

Your grade: 83.33%

Your latest: 83.33% • Your highest: 83.33% • We keep your highest score. Review your overall course grades [here](#).

1. ¿Qué característica distingue a la representación de textos con Bolsa de Palabras (BOW)?

1 / 1 point

BOW considera el orden de las palabras, lo cual es esencial para capturar el contexto y las relaciones semánticas entre ellas.

BOW ignora el orden de las palabras, centrándose en la presencia o ausencia de términos.

En BOW, frecuencias absolutas asignan mayor peso a las palabras menos frecuentes.

☑ Correct

Correcto. La representación con base en la bolsa de palabras (BOW) no considera el orden de las palabras y se enfoca en la presencia o ausencia de términos.

2. En el pipeline de transformación de textos a un formato vectorial utilizando BOW, el diccionario de palabras vacías se utiliza para la reducción de la dimensionalidad. Indique si esta afirmación es verdadera o falsa.

1 / 1 point

Falsa

Verdadera

☑ Correct

Correcto. El diccionario de palabras vacía se utiliza como base para "sacar" de la bolsa aquellas palabras menos informativas, lo que resulta en una reducción de la dimensionalidad.

3. ¿Cuál es la diferencia clave entre stemming y lematización en la preparación de textos?

1 / 1 point

La lematización reduce las palabras a sus formas básicas, mientras que el stemming elimina afijos.

El stemming reduce las palabras a sus formas básicas, mientras que la lematización elimina afijos.

☑ Correct

Correcto. La lematización busca reducir las palabras a su forma base (lemas), considerando la morfología y el significado, mientras que el stemming elimina afijos (prefijos y sufijos) para obtener una forma truncada de la palabra.

4. ¿Cuál es el propósito principal del pesado de términos por TF-IDF en la representación de textos?

1 point

Asignar mayor peso a las palabras más frecuentes en un documento.

Penalizar términos que aparecen con poca frecuencia en el conjunto de documentos.

Destacar palabras relevantes para un documento específico.

⊗ Incorrect

Revisa tu respuesta. TF-IDF asigna mayor peso a las palabras que son importantes en un documento específico, al considerar tanto la frecuencia absoluta como el número de documentos que las contienen.

5. ¿Por qué es importante la reducción de dimensionalidad en el análisis de texto?

1 / 1 point

Para visualizar los datos en un espacio de baja dimensión

