Aprendizaje Automático Tarea #2 Clasificador Bayesiano Ingenuo

① Generos: Un programa de salud gubernamental desea clasificar los registros de las personas en géneros femenino (F) y masculino (M) a partir de los atributos nombre, estatura y peso. Se cuentan con los sig. registros.

Nombre	Estatura (m)	Peso (kg)	Género
Denis	1.72	75.3	M
Guadalupe	1.82	81.6	M
Alex	1.80	86.1	M
Alex	1.70	77.1	M
Cris	1.73	78.2	M
Juan	1.80	74.8	M
Juan	1.80	74.3	M

Denis	1.50	50.5	F
Alex	1.52	45.3	F
Cris	1.62	61.2	F
Rene	1.67	68.0	F
Guadalupe	1.65	58.9	F
Guadalupe	1.75	68.0	F

Entrena un clasificador bayesiano ingenuo usando estimacion por máxima verosimilitud y otro usando estimacion por máximo a posteriori. Reporta los parametros que obtuviste en ambos casos y usa los clasificadores entrenados para predecir la clase de los sig. vectores: $x_1 = (\text{Rene}, 1.68, 65)$, $x_2 = (\text{Guadalupe}, 1.75, 80)$, $x_3 = (\text{Denis}, 1.80, 79)$ $x_4 = (\text{Alex}, 1.90, 85)$ y $x_5 = (\text{Cris}, 1.65, 70)$. Describe detalladamente el procedimiento que seguiste tanto en el entrenamiento como en la predicción y discutir los resultados obtenidos. Para el entrenamiento del clasificador por máximo a posteriori considerar los sig valores para las distribuciones correspondientes

Género	Nombre α_k	Estatura μ_0	Estatura σ_0^2	Estatura σ^2	Peso μ_0	Peso σ_0^2	Peso σ^2
M	$2 \; \forall k$	1.7	0.3	0.0020	85.5	17.0	15.76
F	$2 \; \forall k$	1.5	0.1	0.0074	70.3	85.0	71.00

A ojometro x_2, x_3, x_4 pertenecen a la clase M; x_5 y x_1 podrian ser F. Sólo de ver los datos. Veamos si es "verdad".

- Estimación por máxima verosimilitud:

<u>Def.</u> (verosimilitud - Muestra observada) Sea $X_1, \dots, X_n$ muestra aleatoria (m.a.) de un modelo con densidad $f_X(x;\theta)$ y $\theta \in \Theta$ un parámetro desconocido, definimos la función de verosimilitud $\ell : \Theta \to \mathbb{R}$ como:

$$\ell(\theta; \underline{x}) := f_{\underline{X}}(x_1, \dots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i; \theta)$$

Primero asumimos una distribución de probabilidad para cada uno de los atributos. Para este caso la estatura y el peso las asumimos como distribuciones normales $N(\theta, \sigma^2)$. Mientras que el nombre lo modelamos como una categorica $Cat(x; \vec{q})$.

Por estabilidad numerica usaremos la Estatura en cm

Entrenamiento: Tenemos la siguiente tabla: (calculada en excel)

Género	μ_P (media peso) (kg)	σ_P^2	μ_A (media altura) (cm)	σ_A^2
M	78.2	115.7657	1.76	2.02×10^{-3}
F	58.6	171.009	1.61.8	7.44×10^{-3}

Ahora bien; El clasificador bayesiano está dado por

$$\hat{y} = \underset{k\in\{1,\dots,K\}}{\text{argmax}}\left[P(C_k)\prod_{i=1}^{n} p(x_i|C_k)\right].$$

Para el caso de las características categoricas tenemos que:

Características: $[\vec{x}]_\alpha \in \{f_1, f_2, \dots, f_{K_\alpha}\}$: cada característica α cae dentro de una de las k_α categorías.

$$P(x_\alpha = j | y = c) = [\theta_{jc}]_\alpha \quad y \quad \sum_{j=1}^{k_\alpha}[\theta_{jc}]_\alpha = 1$$

donde $[\theta_{jc}]_\alpha$ es la probabilidad de que la caract. α tenga el valor j dado que la etiqueta es c y la restriccion indica que x_α debe tener una de las categorias $\{1,\dots,k_\alpha\}$

La estimacion de parametros se hace como:

$$[\hat{\theta}_{jc}] = \frac{\sum_{i=1}^{n} I(y_i = c)\, I(x_{i\alpha} = j) + L}{\sum_{i=1}^{n} I(y_i = c) + Lk_\alpha}$$

