

Comenzado el	sábado, 26 abril 2025, 9:48 PM
Estado	Finalizado
Finalizado en	sábado, 26 abril 2025, 10:18 PM
Tiempo empleado	30 minutos 20 segundos
Calificación	100,00 de 100,00

Pregunta 1

Correcta

Se puntúa 5,00 sobre 5,00

A pesar de la versatilidad de los LLMs, ¿cuál sigue siendo el enfoque recomendado para garantizar la robustez y calidad de las predicciones en tareas específicas como el análisis de sentimientos?

☐

a. Usar únicamente prompts sin ningún ajuste posterior.

☐

b. Confiar en el autoaprendizaje supervisado sin datos manuales.

☒

c. Tomar un LLM pre-entrenado y hacerle fine-tuning con un conjunto de datos etiquetado para esa tarea. ✓

☐

d. Entrenar un modelo pequeño desde cero para cada caso.

Respuesta correcta

La respuesta correcta es:

Tomar un LLM pre-entrenado y hacerle fine-tuning con un conjunto de datos etiquetado para esa tarea.

Pregunta 2

Correcta

Se puntúa 5,00 sobre 5,00

¿Para qué otros tipos de datos, además del texto, se pueden utilizar los Embeddings?

☐

a. Solo para imágenes en visión por computador.

☒

b. Para datos **discretos** como genes, jugadas de ajedrez o variables categóricas, donde interesa capturar relaciones latentes. ✓

☐

c. Únicamente para audio y vídeo.

☐

d. Para series temporales continuas exclusivamente.

Respuesta correcta

La respuesta correcta es:

Para datos **discretos** como genes, jugadas de ajedrez o variables categóricas, donde interesa capturar relaciones latentes.

Pregunta 3

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué hito o avance reciente en PNL se menciona como un indicador clave de esta revolución y por qué es significativo?

- ☒ a. Publicación de GPT-3 en mayo de 2020 ✓
- ☐ b. Lanzamiento de GPT-2 en 2019
- ☐ c. Publicación de BERT en 2018
- ☐ d. Presentación de LLaMA por Meta

Respuesta correcta

La respuesta correcta es:

Publicación de GPT-3 en mayo de 2020

Pregunta 4

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué se espera de los sistemas futuros de procesamiento del habla?

Elija las opciones correctas.

- ☒ a. Mayor robustez ante ruido y variabilidad lingüística ✓
- ☐ b. Reducción de todos los sistemas a texto plano
- ☒ c. Integración con visión por computadora ✓
- ☒ d. Personalización y conciencia del contexto ✓

Respuesta correcta

Las respuestas correctas son:

Mayor robustez ante ruido y variabilidad lingüística,

Integración con visión por computadora,

Personalización y conciencia del contexto

Pregunta 5

Correcta

Se puntúa 5,00 sobre 5,00

¿Cuál es el problema de representar palabras asignando una etiqueta numérica única a cada una?

- ☒ a. Que introduce relaciones numéricas artificiales entre palabras que no existen realmente. ✓
- ☐ b. Que no permite procesar texto largo
- ☐ c. Que las etiquetas se repiten frecuentemente.
- ☐ d. Que solo se puede usar con pocos idiomas.

Respuesta correcta

La respuesta correcta es:

Que introduce relaciones numéricas artificiales entre palabras que no existen realmente.

Pregunta 6

Correcta

Se puntúa 5,00 sobre 5,00

¿Cómo aprenden las redes neuronales a generar embeddings significativos?

- ☐ a. Mediante preprocesamiento estático fuera del modelo.
- ☐ b. Usando reglas lingüísticas codificadas a mano.
- ☒ c. Ajustando la capa de embedding durante el **entrenamiento supervisado** en una tarea específica. ✓
- ☐ d. Con clustering K-means sobre un corpus ya vectorizado.

Respuesta correcta

La respuesta correcta es:

Ajustando la capa de embedding durante el **entrenamiento supervisado** en una tarea específica.

Pregunta 7

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué es la "tokenización" y cuáles son algunos de los diferentes niveles en los que se puede realizar?

- ☐ a. Es eliminar artículos y preposiciones, a nivel semántico.
- ☐ b. Es traducir las palabras a otro idioma, a nivel de texto completo.
- ☒ c. Es dividir el texto en unidades pequeñas llamadas tokens: palabras, caracteres o subpalabras. ✓
- ☐ d. Es dividir texto en líneas y párrafos, y se hace únicamente a nivel de oraciones.

Respuesta correcta

La respuesta correcta es:

Es dividir el texto en unidades pequeñas llamadas tokens: palabras, caracteres o subpalabras.

Pregunta 8

Correcta

Se puntúa 5,00 sobre 5,00

¿Cuáles son ejemplos actuales de aplicaciones del procesamiento del habla?

Elija las opciones correctas.

- ☒ a. Transcripción automática en consultas médicas ✓
- ☒ b. Interfaces de control por voz en automóviles ✓
- ☐ c. Juegos de mesa tradicionales
- ☒ d. Asistentes virtuales como Alexa o Google Assistant ✓

Respuesta correcta

Las respuestas correctas son:

Asistentes virtuales como Alexa o Google Assistant,

Transcripción automática en consultas médicas,

Interfaces de control por voz en automóviles

Pregunta 9

Correcta

Se puntúa 5,00 sobre 5,00

¿Cómo se entrena un embedding para que sea lo más universal posible?

- ☐ a. Con texto traducido automáticamente; amplía los idiomas.
- ☐ b. Con un corpus pequeño y muy específico de un dominio.
- ☒ c. Con una **tarea genérica** (ej. predecir la siguiente palabra) y **mucho texto**, garantizando que capture relaciones amplias. ✓
- ☐ d. Con fine-tuning exclusivo en la tarea final sin pre-entrenamiento.

Respuesta correcta

La respuesta correcta es:

Con una **tarea genérica** (ej. predecir la siguiente palabra) y **mucho texto**, garantizando que capture relaciones amplias.

Pregunta 10

Correcta

Se puntúa 5,00 sobre 5,00

¿Cuál es la limitación del “one-hot encoding” y qué información valiosa se pierde con esta representación?

- ☐ a. No permite representar palabras compuestas; solo tokeniza una por una.
- ☒ b. Asume que todas las palabras son igualmente diferentes, perdiendo la **similitud semántica** entre términos relacionados. ✓
- ☐ c. No sirve para idiomas que no usen alfabeto latino.
- ☐ d. Ocupa demasiado espacio en disco; no hay problemas semánticos.

Respuesta correcta

La respuesta correcta es:

Asume que todas las palabras son igualmente diferentes, perdiendo la **similitud semántica** entre términos relacionados.

Pregunta 11

Correcta

Se puntúa 5,00 sobre 5,00

¿Cómo se puede utilizar un LLM pre-entrenado como GPT-3 para realizar tareas como el análisis de sentimientos o la clasificación sin un entrenamiento explícito para esas tareas?

- ☐ a. Re-entrenando todo el modelo desde cero con un nuevo corpus etiquetado.
- ☐ b. Modificando la arquitectura interna del Transformer manualmente.
- ☒ c. Diseñando un prompt que describa la tarea y ejemplos, de modo que el modelo “entienda” y complete según esa dinámica. ✓
- ☐ d. Con algoritmos de clustering interno que el usuario no controla.

Respuesta correcta

La respuesta correcta es:

Diseñando un prompt que describa la tarea y ejemplos, de modo que el modelo “entienda” y complete según esa dinámica.

Pregunta 12

Correcta

Se puntúa 5,00 sobre 5,00

¿Cómo se extiende la aplicación del PNL más allá del simple texto, que apareció como una novedad que no se había tenido en cuenta?

- ☐ a. Se aplica a reconocimiento de imágenes
- ☐ b. Se integra en simulaciones de vídeo
- ☐ c. Se usa en análisis de audio musical
- ☒ d. Se utiliza para entender, generar y traducir código de programación ✓

Respuesta correcta

La respuesta correcta es:

Se utiliza para entender, generar y traducir código de programación

Pregunta 13

Correcta

Se puntúa 5,00 sobre 5,00

¿Cómo funcionaba el autoaprendizaje supervisado en los primeros modelos de lenguaje basados en Transformers y cuál fue el cambio de paradigma que supuso?

