Examen de Aprendizaje Automático Avanzado

DSIC, Universitat Politècnica de València, 7 de junio de 2024

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 en bits de la distribución "color de ojos de las personas que no tienen el pelo castaño" es:

A) 0.72 bits

B) 1.92 bits

C) 0.88 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 C Se tiene un distribución normal p de media y varianza conocidas, $N(\mu, \sigma^2)$, con $\sigma \neq 0$. Dado el evento $s_0 = p(x \leq \mu)$ y el evento $s_1 = p(x > \mu)$ indica qué afirmación es cierta:

A) La entropía del evento s_0 es nula.

B) La entropía del evento s_1 es nula.

C) La entropía de los eventos s_0 y s_1 es no nula.

D) La entropía de los eventos s_0 y s_1 es nula.

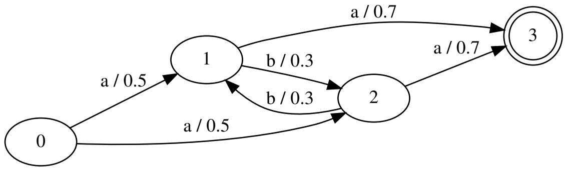
3 A Dado el grafo de la figura, indica qué afirmación es correcta:

A) La entropía sentencial es menor que la entropía derivacional.

B) La entropía sentencial es mayor que la entropía derivacional.

C) La entropía sentencial es igual que la entropía derivacional.

D) La entropía sentencial y la entropía derivacional no se pueden calcular.



4 B Dado el grafo de la pregunta anterior, y la cadena $w = aba$ indica qué afirmación es correcta:

A) $H_{\mathcal{A}}(\theta|w) = .34$ bits. B) $H_{\mathcal{A}}(\theta|w) = .68$ bits. C) $H_{\mathcal{A}}(\theta|w) = .47$ bits. D) $H_{\mathcal{A}}(\theta|w) = .87$ bits.

5 C Dado un modelo M que se desea utilizar para representar una muestra de datos S, indica qué afirmación es cierta. La estimación por máxima-verosimilitud

A) maximiza la distancia de Kullback-Leibler entre la distribución representada por el modelo y la distribución empírica de la muestra.

B) no afecta en nada a la distancia de Kullback-Leibler entre la distribución representada por el modelo y la distribución empírica de la muestra.

C) minimiza la distancia de Kullback-Leibler entre la distribución representada por el modelo y la distribución empírica de la muestra.

D) provoca la divergencia de la distancia de Kullback-Leibler entre la distribución representada por el modelo.

6 C En el marco de la máxima entropía, los valores λ asociados a las características pueden

A) No pueden pueden tomar valores nulos.

B) Representan el peso final de cada muestra de entrenamiento.

C) Representan el peso final de cada característica.

D) Representan el peso final de cada clase.

7 D Elige la afirmación correcta. En el marco de la máxima entropía, el algoritmo IIS :

A) solo puede alcanzar un óptimo relativo de la función optimizada.

B) no alcanza un óptimo absoluto de la función optimizada.

C) alcanza un punto de silla de función optimizada.

D) ninguna de la anteriores.

8 A El algoritmo EM se basa:

A) En la valor esperado del logaritmo de la distribución conjunta con la distribución de las posterioris de la variable no observada.

B) Solo en el conteo de los eventos observados.

C) Solo en el conteo de los eventos no observados.

D) En la suma de la probabilidad *apriori* de la variable observada y la variable no observada.

9 $\boxed{A}$ Los parámetros de un modelo estimado con el algoritmo EM regularizado:

A) Puede coincidir con el algoritmo EM sin regularizar cuando $\gamma = 1$.
B) Siempre coinciden con el algoritmo EM sin regularizar cuando $\gamma = 1$.
C) Nunca pueden coincidir con el algoritmo EM sin regularizar.
D) Ninguna de las anteriores.

10 $\boxed{B}$ El algoritmo de optimización por el criterio MMI optimiza los parámetros de un modelo según la expresión que puede verse bajo.

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

Indica qué afirmación es cierta.

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

11 $\boxed{D}$ 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 $\boxed{C}$ 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 *posterioris* $p(C \mid X)$ 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 mínima 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