Test de Sistema Inteligentes - MUIINF

ETSINF, Universitat Politècnica de València, Junio de 2015

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

Cuestiones(60 minutos, sin apuntes)

Marca cada recuadro con una única opción entre las dadas.

$\boxed{\text{A}}$ En el marco de la máxima entropía:

A) Se pueden abordar problemas de clasificación de reconocimiento de formas.
B) Sólo se pueden abordar problemas de clasificación en dos clases.
C) No se pueden abordar problemas de clasificación de reconocimiento de formas.
D) Ninguna de las anteriores.

$\boxed{\text{D}}$ En el marco de la máxima entropía, las funciones de distribución de probabilidad condicional son de la forma:

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

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

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

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

A) $\delta_i = \log \frac{\widetilde{p}(x,y)}{p_\lambda(x,y)}$.
B) $\delta_i = \frac{1}{M} \log \frac{\widetilde{p}(x,y)}{p_\lambda(x,y)}$ donde $M = f^\#(x,y)$.
C) $\delta_i = \frac{1}{M} \log \frac{\widetilde{p}(f_i)}{p_\lambda(f_i)}$ donde $M = f^\#(x,y)$.
D) $\delta_i = \log \frac{\widetilde{p}(f_i)}{p_\lambda(f_i)}$.

$\boxed{\text{C}}$ 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 todos los valores λ asociado a cada característica 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 la clase A.

B) En la clase con mayor probabilidad *a priori*.

C) En cualquiera de las clases puesto que la probabilidad a *a posteriori* es la misma para todas la clases.

D) No se puede clasificar.

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

A) Son siempre positivos.
B) Son siempre negativos.
C) Pueden tomar valores positivos, nulos o negativos.
D) Nunca toman valores nulos.

B C D Para abordar el problema de la traducción estadística mediante la aproximación directa, la expresión utilizada es:

A) $\widehat{e} = \arg\max_e P(e)P(f|e)$.
B) $\widehat{e} = \arg\max_e P(f)P(e|f)$.
C) $\widehat{e} = \arg\max_e P(e|f)$.
D) $\widehat{e} = \arg\max_e P(f, e)$.

D Con un modelo de lenguaje de n-gramas la probabilidad de una cadena y se aproxima como:

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

B En traducción estadística, un alineamiento

A) es una función que permite predecir una palabra dada las n anteriores.
B) es una función que relaciona palabras de las cadenas de entrada y salida.
C) es una función que relaciona una palabra de la cadena de entrada con la cadena vacía en la cadena de salida.
D) es una función que relaciona segmentos de palabras de la cadena de entrada con palabras de la cadena de salida.

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

A El modelo 1 de IBM de traducción estadística:

A) Se utiliza para obtener alineamientos entre palabras de las cadenas de entrada y salida.
B) Obtiene un alineamiento entre palabras que siempre es simétrico, es decir, que el alineamiento de x contra y es igual que el de y contra x.
C) Es un modelo que no está definido.
D) Es el modelo de traducción más complejo de los modelos de IBM.

B En traducción estadística, el BLEU se define como:

A) $\text{BLEU} = \exp\left(\sum_{n=1}^{N} w_n \log P_n\right)$
B) $\text{BLEU} = \text{BP}\exp\left(\sum_{n=1}^{N} w_n \log P_n\right)$
C) $\text{BLEU} = \text{BP}\exp\left(\prod_{n=1}^{N} w_n \log P_n\right)$
D) $\text{BLEU} = \exp\left(\sum_{n=1}^{N} w_n P_n\right)$

A Dada la frase de referencia "*la casa hoy*" y la frase "*se casa hoy*" producida por un sistema de traducción estadística, y suponiendo que $\text{BP} = 1$, y w_n es equiprobable, el BLEU con precisión de n-gramas hasta $n = 2$ es:

A) 0,58.
B) 0,20.
C) 0,50.
D) 0,90.

C El paquete de traducción estadística MOSES

A) Solo se utiliza para aprender un modelo de lenguaje.
B) Es una paquete que solo se utiliza para traducción estadística basada en palabras.
C) Permite aprender la tablas de traducción de un sistema de traducción estadística.
D) Ninguna de la anteriores.