Test de Sistema Inteligentes - MUIINF

ETSINF, Universitat Politècnica de València, 14 de Junio de 2016

Apellido: | | **Nombre:** | |

Cuestiones(60 minutos, sin apuntes)

Marca cada recuadro con una única opción entre las dadas.

$\boxed{\text{B}}$ En el marco de la máxima entropía, la función

$$p(y|x) = \frac{1}{Z(x)} \exp(\sum_i \lambda_i f_i(x,y))$$

donde $Z(x) = \sum_y \exp(\sum_i \lambda_i f_i(x,y))$ es:

A) la función obtenida para clasificar una vez se deriva la función a optimizar respecto a los multiplicadores de Lagrange.

B) la función obtenida para clasificar una vez se deriva la función a optimizar respecto a las probabilidades *a posteriori* del modelo.

C) la función a optimizar.

D) la función obtenida para clasificar una vez se deriva la función a optimizar respecto a las restriciones.

$\boxed{\text{D}}$ En el marco de la máxima entropía, la expresión $\delta_i = \frac{1}{M} \log \frac{\tilde{p}(f_i)}{p_\lambda(f_i)}$

A) se utiliza en para clasificar una muestra según la expresión $p(y|x)$
B) se utiliza en para clasificar una muestra según la expresión $p(y,x)$
C) se utiliza en el algoritmo de aprendizaje IIS para optimizar las características $f_i(y,x)$.
D) se utiliza en el algoritmo de aprendizaje IIS para optimizar los valores λ asociados a las características

$\boxed{\text{C}}$ En el marco de la máxima entropía, la expresión $\tilde{p}(f_i) = \sum_{x,y} \tilde{p}(x,y) f_i(x,y)$ representa:

A) el valor esperado de la distribución empírica $\tilde{p}(x,y)$ de acuerdo con la característica $f_i(x,y)$.
B) el producto esperado de $f_i(x,y)$ y $\tilde{p}(x,y)$.
C) el valor esperado de la característica $f_i(x,y)$ de acuerdo con la distribución empírica $\tilde{p}(x,y)$
D) el valor normalizado de la característica $f_i(x,y)$.

$\boxed{\text{C}}$ En el algoritmo IIS el incremento δ_i a aplicar a cada λ_i en cada iteración es función de los valores:

A) $\tilde{p}(f_i) = \sum_{x,y} \tilde{p}(x|y) f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} p_\lambda(y|x) f_i(x,y)$.

B) $\tilde{p}(f_i) = \sum_{x,y} \tilde{p}(x|y) f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} \tilde{p}(x) p_\lambda(y|x) f_i(x,y)$.

C) $\tilde{p}(f_i) = \sum_{x,y} \tilde{p}(x)\tilde{p}(y|x) f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} \tilde{p}(x) p_\lambda(y|x) f_i(x,y)$.

D) $\tilde{p}(f_i) = \sum_{x,y} \tilde{p}(x,y) f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} \tilde{p}(x) p_\lambda(x,y) f_i(x,y)$.

$\boxed{\text{C}}$ Sea un problema de clasificación en cuatro clases A, B, C y D tal que la clasificación se realiza a partir de 3 características c_0, c_1 y c_2. Se dispone de un modelo entrenado por Máxima Entropía cuyas características son del tipo:

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

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

Suponiendo que $\lambda_{A,c_0} = \lambda_{C,c_1} = 1, \lambda_{B,c_1} = \lambda_{D,c_2} = -1$ y el resto de valores λ son 0, indica cuál sería la clase en la que se clasificaría una muestra que tuviese las características c_1 y c_2.

A) En A.

B) En B.

C) En C.

D) En D.

$\boxed{\text{A}}$ En el marco de la máxima entropía, los valores $f_i(x,y)$:

A) Son siempre enteros.
B) Son siempre reales.
C) Son siempre vectores de valores reales.
D) Nunca toman valores nulos.

$\boxed{\text{A}}$ En la aproximación inversa a traducción estadística mediante la expresión $\widehat{e} = \arg\max_e P(e)P(f|e)$

A) $P(e)$ representa la probabilidad *a priori* de la traducción e de la cadena f y se denomina modelo de lenguaje.
B) $P(f|e)$ representa la probabilidad *a posteriori* de la traducción e de la cadena f y se denomina modelo de lenguaje.
C) $P(e)$ representa la probabilidad *a priori* del modelo de traducción.
D) Ninguna de las anteriores

$\boxed{\text{A}}$ Con un modelo de lenguaje de n-gramas la probabilidad de una cadena y se aproxima como:

A) $P(y) = P(y_1) \prod_{i=2}^{|y|} P(y_i|y_{i-n+1}..y_{i-1})$.
B) $P(y) = P(y_1) \prod_{i=2}^{|y|} P(y_i, y_{i-n+1}..y_{i-1})$.
C) $P(y) = P(y_1) \prod_{i=2}^{|y|} P(y_i|y_1..y_{i-1})$.
D) $P(y) = P(y_1) \prod_{i=2}^{|y|} P(y_{i-n+1})$.

$\boxed{\text{C}}$ En traducción estadística, el modelo de lenguaje

A) Se aprende a partir de pares de entrada (e, f), donde e es una frase en la lengua origen y f es su traducción en la lengua destino.
B) Se aprende a partir de cadenas en la lengua origen.
C) Se aprende a partir de cadenas en la lengua destino.
D) Se define manualmente.

$\boxed{\text{D}}$ En traducción estadística, el problema de la búsqueda con un modelo log-lineal 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) $\widehat{y} = \arg\max_y \sum_{k=1}^{K} \lambda_k h_k(x, y)$.

$\boxed{\text{A}}$ Dada la frase de referencia "*éramos dos antiguos amigos*" y la frase "*éramos dos amigos antiguos*" 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) 0,42.
C) 0,61.
D) 0,70.

$\boxed{\text{B}}$ El paquete de traducción estadística GIZA

A) Permite aprender el modelo de lenguaje utilizado para traducción.
B) Permite aprender los modelos de alineamiento de los modelos de IBM.
C) Permite aprender los pesos del modelo log-lineal de traducción.
D) Permite aprender los modelos de alineamiento de los modelos de IBM y los pesos del modelo log-lineal de traducción.