



# **U2. Sistemas expertos y Sistemas automáticos**

**Lab 2: Modelos de redes bayesianas.**

# ÍNDICE



1. Sistemas expertos.
2. Tipos de sistemas expertos.
3. Ventajas vs Desventajas.
4. Actualidad.
5. Automatismos.
6. Lab1: Explorando sistemas de automatización
- 7. Lab 2: Modelos de redes bayesianas.**
8. Lab 3: Selenium.

## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.

El clasificador Näive Bayes utiliza las probabilidades a-priori para estimar la probabilidad de los eventos futuros.



Reverendo Thomas Bayes

## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.



- Este clasificador utiliza datos de entrenamiento para calcular la probabilidad observada de cada evento en función del vector de características.
- Cuando el clasificador es utilizado con datos sin etiquetar, utiliza las probabilidades observadas para estimar la clase más probable.

## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.



- Se utiliza principalmente para: clasificar texto, detección de intrusos en redes, diagnósticos médicos, etc.
- Si todos los eventos fueran independientes, sería imposible predecir ningún evento con los datos observados por el otro.
- Los eventos dependientes son la base del modelado predictivo.

*Ejemplo, la presencia de nubes suele ser un evento predictivo de un día lluvioso. La presencia de la palabra viagra suele ser un evento Predictivo de spam.*

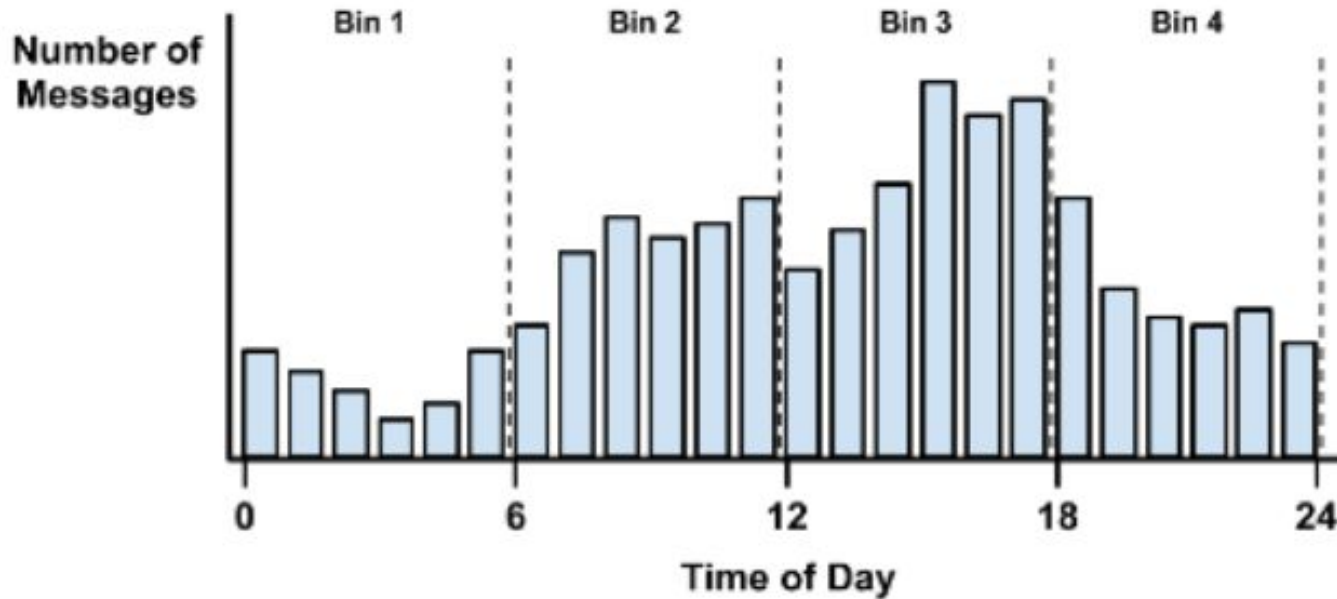
## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.



Debido a que NB utiliza tablas de frecuencias para aprender de los datos, cada variable debe de ser categórica y no puede utilizar directamente variables numéricas.

Una solución sencilla es discretizar las variables numéricas. Este método es ideal cuando hay grandes cantidades de datos. La forma más común de discretizar es explorar los datos para ver los puntos de corte en la distribución de los datos.

## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.



## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.



- La discretización siempre se traduce en una reducción de la información ya que la granularidad inicial se reduce.
- Es importante mantener un balance en el número de bins. Con pocas se pierde mucha información y con muchas es muy costoso.



## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.



- Simple, rápido y efectivo.
- Funciona bien con datos noisy y missing.
- Requiere de pocos ejemplos para entrenar, pero puede usar muchos.

## 7. Lab 2: Modelos de redes bayesianas: Naive Bayes.



No es lo ideal para datasets con un gran número de variables numéricas.

Es naïve porque trata a todas las variables como independientes e igualmente importantes.

- Esto no es así en el mundo real. Por ejemplo, la persona que envía el email puede ser más importante que el propio texto.
- Además la aparición de las palabras no son independientes. Si aparece la palabra viagra, es muy probable que la palabra droga aparezca cerca.

## 5. Lab 2: Modelos de redes bayesianas: Naive Bayes.



### Actividad

Vamos a descargarnos el notebook “Lab1: Modelos de redes bayesianas Naive Bayes” de Aules y cargarlo en Google Colab.

Es un notebook guiado. Donde vais a tener que rellenar cada apartado, una vez terminado realizaremos una corrección en conjunto.

Respecto al dataset:

- Class Name: 2 (democrat, republican) -> (1,0)
- Resto: Handicapped-infants: 2 (y,n) -> (1,0)

## 5. Lab 2: Modelos de redes bayesianas: Naive Bayes.



Los resultados los anotaremos en el siguiente excel, después haremos un debate con los resultados obtenidos.

[https://docs.google.com/spreadsheets/d/1ETLQQfnlvqDw7QCDNzKW\\_GcrQ2bxDN2rEdCXwd4jxjM/edit#gid=0](https://docs.google.com/spreadsheets/d/1ETLQQfnlvqDw7QCDNzKW_GcrQ2bxDN2rEdCXwd4jxjM/edit#gid=0)