

Examen de Traducción Automática

MIARFID, Universitat Politècnica de València, 10 de febrero de 2020

Apellidos:

Nombre:

Cuestiones

Marca cada recuadro con una única opción de entre las dadas. Cada 3 errores restan una respuesta correcta. Las respuestas en blanco ni suman ni restan. Tiempo de duración del examen: 1 hora.

- 1 ☒ A Supongamos que se dispone de un modelo de traducción estadístico basado en el modelo M1 con un vocabulario de 5 palabras fuente (E1, E2, E3, E4 y E5) y 4 palabras destino (F1, F2, F3 y F4) y que viene dado por la siguiente tabla de probabilidades $p(y | x)$ donde $y \in \{F1, F2, F3, F4\}$ y $x \in \{E0, E1, E2, E3, E4, E5\}$ (E0 es la palabra nula):

M1	E0	E1	E2	E3	E4	E5
F1	0.2	0.1	0.4	0.0	0.5	0.3
F2	0.2	0.2	0.3	0.0	0.5	0.0
F3	0.3	0.3	0.2	0.5	0.0	0.7
F4	0.3	0.4	0.1	0.5	0.0	0.0

¿Cuál es la probabilidad $p(y | x)$ de que la frase $x = \text{"E1 E5"}$ se traduzca como $y = \text{"F1 F3"}$ según el modelo M1?. Si $I = |y|$ y $J = |x|$, asumir que $p(I | x) = 1$ si $I = J$ y 0 en otro caso.

- A) 0.087
B) 0.78
C) 0.333
D) 0.0
- 2 ☒ D En la traducción estadística basada en palabras, indicar qué afirmación es incorrecta.
- A) El modelo M2 y HMM tienen en común el diccionario estadístico.
B) El modelo M2 es de orden cero y el HMM de primer orden.
C) El coste de calcular $\Pr(y | x)$ es $O(|y| \cdot |x|)$ con el modelo M2 y de $O(|y| \cdot |x|^2)$ con el modelo HMM.
D) El coste de calcular $\Pr(y | x)$ es $O(|y| \cdot |x|)$ con el modelo HMM y de $O(|y| \cdot |x|^2)$ con el modelo M2.
- 3 ☒ D En el entrenamiento de los modelos estadísticos basados en segmentos o "phrases", indicar qué afirmación es correcta:
- A) Los pesos del modelo log-lineal se estiman al mismo tiempo que el resto de parámetros.
B) Para extraer los segmentos o "phrases" hay que aplicar el algoritmo MERT.
C) El modelo de lenguaje solo se puede aprender a partir del conjunto de datos bilingües que se utiliza para obtener los segmentos of "phrases".
D) La aplicación de los algoritmos MERT o MIRA para aprender los pesos del modelo log-lineal permite mejorar el resultado de la traducción con respecto a no aplicarlos.
- 4 ☒ B En el proceso de traducción con modelos basados en "phrases" mediante el algoritmo "stack decoding" indicar qué afirmación es incorrecta:
- A) Hay tantas pilas como tamaños de los conjuntos de cobertura de las palabras de entrada ya traducidas.
B) En general, se van expandiendo los prefijos de la entrada.
C) En cada iteración, se selecciona la mejor hipótesis de cada pila.
D) Se van expandiendo los prefijos de la salida.
- 5 ☒ A Indicar qué afirmación es falsa, en general, con respecto a la generación de "sentence-embeddings":
- A) Siempre se obtiene el embedding de una oración mediante el conteo de las palabras que forma parte de la oración.
B) Se puede obtener el embedding de una oración mediante la suma de los embedding de las palabras que la componen
C) Se puede obtener el embedding de una oración mediante la media de los embedding de las palabras que la componen
D) Se puede utilizar una red recurrente y el sentence-embeddings es el estado final de la red recurrente.

