

# UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

## Ingeniería en Desarrollo y Gestión de Software



### Extracción de Conocimientos de Bases de Datos

### Diagnóstico Unidad 3: Análisis Supervisado

**IDGS91N**

PRESENTAN:

Giselle Cantú Chávez

NOMBRE DEL DOCENTE:

Ing. Luis Enrique Mascote Cano

Chihuahua, Chih., 29 de noviembre de 2025

## Índice

1. ¿Cuál no es un algoritmo de regresión? .....	3
2. En clasificación, ¿qué métrica mide la proporción de verdaderos positivos correctamente identificados? .....	3
3. ¿Qué técnica ayuda a evitar el sobreajuste ajustando penalizaciones? .....	3
4. ¿Cuál es el propósito de GridSearchCV? .....	3
5. ¿Qué representa la curva ROC? .....	4
6. Respuesta corta — Menciona un caso práctico donde usarías regresión. ....	4
7. Respuesta corta — ¿Por qué es útil la validación cruzada? .....	4
8. Respuesta corta — Explica brevemente qué es una matriz de confusión. ....	5

**1. ¿Cuál no es un algoritmo de regresión?**

- A) Regresión lineal
- B) K-Means
- C) Lasso
- D) Ridge

**Respuesta: B) K-Means**

**2. En clasificación, ¿qué métrica mide la proporción de verdaderos positivos correctamente identificados?**

- A) Precision
- B) Recall
- C) Accuracy
- D) MSE

**Respuesta: B) Recall**

**3. ¿Qué técnica ayuda a evitar el sobreajuste ajustando penalizaciones?**

- A) Cross-validation
- B) Regularización
- C) PCA
- D) DBSCAN

**Respuesta: B) Regularización**

**4. ¿Cuál es el propósito de GridSearchCV?**

- A) Ajustar hiperparámetros
- B) Dibujar curvas ROC
- C) Preprocesar datos
- D) Reducir dimensionalidad

**Respuesta: A) Ajustar hiperparámetros**

**5. ¿Qué representa la curva ROC?**

- A) Relación entre precisión y recall
- B) Tasa de verdaderos positivos vs. tasa de falsos positivos
- C) Error vs. tiempo
- D) Correlación de variables

**Respuesta: B) Tasa de verdaderos positivos vs. tasa de falsos positivos**

**6. Respuesta corta — Menciona un caso práctico donde usarías regresión.**

**Respuesta:**

Usaría regresión cuando necesito predecir un valor numérico continuo. Por ejemplo, estimar las ventas del mes siguiente a partir de variables como publicidad, estacionalidad y número de clientes activos.

**7. Respuesta corta — ¿Por qué es útil la validación cruzada?**

**Respuesta:**

Sirve para evaluar el modelo de forma más confiable. En vez de entrenar y probar una sola vez, lo prueba en varias particiones de los datos y así sé si el modelo realmente generaliza y no solo memorizó un conjunto específico.

**8. Respuesta corta — Explica brevemente qué es una matriz de confusión.**

**Respuesta:**

Es una tabla que resume los aciertos y errores del modelo clasificando datos. Muestra cuántos verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos tuvo, y con eso puedo entender exactamente dónde está fallando.