

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

Tecnologías de la información



Extracción de Conocimiento en Bases de Datos

Diagnostico Unidad 4: Análisis no supervisado

IDGS91N

Presenta:

Carlos Isaac Parra Aguirre

Docente:

Enrique Mascote

30 de November de 2025

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 los datos de otros grupos.

6. Menciona un escenario donde usarías reducción de dimensionalidad.

Para visualizar datos de muchas variables (por ejemplo, 50 atributos) en un plano 2D usando PCA o t-SNE.

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

¿?