

Examen de Traducción Automática

MIARFID, Universitat Politècnica de València, 11 de febrero de 2022

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.

- 1 ☒ D En la traducción estadística basada en palabras, indicar qué afirmación es incorrecta cuando se define el denominado modelo 1.
- A) La distribución de probabilidad de las palabras en la lengua destino que son traducción de una palabra en el lenguaje fuente (diccionario estadístico) viene dada normalmente por una tabla.
 - B) La distribución de probabilidad de alinear una posición de la frase fuente con una posición dada de la traducción es equiprobable.
 - C) En la aproximación por Viterbi con el modelo 1 el alineamiento con una palabra de la traducción se obtiene buscando la palabra de la frase fuente que maximice la probabilidad del diccionario estadístico.
 - D) Los parámetros a estimar a partir de un corpus paralelo son los que definen el diccionario estadístico y la probabilidad de alineamiento.
- 2 ☒ A En un modelo de traducción basado en alineamientos estadísticos, indicar qué afirmación es falsa (I es el tamaño de la traducción y J es el tamaño de la frase a traducir):
- A) Si los alineamientos se definen con una función $a : \{1, \dots, I\} \rightarrow \{1, \dots, J\}$, el número de posibles alineamientos es lineal con I .
 - B) Si los alineamientos se definen con una relación $a \subseteq \{1, \dots, J\} \times \{1, \dots, I\}$, el número de posibles alineamientos es excesivamente grande.
 - C) Si los alineamientos se definen con una función $a : \{1, \dots, I\} \rightarrow \{1, \dots, J\}$, el número de posibles alineamientos es inferior asintóticamente del que se obtiene si se define como una relación.
 - D) Si los alineamientos se definen con una función $a : \{1, \dots, I\} \rightarrow \{1, \dots, J\}$, el número de posibles alineamientos es exponencial con I .
- 3 ☒ D En un sistema de traducción estadística basado en “phrases” y en un modelo log-lineal, indicar qué afirmación es correcta:
- A) En la construcción de segmentos bilingües (“phrases”) la técnica menos utilizada para obtener los alineamientos es la denominada “grow-diagonal”.
 - B) Solo hay que estimar los pesos del modelo log-lineal.
 - C) El número de segmentos en la frase fuente puede ser distinta del número de segmentos en la traducción.
 - D) Normalmente la asignación de probabilidades a los segmentos bilingües de “phrases” se obtiene por conteo a partir del conjunto de entrenamiento.
- 4 ☒ D En un sistema de traducción estadística basado en “phrases” y en un modelo log-lineal, indicar qué afirmación es correcta:
- A) La diferencia entre los alineamientos monótonos y no monótonos es que el número de segmento es diferente.
 - B) La diferencia entre los alineamientos monótonos y no monótonos es que la probabilidad que da el sistema de traducción monótono es siempre inferior al que da el sistema de traducción no monótono.
 - C) Una diferencia entre los alineamientos monótonos y no monótonos es que la probabilidad de alineamiento entre segmentos es equiprobable en la traducción no monótona y no lo es en la monótona.
 - D) Una diferencia entre los alineamientos monótonos y no monótonos es que la probabilidad de alineamiento entre segmentos es equiprobable en la traducción monótona y no lo es en la no monótona.
- 5 ☒ D En el proceso de inferencia con un sistema de traducción estadística basado en “phrases” y en un modelo log-lineal, indicar qué afirmación es correcta:
- A) La estrategia básica más utilizada está basada en recorrer la frase fuente y generar la palabra de mayor probabilidad de cada palabra fuente.
 - B) La estrategia básica más utilizada está basada en el uso de múltiples colas de prioridad o “stacks”, una por cada talla del conjunto de palabras traducidas en el lenguaje destino.
 - C) La información almacenada en las colas de prioridad o “stacks” solo contiene las palabras que ya han sido traducidas y la probabilidad que da el modelo de lenguaje de entrada para la parte de la frase fuente utilizada hasta ese momento.
 - D) La estrategia básica más utilizada está basada en el uso de múltiples colas de prioridad o “stacks”, una por cada talla del conjunto de palabras fuentes ya traducidas.

