

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



**DIAGNOSTICO UNIDAD 3
ANÁLISIS SUPERVISADO
EXTRACCIÓN DE CONOCIMIENTO EN BASES DE DATOS**

PRESENTA:

KARLA ALEJANDRA DE LA CRUZ ZEA

DOCENTE:

ING. LUIS ENRIQUE MASCOTE CANO

13 de noviembre de 2025

Formato: 8 preguntas (5 opción múltiple, 3 respuesta corta)

1. ¿Cuál no es un algoritmo de regresión?
 - A) Regresión lineal
 - B) K-Means
 - C) Lasso
 - D) Ridge
2. En clasificación, ¿qué métrica mide la proporción de verdaderos positivos correctamente identificados?
 - A) Precision
 - B) Recall
 - C) Accuracy
 - D) MSE
3. ¿Qué técnica ayuda a evitar el sobreajuste ajustando penalizaciones?
 - A) Cross-validation
 - B) Regularización
 - C) PCA
 - D) DBSCAN
4. ¿Cuál es el propósito de GridSearchCV?
 - A) Ajustar hiperparámetros
 - B) Dibujar curvas ROC
 - C) Preprocesar datos
 - D) Reducir dimensionalidad
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 Corta

6. Menciona un caso práctico donde usarías regresión.

Para predecir el precio de casas en función de variables como tamaño, ubicación y número de habitaciones.

7. ¿Por qué es útil la validación cruzada?

Porque permite medir el desempeño del modelo en diferentes tipos de datos.

8. Explica brevemente qué es una matriz de confusión.

Es una tabla que compara las predicciones de un modelo de clasificación con los valores reales, mostrando cuántos casos fueron verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos.