Test de Sistema Inteligentes - MUIINF

ETSINF, Universitat Politècnica de València, 13 de Junio de 2017

Apellido: **Nombre:**

Cuestiones(60 minutos, sin apuntes)

Marca cada recuadro con una única opción entre las dadas.

C En el marco de la máxima entropía, la expresión $\widetilde{p}(f) = \sum_{x,y} \widetilde{p}(x,y) f(x,y)$ representa

A) Una restricción.

B) Una distribución condicional de una característica.

C) El valor esperado de una característica de acuerdo con una distribución empírica.

D) No es una expresion que se utilice en máxima entropía.

C Dada la expresión $\delta_i = \frac{1}{M} \log \frac{\widetilde{p}(f_i)}{p_\lambda(f_i)}$ utilizada para actualizar el valor λ_i asociado a la característica i-ésima en un modelo entrado por máxima entropía:

A) Dicho valor es $1/M$ cuando el valor esperado de la característica i-ésima con la distribución emprírica es 0.
B) Dicho valor es $1/M$ cuando el valor esperado de la característica i-ésima con la distribución emprírica coincide con el valor esperado de la característica de acuerdo con la distribución $\widetilde{p}(x)p_\lambda(y|x)$.
C) Dicho valor puede ser $1/M$.
D) Ninguna de las anteriores.

D En el marco de la máxima entropía, la expresión $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(x,y) f_i(x,y)$ se puede expresar también como:

A) $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(x|y) f_i(x,y)$.
B) $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(y|x) f_i(x,y)$.
C) $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(y)\widetilde{p}(y|x) f_i(x,y)$.
D) $\widetilde{p}(f_i) = \frac{1}{|\mathcal{M}|} \sum_{x,y} N((x,y),\mathcal{M}) f_i(x,y)$, donde $N((x,y),\mathcal{M})$ representa el número de veces que el par (x,y) ha aparecido en la muestra $\mathcal{M}$, y $|\mathcal{M}|$ es la talla de dicha muestra.

B Sea un problema de clasificación en 3 clases A, B y C tal que la clasificación se realiza a partir de 2 caracterísSticas 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 $\lambda_{A,c_0} = \lambda_{B,c_0} = \lambda_{C,c_0} = 3, \lambda_{A,c_1} = \lambda_{B,c_1} = \lambda_{C,c_1} = -3$ 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 cualquiera de ellas.

C) En A o B.

D) En C.

A,B En el ejercicio anterior, indica cuál de los siguientes cambios en los *lambdas* provocaría que una muestra con las características c_0 y c_1 se clasificase en la clase A.

A) $\lambda_{B,c_1} = \lambda_{C,c_1} = -4$ y el resto como están.
B) $\lambda_{A,c_1} = -2$ y el resto como están.
C) $\lambda_{A,c_0} = -3$ y el resto como están.
D) $\lambda_{B,c_0} = \lambda_{C,c_0} = 4$ y el resto como están.

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

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

C Supongamos que tenemos un modelo de lenguaje de 1-gramas con un vocabulario compuesto por n palabras donde todos los parámetros son equiprobables. La probabilidad de la cadena *"a b c d"* será:

A) $1/n$.
B) n.
C) $1/n^4$.
D) $4/n$.

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

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

A En la aproximación inversa a la traducción estadística, el modelo de lenguaje

A) Se aprende a partir de cadenas en la lengua destino.
B) Se aprende a partir de cadenas en la lengua origen.
C) Se aprende con pares de cadenas de ambas lenguas.
D) Se define manualmente.

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

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

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 $\text{BLEU} = \text{BP} \exp\left(\sum_{n=1}^{N} w_n \log P_n\right)$ 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 evaluan 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 refenrencia.
C) Que los dos sistemas han generado las mismas palabras que la frase de refenrencia y en el mismo orden.
D) Ninguna de las anteriores.