- 6 ☒ D Un traductor neuronal “encoder-decoder” contiene un número elevado de parámetros que hay que estimar, indicar qué afirmación es falsa:
- A) En general, todos los parámetros, incluidos los “word-embeddings” se estiman conjuntamente con el resto de parámetros.
 - B) En algunas partes del traductor hay que utilizar la función “softmax” para obtener distribuciones de probabilidad y el coste para su manipulación es elevado.
 - C) Los algoritmos de optimización más importante utilizados para el entrenamiento son: “Stochastic gradient descent with momentum”, “Adaptive Gradient”, “Adadelata” y “Adam”.
 - D) En algunas partes del traductor hay que utilizar la función de activación lineal para obtener distribuciones de probabilidad y su coste es elevado.
- 7 ☒ A El traductor neuronal “encoder-decoder” que mejores prestaciones está dando es el basado en un modelo de atención, indicar qué afirmación es falsa:
- A) El modelo de atención permite dar más peso a las palabras traducidas hasta ese momento de la frase destino para producir la siguiente palabra en la fases destino.
 - B) El modelo de atención permite dar más peso a las palabras de la frase fuente para producir una palabra en la fases destino.
 - C) De entre los modelos de traductores neuronales “encoder-decoder” basados en modelos de atención se encuentran los modelos recurrentes del tipo LSTM o GRUs y el Transformer.
 - D) En general, un traductor neuronal “encoder-decoder” basado en LSTMs posee un modelo de atención entre las representaciones de entrada y salida y el basado en Transformer posee además dos modelos de auto-atención.
- 8 ☒ A En un modelo de traducción neuronal basado en LSTMs y un modelo de atención, inidicar qué afirmación es cierta:
- A) El modelo de atención está basado en calcular el grado de semejanza entre un estado del LSTM asociado a una posición de la entrada y el estado del LSTM de salida en la posición anterior al de la palabra que se desea predecir.
 - B) Cada estado del LSTM de entrada hay que multiplicarlo por una matriz y aplicar una función “softmax” antes de aplicar el modelo de atención.
 - C) Los modelos de traducción neuronal basado en LSTMs bidireccionales están basados en dos LSTMs que analizan la frase fuente en los dos sentidos y en dos LSTMs que analizan la frase traducida también en los dos sentidos.
 - D) Los modelos de traducción neuronal basado en LSTMs bidireccionales están basados en un LSTM que analizan la frase fuente y en dos LSTMs que analizan la frase traducida también en los dos sentidos.
- 9 ☒ C En el proceso de inferencia o traducción utilizando un traductor neuronal “encoder-decoder” basado en un modelo de atención, indicar qué afirmación es falsa:
- A) Si se utiliza una estrategia voraz, la inferencia es más rápida que si se utiliza la búsqueda en haz pero de peor calidad en general.
 - B) Si se utiliza una estrategia de búsqueda en árbol sin poda se obtiene una solución óptima con respecto al modelo utilizado.
 - C) En general, si se utiliza una estrategia de búsqueda en árbol con la poda basada en la búsqueda en haz la inferencia es más rápida que si se utiliza la búsqueda voraz pero de mejor calidad.
 - D) Si en la estrategia de búsqueda en haz se restringe el haz a una hipótesis sobreviviente en cada itereación, entonces es equivalente a la búsqueda voraz.
- 10 ☒ C En un modelo de traducción neuronal basado en Transformer, indicar qué afirmación es cierta:
- A) El uso de modelos de atención con múltiples cabezas (“multi-head attention”) permite un entrenamiento más rápido del Transformer.
 - B) El uso de conexiones residuales en el Transformer permite una inferencia más rápida.
 - C) El decodificador del Transformer es normalmente autoregresivo pero el codificador no.
 - D) Normalmente las redes “feed-forward” utilizadas tanto en el codificador como en el decodificador del Transformer son de dos capas ocultas.