- ☒ a. Enmascarando palabras o prediciendo la siguiente palabra en texto no etiquetado, eliminando la dependencia de anotaciones manuales. ✓
- ☐ b. Usando sólo ejemplos etiquetados manualmente, pero de menor tamaño.
- ☐ c. Aplicando clustering de frases, aunque sin predicción de palabras.
- ☐ d. Mezclando audio y texto para crear anotaciones automáticas.

Respuesta correcta

La respuesta correcta es:

Enmascarando palabras o prediciendo la siguiente palabra en texto no etiquetado, eliminando la dependencia de anotaciones manuales.

Pregunta 14

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué factores afectan la variabilidad del habla humana y su reconocimiento automático?

Elija las opciones correctas.

- ☒ a. Ruido ambiental ✓
- ☒ b. Acentos regionales ✓
- ☒ c. Estilos de habla (formal, coloquial) ✓
- ☐ d. Tamaño del dispositivo

Respuesta correcta

Las respuestas correctas son:

Acentos regionales,

Estilos de habla (formal, coloquial),

Ruido ambiental

Pregunta 15

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué componente del aparato fonador genera la vibración inicial en la producción del habla?

- ☐ a. Lengua
- ☒ b. Cuerdas vocales en la laringe ✓
- ☐ c. Labios
- ☐ d. Paladar

Respuesta correcta

La respuesta correcta es:

Cuerdas vocales en la laringe

Pregunta 16

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué capacidades sorprendentes emergieron en modelos como GPT-3 a pesar de haber sido entrenados únicamente para la tarea de autocompletar texto?

- ☐ a. Clasificar automáticamente sentimientos con etiquetas manuales.
- ☐ b. Solo completar fragmentos de texto sin comprender contenido.
- ☐ c. Reconocimiento de imágenes y vídeo en streaming.
- ☒ d. Traducción de idiomas, generación de código y simulación de diálogo, sin entrenamiento específico en esas tareas. ✓

Respuesta correcta

La respuesta correcta es:

Traducción de idiomas, generación de código y simulación de diálogo, sin entrenamiento específico en esas tareas.

Pregunta 17

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué técnica convierte una señal analógica de habla en una forma digital?

- ☐ a. Compresión de audio
- ☐ b. Amplificación de frecuencias
- ☐ c. Tokenización fonética
- ☒ d. Muestreo y cuantificación (digitalización) ✓

Respuesta correcta

La respuesta correcta es:

Muestreo y cuantificación (digitalización)

Pregunta 18

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué es un “embedding pre-entrenado” y cuáles son sus ventajas?

- ☐ a. Un embedding creado con TF-IDF; mejora la interpretabilidad.
- ☐ b. Un embedding que solo codifica caracteres, no palabras.
- ☒ c. Un embedding aprendido en un corpus grande y genérico; aporta **conocimiento semántico previo** y ahorra datos etiquetados. ✓
- ☐ d. Un embedding generado aleatoriamente y luego ajustado; acelera el entrenamiento.

Respuesta correcta

La respuesta correcta es:

Un embedding aprendido en un corpus grande y genérico; aporta **conocimiento semántico previo** y ahorra datos etiquetados.

Pregunta 19

Correcta

Se puntúa 5,00 sobre 5,00

¿Cómo se introduce la noción de similitud semántica en el procesamiento del lenguaje natural para las redes neuronales?

- ☐ a. Con TF-IDF, valorando la frecuencia inversa de documentos.
- ☐ b. Con análisis de dependencias sintácticas.
- ☐ c. Con bag-of-words y ponderaciones manuales.
- ☒ d. Con **Embeddings**, vectores densos que acercan palabras con significados similares. ✓

Respuesta correcta

La respuesta correcta es:

Con **Embeddings**, vectores densos que acercan palabras con significados similares.

Pregunta 20

Correcta

Se puntúa 5,00 sobre 5,00

¿Qué incertidumbre o pregunta fundamental persiste a pesar de los impresionantes avances en PNL?

- ☐ a. Si verdaderamente el PNL funcionará en todos los idiomas minoritarios
- ☒ b. Si los modelos realmente “entienden” o solo memorizan patrones ✓
- ☐ c. Si existirá un límite físico al tamaño de los modelos
- ☐ d. Si la capacidad computacional será suficiente para procesarlos

Respuesta correcta

La respuesta correcta es:

Si los modelos realmente “entienden” o solo memorizan patrones

