Test de Sistema Inteligentes - MUIINF - Recuperación

ETSINF, Universitat Politècnica de València, 23 de Junio de 2015

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

Cuestiones(60 minutos, sin apuntes)

Marca cada recuadro con una única opción entre las dadas.

$\boxed{\text{C}}$ En el marco de la máxima entropía, la función que se optimiza es:

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

$\boxed{\text{B}}$ En el marco de la máxima entropía, la optimización de las funciones de distribución de probabilidad condicional $p(y|x) = \frac{1}{Z(x)}\exp(\sum_i \lambda_i f_i(x,y))$ donde $Z(x) = \sum_y \exp(\sum_i \lambda_i f_i(x,y))$

A) da como resultado la optimización del espacio de búsqueda de soluciones.
B) da como resultado la optimización de la entropia condicional.
C) da como resultado la optimización de las restriciones del problema.
D) da como resultado la optimización de las características.

$\boxed{\text{C}}$ En el marco de la máxima entropía, el algoritmo IIS:

A) Se utiliza para estimar los valores esperados empírico de las características.
C) No se utiliza en el marco de máxima entropia.
B) Se utiliza para estimar los *multiplicadores* de Lagrange.
D) Se utiliza para ajustar las muestras de aprendizaje.

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

A) $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(x,y)f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} \widetilde{p}(x)p_\lambda(y|x)f_i(x,y)$.
B) $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(x)f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} p_\lambda(y|x)f_i(x,y)$.
C) $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(y|x)f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} \widetilde{p}(x)p_\lambda(y|x)f_i(x,y)$.
D) $\widetilde{p}(f_i) = \sum_{x,y} \widetilde{p}(x,y)f_i(x,y)$ y $p_\lambda(f_i) = \sum_{x,y} \widetilde{p}(x)p_\lambda(x,y)f_i(x,y)$.

$\boxed{\text{B}}$ 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 $\lambda_{A,c_0} = \lambda_{C,c_2} = 1, \lambda_{B,c_1} = \lambda_{D,c_2} = -1$ 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_0 y c_2.

A) En B.
B) En A o C.
C) En D.
D) En A o B.

$\boxed{\text{B}}$ En el marco de la máxima entropía, los valores λ:

A) Son siempre enteros.
B) Son siempre reales.
C) Pueden tomar valores imaginarios.
D) Nunca toman valores nulos.

$\boxed{\text{C}}$ Para abordar el problema de la traducción estadística mediante la aproximación inversa, la expresión utilizada es:

A) $\widehat{e} = \arg \max_e P(e)P(f|e)$ donde $P(e)$ representa la probabilidad del modelo de alineamiento.
B) $\widehat{e} = \arg \max_e P(e)P(f|e)$ donde $P(e)$ representa la probabilidad del modelo de traducción.
C) $\widehat{e} = \arg \max_e P(e)P(f|e)$ donde $P(e)$ representa la probabilidad del modelo de lenguaje.
D) $\widehat{e} = \arg \max_e P(e)P(f|e)$ donde $P(e)$ representa la probabilidad del modelo de búsqueda.

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

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

A) Amplia el espacio de búsqueda.
B) Se combina con el modelo de traducción para ampliar el espacio de búsqueda.
C) Restringe el espacio de búsqueda.
D) No aporta ninguna información.

$\boxed{\text{D}}$ 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) $\widehat{y} = \arg \max_y \sum_{k=1}^{K} \lambda_k h_k(x, y)$.

$\boxed{\text{C}}$ Dada la frase de referencia "*tenía dos amigas cerca*" y la frase "*había dos amigas dos*" 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) 0,58.
B) 0,52.
C) 0,41.
D) 0,60.

$\boxed{\text{B}}$ 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.