

Métodos Bayesianos I

García Prado, Sergio
sergio@garciparedes.me

20 de abril de 2017

Resumen

[TODO]

1. SE SABE QUE UN 1 % DE LAS MUJERES DE 40 AÑOS QUE PARTICIPAN EN UN EXAMEN RUTINARIO TIENEN CÁNCER DE MAMA. TAMBIÉN SE SABE QUE UN 80 % DE LAS MUJERES QUE TIENEN CÁNCER DE MAMA, DARÁN POSITIVO AL HACERSE UNA MAMOGRAFÍA. SIN EMBARGO, UN 9,6 % DE LAS MUJERES QUE NO TIENEN CÁNCER DE MAMA DARÁN POSITIVO EN UNA MAMOGRAFÍA. EN ESTE CONTEXTO UNA MUJER DE 40 AÑOS SE SOMETE A UN EXAMEN RUTINARIO Y SU MAMOGRAFÍA DA POSITIVO. ¿CUÁL ES LA PROBABILIDAD DE QUE REALMENTE TENGA CÁNCER DE MAMA?

[TODO]

2. DADAS DOS VARIABLES ALEATORIAS DISCRETAS, X E Y , Y DADA SU DISTRIBUCIÓN DE PROBABILIDAD CONJUNTA QUE APARECE EN LA TABLA, SE PIDE:

- 2.1. ¿CUMPLE LA DISTRIBUCIÓN CONJUNTA LAS PROPIEDADES DE UNA DISTRIBUCIÓN DE PROBABILIDADES?

[TODO]

- 2.2. ¿CUÁL ES LA PROBABILIDAD DE $Pr(X = x_1)$?

[TODO]

- 2.3. ¿CUÁLES SON LAS DISTRIBUCIONES MARGINALES DE $Pr(X = x)$ Y $Pr(Y = y)$?

[TODO]

	x_1	x_2	x_3	x_4	$Pr(Y)$
y_1	2/16	1/16	1/16	1/16	5/16
y_2	1/16	2/16	2/16	1/16	6/16
y_3	1/16	1/16	1/16	0	3/16
y_4	0	2/16	0	0	2/16
$Pr(X)$	4/16	6/16	4/16	2/16	16/16

Tabla 1: Frecuencias relativas de la distribución de probabilidad conjunta de X e Y

2.4. ¿VERIFICAN LAS DISTRIBUCIONES MARGINALES LAS PROPIEDADES DE UNA DISTRIBUCIÓN DE PROBABILIDADES?

[TODO]

3. UTILIZANDO EL CONJUNTO DE DATOS *weather-nominal-practica* QUE SE PROPORCIONA, DETERMINAR LA CLASIFICACIÓN NAIVE BAYES DE LAS SIGUIENTES INSTANCIAS, UTILIZANDO LA ESTIMACIÓN DE MÁXIMA VEROSIMILITUD (FRECUENCIAL)

$$x_1 = \langle \text{sunny}, \text{cool}, \text{normal}, \text{false} \rangle \quad (1)$$

$$x_2 = \langle \text{overcast}, \text{mild}, \text{high}, \text{true} \rangle \quad (2)$$

[TODO]

4. UTILIZANDO WEKA Y EL CLASIFICADOR NAIVEBAYES DETERMINAR LA CLASIFICACIÓN DE LOS EJEMPLOS ANTERIORES, ¿COINCIDE CON LA CLASIFICACIÓN CALCULADA EN EL EJERCICIO ANTERIOR?

[TODO]

5. ENTRENAR CON WEKA, UN CLASIFICADOR NAIVE BAYES PARA EL CONJUNTO DE DATOS *weather-nominal*

5.1. ESTIMAR LA TASA DE ERROR COMETIDA POR EL CLASIFICADOR UTILIZANDO VALIDACIÓN CRUZADA DE 10 PARTICIONES

[TODO]

5.2. EXAMINAR LA SALIDA PROPORCIONADA POR EL EXPLORER Y DETERMINAR CÓMO ESTÁ ESTIMANDO ESTA IMPLEMENTACIÓN DE *Naive Bayes* LOS PARÁMETROS DEL CLASIFICADOR

[TODO]

6. EL CONJUNTO DE DATOS *weather-nominal-X6* SE HA GENERADO REPITIENDO CADA INSTANCIA DEL CONJUNTO *weather-nominal* SEIS VECES. ENTRENAR CON WEKA UN CLASIFICADOR NAIVE BAYES PARA ESTE CONJUNTO DE DATOS:

6.1. ESTIMAR LA TASA DE ERROR COMETIDA POR EL CLASIFICADOR UTILIZANDO VALIDACIÓN CRUZADA DE 10 PARTICIONES

[TODO]

6.2. COMPARE ESTA TASA DE ERROR CON LA ESTIMADA EN EL EJERCICIO ANTERIOR Y DISCUTA LOS RESULTADOS

[TODO]

REFERENCIAS

- [CCAG17] Teodoro Calonge Cano and Carlos Javier Alonso González. Técnicas de Aprendizaje Automático, 2016/17.
- [GP17] Sergio García Prado. Métodos bayesianos 1. <https://github.com/garciparedes/machine-learning-bayesian-1>, 2017.
- [too] Weka. <http://www.cs.waikato.ac.nz/ml/weka/>.