

Introducción

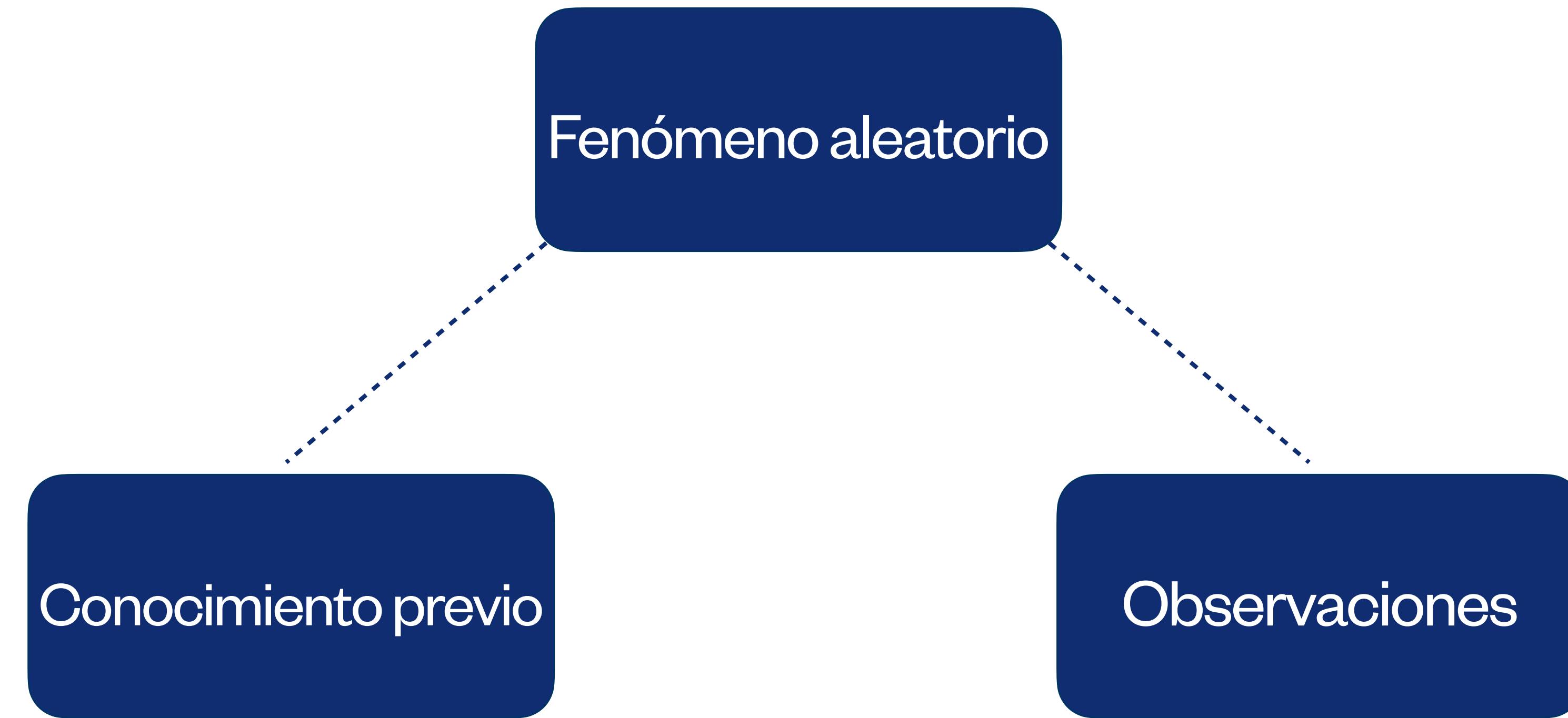
José A. Perusquía Cortés

Estadística Bayesiana, Semestre 2026-I



Estadística

- La estadística se puede definir como el estudio de fenómenos aleatorios



- Si se cuenta con toda la información bastará con un **análisis exploratorio** de lo contrario se necesitará un proceso de inferencia

Proceso de inferencia

- La única forma científica de hacer inferencia con base en una muestra es a partir de la probabilidad

$$x \longrightarrow X \longrightarrow \mathbb{P}(X = x)$$

- Inferencia paramétrica: Se asume que $\mathbb{P}(X = x) = p(x | \theta)$, donde θ es desconocido
- Inferencia no paramétrica: La forma funcional es desconocida

Enfoque clásico

- Se asume que θ es una **constante fija** que se aproxima a partir de un estimador $T(\mathbf{X})$ que se elige a partir de propiedades como:
 - Insesgado
 - Consistente
 - Eficiente
 - UMVUE
 - Etc.
- ¡Los datos se consideran aleatorios a pesar de haber sido observados!



