

**UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA
DESARROLLO Y GESTIÓN DE SOFTWARE**



**Diagnóstico Unidad 4: Análisis no supervisado
EXTRACCIÓN DE CONOCIMIENTO EN BASES DE DATOS**

PRESENTA:

KARLA ALEJANDRA DE LA CRUZ ZEA

DOCENTE:

ING. LUIS ENRIQUE MASCOTE CANO

29 de noviembre de 2025

Formato: 7 preguntas (4 opción múltiple, 3 respuesta corta)

1. ¿Cuál no es algoritmo de clustering?
 - A) K-Means
 - B) DBSCAN
 - C) PCA
 - D) Jerárquico
2. PCA sirve para:
 - A) Clasificar imágenes
 - B) Reducir dimensiones
 - C) Detectar outliers
 - D) Ajustar hiperparámetros
3. Silhouette score mide:
 - A) Coherencia interna de clusters
 - B) Precisión del modelo
 - C) MSE
 - D) Número de clusters
4. ¿Qué tipo de datos requiere t-SNE?
 - A) Tablas relacionales
 - B) Datos de alta dimensión
 - C) Series de tiempo
 - D) Texto puro

Respuesta Corta

5. Define en una frase qué es un cluster.

Un cluster es un grupo de datos similares entre sí y diferentes de otros grupos.

6. Menciona un escenario donde usarías reducción de dimensionalidad.

Cuando tengo muchas variables y quiero visualizar datos en 2D o 3D, como en análisis de clientes o imágenes.

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

Para evitar que los puntos extremos deformen los grupos y produzcan clusters incorrectos.