

Test de Sistemas Inteligentes - Bloque 2 - MUIINF

ETSINF, Universitat Politècnica de València, 11 de Junio de 2019

Apellidos:

Nombre:

Cuestiones (60 minutos, sin apuntes)

Escribe en cada recuadro la opción correcta entre las dadas.

- ☐ C En el marco de la máxima entropía, si el valor esperado de una característica (*feature*) f es 0 ($\tilde{p}(f) = \sum_{x,y} \tilde{p}(x,y)f(x,y) = 0$), entonces significa que:
- A) no coincide con su valor esperado de acuerdo con la expresión $\sum_{x,y} \tilde{p}(x)p_{\lambda}(y|x)f_i(x,y)$.
 - B) $\tilde{p}(x,y)$ es un valor no nulo.
 - C) la característica no ha aparecido en la muestra.
 - D) $\tilde{p}(x,y)$ es un valor negativo.
- ☐ C En el marco de la máxima entropía, si una característica aparece el mismo número de veces asociado a dos clases, entonces los valores λ asociados a dicha característica para ambas clases
- A) serán idénticos.
 - B) serán nulos para ambas clases.
 - C) pueden ser diferentes.
 - D) ninguna de la anteriores.
- ☐ A Dada la muestra $\mathcal{M} = \{(c_0, (f_0, f_1)), (c_0, (f_0, f_2)), (c_1, (f_3, f_1)), (c_1, (f_3, f_2))\}$, donde c_i , ($0 \leq i \leq 1$) son etiquetas de clase y f_i , ($0 \leq i \leq 3$) 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 o c_1 indistintamente.
 - B) que solo tiene dos características que son f_1 y f_2 , entonces se clasificará en la clase c_0 .
 - C) que solo tiene dos características que son f_0 y f_2 , entonces $p(c_0|(f_0, f_2)) = 1,0$ y se clasificará en la clase c_0 .
 - D) siempre se clasificará en la clase c_0 .
- ☐ B Dada la expresión $\delta_i = \frac{1}{M} \log \frac{\tilde{p}(f_i)}{p_{\lambda}(f_i)}$ utilizada para actualizar el valor λ_i asociado a la característica i -ésima en un modelo entrenado por máxima entropía con el algoritmo IIS,
- A) dicha expresión no se utiliza en el proceso de estimación por máxima entropía.
 - B) el valor $\tilde{p}(f_i)$ es constante en todas la iteraciones.
 - C) M es una constante universal que vale e .
 - D) $p_{\lambda}(f_i)$ solo es necesario calcularlo en la primera iteración del algoritmo.
- ☐ D Dada la expresión $\delta_i = \frac{1}{M} \log \frac{\tilde{p}(f_i)}{p_{\lambda}(f_i)}$ utilizada para actualizar el valor λ_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 el numerador hasta que valen 1,0.
 - C) dicha expresión modifica M hasta anular la parte derecha de la igualdad.
 - D) dicha expresión altera iterativamente el denominador y lo va aproximando al valor empírico del numerador.
- ☐ A 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 $\lambda_{A,c_0} = \lambda_{B,c_0} = \lambda_{C,c_0} = \lambda_{A,c_1} = \lambda_{B,c_1} = \lambda_{C,c_1} = 1$ indica cuál sería la clase en la que se clasificaría una muestra que tuviese las características c_0 y c_1 .
- A) En cualquiera de ellas.
 - B) Dependería de la probabilidad a priori de la clase.
 - C) En A o B, pero nunca en C.
 - D) En C.
- ☐ A En el ejercicio anterior, si las probabilidades *a priori* de A, B y C fuesen, respectivamente, 0,3, 0,1 y 0,6, y todos los valores $\lambda = 0$, entonces la nueva clase en la que se clasificaría la muestra sería
- A) en cualquiera de ellas.
 - B) en la clase C.
 - C) en A o B, pero no en C.
 - D) no se puede clasificar.