- 6 ☒ A En el aprendizaje de redes neuronales simples indicar qué afirmación es correcta:
- A) Si se pretenden implementar distribuciones de probabilidad en la salida, es conveniente utilizar la función softmax para las capas de salidas y la minimización de la entropía cruzada en el entrenamiento.
 - B) Si se pretenden implementar distribuciones de probabilidad en la salida, es necesario utilizar la función sigmoid para las capas de salidas y la minimización de la entropía cruzada en el entrenamiento.
 - C) Si se pretenden implementar distribuciones de probabilidad en la salida, es necesario utilizar la función relu para las capas de salidas y la minimización de la entropía cruzada en el entrenamiento.
 - D) Si se pretenden implementar distribuciones de probabilidad en la salida, es necesario utilizar la función sigmoid para las capas de salidas y la minimización del error cuadrático medio en el entrenamiento.
- 7 ☒ C Indicar qué afirmación es correcta cuando se utilizan “Long Short-Term Memories” en la traducción neuronal:
- A) Las “Long Short-Term Memories” permiten aprender los modelos de forma mas rápida que con otros modelos.
 - B) Las “Long Short-Term Memories” permiten aprender los modelos sin necesidad de usar el algoritmo “back-propagation through time algorithm”
 - C) Las “Long Short-Term Memories” permiten dar cuenta de parte de la información pasada.
 - D) Las “Long Short-Term Memories” permiten aprender los modelos usando dependencias de segundo orden
- 8 ☒ A Los modelos de traducción neuronal basados en “Transformer” están considerados como el estado del arte en este área, indicar qué afirmación es correcta:
- A) Es un modelo no recurrente auto-regresivo basado en modelos de auto-atención entre las palabras de entrada, entre las palabras de salida y un modelo atención cruzada entre las palabras de entrada y salida.
 - B) Es un modelo no recurrente auto-regresivo basado solo en modelos de auto-atención entre las palabras de entrada y entre las palabras de salida.
 - C) Es un modelo recurrente auto-regresivo basado solo en modelos de atención cruzada entre las palabras de entrada y salida.
 - D) Es un modelo recurrente auto-regresivo basado en modelos de auto-atención entre las palabras de entrada, entre las palabras de salida y un modelo atención cruzada entre las palabras de entrada y salida.
- 9 ☒ A En los modelos recurrentes de traducción neuronal basados en modelos de atención, indicar qué afirmación es correcta:
- A) Cada palabra de salida se genera a partir de la última palabra traducida, de una función de las palabras de entrada y de la historia.
 - B) Cada palabra de salida se genera en función solo de la última generada, de una palabra de entrada y de la historia.
 - C) Cada palabra de salida se genera en función solo de las palabras traducidas hasta ese momento.
 - D) Cada palabra de salida se genera en función solo de la palabras de entrada afectadas por un peso.
- 10 ☒ B En un sistema de traducción interactiva cada vez que el usuario corrige una palabras traducida, el sistema reacciona generando una nueva traducción. Indicar qué afirmación es falsa:
- A) El sistema de traducción devuelve la traducción de mayor probabilidad compatible con el prefijo validado por el usuario y por la palabra traducida corregida.
 - B) El sistema de traducción devuelve una traducción de mayor probabilidad que la previamente generada.
 - C) El sistema de traducción puede devolver una traducción de peor calidad que la previamente generada.
 - D) El sistema de traducción puede devolver una traducción de mayor TER que la previamente generada.
- 11 ☒ A La post-edición y la traducción interactiva son dos aproximaciones que permiten obtener traducciones de alta calidad y en menos tiempo que la traducción realizada solamente por humanos. Indicar qué afirmación de las siguientes es falsa:
- A) La post-edición permite obtener muestras bilingües para re-entrenar los modelos de traducción, pero la traducción interactiva no.
 - B) Tanto la post-edición como la traducción interactiva permiten re-entrenar los modelos de traducción con las traducciones validadas por el humano.
 - C) En la post-edición el sistema de traducción solo genera una sola traducción para cada oración de entrada mientras que en la traducción interactiva el sistema de traducción genera una hipótesis cada vez que el humano introduce una corrección.
 - D) Tanto los modelos de traducción neuronal como los púramente estadísticos pueden ser utilizados en la post-edición y en la traducción interactiva.

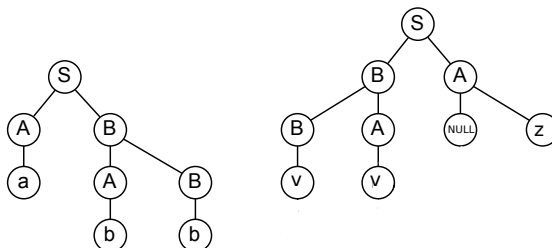
12 **A** Dada la SITG $G = \{(p/2, S \rightarrow [SS]), (1 - p, S \rightarrow \langle SS \rangle), (p/2, S \rightarrow a/x)\}$ indica la probabilidad de par (aaa, xx) :

- A) No es posible calcular dicha probabilidad.
- B) $p^3/8$
- C) $(1 - p/2) p^2/4$
- D) $(1 - p/2) p^3/8$

13 **B** Dadas las siguientes tablas de reordenamiento, inserción y traducción asociadas a un traductor “tree-to-string”:

<i>r</i> -table			<i>n</i> -table				<i>t</i> -table			
original order	reorder	P(reorder)	parent	S	A	B	w	p(ins-w)	a	b
A B	B A	1.0	node	A	A	A	z	0.6	z	0.6
B A	A B	0.6	P(None)	0.6	0.5	0.4	v	0.4	NULL	0.4
	B A	0.4	P(Left)	0.3	0.3	0.4			v	0.3
			P(right)	0.1	0.2	0.2			NULL	0.7

indica la probabilidad de convertir el árbol de la izquierda en el árbol de la derecha:



- A) 0.00216
- B) > 0.00216
- C) < 0.00216
- D) < 0.00316

14 **C** Dadas las tablas de reordenamiento, inserción y traducción asociadas al traductor “tree-to-string” de la pregunta anterior y el árbol de la izquierda, ¿es posible obtener la cadena “vzvzz”?:

- A) No es posible.
- B) Habría que cambiar las probabilidades.
- C) Si es posible, y se puede obtener con varios árboles.
- D) Si, y se puede obtener con un solo árbol.

15 **B** Sean las tablas de reordenamiento, inserción y traducción asociadas al traductor “tree-to-string” de la pregunta que aparece dos cuestiones más arriba, el árbol de la izquierda, y la cadena de salida “vzvzzz”. Si estimamos los parámetros del modelo con dicha cadena indica cuál de la siguientes opciones es correcta:

- A) Cambian todos los parámetros.
- B) No es posible realizar la estimación.
- C) Solo cambian los parámetros de la *n*-tabla.
- D) Solo cambian los parámetros de la *r*-tabla y de la *n*-tabla.