

# Test de Sistemas Inteligentes - Bloque 2 - MUIINF

ETSINF, Universitat Politècnica de València, 18 de Junio de 2018

Apellidos:

Nombre:

## Cuestiones (60 minutos, sin apuntes)

Escribe en cada recuadro la opción correcta entre las dadas.

- B** En el marco de la máxima entropía, la expresión  $\tilde{p}(f) = \sum_{x,y} \tilde{p}(x,y) f(x,y)$
- A) es un valor esperado y por tanto está necesariamente entre 0,0 y 1,0.
  - B) es un valor esperado de una característica que es necesariamente mayor o igual que 0,0.
  - C) es un valor esperado de una característica que puede ser negativo, cero o positivo.
  - D) no es una expresión que se utilice en máxima entropía.
- D** Dada la muestra  $\mathcal{M} = \{(c_0, (f_0, f_1)), (c_0, (f_0, f_2)), (c_1, (f_1, f_2)), (c_1, (f_1, f_2))\}$ , donde  $c_i$ , ( $0 \leq i \leq 1$ ) son etiquetas de clase y  $f_i$ , ( $0 \leq i \leq 2$ ) son características, si se estima un modelo de clasificación por máxima entropía, entonces si tenemos una muestra por clasificar:
- A) que solo tiene una característica que es  $f_2$ , entonces se clasificará en la clase  $c_0$ .
  - B) que solo tiene dos características que son  $f_1$  y  $f_1$ , entonces se clasificará en la clase  $c_0$ .
  - C) que solo tiene dos características que son  $f_0$  y  $f_2$ , entonces se clasificará en la clase  $f_1$ .
  - D) que solo tiene una característica que es  $f_2$ , entonces se clasificará en la clase  $c_1$ .
- C** En el marco de la máxima entropía, la expresión  $\sum_{x,y} \tilde{p}(x) p_\lambda(y|x) f_i(x,y) = \sum_{x,y} \tilde{p}(x,y) f_i(x,y)$  representa:
- A) una desigualdad notable.
  - B) una restricción que debe satisfacer una sola característica.
  - C) una restricción que debe satisfacer la distribución que se estima.
  - D) una restricción que debe satisfacer la distribución empírica.
- C** Dada la expresión  $\delta_i = \frac{1}{M} \log \frac{\tilde{p}(f_i)}{p_\lambda(f_i)}$  utilizada para actualizar el valor  $\lambda_i$  asociado a la característica  $i$ -ésima en un modelo entrenado por máxima entropía,
- A) dicha expresión no se utiliza en el proceso de estimación por máxima entropía.
  - B) dicha expresión se obtiene después de resolver el problema de optimización *primal*, pero no el problema de optimización *dual*.
  - C) dicha expresión se obtiene después de resolver el problema de optimización *dual*.
  - D) dicha expresión se obtiene después de resolver el problema de optimización *dual*, y después el problema de optimización *dual*.
- B** Dada la expresión  $\delta_i = \frac{1}{M} \log \frac{\tilde{p}(f_i)}{p_\lambda(f_i)}$  utilizada para actualizar el valor  $\lambda_i$  asociado a la característica  $i$ -ésima en un modelo entrenado por máxima entropía:
- A) dicha expresión altera iterativamente el numerador y lo va aproximando al valor empírico del denominador.
  - B) dicha expresión altera iterativamente el denominador y lo va aproximando al valor empírico del numerador.
  - C) dicha expresión altera iterativamente el denominador y el numerador hasta que valen 1,0.
  - D) dicha expresión modifica  $M$  hasta anular la parte derecha de la igualdad.
- D** Sea un problema de clasificación en 3 clases  $A$ ,  $B$  y  $C$  tal que la clasificación se realiza a partir de 2 características  $c_0$  y  $c_1$ . Se dispone de un modelo entrenado por máxima entropía cuyas características son del tipo:

$$f(x,y) = \begin{cases} 1 & \text{si } y = S \text{ y la característica } c_j \text{ está presente en } x \\ 0 & \text{en otro caso} \end{cases}$$

donde  $S \in \{A, B, C\}$ .