Ejemplo 1

- Sea x_1, \dots, x_n una m.a. de distribución $Ber(\theta)$. ¿Cuál es el UMVUE para θ^2 ?
- El UMVUE es

$$\frac{\sum_i X_i (\sum_i X_i - 1)}{n(n-1)}$$

- ¿Qué sucede si se observa $x_1 = 1, x_2 = 0, \dots, x_n = 0$?

El estimador es igual a cero. ¡No es posible!



Ejemplo 2

- Suponer que se lanza una moneda 10 veces con 5 águilas
- Suponer que se lanza una tachuela 20 veces y se observa que cae de cabeza 10 veces
- ¿Cuál es el estimador máximo verosímil para la probabilidad de caer águila? Y ¿cuál sería el estimador máximo verosímil para la probabilidad de que la tachuela caiga de cabeza?
- Los estimadores máximo verosímiles son los mismos. ¿Esto es lógico?

No se toman en cuenta los conocimientos previos del experimento



Reflexión filosófica

- La inferencia clásica responde la pregunta

¿Qué valor de θ hace más plausibles los datos?

- Deberíamos de preguntarnos

¿Qué valor de θ es más probable dados los datos?



*¿Habrá alguna forma más coherente y libre
de contradicciones de hacer estadística?*

Sí, con el enfoque bayesiano



Teorema de Bayes

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, $\{E_i\}$ una partición finita o numerable de Ω y $A \in \mathcal{F}$ tal que $\mathbb{P}(A) > 0$. Entonces,

$$\mathbb{P}(E_i | A) = \frac{\mathbb{P}(A | E_i) \cdot \mathbb{P}(E_i)}{\sum_j \mathbb{P}(A | E_j) \cdot \mathbb{P}(E_j)}$$

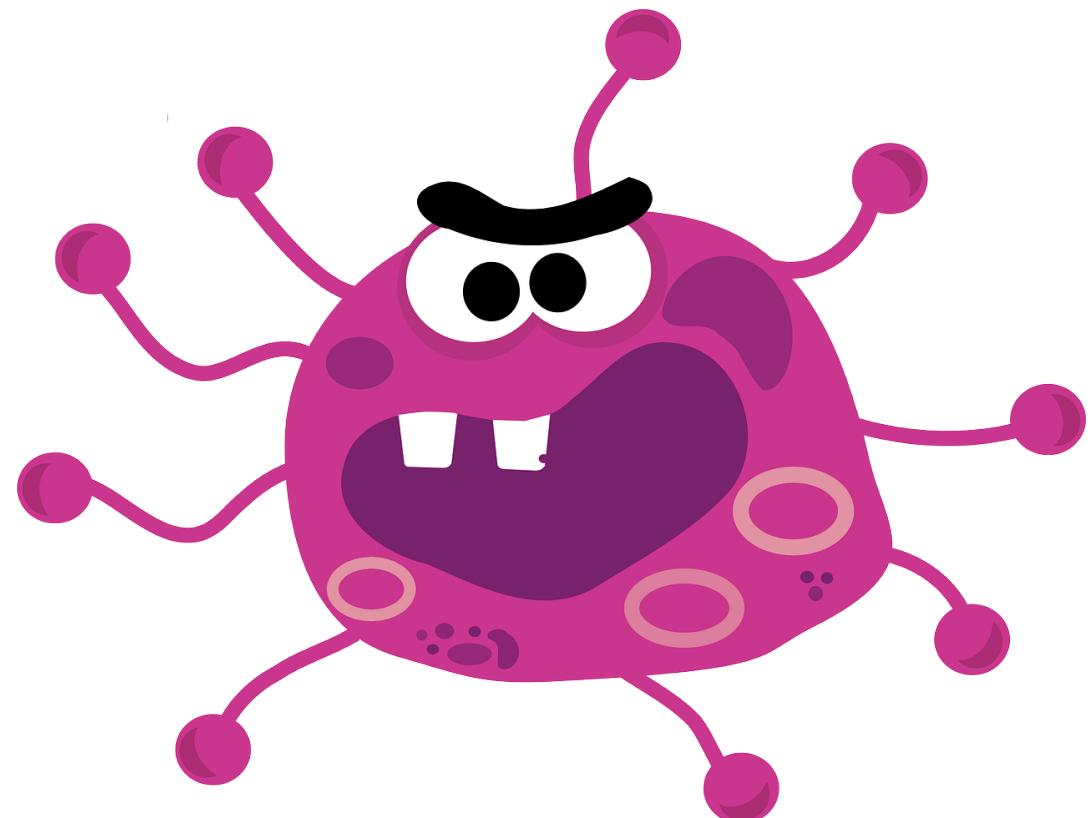
Ejemplo 3. Problema de Monty Hall.

