

Comenzado el	martes, 11 de febrero de 2025, 16:22
Estado	Finalizado
Finalizado en	martes, 11 de febrero de 2025, 16:40
Tiempo empleado	17 minutos 43 segundos
Puntos	40,50/46,00
Calificación	8,80 de 10,00 (88,04%)

Pregunta 1

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de las siguientes es una posible aplicación del PLN?

- ☒ a. Clasificación de notas médicas de profesionales sanitarios ✓
- ☐ b. Análisis financiero
- ☐ c. Diseño gráfico
- ☐ d. Desarrollo de videojuegos

La respuesta correcta es: Clasificación de notas médicas de profesionales sanitarios

Pregunta 2

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es el modelo de Bolsa de Palabras (Bag of Words) en PLN?

- ☒ a. Una técnica que convierte un texto en una colección que contiene la frecuencia de aparición de cada palabra ✓
- ☐ b. Un método para traducir palabras a sus raíces
- ☐ c. Un modelo que considera el orden de las palabras en un texto
- ☐ d. Una técnica para eliminar palabras comunes de un texto

La respuesta correcta es: Una técnica que convierte un texto en una colección que contiene la frecuencia de aparición de cada palabra

Pregunta 3

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es la diferencia clave entre stemming y lematización?

- ☐ a. El stemming considera el contexto de la palabra, mientras que la lematización no
- ☐ b. El stemming siempre produce palabras válidas, mientras que la lematización puede no hacerlo
- ☐ c. No hay diferencia, son sinónimos.
- ☒ d. La lematización considera el contexto y la morfología de la palabra, mientras que el stemming aplica reglas más simples y directas ✓

La respuesta correcta es: La lematización considera el contexto y la morfología de la palabra, mientras que el stemming aplica reglas más simples y directas

Pregunta 4

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué mide la técnica TF-IDF en el análisis de textos?

- ☐ a. La longitud promedio de las oraciones en un texto
- ☒ b. La frecuencia de una palabra en un documento y su importancia en todo el corpus ✓
- ☐ c. La complejidad gramatical de un texto
- ☐ d. La cantidad de stop words en un documento

La respuesta correcta es: La frecuencia de una palabra en un documento y su importancia en todo el corpus

Pregunta 5

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué son los word embeddings en PLN?

- ☒ a. Representaciones vectoriales densas de palabras que capturan su significado semántico y contexto ✓
- ☐ b. Listas de palabras ordenadas alfabéticamente
- ☐ c. Métodos para eliminar palabras irrelevantes en un análisis
- ☐ d. Técnicas para contar la frecuencia de palabras en un texto

La respuesta correcta es: Representaciones vectoriales densas de palabras que capturan su significado semántico y contexto

Pregunta 6

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué son los N-gramas en el análisis de texto?

- ☐ a. Palabras que aparecen con mayor frecuencia en un corpus
- ☐ b. Técnicas para reducir palabras a su forma base
- ☐ c. Métodos para eliminar palabras comunes en un texto
- ☒ d. Secuencias contiguas de 'n' elementos de un texto, que pueden ser caracteres, sílabas, palabras o frases ✓

La respuesta correcta es: Secuencias contiguas de 'n' elementos de un texto, que pueden ser caracteres, sílabas, palabras o frases

Pregunta 7

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es el reconocimiento de entidades nombradas (NER) en PLN?

- ☐ a. Un método para contar la frecuencia de palabras en un documento
- ☐ b. Una técnica para traducir texto entre diferentes idiomas
- ☒ c. Una técnica que identifica y clasifica entidades mencionadas en un texto en categorías predefinidas como nombres de personas, organizaciones, lugares, etc. ✓
- ☐ d. Un proceso para eliminar palabras irrelevantes en un análisis

La respuesta correcta es: Una técnica que identifica y clasifica entidades mencionadas en un texto en categorías predefinidas como nombres de personas, organizaciones, lugares, etc.

Pregunta 8

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es el propósito del análisis de sentimientos en el procesamiento de texto?

- ☐ a. Eliminar palabras irrelevantes en las opiniones de usuarios de servicios.
- ☐ b. Contar la frecuencia de palabras en un documento para determinar enfermedades mentales
- ☒ c. Determinar la polaridad y subjetividad del texto para entender las percepciones y emociones de los usuarios hacia productos y servicios ✓
- ☐ d. Analizar textos para mejorar el estado anímico de las personas que lo escriben.