Donde $x_{i\alpha} = [\vec{x}_i]_\alpha$ y L es un parametro de suavizado. Si $L=0$ tenemos MVE, si $L>0$ tenemos MAP. Si $L=1$ obtendremos el suavizado de Laplace

En palabras: $[\hat{\theta}_{jc}] = \dfrac{\#\text{ muestras con etiqueta } c \text{ que tienen la caract } \alpha \text{ con valor } j}{\#\text{ muestras con la etiqueta } c}$

Predicción: $\underset{y}{\text{argmax}}\,(P(y=c|\bar{x})) \propto \underset{y}{\text{argmax}}\, \hat{\pi}_c \prod_{\alpha=1}^{d}[\hat{\theta}_{jc}]_\alpha$

Extracto tomado de: www.cs.cornell.edu/courses/cs4780/2018sp/lectures/lecturenote05.html

Dadas las condiciones de independencia de Naive Bayes para calcular $\hat{y}$ basta con calcular los estimadores de cada distribucion y despues multiplicarlas entre sí

- Estimacion para la normal de las estaturas:

Sabemos que para la normal $\hat{\mu}_{MVE} = \frac{1}{n}\sum_{i=1}^{n} x^{(i)}$ y $\hat{\sigma}^2_{MVE} = \frac{1}{n}\sum_{i=1}^{n}(x^{(i)} - \hat{\mu}_{EMV})^2$.

$\hat{\mu}^{(hombre)}_{MVE_A} = \frac{1}{7}(172+182+180+170+173+180+180) = 176.71 \approx 177$. → ya lo había calculado

$\hat{\mu}^{(mujer)}_{MVE_A} = \frac{1}{6}(150+152+162+167+165+175) = 161.8 \approx 162$

$$\hat{\sigma}^{2(h)}_{MVE_A} = \frac{1}{7}\left[(172-177)^2+(182-177)^2+(180-177)^2+(170-177)^2+(173-177)^2+(180-177)^2+(180-177)^2\right]$$

$$= \frac{1}{7}[25+25+9+49+16+9+9] = \frac{142}{7} = 20.28 \approx 20$$

2 $\hat{\sigma}^{2(mujer)}_{MVE_A} = \frac{1}{6}\left[(150-162)^2+(152-162)^2+(162-162)^2+(167-162)^2+(165-162)^2+(175-162)^2\right]$

$= \frac{1}{6}[144+100+0+25+9+169] = 74.5 \approx 74$

- Estimacion para la normal de los pesos:

$\hat{\mu}^{(hombre)}_{MVE_P} = \frac{1}{7}(75.3+81.6+86.1+77.1+78.2+74.8+74.3) = 78.2$

$\hat{\sigma}^{2(hombre)}_{MVE_P} = \frac{1}{7}\left[(75.3-78.2)^2+(81.6-78.2)^2+(86.1-78.2)^2+(77.1-78.2)^2+(78.2-78.2)^2+\ldots\right.$

$\left.\ldots+(74.8-78.2)^2+(74.3-78.2)^2\right] = \frac{1}{7}[8.41+11.56+62.41+1.21+0+11.56+15.21]$

$= \frac{1}{7}[110.36] = 15.766 = 15.8$

$\hat{\mu}^{(mujer)}_{MVE_P} = \frac{1}{6}(50.5+45.3+61.2+68.0+58.9+68.0)] = 58.65 = 58.6$

$\hat{\sigma}^{2(mujer)}_{MVE_P} = \frac{1}{6}\left[(50.5-58.6)^2+(45.3-58.6)^2+(61.2-58.6)^2+(68.0-58.6)^2+(58.9-58.6)^2+(68.0-58.6)^2\right]$

$= \frac{1}{6}[(65.61+176.89+6.76+88.36+0.09+88.36)] = \frac{1}{6}(426.07) = 71.011 = 71.0$

- Estimacion para la categorica:

$L=0$ ya que es la estimacion máxima verosimilitud

$[\hat{\theta}_{hombre\,Denis}] = \frac{1}{2}$; $[\hat{\theta}_{hombre\,Gad}] = \frac{1}{3}$; $[\hat{\theta}_{hAle}] = \frac{2}{3}$; $[\hat{\theta}_{hCris}] = \frac{1}{2}$

$[\hat{\theta}_{hJuan}] = \frac{2}{2} = 1$; $[\hat{\theta}_{mDenis}] = \frac{1}{2}$; $[\hat{\theta}_{mAle}] = \frac{1}{3}$; $[\hat{\theta}_{mCris}] = \frac{1}{2}$; $[\hat{\theta}_{mRene}] = 1$

$[\hat{\theta}_{mGad}] = \frac{2}{3}$; $[\hat{\theta}_{mJuan}] = [\hat{\theta}_{hRene}] = 0$ ya que según los datos no hay mujeres que se llamen Juan y no hay hombres que se llamen Rene

Y con esto ya podemos construir nuestro clasificador bayesiano

$$P(hombre\,|\,\vec{x}) \propto \left[\frac{1}{\sqrt{2\pi\hat{\sigma}^{2(h)}_{MVE_P}}}\exp\left(-\frac{(x_i^{(P)}-\hat{\mu}_{MVE_P})^2}{2\hat{\sigma}^{2(h)}_{MVE_P}}\right)\right]\left[\frac{1}{\sqrt{2\pi\hat{\sigma}^{2(h)}_{MVE_A}}}\exp\left(-\frac{(x_i^{(A)}-\hat{\mu}_{MVE_P})^2}{2\hat{\sigma}^{2(h)}_{MVE_P}}\right)\right]*\ldots$$

$$\ldots*\left[\hat{\theta}_{hombre\,x_i^{(cat)}}\right]\left[P(hombre)\right] \quad \ldots(1)$$

$$P(mujer\,|\,x_i) \propto \left[\frac{1}{\sqrt{2\pi\hat{\sigma}^{2(m)}_{MVEP}}}\exp\left(-\frac{(x_i^{(P)}-\hat{\mu}_{MVEP})^2}{2\hat{\sigma}^{2(m)}_{MVEP}}\right)\right]\left[\frac{1}{\sqrt{2\pi\hat{\sigma}^{2(m)}_{MVEA}}}\exp\left(-\frac{(x_i^{(A)}-\hat{\mu}_{MVEA})^2}{2\hat{\sigma}^{2}_{MVEA}}\right)\right]*\ldots$$

$$\ldots*\left[\hat{\theta}_{mujer\,x_i^{(cat)}}\right]\left[P(mujer)\right] \quad \ldots(2)$$

¿Topaste ese abuso de notacion? :)

$x_i^{(\Omega)}$ es la componente correspondiente a Ω del vector de entrada $\vec{x}$

(No soy bueno con la rotación) :(

Testing: Ahora sí, sea la entrada $\vec{x}_1 = (\text{Rene}, 168, 65)$ $\hat{\theta}_{\text{hombre Rene}}$

$$P(h|\vec{x}_1) = \left[\frac{1}{\sqrt{2\pi(15.8)}} \exp\left(\frac{-(65-78.2)^2}{2(15.8)}\right)\right]\left[\frac{1}{\sqrt{2\pi(20)}} \exp\left(\frac{-(1.68-177)^2}{2(20)}\right)\right]\left[0\right]\left[\frac{7}{13}\right] = 0$$

a partir de aquí no anotaré todo, porque si no acabo hasta mañana

$$P(m|\vec{x}_1) = \left[(0.047)\exp\left(\frac{-(65-58.6)^2}{142}\right)\right]\left[(0.046)\exp\left(\frac{-(168-162)^2}{148}\right)\right]\left[1\right]\left[\frac{6}{13}\right] = \ldots$$

$$= \left[(0.047)(0.749)(0.046)(0.784)(1)(0.462) = 5.86 \times 10^{-4}\right.$$

Por lo que $\vec{x}_1$ es mujer

$\vec{x}_2 = (\text{Guadalupe}, 1.75, 80)$

$P(h|\vec{x}_2) = 0.00131$

$P(m|\vec{x}_2) = 8.57 \times 10^{-6}$

Por lo que $\vec{x}_2$ es hombre

$\vec{x}_3 = (\text{Denis}, 180, 79)$

$P(h|\vec{x}_3) = 0.001886$

$P(m|\vec{x}_3) = 3.028 \times 10^{-6}$

Por lo que $\vec{x}_3$ es hombre

$\vec{x}_4 = (\text{Alex}, 190, 85)$

$P(h|\vec{x}_4) = 1.08 \times 10^{-5}$

$P(m|\vec{x}_4) = 1.248 \times 10^{-8}$

Por lo que $\vec{x}_4$ es hombre

$\vec{x}_5 = (\text{Cris}, 165, 70)$

$P(h|\vec{x}_5) = 7.843 \times 10^{-6}$

$P(m|\vec{x}_5) = 1.909 \times 10^{-4}$

Por lo que $\vec{x}_5$ es mujer

Veamos que para este caso la clasificacion está muy restringida por el parámetro $[\hat{\theta}_{jc}]$ ya que si al clasificador se le pasa un vector $\vec{x}$ tal que $x_i^{(cat)}$ contiene un nombre que no estaba en nuestros datos de entrenamiento el clasificador no dará ninguna clasificacion ya que $P(h|\vec{x}) = P(m|\vec{x}) = 0$ pues $[\hat{\theta}_{hc}] = [\hat{\theta}_{mc}] = 0$

Es decir obtenemos una indeterminacion, luego nuestro clasificador sólo servirá para entradas $\vec{x}$ cuyo $x_i^{(cat)}$ nuestro clasificador conoce. Por otro lado si alguna de las clases tiene un nombre que la otra no (como Rene o Juan) entonces el clasificador asignará cualquier vector de entrada, cuyo nombre coincide con el que ya se mencionó, a la clase que tiene dicho nombre sin importar los otros parámetros

Por ejemplo si yo le doy $\vec{x} = (\text{Rene}, 177, 78.2)$ el clasificador dirá que es mujer aún cuando su altura y peso sean los característicos de un hombre.

Para esta parte concluimos que la variable categorica es muy mala para este modelo ya que si queremos realizar una buena clasificacion necesitamos tener datos para todos los nombres, y que en el caso de nombres que pueden ser de hombre y mujer (como Guadalupe) se tengan en ambas clases. Si es el caso, entonces es buena está variable ya que permitiria ahorrar calculos para cuando las entradas tengan nombres únicos de cada clase (como Juan y Juana) Aunque esto no asegura que la clasificacion sea perfecta.

Con respecto a los parámetros no hay mucho que discutir. Los datos de entrenamiento muestran una clara division de las clases por ésta razon pude adivinar las salidas sin hacer calculos sería más interesante ver valores más cercanos.

El suavizado de Laplace arreglará uno de los problemas de la categorica (ejemplo de $\vec{x} = (\text{Rene}, 177, 78.2)$ pero no la falta de datos en cuanto a los nombres. veamoslo.

Dado que ahora hacemos el maximo a posteriori MAP nuestros estimadores serán:

$$[\hat{\theta}_{jc}] = \frac{\sum_{i=1}^{n} I(y_i = c) I(x_{ia} = j) + L}{\sum I(y_i = c) + L k_a}$$ donde $L = \alpha_k$ de la tabla; $L = 1$ $\forall$ Categorica

$$\hat{\mu} = \frac{\sigma_0^2 \left(\sum_{i=1}^{n} x^{(i)}\right) + \sigma^2 \mu_0}{\sigma_0^2 n + \sigma^2}$$

En la etapa de aprendizaje basta calcular $\hat{\mu}$ para la altura y peso y $[\hat{\theta}_{jc}]$ para la categorica, esto se hace usando la tabla y los datos.

Para la altura:

$$\hat{\mu}^{(hombre)}_{MAPA} = \frac{(0.002)(1.7) + 0.3(12.37)}{(0.3)(7) + 0.002} = \frac{0.0034 + 3.711}{2.1 + 0.002} = \frac{3.7144}{2.102} = 1.767^m = 176.7\text{ cm}$$

$$\hat{\mu}^{(mujer)}_{MAPA} = \frac{(0.1)(9.71) + (0.0074)(1.5)}{(0.1)(6) + 0.0074} = \frac{0.971 + 0.0111}{0.6 +} = \frac{0.9821}{0.6074} = 1.617^m = 161.7\text{ cm}$$

Para el Peso:

$$\hat{\mu}^{(hombre)}_{MAPP} = \frac{(17)(547.4) + (15.76)(85.5)}{(17)(7) + 15.76} = 79.054 = 79.05\text{ kg}$$

$$\hat{\mu}^{(mujer)}_{MAPP} = \frac{(85)(351.9) + (71)(70.3)}{(85)(6) + 71} = 60.074 = 60.07\text{ kg}$$

El MAP de la categorica es $[\hat{\theta}]_{jc} = \dfrac{C_k + \alpha_k - 1}{n + \sum_{k=1}^{K} \alpha_k - K}$

y para este caso $K = 5$; $\sum_{k=1}^{5} \alpha_k = (2)5 = 10$

$[\hat{\theta}_{hDenis}]_{MAP} = \dfrac{1+2-1}{2+10-5} = \dfrac{2}{7}$; $[\hat{\theta}_{hGuad}]_{MAP} = \dfrac{1+2-1}{3+10-5} = \dfrac{2}{8}$; $[\hat{\theta}_{hAlex}]_{MAP} = \dfrac{2+2-1}{3+10-5} = \dfrac{3}{8}$