En un programa televisivo un concursante elige una puerta con la posibilidad de ganar un auto. El presentador abrirá inmediatamente una puerta conteniendo una cabra y le dará la oportunidad de cambiar puerta al concursante. ¿Cuál es la mejor decisión para el concursante?



Ejemplo 4. Aplicación a epidemiología.

Una cierta enfermedad tiene una incidencia del 2%. Si la tasa de falsos negativos de una prueba es del 10% y la tasa de faltos positivos es del 1%, ¿Cuál es la probabilidad de que una persona que tenga un test positivo realmente tenga la enfermedad?



Ejemplo 5.

Suponer que se tiene un urna con 100 panditas de las cuales $100 - n$ son rojos y n son verdes, en donde n se distribuye uniformemente en $\{0, 1, \dots, 100\}$. Si en el primer intento se saca un pandita verde y no se regresa a la urna, entonces el segundo pandita:

1. Es más probable que sea verde.
2. Es más probable que sea rojo.
3. Son equiprobables.



Naive Bayes

- Un conjunto de modelos bayesianos de aprendizaje supervisado y utilizado en gran medida para la clasificación de textos
- Dado un conjunto de observaciones $\{\mathbf{x}_i\}_{i=1}^N$ para los cuales **se conoce el grupo al que pertenecen**, denotado por $\{y_i\}_{i=1}^N$ se asume que:

$$\mathbb{P}(\mathbf{x}_i | y_i) = \prod_{j=1}^M \mathbb{P}(x_{ij} | y_i)$$

- En otras palabras, dado el grupo los atributos son **condicionalmente independientes**

Naive Bayes

- Para una nueva observación \mathbf{x}_{N+1} se busca encontrar el grupo al que pertenece, calculando para toda j

$$\mathbb{P}(y_j | \mathbf{x}_{N+1}) \propto \mathbb{P}(\mathbf{x}_{N+1} | y_j) \mathbb{P}(y_j)$$

- Se elige el grupo que tenga la **mayor probabilidad**
- Para evitar posibles problemas numéricos se considera en su lugar una transformación logarítmica, esto es,

$$\log(\mathbb{P}(y_j | \mathbf{x}_{N+1})) \propto \log(\mathbb{P}(\mathbf{x}_{N+1} | y_j)) + \log(\mathbb{P}(y_j))$$

Naive Bayes

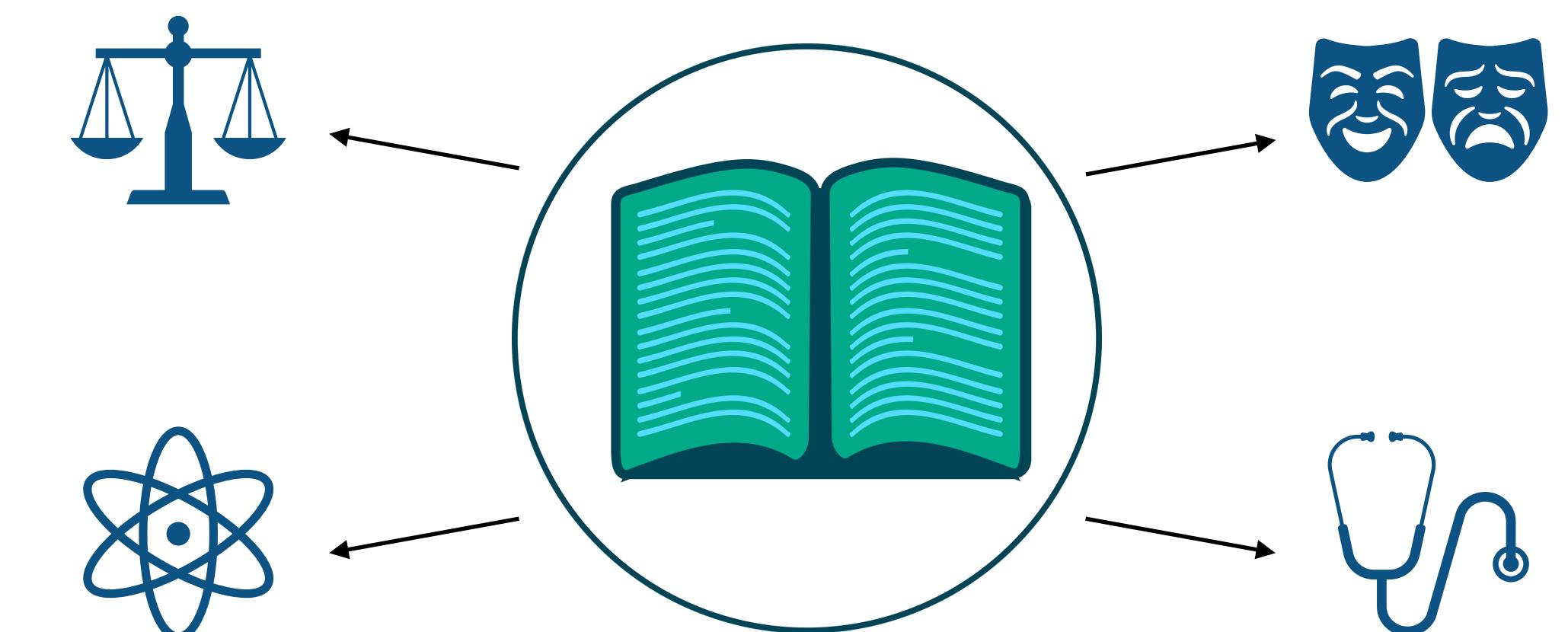
- Dependiendo del tipo de atributos se tiene:

