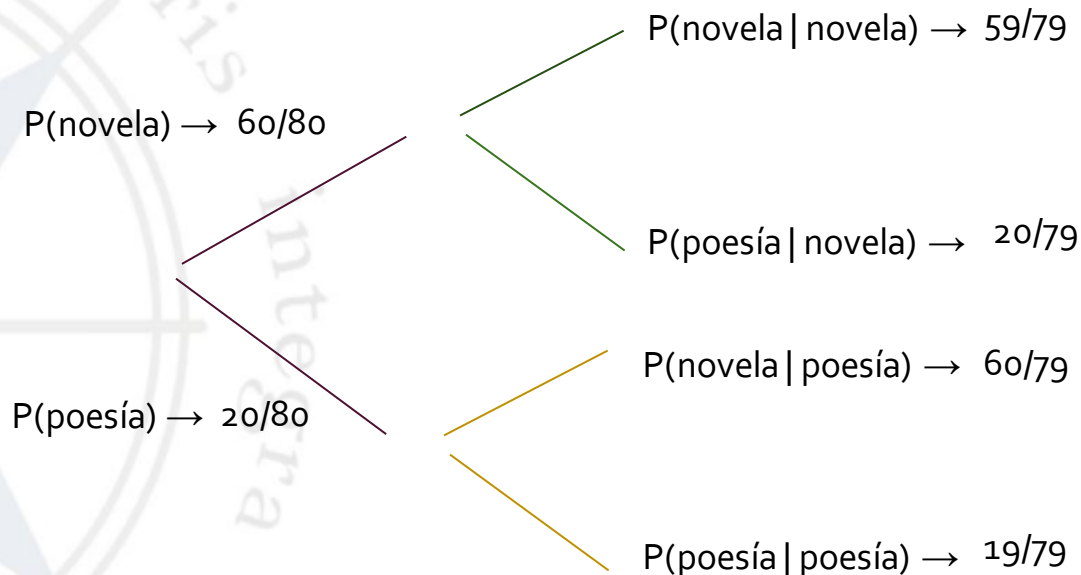
$$P(\text{novela}) \rightarrow 60/80$$

$$P(\text{poesía}) \rightarrow 20/80$$



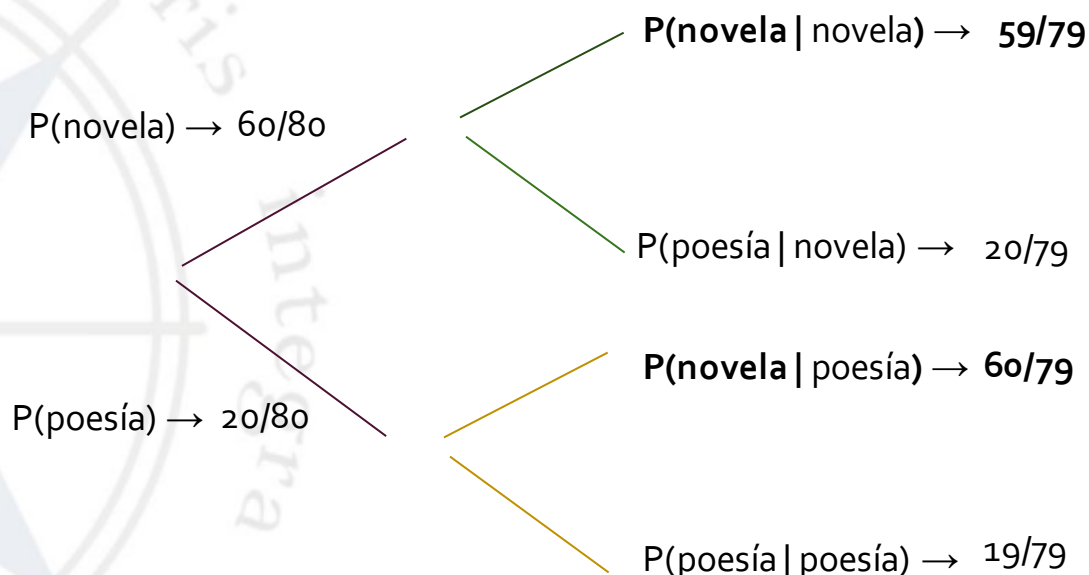
## 2. Teorema de *Bayes*

- **Ejemplo:** en una estantería hay **60** novelas y **20** libros de poesía.  
Una persona **A** elige un libro al azar de la estantería y se lo lleva.  
A continuación otra persona **B** elige otro libro al azar.
- ¿Cuál es la probabilidad de que el libro seleccionado por **B** sea una novela?



## 2. Teorema de Bayes

- **Ejemplo:** en una estantería hay **60** novelas y **20** libros de poesía.  
Una persona **A** elige un libro al azar de la estantería y se lo lleva.  
A continuación otra persona **B** elige otro libro al azar.
- ¿Cuál es la probabilidad de que el libro seleccionado por **B** sea una novela?



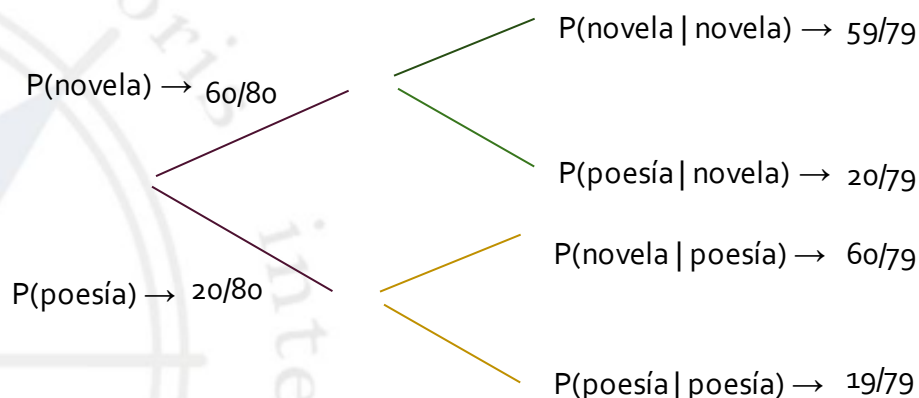
$P(\text{B escoge novela}) =$

$$60/80 * 59/79 + 20/80 * 60/79 = 3/4$$



