

Ejercicio 1

Determine si las siguientes afirmaciones sobre los valores atípicos son verdaderas o falsas:

- El análisis de boxplot permite detectar outliers multivariados. _____
- Isolation Forest es un método basado en densidad que permite detectar tanto outliers univariados como multivariados. _____
- La única forma de tratar valores atípicos es eliminar los valores más altos y más bajos de una variable. _____
- Si se calcula el z-score para un conjunto de datos y se quieren eliminar los outliers deben eliminarse los valores dentro del rango -3 y 3. _____

Ejercicio 2

Determinar si las siguientes afirmaciones son verdaderas o falsas:

- El coeficiente de correlación de pearson mide cuán relacionadas están dos variables sea o no de forma lineal. _____
- El método de regresión logística se utiliza para resolver problemas de clasificación binaria. _____
- La regresión lineal simple consiste en predecir un valor numérico en rango continuo en base a más de una variable predictora. _____

Ejercicio 3

Determinar cuáles de las siguientes opciones son características que comparten los métodos de Bagging y Boosting. Indicar verdadero o falso:

- Se basan en entrenar varios modelos, y luego tomar una única decisión. _____
- Cada modelo se entrena en paralelo. _____
- Cada modelo es influenciado por la performance de los otros. _____
- Utilizan datasets de entrenamiento generados mediante Bootstrap. _____

Ejercicio 4

Se entrenó un modelo de clasificación (MODELO A) para detectar personas enfermas (clase 1) y sanas (clase 0). Luego se evaluó el modelo en los datos de test y se obtuvo la siguiente matriz de confusión:

True \ Predicted	0	1
0	135	36
1	38	111

Luego se entrenó un segundo modelo de clasificación (MODELO B), se evaluó sobre los mismos datos de test y se obtuvieron los siguientes resultados:

True Positive: 49
True Negative: 130
False Positive: 41
False Negative: 100

¿Qué modelo (A o B) elegiría en términos de Accuracy?: _____

- Accuracy Modelo A _____
- Accuracy Modelo B _____