Gaussian naive Bayes para datos continuos.

Multinomial naive Bayes para frecuencias.

Bernoulli naive Bayes para atributos dicotómicos.

- Para la clasificación de textos nos centramos (principalmente) en el clasificador multinomial o Bernoulli.



Detección de spam

- El *spam* es cualquier correo electrónico no solicitado con una gran variedad de fines, desde:
 - Propaganda comercial
 - Robo de identidad
 - Difusión de virus informáticos
- El nombre Spam originalmente se refiere a un producto ultraprocesado
- El uso de spam como correo no solicitado se debe a Monty Python's Flying Circus
- Los correos legítimos se conocen como *ham*



Detección de spam

¿Cómo sabemos cuando algo es spam?

 Spotify
recordatorio !#LPxqvS
To: José A. Perusquía

Junk - Ciencias 7 August 2025 at 13:19

Actualiza tu método de pago

Hola,

No pudimos procesar tu último pago de Spotify Premium. Para evitar la interrupción del servicio, te pedimos que actualices tu método de pago en las próximas **48 horas**.

Si no realizas la actualización dentro de este plazo, tu cuenta pasará automáticamente al plan gratuito, perdiendo acceso a beneficios como escuchar sin anuncios, descargas offline y saltos ilimitados.

Actualizar pago

Es un proceso rápido y seguro que te permite seguir disfrutando tu música sin interrupciones.

Gracias por ser parte de Spotify.
Equipo de Spotify México

© 2025 Spotify AB. Este mensaje es informativo y fue enviado a la dirección registrada en tu cuenta.
Consulta nuestros términos en: [spotify.com/mx/legal](https://www.spotify.com/mx/legal)

 Aisha Gaddafi
Blessed new week to you my dear, your kind response is needed.
To: undisclosed-recipients;,, Bcc: José A. Perusquía,
Reply-To: aishagaddafisdaughter@gmail.com

Junk - Ciencias 21 July 2025 at 07:03

Details

Blessed new week to you my dear,

My name is Mrs. Aisha Gaddafi, the only daughter of late Colonel Muammar Gaddafi Of Libya and I am currently living as a refugee in Oman. I urgently need your help and personal assistance to help me receive funds worth \$27.500.000.00 US Dollars for an investment project in your country.

If you are willing to help, please reply promptly through my person email below:
aishagaddafisdaughter@gmail.com

Thank you,
Mrs. Aisha Gaddafi

Detección de spam

- La detección de spam se plantea como un problema de clasificación de textos
- Dado un conjunto de entrenamiento se obtienen a partir de las frecuencias:
 - $P(\text{Spam})$ y $P(\text{Ham})$
 - $P(\omega_i | \text{Spam})$ y $P(\omega_i | \text{Ham})$ para cada palabra ω_i en el vocabulario
 - Palabras que no se hayan visto en el entrenamiento se omiten



Detección de spam

- Considerando el modelo multinomial se tiene que

$$\mathbb{P}(\omega_i | \text{Spam}) = \frac{\#(\omega_i, \text{Spam})}{\sum_i \#(\omega_i, \text{Spam})}$$

- Para evitar probabilidades nulas se realiza el **suavizamiento de Laplace**

$$\mathbb{P}(\omega_i | \text{Spam}) = \frac{\#(\omega_i, \text{Spam}) + \alpha}{\sum_i (\#(\omega_i, \text{Spam}) + \alpha)}$$

- Por lo general $\alpha = 1$