## 2. Teorema de Bayes

- **Ejemplo:** en una estantería hay **60** novelas y **20** libros de poesía. Una persona **A** elige un libro al azar de la estantería y se lo lleva. A continuación otra persona **B** elige otro libro al azar.
- ¿Cuál es la probabilidad de que el libro seleccionado por **B** sea una novela?



**$P(\text{B escoge novela}) =$**

$$60/80 * 59/79 + 20/80 * 60/79 = 3/4$$

La probabilidad en este caso se da entre sucesos **dependientes**





### 3. Caso práctico

- Consideremos el caso de dos compañeros que trabajan en la misma oficina: **Alicia** y **Bruno**.
- Sabemos que:
  - Alicia viene a la oficina **3** días a la semana
  - Bruno viene a la oficina **1** día a la semana
  - ambos solo trabajan **4** días de los 5 laborables
- Probabilidad de “encontrarnos con” (suceso) **Alicia** o **Bruno**:
  - $P(\text{Alicia}) = 3/4 = \mathbf{0.75}$
  - $P(\text{Bruno}) = 1/4 = \mathbf{0.25}$



### 3. Caso práctico

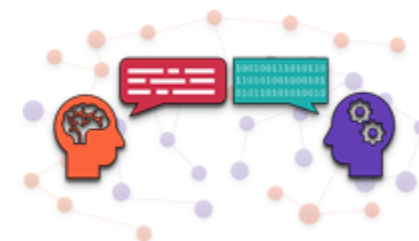
- Situación hipotética:

- Imaginemos que vemos pasar a una persona vestida con una **chaqueta roja** (suceso) y sabemos que ...

