

# **UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA**

## **DESARROLLO Y GESTION DE SOFTWARE**



### **EXTRACCION DE CONOCIMIENTO DE BASE DE DATOS**

### **DIAGNOSTICO UNIDAD III: ANALISIS SUPERVISADO**

**DOCENTE:**

**ENRIQUE MASCOTE**

**PRESENTA:**

**DARON TARÍN GONZÁLEZ**

**GRUPO:**

**IDGS91N**

Chihuahua, Chih., 30 de noviembre de 2025

¿Cuál no es un algoritmo de regresión?

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

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

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

¿Qué técnica ayuda a evitar el sobreajuste ajustando penalizaciones?

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

¿Cuál es el propósito de GridSearchCV?

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

¿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

Menciona un caso práctico donde usarías regresión.

Para estimar los rendimientos de una inversión tomando en cuenta su precio, dividendos y crecimiento.

¿Por qué es útil la validación cruzada?

Evaluá el rendimiento de forma exhaustiva al entrenarlo y probarlo en diferentes conjuntos de datos, evitando el desajuste.

Explica brevemente qué es una matriz de confusión.

Es una matriz que facilita observar cómo se comporta un modelo de clasificación al contrastar sus predicciones con los datos reales, organizando los resultados en verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos.