

Universidad Rafael Landívar
 Inteligencia Artificial
 Primer Semestre 2025

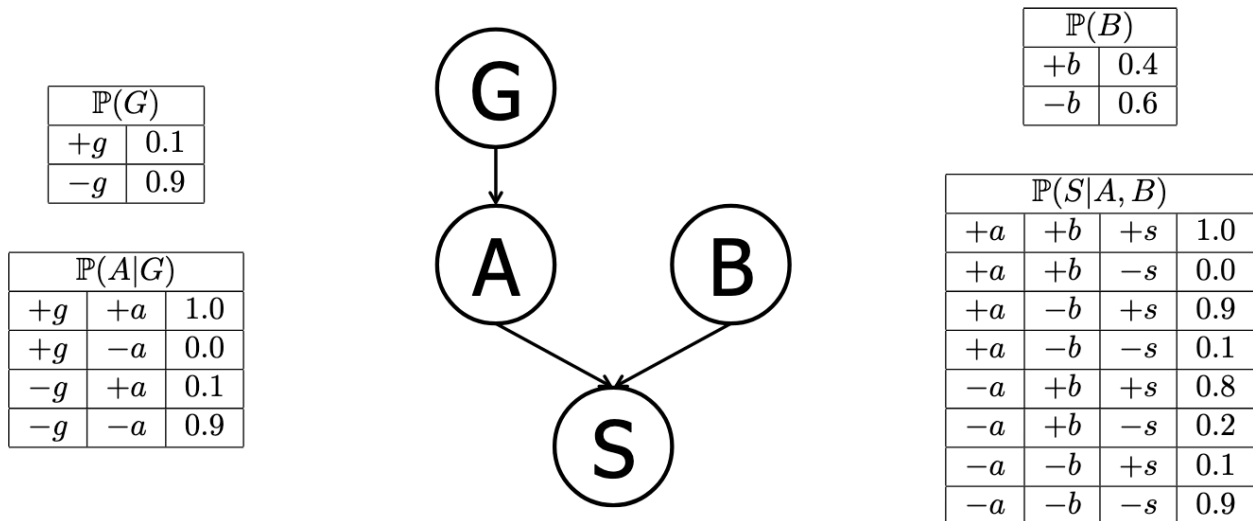
Hoja de Trabajo 8

Descripción

Esta actividad se divide en dos partes: la primera aborda un problema de **redes bayesianas**, y la segunda aplica el clasificador **Naïve Bayes** para la predicción de etiquetas de clase.

Serie 1. Redes Bayesianas (50 puntos)

Suponga la existencia de un paciente que tiene síntomas (S) los cuales pueden ser causados por dos diferentes enfermedades (A y B). En esta línea es un hecho conocido que la variación en el gen G tiene un gran impacto en la aparición de la enfermedad A . Este escenario se encuentra representado en la siguiente red bayesiana incluyendo sus tablas de probabilidad.



Para esto, se le solicita que responda a las siguientes interrogantes utilizando su conocimiento de modelos de redes bayesianas.

- (5 Puntos) ¿Cuál es la probabilidad de que el paciente presente los síntomas, padezca ambas enfermedades y sea afectado por el gen G ?
- (5 Puntos) ¿Cuál es la probabilidad de que el paciente tenga la enfermedad A ?
- (10 Puntos) ¿Cuál es la probabilidad de que el paciente tenga la enfermedad A dado que ya tiene la enfermedad B ?
- (10 Puntos) ¿Cuál es la probabilidad de que un paciente tenga la enfermedad A dado que tiene los síntomas S y la enfermedad B ?
- (10 Puntos) ¿Cuál es la probabilidad de que un paciente fue afectado por el gen G dado que presenta la enfermedad A ?
- (10 Puntos) ¿Cuál es la probabilidad de que un paciente fue afectado por el gen G dado que presenta la enfermedad B ?

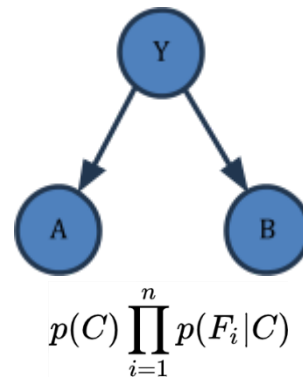
Serie 2. Naïve Bayes (50 puntos)

En este ejercicio, resolverá un caso particular de las redes bayesianas: el clasificador Naïve Bayes, que se utilizará para predecir la etiqueta de clase Y en función de las variables de entrada A y B . Todas las variables son binarias (0 o 1). En algunos casos, deberá aplicar conceptos de Máxima Verosimilitud y Suavizamiento de Laplace para estimar las distribuciones de probabilidad.

Datos de Entrenamiento:

Se tienen 10 ejemplos de entrenamiento:

A	B	Y
1	1	1
1	0	1
1	0	0
1	1	0
0	1	0
1	1	1
0	1	1
1	0	0
1	1	0
1	1	0



La estimación de Máxima Verosimilitud se basa en la frecuencia observada en los datos.

Preguntas:

- (20 puntos)** Calcule las CPTs necesarias para construir la tabla de probabilidades:
 - Calcule la probabilidad marginal $P(Y)$:
 - $P(Y=0)$
 - $P(Y=1)$
 - Calcule la distribución condicional $P(A|Y)$:
 - $P(A=0|Y=0), P(A=1|Y=0)$
 - $P(A=0|Y=1), P(A=1|Y=1)$
 - Calcule la distribución condicional $P(B|Y)$:
 - $P(B=0|Y=0), P(B=1|Y=0)$
 - $P(B=0|Y=1), P(B=1|Y=1)$
- (20 puntos)** Clasificación con Naïve Bayes

Se sabe que $A=1$ y $B=1$, identifique la etiqueta aplicando Naïve Bayes:

 - Prediga la clase más probable (Aplica *Maximun Likelihood*).
 - $P(Y=0|A=1, B=1)$ (Normaliza)
 - $P(Y=1|A=1, B=1)$ (Normaliza)

3. (10 puntos) **Aplicación de Suavizamiento de Laplace**

Lea el complemento de esta hoja de trabajo.

Para evitar probabilidades de cero, aplica el suavizamiento de Laplace con $k=2$ y recalcula la tabla $P(A|Y)$.

3.1 Calcule los nuevos valores de $P(A|Y)$ con Laplace Smoothing.

Complemento:

La **fórmula de suavizado de Laplace** en el contexto del **Clasificador Naïve Bayes** se usa para evitar probabilidades nulas cuando una categoría no ha aparecido en los datos de entrenamiento.

$$P(clase) = \frac{Frecuencia(clase)+k}{Total\ de\ características\ en\ la\ clase + kV}$$

Donde:

- k es el parámetro de suavizado (si $k = 1$, se trata del suavizado de Laplace clásico).
- V es el número total de valores distintos que puede tomar la característica.

Este ajuste permite usar valores distintos de k para controlar la cantidad de suavizado aplicado.

Este método evita que la probabilidad condicional sea cero cuando una combinación no aparece en los datos de entrenamiento, lo que podría hacer que la probabilidad total de una clase se vuelva cero en el cálculo de **Naïve Bayes**.