La respuesta correcta es: Determinar la polaridad y subjetividad del texto para entender las percepciones y emociones de los usuarios hacia productos y servicios

Pregunta 9

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué técnica se utiliza para reducir la dimensionalidad de los datos y facilitar la visualización de embeddings?

- ☒ a. t-SNE (t-Distributed Stochastic Neighbor Embedding) ✓
- ☐ b. Stemming
- ☐ c. Tokenización
- ☐ d. Lematización

La respuesta correcta es: t-SNE (t-Distributed Stochastic Neighbor Embedding)

Pregunta 10

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué técnica de codificación asigna un vector binario a cada palabra, donde solo una posición tiene el valor 1?

- ☐ a. Bag of Words
- ☐ b. Embeddings
- ☐ c. TF-IDF
- ☒ d. One-hot encoding ✓

La respuesta correcta es: One-hot encoding

Pregunta 11

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es una limitación principal del one-hot encoding en PLN?

- ☒ a. No captura relaciones semánticas entre palabras ✓
- ☐ b. Genera vectores densos
- ☐ c. No es compatible con NLTK
- ☐ d. Es difícil de implementar

La respuesta correcta es: No captura relaciones semánticas entre palabras

Pregunta 12

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de las siguientes librerías en Python es más comúnmente utilizada para PLN?

- ☐ a. Matplotlib
- ☒ b. NLTK ✓
- ☐ c. NumPy
- ☐ d. Pandas

La respuesta correcta es: NLTK

Pregunta 13

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es el preprocesamiento de texto en el contexto de la clasificación de textos?

- ☒ a. El proceso de limpiar y preparar textos: tokenización, stop words, lematización... ✓
- ☐ b. La etapa de recopilación de datos de texto
- ☐ c. La etapa en la que se etiquetan las clases.
- ☐ d. La fase de evaluación de modelos

La respuesta correcta es: El proceso de limpiar y preparar textos: tokenización, stop words, lematización...

Pregunta 14

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la tokenización en el preprocesamiento de texto?

- ☐ a. La traducción de texto a otro idioma
- ☐ b. La eliminación de palabras irrelevantes de un texto
- ☐ c. La conversión de palabras a sus formas base
- ☒ d. La división de un texto en unidades más pequeñas, como palabras o frases ✓

La respuesta correcta es: La división de un texto en unidades más pequeñas, como palabras o frases

Pregunta 15

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es la principal desventaja del modelo de Bolsa de Palabras?

- ☐ a. No puede manejar grandes volúmenes de datos
- ☒ b. No captura el contexto o el orden de las palabras en un texto ✓
- ☐ c. No es compatible con bibliotecas de Python
- ☐ d. Es demasiado complejo computacionalmente

La respuesta correcta es: No captura el contexto o el orden de las palabras en un texto

Pregunta 16

Sin contestar

Se puntúa como 0 sobre 1,00

¿Cuál es el propósito de la técnica de "padding" en redes neuronales para PLN?

- ☐ a. Convertir palabras en vectores
- ☐ b. Asegurar que todas las secuencias tengan la misma longitud
- ☐ c. Normalizar palabras
- ☐ d. Acelerar el entrenamiento

La respuesta correcta es: Asegurar que todas las secuencias tengan la misma longitud

Pregunta 17

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es Hugging Face en el contexto del Procesamiento de Lenguaje Natural?

- ☐ a. Un software de visualización de datos de PLN
- ☐ b. Un framework exclusivo para la traducción automática de textos
- ☒ c. Una plataforma y biblioteca que proporciona modelos de aprendizaje profundo para PLN, como BERT y GPT ✓
- ☐ d. Una base de datos de textos anotados para entrenar modelos

La respuesta correcta es: Una plataforma y biblioteca que proporciona modelos de aprendizaje profundo para PLN, como BERT y GPT

Pregunta 18

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué técnica ayuda a reducir la dimensionalidad de los vectores de palabras?

- ☒ a. Embeddings ✓
- ☐ b. PCA
- ☐ c. Tokenización
- ☐ d. TF-IDF

La respuesta correcta es: Embeddings

Pregunta 19

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué técnica se usa para reducir una palabra a su forma base eliminando sufijos?