- **Alicia** viste de rojo **2** veces en días laborables
- **Bruno** viste de rojo **3** veces en días laborables

- probabilidad de que **Alicia** vista de rojo:  $P(\text{Rojo}|\text{Alicia}) = 2/5 = 0.4$

- probabilidad de que **Bruno** vista de rojo:  $P(\text{Rojo}|\text{Bruno}) = 3/5 = 0.6$





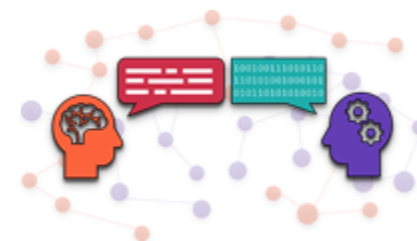
**CIPFP Mislata**  
Centre Integrat Públic  
Formació Professional Superior

### 3. Caso práctico

## Pregunta

¿A quién vimos pasar en la oficina?

¿A Alicia o a Bruno?



### 3. Caso práctico

#### Conocido

Probabilidad de que  
Alicia vista de rojo

Probabilidad de que  
Bruno vista de rojo

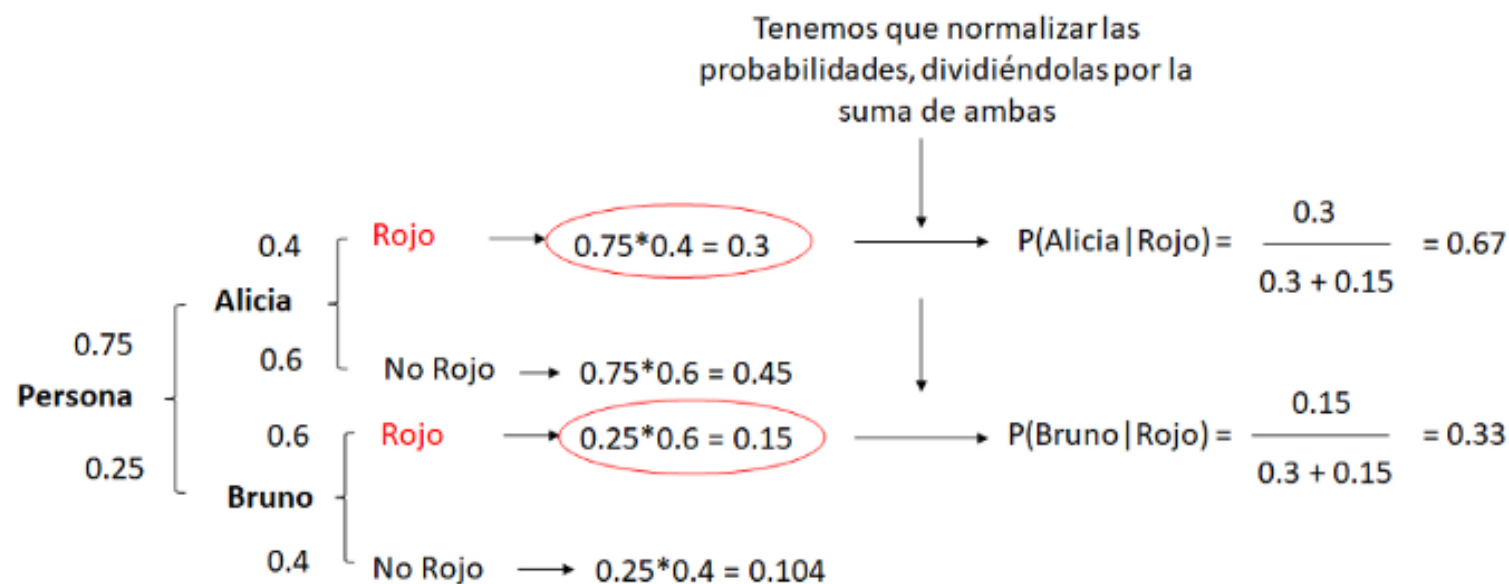
#### Supuesto

Probabilidad de que  
la persona que viste  
de rojo sea Alicia

Probabilidad de que  
la persona que viste  
de rojo sea Bruno

