

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:

ANGEL RICARDO CHAVEZ ZARAGOZA

GRUPO:

IDGS91N

Chihuahua, Chih., 29 de November 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 predecir el precio de una vivienda basándose en variables como su tamaño, número de habitaciones y ubicación.

¿Por qué es útil la validación cruzada?

Porque permite evaluar el rendimiento del modelo de forma robusta al entrenarlo y probarlo en diferentes subconjuntos de datos, ayudando a detectar y evitar el sobreajuste.

Explica brevemente qué es una matriz de confusión.

Es una tabla que permite visualizar el desempeño de un modelo de clasificación comparando sus predicciones con los valores reales, dividiendo los resultados en Verdaderos Positivos, Verdaderos Negativos, Falsos Positivos y Falsos Negativos.