$[\hat{\theta}_{hCris}]_{MAP} = \dfrac{1+2-1}{2+10-5} = \dfrac{2}{7}$; $[\hat{\theta}_{hJuan}]_{MAP} = \dfrac{2+2-1}{2+10-5} = \dfrac{3}{7}$; $[\hat{\theta}_{hRene}]_{MAP} = \dfrac{0+2-1}{1+10-5} = \dfrac{1}{6}$

$[\hat{\theta}_{mDenis}]_{MAP} = \dfrac{1+2-1}{2+10-5} = \dfrac{2}{7}$; $[\hat{\theta}_{mGuad}]_{MAP} = \dfrac{2+2-1}{3+10-5} = \dfrac{3}{8}$; $[\hat{\theta}_{mAlex}]_{MAP} = \dfrac{1+2-1}{3+10-5} = \dfrac{2}{8}$

$[\hat{\theta}_{mCris}]_{MAP} = \dfrac{1+2-1}{2+10-5} = \dfrac{2}{7}$; $[\hat{\theta}_{mRene}]_{MAP} = \dfrac{1+2-1}{1+10-5} = \dfrac{2}{6}$

Aquí termina el entrenamiento

Las funciones predictivas serán:

$$P(\text{hombre}|\vec{x}) = \mathcal{N}^{(h)}_{Peso}(\mu_p, \sigma_p^2) * \mathcal{N}^{(h)}_{Altura}(\mu_p, \sigma_p^2) * [\hat{\theta}_{hj}]_{MAP}$$

$\mathcal{N}$ denota la curva normal o gaussiana

y $\mu_p = \dfrac{1}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}} \left[\dfrac{\mu_0}{\sigma_0^2} + \dfrac{\sum_{i=1}^{n} x^{(i)}}{\sigma^2} \right]$; $\sigma_p^2 = \left[\dfrac{1}{\sigma_0^2} + \dfrac{n}{\sigma^2} \right]^{-1}$

$$P(\text{Mujer}|\vec{x}) = \mathcal{N}^{(m)}_{Peso}(\mu_p, \sigma_p^2) * \mathcal{N}^{(m)}_{altura}(\mu_p, \sigma_p^2) * [\hat{\theta}_{mj}]_{MAP}$$

Con $\vec{x}_1 = (\text{Rene}, 168, 65)$ tenemos:

$$P(\text{hombre}|\vec{x}_1) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sigma_{0p}^2}+\frac{n}{\sigma_p^2}\right)^{1/2} \exp\left(-\frac{\left(x-\left[\left(\frac{1}{\sigma_{0p}^2}+\frac{n}{\sigma_p^2}\right)^{-1}\left(\frac{\mu_{0p}}{\sigma_{0p}^2}+\frac{547.4}{\sigma_p^2}\right)\right]\right)^2}{2}\left(\frac{1}{\sigma_0^2}+\frac{n}{\sigma_p^2}\right)\right) * \cdots$$

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sigma_{0A}^2}+\frac{n}{\sigma_A^2}\right)^{1/2} \exp\left(-\frac{\left(x-\left[\left(\frac{1}{\sigma_{0A}^2}+\frac{n}{\sigma_A^2}\right)^{-1}\left(\frac{\mu_{0A}}{\sigma_{0A}^2}+\frac{1237}{\sigma_A^2}\right)\right]\right)^2}{2}\left(\frac{1}{\sigma_0^2}+\frac{n}{\sigma_A^2}\right)\right) * \cdots$$

$$\frac{\alpha_k + c_k}{n+\sum_{k=1}^{k}\alpha_k}$$

pero: $\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2} = \frac{\sigma^2+n\sigma_0^2}{\sigma_0^2\sigma^2} \iff \frac{1}{\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}} = \frac{\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}$

ahora $\frac{\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}\left[\frac{\mu}{\sigma_0^2}+\frac{\sum_{i=1}^{n}x^{(i)}}{\sigma^2}\right] = \frac{\mu\cancel{\sigma_0^2}\sigma^2}{(\sigma^2+n\sigma_0^2)\cancel{\sigma_0^2}} + \frac{\sigma_0^2\cancel{\sigma^2}\sum x^{(i)}}{\cancel{\sigma^2}(\sigma^2+n\sigma_0^2)} = \frac{\mu\sigma^2}{\sigma^2+n\sigma_0^2}+\frac{\sigma_0^2\sum x^{(i)}}{\sigma^2+n\sigma_0^2} = \hat{\mu}$

que ya habíamos calculado

falta calcular $\frac{\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2} = \sigma_p^2$ para el peso y altura esto para simplificar las cuentas, entonces

$$\sigma_{P_{\text{Altura}}}^{(h)2} = \frac{(0.0020)(0.3)}{(0.0020)+7(0.3)} = \frac{0.0006}{2.102} = 2.854\times10^{-4} = 0.0002.854$$

$$\sigma_{P_{\text{Altura}}}^{(m)2} = \frac{(0.1)(0.0074)}{(0.0074)+6(0.1)} = \frac{0.00074}{0.6074} = 0.001218$$

$$\sigma_{P_{\text{Peso}}}^{2(h)} = \frac{(17.0)(15.76)}{15.76+7(17)} = \frac{267.92}{134.76} = 1.9881$$

$$\sigma_{P_{\text{peso}}}^{2(m)} = \frac{(85.0)(71)}{71+6(85)} = \frac{6035.0}{581.0} = 10.387$$

Ahora sí sustituyendo valores:

$$P(\text{hombre}|\vec{x}_1) = \frac{1}{2\pi\sqrt{2.854\times10^{-4}}}\exp\left(-\frac{(1.68-1.77)^2}{2(2.854\times10^{-4})}\right) * \frac{1}{\sqrt{1.9881}}\exp\left(-\frac{(65-79)^2}{2(1.9881)}\right) * \frac{1}{6}$$

$= 2.9923\times10^{-28}$

$$P(\text{mujer}|\vec{x}_1) = \frac{1}{2\pi\sqrt{0.001218}}\exp\left(-\frac{(1.68-1.62)^2}{2(0.001218)}\right) * \frac{1}{\sqrt{10.387}}\exp\left(-\frac{(65-60)^2}{2(10.387)}\right) * \frac{2}{6} = 0.03227$$

$\therefore$ se clasifica como mujer

Con $\vec{x}_2 = (\text{Guadalupe}, 1.75, 80)$ → Ya no pondré el paso a paso porque me canso

$P(\text{mujer} \mid \vec{x}_2) = 2.207 \times 10^{-12}$

$P(\text{hombre} \mid \vec{x}_2) = 0.64$

∴ $\vec{x}_2$ lo clasificamos como hombre

Con $\vec{x}_3 = (\text{Denis}, 1.80, 79)$

$P(\text{mujer} \mid \vec{x}_3) = 1.898 \times 10^{-14}$

$P(\text{hombre} \mid \vec{x}_3) = 0.394$

∴ $\vec{x}_3$ lo clasificamos como hombre

Con $\vec{x}_4 = (\text{Alex}, 1.90, 85)$

$P(\text{mujer} \mid \vec{x}_4) = 3.137 \times 10^{-28}$

$P(\text{hombre} \mid \vec{x}_4) = 4.0593 \times 10^{-17}$

∴ $\vec{x}_4$ lo clasificamos como hombre

Con $\vec{x}_5 = (\text{Cris}, 165, 70)$

$P(\text{mujer} \mid \vec{x}_5) = 0.00226$

$P(\text{hombre} \mid \vec{x}_5) = 3.0002 \times 10^{-20}$

∴ $\vec{x}_5$ lo clasificamos como mujer

Vemos que la clasificacion en sí no cambia para ambos casos, sin embargo lo que sí cambia es $P(C \mid \vec{x}_i)$ ya que se considera el MAP en este caso.
Este resultado (que la clasificacion no cambie) era esperado ya que los vectores de entrada tienen tendencias muy faciles de detectar en el caso de los hombres la altura es mayor que 1.75 que se considran raros en las mujeres. Lo mismo para el peso. Lo unico extra a agregar es lo que ya se discutió considerar. MAP hace que la categorica no sea decisiva y permita clasificar personas con un nombre especifico en una clase tal que los datos de entrenamiento no contenga ese nombre tal clase un ejemplo es el caso con $\vec{x}_1$ donde Rene hombre no está en los datos de entrenamiento pero pasa que $P(\text{hombre} \mid \vec{x}_1) > 0$. Concluimos con un aprendí mucho xD