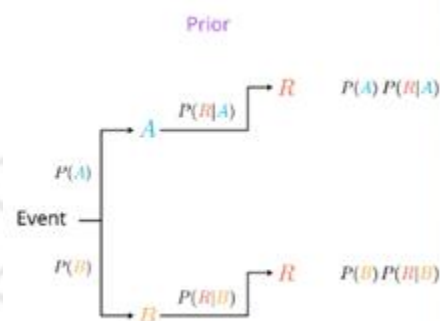
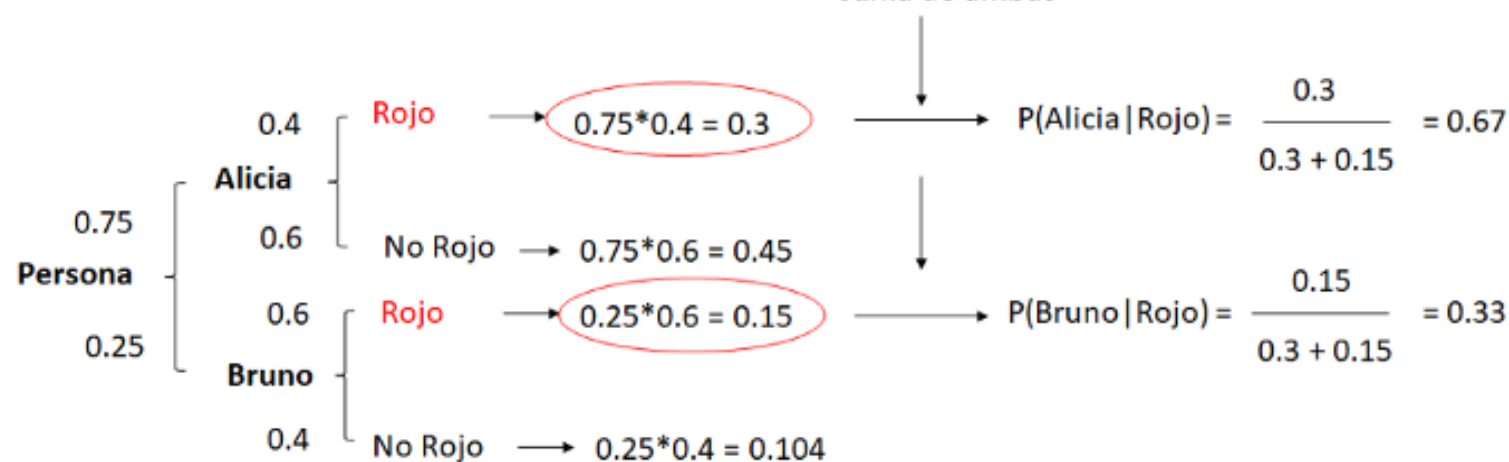


### 3. Caso práctico



# 3. Caso práctico

Tenemos que normalizar las  
probabilidades, dividiéndolas por la  
suma de ambas



Posterior

$$P(A|R) = \frac{P(A)P(R|A)}{P(A)P(R|A) + P(B)P(R|B)}$$

$$P(B|R) = \frac{P(B)P(R|B)}{P(A)P(R|A) + P(B)P(R|B)}$$





## 4. Modelo Naïve Bayes

- Una utilidad de Naive Bayes se da en el contexto de la **clasificación binaria**, donde tenemos dos clases (positiva y negativa)

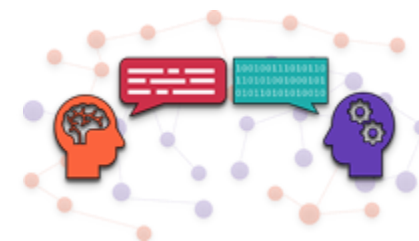


## 4. Modelo Naïve Bayes

- Una utilidad de Naive Bayes se da en el contexto de la **clasificación binaria**, donde tenemos dos clases (positiva y negativa)
- La matriz de confusión es una forma de representar el conjunto de posibilidades entre dos clases de eventos: uno "real" y otro que predecimos.