B,C Supongamos que tenemos un modelo de lenguaje de 3-gramas con un vocabulario compuesto por n palabras donde todos los parámetros son equiprobables. Si no se utiliza ningún tipo de suavizado, entonces la probabilidad de la cadena “ $a b c d$ ” será:

- A) $1/n^3$. B) $1/n^2$. C) $1/n^4$. D) $4/n$.

C Sea el siguiente conjunto de cadenas: $\{a a b a, b b b b a, a a a b b a\}$. Si estimamos un 3-grama con estas muestras entonces tenemos que:

- A) $P(a|bb) = 9/16$. B). $P(b|bb) = 1/3$. C) $P(a|bb) = P(b|bb)$. D) $P(b|bb) = 2/3$.

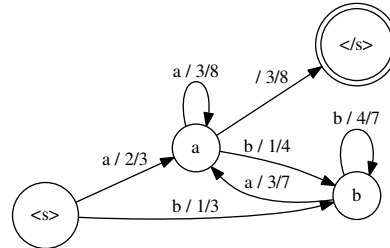
D Con la muestra de la pregunta anterior hemos generado un 2-grama y lo hemos transformado en un autómata estocástico sin suavizar como el que puede verse bajo. Dicho automata: **(ANULADA)**

A) es incorrecto porque la probabilidad de la arista $a \rightarrow b$ debería ser mayor que 0,5.

B) le sobra la arista $b \xrightarrow{a/1/2} a$

C) le falta la arista $ab \xrightarrow{b/1/2} bb$

D) es correcto.



B En el PFA de la pregunta anterior, el análisis de la cadena “ $a a a b b b a$ ” daría una probabilidad:

- A) nula porque no aparece en la muestra de aprendizaje. B) $3^2/2^{13}$. C) 0.05. D) 1.0.

B El PFA que aparece arriba: **(ANULADA)**

A) no puede aceptar cadenas de longitud mayor que 1 que empiecen con “ b ”.

B) no puede aceptar cadenas que contengan la subcadena “ $a b a b$ ”.

C) puede aceptar la cadena “ b ”.

D) puede aceptar cadenas que contengan solo b 's.

D En el PFA de la pregunta anterior, el análisis de la cadena “ $a a a b c b b$ ”:

A) lo realizaría correctamente.

B) asignaría una probabilidad no nula a la cadena pero muy pequeña.

C) no es posible si no se utiliza un 1-grama.

D) exigiría introducir un símbolo (por ejemplo $\langle \text{unk} \rangle$) y las correspondientes transiciones para aglutinar todos los términos fuera del vocabulario.

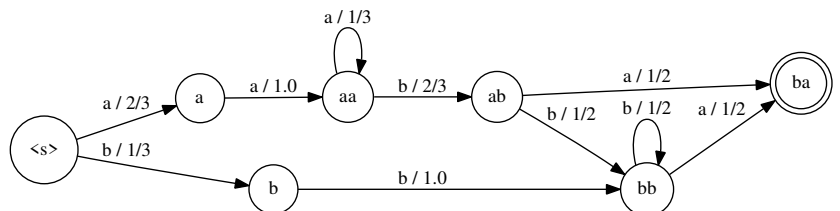
C Con la muestra $\{a a b a, b b b b a, a a a b b a\}$ se ha generado un 3-grama y se ha transformado en un autómata estocástico sin suavizar como el que puede verse bajo. El análisis de la cadena “ $a a a b b b a$ ” daría una probabilidad:

A) igual que la que le asigna el 2-grama.

B) no se puede analizar con este autómata.

C) mayor que la que le asigna el 2-grama.

D) 0.19



C Si a la muestra de la pregunta anterior le añadimos la cadena “ $b b b$ ”, el autómata estocástico sin suavizar obtenido con un 3-grama:

A) cambian únicamente las probabilidades.

B) no cambia.

C) cambian sus transiciones y sus probabilidades.

D) cambian únicamente las transiciones.