

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

TECNOLOGÍAS DE LA INFORMACIÓN



EXTRACCIÓN DE CONOCIMIENTO EN BASES DE DATOS

DIAGNOSTICO UNIDAD 4: ANÁLISIS NO SUPERVISADO

DOCENTE

ING. LUIS ENRIQUE MASCOTE CANO

ALUMNO

MILDRED VILLASEÑOR RUIZ

29 DE NOVIEMBRE DEL 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.

Es un grupo de datos similares entre sí y diferentes de otros grupos.
6. Menciona un escenario donde usarías reducción de dimensionalidad.

Cuando un conjunto de datos tiene muchas columnas (alta dimensionalidad) y necesito visualizarlo o mejorar el rendimiento de un modelo, por ejemplo datos de imágenes o sensores.
7. ¿Para qué sirve detectar outliers antes del clustering?

Para evitar que valores atípicos distorsionen la formación de clusters y afectan su calidad.