Prediction \ Class	True	False
True	True Positive	False Positive
False	False Negative	True Negative





## 4. Modelo Naïve Bayes

- Una utilidad de Naive Bayes se da en el contexto de la **clasificación binaria**, donde tenemos dos clases (positiva y negativa)
- La **matriz de confusión** es una forma de representar el conjunto de posibilidades entre dos clases de eventos: uno "real" y otro que predecimos.
- Métricas disponibles:
  - **exactitud**
  - **precisión**
  - **sensibilidad**



Prediction \ Class	True	False
True	True Positive	False Positive
False	False Negative	True Negative



Rango de 0 a 1 → **cuánto más cerca a 1 mayor bondad del modelo**

(es decir, de su fiabilidad en las predicciones que elabora)



## 4. Modelo Naïve Bayes

Supongamos que estamos evaluando a **100 pacientes** para una enfermedad y sabemos que la prevalencia de la enfermedad en la población es del **5%** ( $P(\text{enfermedad}) = 0.05$ ) y la no enfermedad es del **95%** ( $P(\text{no enfermedad}) = 0.95$ ).



## 4. Modelo Naïve Bayes

### Probabilidades a Priori:

- $P(\text{enfermedad}) = 0.05$
- $P(\text{no enfermedad}) = 0.95$

### Matriz de Confusión (Resultados del Modelo):

Supongamos que nuestro modelo predice:

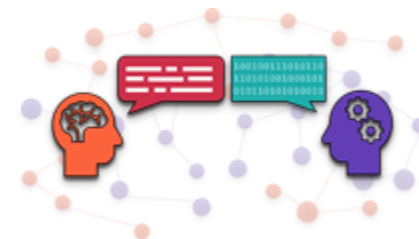
- TP (Verdaderos Positivos) = 4
- FP (Falsos Positivos) = 1
- TN (Verdaderos Negativos) = 90
- FN (Falsos Negativos) = 5



## 4. Modelo Naïve Bayes

¿Cómo podemos interpretar este caso desde Naïve Bayes?

- **$P(\text{enfermedad}|\text{predicción})$**  es la probabilidad de que un paciente tenga la enfermedad dado que el modelo lo predijo como positivo.
- **$P(\text{no enfermedad}|\text{predicción})$**  sería la probabilidad de que un paciente no tenga la enfermedad dado que el modelo lo predijo como negativo.



## 4. Modelo Naïve Bayes

### > Cálculo de la predicción

P(predicción)

... se calcularía como ...

$P(\text{enfermedad}) * P(\text{predicción}|\text{enfermedad}) + P(\text{no enfermedad}) * P(\text{predicción}|\text{no enfermedad})$

... es decir ...

$P(p) = P(e) * P(p/e) + P(ne) * P(p/ne)$



## 5. Pros y contras

Los puntos débiles principales son:

- **Suposición de independencia:** el término "Naive" se refiere a la suposición de independencia condicional entre las características, lo cual es una simplificación que permite que el algoritmo sea más eficiente.
- **Sensibilidad a características irrelevantes:** cuando el conjunto de datos de prueba tiene una característica que no ha sido observada en el conjunto de entrenamiento, el modelo le asignará una probabilidad de cero y será inútil realizar predicciones.



## 5. Pros y contras

Los puntos fuertes principales son:

- **Eficiencia y simplicidad:** Naïve Bayes representa una manera fácil y rápida de predecir clases.
- **Buena generalización:** en los casos en que sea apropiada una presunción de independencia, el algoritmo se comporta mejor que otros modelos de clasificación, incluso con menos datos de entrenamiento.