- ☒ a. Stemming ✓
- ☐ b. Normalización
- ☐ c. Tokenización
- ☐ d. Vectores de ocurrencia

La respuesta correcta es: Stemming

Pregunta 20

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de las siguientes técnicas es un modelo basado en transformers para PLN?

- ☐ a. BoW
- ☐ b. Word2Vec
- ☐ c. TF-IDF
- ☒ d. BERT ✓

La respuesta correcta es: BERT

Pregunta 21

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué módulo de NLTK proporciona herramientas para la lematización?

- ☐ a. nltk.corpus
- ☐ b. nltk.chunk
- ☒ c. nltk.stem ✓
- ☐ d. nltk.tokenize

La respuesta correcta es: nltk.stem

Pregunta 22

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es la diferencia entre stemming y lematización?

- ☐ a. Stemming es más lento que lematización
- ☒ b. Lematización usa diccionarios y reglas gramaticales, mientras que stemming corta palabras sin contexto ✓
- ☐ c. No hay diferencias
- ☐ d. Stemming es más preciso que lematización

La respuesta correcta es: Lematización usa diccionarios y reglas gramaticales, mientras que stemming corta palabras sin contexto

Pregunta 23

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es una ventaja del stemming sobre la lematización?

- ☒ a. Es más rápido en términos de procesamiento ✓
- ☐ b. Considera el contexto de la palabra
- ☐ c. Siempre produce palabras válidas
- ☐ d. Es más preciso en la reducción de palabras

La respuesta correcta es: Es más rápido en términos de procesamiento

Pregunta 24

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es una ventaja de la lematización sobre el stemming?

- ☐ a. Es más rápida en términos de procesamiento
- ☐ b. No requiere bibliotecas adicionales
- ☒ c. Produce palabras que son siempre válidas y significativas ✓
- ☐ d. Aplica reglas más simples y directas

La respuesta correcta es: Produce palabras que son siempre válidas y significativas

Pregunta 25

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué resultado se obtiene al aplicar el stemming con PorterStemmer a la palabra "caminando"?

- ☐ a. caminando
- ☐ b. caminar
- ☐ c. camina
- ☒ d. camin ✓

La respuesta correcta es: camin

Pregunta 26

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué resultado se obtiene al aplicar la lematización con WordNetLemmatizer a la palabra "caminando" especificando que es un verbo?

- ☐ a. camina
- ☐ b. camin
- ☒ c. caminar ✓
- ☐ d. caminando

La respuesta correcta es: caminar

Pregunta 27

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué mide la técnica TF-IDF en el análisis de textos?

- ☐ a. Un método de tokenización
- ☐ b. Un tipo de embedding
- ☐ c. Un algoritmo de stemming
- ☒ d. Una métrica que combina la frecuencia de palabras y su relevancia en un corpus ✓

La respuesta correcta es: Una métrica que combina la frecuencia de palabras y su relevancia en un corpus

Pregunta 28

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué técnica se utiliza para convertir datos de texto en una representación numérica que los algoritmos de aprendizaje automático pueden procesar?

- ☐ a. Stemming
- ☐ b. Lematización
- ☒ c. Modelo de Bolsa de Palabras (Bag of Words) ✓
- ☐ d. Tokenización

La respuesta correcta es: Modelo de Bolsa de Palabras (Bag of Words)

Pregunta 29

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué técnica se usa para convertir una palabra en su forma canónica, teniendo en cuenta su significado?

- ☐ a. Stemming
- ☐ b. Tokenización
- ☐ c. Expansión léxica
- ☒ d. Lematización ✓

La respuesta correcta es: Lematización

Pregunta 30

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál de los siguientes es un caso de uso común de la clasificación de textos?

- ☐ a. Traducción automática de idiomas
- ☒ b. Clasificación de correos electrónicos en categorías de spam y ham ✓
- ☐ c. Generación de poesía
- ☐ d. Reconocimiento de voz

La respuesta correcta es: Clasificación de correos electrónicos en categorías de spam y ham

Pregunta 31

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la frecuencia inversa de documento (IDF) en TF-IDF?

- ☐ a. Es un método de eliminación de stop words
- ☐ b. Es un método de traducción automática
- ☒ c. Penaliza las palabras que aparecen frecuentemente en muchos documentos ✓
- ☐ d. Mide cuántas veces aparece una palabra en un documento

La respuesta correcta es: Penaliza las palabras que aparecen frecuentemente en muchos documentos

Pregunta 32

Sin contestar

Sin calificar

¿Cuál es la salida esperada del siguiente código BoW en Python?

```
from sklearn.feature_extraction.text
import CountVectorizer

corpus = ["El gato juega en el jardín"]

vectorizer = CountVectorizer()

X = vectorizer.fit_transform(corpus)

print(X.toarray())
```

- ☐ a. `[[1, 2, 1, 1, 1]]`
- ☐ b. `[[0, 1, 1, 1, 1]]`
- ☐ c. `[[1, 1, 1, 1, 1]]`
- ☐ d. `[[1, 1, 1, 1]]`

La respuesta correcta es: `[[1, 1, 1, 1, 1]]`

Pregunta 33

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es el resultado de aplicar stemming con PorterStemmer en la palabra "running"?

- ☐ a. runs
- ☐ b. running
- ☒ c. run ✓
- ☐ d. runn

La respuesta correcta es: run

Pregunta 34

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué representa un modelo Bag of Words (BoW)?

- ☐ a. Un sistema de categorización de palabras
- ☒ b. La cantidad de veces que aparece cada palabra en un texto ✓
- ☐ c. La gramática y estructura de las oraciones
- ☐ d. La relación semántica entre palabras en un texto

La respuesta correcta es: La cantidad de veces que aparece cada palabra en un texto

Pregunta 35

Sin contestar

Se puntúa como 0 sobre 1,00

¿Cuál de las siguientes es una ventaja del modelo Bag of Words (BoW)?

- ☐ a. Es simple y fácil de implementar, permitiendo representar texto en un formato numérico básico
- ☐ b. Captura el contexto y el significado de las palabras en una oración
- ☐ c. Es más eficiente en términos computacionales que modelos basados en embeddings
- ☐ d. Requiere menos preprocesamiento que Word2Vec y Transformers

La respuesta correcta es: Es simple y fácil de implementar, permitiendo representar texto en un formato numérico básico

Pregunta 36

Correcta

Se puntúa 1,00 sobre 1,00

¿Cuál es una ventaja del modelo Word2Vec sobre BoW?

- ☐ a. Word2Vec es más fácil de implementar
- ☐ b. BoW captura mejor el contexto de las palabras
- ☐ c. Word2Vec usa menos recursos computacionales
- ☒ d. Word2Vec captura el significado semántico y la relación entre palabras ✓

La respuesta correcta es: Word2Vec captura el significado semántico y la relación entre palabras

Pregunta 37

Incorrecta

Se puntúa -0,25 sobre 1,00

¿Qué modelo de representación de palabras usa redes neuronales para aprender relaciones semánticas?

- ☐ a. CountVectorizer
- ☐ b. TF-IDF
- ☐ c. Word2Vec
- ☒ d. BoW ✗

La respuesta correcta es: Word2Vec

Pregunta 38

Sin contestar

Se puntúa como 0 sobre 1,00

¿Cuál de los siguientes métodos de tokenización en NLTK divide correctamente una oración en palabras?

- ☐ a. TextBlob
- ☐ b. sent_tokenize
- ☐ c. split
- ☐ d. word_tokenize

La respuesta correcta es: word_tokenize

Pregunta 39

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la normalización de texto en PLN?

- ☐ a. Contar la frecuencia de las palabras
- ☒ b. Convertir todas las palabras a minúsculas y eliminar caracteres especiales ✓
- ☐ c. Traducir texto de un idioma a otro
- ☐ d. Tokenizar un documento en palabras individuales

La respuesta correcta es: Convertir todas las palabras a minúsculas y eliminar caracteres especiales

Pregunta 40

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué tecnología es utilizada para tareas de generación de texto?

- ☐ a. Word2Vec
- ☒ b. GPT ✓
- ☐ c. LDA
- ☐ d. BERT

La respuesta correcta es: GPT

Pregunta 41

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es ALIA?

