

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA
TECNOLOGÍAS DE LA INFORMACIÓN



Extracción de Conocimiento en Bases de Datos

DOCENTE:

Enrique Mascote

Diagnostico4

PRESENTA:

Iván Eduardo Martínez Martínez

Grupo:

IDGS91N

Opción múltiple

1. ¿Cuál no es algoritmo de clustering?

Respuesta: C) PCA

2. PCA sirve para:

Respuesta: B) Reducir dimensiones

3. Silhouette score mide:

Respuesta: A) Coherencia interna de clusters

4. ¿Qué tipo de datos requiere t-SNE?

Respuesta: B) Datos de alta dimensión

5. Define en una frase qué es un cluster.

Un *cluster* es un grupo de datos que comparten características similares entre sí y se diferencian de otros grupos.

6. Menciona un escenario donde usarías reducción de dimensionalidad.

La usaría cuando tengo un dataset con muchas variables (por ejemplo, imágenes con miles de píxeles) y necesito simplificarlo para visualizarlo o acelerar el entrenamiento de un modelo.

7. ¿Para qué sirve detectar outliers antes del clustering?

Sirve para evitar que los datos atípicos distorsionen la formación de clusters, ya que pueden alterar las distancias y producir agrupamientos incorrectos.