

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

DESARROLLO Y GESTION DE SOFTWARE



EXTRACCION DE CONOCIMIENTO DE BASE DE DATOS

DIAGNOSTICO UNIDAD IV: ANALISIS NO SUPERVISADO

DOCENTE:

ENRIQUE MASCOTE

PRESENTA:

DARON TARÍN GONZÁLEZ

GRUPO:

IDGS91N

Chihuahua, Chih., 30 de noviembre 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 que son más similares entre sí que respecto a otros grupos.

6. Menciona un escenario donde usarías reducción de dimensionalidad.

Usaría reducción de dimensionalidad para visualizar datos de muchas variables (por ejemplo, 300 características) en un espacio de 2D o 3D.

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

Detectar outliers antes del clustering evita que estos puntos extremos distorsionen la formación de los grupos y generen clusters incorrectos.