- 11 ☐ A En la comparación entre los traductores estadísticos y los neuronales, indicar qué afirmación es falsa.
- A) Los modelos estadísticos son más sensibles que los neuronales a la hora de obtener los mejores hiperparámetros de los modelos.
 - B) En general, si se tienen conjuntos de entrenamiento pequeños, los traductores basados en modelos estadísticos generan mejores traducciones que los neuronales.
 - C) En general, si se tienen conjuntos de entrenamiento muy grandes, los traductores basados en modelos neuronales generan mejores traducciones que los basados en modelos estadísticos.
 - D) Los modelos neuronales necesitan más recursos computacionales que los estadísticos.
- 12 ☐ A En un sistema de traducción interactiva (TI) indicar qué afirmación es falsa:
- A) La TI solo sirve para que el sistema busque todos los errores y el humano los corrija.
 - B) En la TI, el sistema puede sugerir en qué partes de la traducción el humano debe fijar su atención para validarlas o corregirlas.
 - C) La TI basada en múltiples segmentos se basa en una búsqueda restringida al espacio de segmentos de traducciones compatibles con el conjunto de segmentos validados por el humano.
 - D) La TI basada en prefijos se basa en una búsqueda restringida al espacio de sufijos compatibles con el prefijo validado por el humano.
- 13 ☐ A La post-edición y la traducción interactiva ofrecen un marco muy adecuado para mejorar los modelos de traducción utilizados, indicar qué afirmación es errónea:
- A) El usuario debe reentrenar los modelos de traducción utilizando conjuntos de entrenamiento distintos de los validados.
 - B) Los modelos se pueden adaptar automáticamente una vez que el usuario ha acabado de corregir/validar cada muestra de test o conjunto de ellas.
 - C) Los modelos se adaptan a las nuevas muestras producidas por el usuario mediante un proceso de aprendizaje “on-line”.
 - D) Tanto en un sistema estadístico como neuronal se pueden adaptar todos los parámetros o un subconjunto de ellos con las muestra recién validadas.
- 14 ☐ A En la traducción interactiva multimodal (TIM) indicar qué afirmación es falsa:
- A) En la TIM solo se puede utilizar una modalidad como el habla o el lápiz electrónico como alternativa al teclado y al ratón.
 - B) En la TIM se puede utilizar el habla, el lápiz electrónico u otros dispositivos como medio para validar y corregir partes de las traducciones generadas por el sistema de traducción.
 - C) Una de las aproximaciones a la TIM consiste en utilizar conocimiento bilingüe de la tarea para mejorar el reconocimiento del habla o de los gestos o de la escritura manuscrita.
 - D) Los errores introducidos por el uso de modalidades como el habla o el lápiz electrónico se pueden corregir con el teclado y el ratón.
- 15 ☐ B Una de las aplicaciones indirectas de la arquitectura “encoder-decoder” surgidas de la traducción automática es la introducción de modelos de lenguaje pre-entrenados. Indicar qué afirmación es correcta:
- A) Los modelos pre-entrenados basados en el codificador o en el decodificador del Transformer no necesitan ser ampliados o modificados en cada tarea en la que se desea aplicarlos.
 - B) Los modelos pre-entrenados basados en el codificador o en el decodificador del Transformer son grandes modelos de lenguaje pre-entrenados con grandes volúmenes de datos monolingües o multilingües.
 - C) Los modelos pre-entrenados basados en el codificador o en el decodificador del Transformer son grandes modelos de lenguaje pre-entrenados solo con grandes volúmenes de datos monolingües.
 - D) Los modelos pre-entrenados basados en el codificador o en el decodificador del Transformer solo se pueden aplicar a aplicaciones de textos.