

Examen de Aprendizaje Automático Avanzado

ETSINF, Universitat Politècnica de València, 10 de junio de 2022

Apellidos:

Nombre:

Cuestiones (tiempo estimado: 60 minutos, 2 puntos)

Marca cada recuadro con una única opción de entre las dadas. Cada tres preguntas incorrectas, cancelan una correcta.

- 1 ☐ B Una población de personas presenta una distribución conjunta en el color de ojos y en el color del pelo como la que se puede ver en la tabla. La entropía de la variable “color de ojos” es:

- A) 2.22 bits
B) 0.88 bits $-0.7 \log_2 0.7 - 0.3 \log_2 0.3$
C) 1.49 bits
D) Ninguna de las anteriores

		pelo		
		negro	castaño	rubio
ojos	castaños	0.2	0.4	0.1
	azules	0.1	0.1	0.1

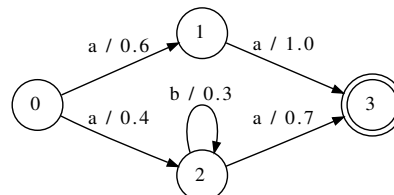
- 2 ☐ B En la tabla de la pregunta anterior, la información mútua entre ambas variables es:

- A) Igual a 0.010 B) Mayor que 0.015 C) Menor que 0.015 D) Ninguna de las anteriores

Solo es necesario calcular un sumando cuya probabilidad conjunta sea mínima. Por ejemplo “pelo negro y ojos azules”: $0.1 \log_2 0.1 / (0.3 * 0.3) = 0.01520$. Los otros sumando incrementan este valor.

- 3 ☐ A Dado el grafo de la figura, indica qué afirmación es correcta:

- A) La entropía sentencial no coincide con la entropía derivacional
B) La entropía sentencial sí coincide con la entropía derivacional
C) La entropía sentencial de la cadena “aba” es nula
D) La entropía derivacional de la cadena “aba” es nula



- 4 ☐ D Dado el grafo de la pregunta anterior, y la cadena “aa” indica qué afirmación es correcta:

- A) La entropía de los caminos que dan cuenta de dicha cadena no se puede calcular
B) La entropía de los caminos que dan cuenta de dicha cadena es negativa
C) La entropía de los caminos que dan cuenta de dicha cadena es máxima
D) La entropía de los caminos que dan cuenta de dicha cadena es mínima

- 5 ☐ A En el marco de la máxima entropía:

- A) Se pueden abordar problemas de clasificación de reconocimiento de formas
B) Sólo se pueden abordar problemas de clasificación en dos clases
C) No se pueden abordar problemas de clasificación de reconocimiento de formas
D) Ninguna de las anteriores

- 6 ☐ D En el marco de la máxima entropía, las funciones de distribución de probabilidad condicional son de la forma:

- A) $p(y, x) = \frac{1}{Z(x)} \exp(\sum_i \lambda_i f_i(x, y))$ donde $Z(x) = \sum_y \exp(\sum_i \lambda_i f_i(x, y))$
B) $p(y|x) = \frac{1}{Z(x)} \exp(\sum_i \lambda_i f_i(x, y))$ donde $Z(x) = \sum_y \exp(\sum_i \lambda_i f_i(x, y))$
C) $p(y|x) = \frac{1}{Z(x)} \exp(\sum_i \lambda_i)$ donde $Z(x) = \sum_y \exp(\sum_i \lambda_i)$
D) $p(y|x) = \frac{1}{Z(x)} \exp(\sum_i \lambda_i f_i(x, y))$ donde $Z(x) = \sum_y \exp(\sum_i \lambda_i f_i(x, y))$

- 7 ☐ B En el marco de la máxima entropía, el algoritmo IIS:

- A) Se utiliza para seleccionar las características
B) Se utiliza para estimar los *multiplicadores* λ de Lagrange
C) No se utiliza en el marco de máxima entropía
D) Se utiliza para ajustar las muestras de aprendizaje

- 8 ☐ C El algoritmo EM se basa:

- A) Solo en el conteo de los eventos no observados
B) Solo en el conteo de los eventos observados
C) En el cálculo de una esperanza que combina la variable observada y la variable no observada
D) En la suma de la probabilidad *a priori* de la variable observada y la variable no observada

9 **D** Los parámetros de un modelo estimado con el algoritmo EM regularizado:

- A) Coinciden siempre con el algoritmo EM sin regularizar
- B) Coinciden con el algoritmo EM sin regularizar cuando $\gamma = 1$
- C) Nunca pueden coincidir con el algoritmo EM sin regularizar
- D) Coinciden con el algoritmo EM sin regularizar cuando $\gamma = 0$

10 **A** El algoritmo de optimización por el criterio MMI optimiza los parámetros de un modelo según la expresión que puede verse bajo. Indica qué afirmación es cierta.

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=1}^n \log \frac{p_{\theta}(\mathbf{x}_i | \mathbf{z}_i)}{p_{\theta}(\mathbf{x}_i)}$$

- A) El denominador se puede calcular marginalizando sobre \mathbf{z}_i
- B) El denominador puede obviarse puesto que no influye en el resultado final
- C) No se puede calcular el denominador puesto que tiene un número exponencial de sumandos
- D) Ninguna de las anteriores

11 **D** La expresión del apartado anterior se puede aproximar como:

A)

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=1}^n \log \frac{p_{\theta}(\mathbf{x}_i | \mathbf{z}_i)}{p_{\theta}(\mathbf{x}_i) p_{\theta}(\mathbf{z}_i)}$$

B)

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=1}^n \log \frac{p_{\theta}(\mathbf{z}_i | \mathbf{x}_i)}{p_{\theta}(\mathbf{x}_i)}$$

C)

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=1}^n \log \frac{p_{\theta}(\mathbf{x}_i, \mathbf{z}_i)}{p_{\theta}(\mathbf{x}_i)}$$

D)

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=1}^n \log \frac{p_{\theta}(\mathbf{x}_i, \mathbf{z}_i)}{p_{\theta}(\mathbf{x}_i) p_{\theta}(\mathbf{z}_i)}$$

12 **B** En un problema de aprendizaje activo tenemos tres muestras $\{x_0, x_1, x_2\}$ sin etiquetar. Las etiquetas posibles son $\{c_0, c_1\}$, y las *posteriors* $p(C | X)$ con un modelo θ son:

	x_0	x_1	x_2
c_0	0.3	0.5	0.8
c_1	0.7	0.5	0.2

Si elegimos para anotar la muestra de máxima entropía entonces la muestra seleccionada sería:

- A) La muestra x_0
- B) La muestra x_1
- C) La muestra x_2
- D) Cualquiera de ellas