

ACTIVIDAD EL ALGORITMO DE BOSQUE DE DECISIÓN (RANDOM FOREST)

Define los términos Underfitting y overfitting, según se plantea en la sección 5.

Overfitting (Sobreajuste): Es un fenómeno donde un modelo coincide con los datos de entrenamiento casi a la perfección, pero no funciona bien en la validación y otros datos nuevos

Underfitting (Infra ajuste): Es cuando un modelo no logra capturar distinciones y patrones importantes en los datos, por lo tanto, se desempeña mal incluso en los datos de entrenamiento.

En tus propias palabras, explica de manera sencilla, la lógica de la regresión Random Forest según se plantea en la sección 6.

El Random Forest (Bosque aleatorio) usa muchos árboles y realiza una predicción promediando las predicciones de cada uno de los árboles, por lo que al final tiene una mejor predicción que un solo árbol

¿Qué ventaja tiene la regresión Random Forest en comparación con la regresión Decision Tree?

Obtiene una predicción mas precisa y trabaja bien con los parámetros predeterminados.

Una de las mejores características de los modelos de Random Forest es que generalmente funcionan de manera razonable incluso sin un ajuste de rendimiento.

¿El error medio absoluto obtenido en la regresión Random Forest, es menor, mayor o igual que el error medio absoluto obtenido en la regresión DecisionTree? Justifica tu respuesta.

Es menor, el MAE del Árbol de Decision es 262000.0 y el MAE obtenido con el Random Forest es 210016.0, ya que el Random Forest nos da una predicción mas precisa esta tiene menos error.