- ☐ a. Un sistema de inteligencia artificial enfocado al análisis de datos meteorológicos en España.
- ☐ b. Un asistente virtual desarrollado por el gobierno para la atención al ciudadano
- ☐ c. Un proyecto europeo de inteligencia artificial centrado en la traducción automática
- ☒ d. Una inteligencia artificial pública en español y lenguas cooficiales, diseñada con código abierto por el gobierno español ✓

La respuesta correcta es: Una inteligencia artificial pública en español y lenguas cooficiales, diseñada con código abierto por el gobierno español

Pregunta 42

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la clasificación de textos en el contexto del Procesamiento de Lenguaje Natural (PLN)?

- ☐ a. Un método para generar texto automáticamente
- ☒ b. La tarea de asignar categorías o etiquetas predefinidas a documentos de texto ✓
- ☐ c. Una técnica para resumir documentos largos
- ☐ d. Un proceso para traducir textos de un idioma a otro

La respuesta correcta es: La tarea de asignar categorías o etiquetas predefinidas a documentos de texto

Pregunta 43

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué hace el siguiente código al utilizar la API de Hugging Face?

```
import requests

API_URL = "https://api-inference.huggingface.co/models/distilgpt2"

headers = {"Authorization": "Bearer TU_TOKEN_AQUI"}

data = {"inputs": "El clima hoy es"}

response = requests.post(API_URL, headers=headers, json=data)

print(response.json())
```

- ☒ a. Generar una continuación del texto enviado usando el modelo DistilGPT-2 ✓
- ☐ b. Devuelve el vector de tokens del texto
- ☐ c. Traduce automáticamente el texto ingresado al inglés usando el modelo DistilGPT-2
- ☐ d. Clasifica el texto en diferentes categorías (DistilGPT-2) según análisis de sentimiento

Respuesta correcta

La respuesta correcta es: Generar una continuación del texto enviado usando el modelo DistilGPT-2

Pregunta 44

Incorrecta

Se puntúa -0,25 sobre 1,00

¿Cuál es la salida de este código?

```
from nltk.tokenize import word_tokenize  
  
text = "¡Hola! ¿Cómo estás?"  
  
tokens = word_tokenize(text)  
  
print(tokens)
```

- ☒ a. ["Hola", "Cómo", "estás"] ✗
- ☐ b. ["¡Hola! ¿Cómo estás?"]
- ☐ c. ["¡Hola!", "Cómo", "estás"]
- ☐ d. ["¡Hola!", "¿", "Cómo", "estás", "?"]

Respuesta incorrecta.

La respuesta correcta es: ["¡Hola!", "¿", "Cómo", "estás", "?"]

Pregunta 45

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la eliminación de stop words en el procesamiento de texto?

- ☐ a. La asignación de etiquetas gramaticales a las palabras
- ☒ b. El proceso de eliminar palabras muy frecuentes en un idioma que no aportan un significado relevante al análisis del texto ✓
- ☐ c. La técnica de reducir palabras a su forma base
- ☐ d. El método de dividir un texto en unidades más pequeñas

La respuesta correcta es: El proceso de eliminar palabras muy frecuentes en un idioma que no aportan un significado relevante al análisis del texto

Pregunta 46

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la lematización en el Procesamiento de Lenguaje Natural?

- ☐ a. Un método para contar la frecuencia de palabras en un texto
- ☒ b. Una técnica que reduce las palabras a su forma base o lema, teniendo en cuenta el contexto y la morfología ✓
- ☐ c. Un proceso que convierte palabras a minúsculas
- ☐ d. Una técnica para dividir un texto en oraciones

La respuesta correcta es: Una técnica que reduce las palabras a su forma base o lema, teniendo en cuenta el contexto y la morfología

Pregunta 47

Correcta

Se puntúa 1,00 sobre 1,00

¿Qué es la tokenización en el contexto del Procesamiento de Lenguaje Natural?

- ☐ a. La asignación de etiquetas gramaticales a las palabras
- ☒ b. La división de un texto en unidades más pequeñas llamadas tokens ✓
- ☐ c. La eliminación de palabras irrelevantes en un texto
- ☐ d. La conversión de palabras a sus formas base

La respuesta correcta es: La división de un texto en unidades más pequeñas llamadas tokens