Suponiendo que  $\lambda_{A,c_0} = \lambda_{B,c_0} = \lambda_{C,c_0} = 3$ ,  $\lambda_{A,c_1} = \lambda_{B,c_1} = \lambda_{C,c_1} = -4$  indica cuál sería la clase en la que se clasificaría una muestra que tuviese las características  $c_0$  y  $c_1$ .

- A) En A.
- B) En B.
- C) En C.
- D) En cualquiera de ellas.

**A** En el ejercicio anterior, si las probabilidades *a priori* de  $A$ ,  $B$  y  $C$  fuesen, respectivamente, 0,3, 0,1 y 0,6, entonces la nueva clase en la que se clasificaría la muestra sería

- A) la A.
- B) la B.
- C) la C.
- D) diferente de clase de la pregunta anterior.

**C** Dada la expresión  $\delta_{A,c_0} = \frac{1}{M} \log \frac{\tilde{p}(f_{A,c_0})}{p_\lambda(f_{A,c_0})}$  utilizada para actualizar el valor  $\lambda_{A,c_0}$  en el ejercicio que aparece dos ejercicios más arriba:

- A) Si  $\tilde{p}(f_{A,c_0}) > p_\lambda(f_{A,c_0})$ , entonces se está penalizando la clasificación en la clase  $A$  de las muestras que tengan las característica  $c_0$ .
- B) Se favorece siempre la clasificación de cualquier muestra en la clase  $A$ .
- C) Si  $\tilde{p}(f_{A,c_0}) > p_\lambda(f_{A,c_0})$ , entonces se está favoreciendo la clasificación en la clase  $A$  de las muestras que tengan las característica  $c_0$ .
- D) Se penaliza siempre la clasificación de cualquier muestra en la clase  $A$ .

**D** Sea el siguiente conjunto de cadenas:  $\{aaba, abbbba, aaabba\}$ . Si estimamos un 3-grama con esta muestra entonces tenemos que

- A)  $P(b|aa) = 1,0$ .
- B)  $P(b|aa) = 0,5$ .
- C)  $P(b|aa) = 1/3$ .
- D)  $P(b|aa) = 2/3$ .

**D** En traducción estadística, el problema de la búsqueda con un modelo log-lineal con  $K$  características utiliza la siguiente expresión:

- A)  $\hat{y} = \arg \max_y \sum_{k=1}^K \lambda_k h_k(x|y)$ .
- B)  $\hat{y} = \arg \max_y \sum_{k=1}^K \lambda_k \log h_k(x|y)$ .
- C)  $\hat{y} = \arg \max_y \sum_{k=1}^K \log h_k(x, y)$ .
- D)  $\hat{y} = \arg \max_y \sum_{k=1}^K \lambda_k h_k(x, y)$ .

**A** Dada la frase de referencia “*éramos dos antiguos amigos*” y la frase “*éramos los antiguos amigos*” producida por un sistema de traducción estadística, y suponiendo que  $BP = 1$ , y  $w_n$  es equiprobable, el  $BLEU = BP \exp \left( \sum_{n=1}^N w_n \log P_n \right)$  con precisión de  $n$ -gramas hasta  $n = 2$  es:

- A) 0,50.
- B) 0,20.
- C) 0,40.
- D) 0,70.

**B** Supongamos que dos sistema de traducción traducen un frase de entrada y cada unos de ellos produce una cadena de salida. Ambas salidas se evalúan con el BLEU con precisión de  $n$ -gramas hasta  $n = 1$ . En ambos casos se obtiene un BLEU igual a 1,0. Eso significa

- A) Que los dos sistemas traducen perfectamente.
- B) Que los dos sistemas han generado las mismas palabras que la frase de referencia.
- C) Que los dos sistemas han generado las mismas palabras que la frase de referencia y en el mismo orden.
- D) Ninguna de las anteriores.