

Preguntas.

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

Sobreaajuste: capturar patrones espurios que no se repetirán en el futuro, lo que conducirá a predicciones menos precisas, o

Underfitting: no capturar patrones relevantes, nuevamente conduce a predicciones menos precisas.

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

Es uno de los muchos algoritmos de aprendizaje automático utilizados para el aprendizaje supervisado, esto significa aprender de los datos etiquetados y hacer predicciones basadas en los patrones aprendidos. RF puede usarse tanto para tareas de clasificación como de regresión.

C) ¿Qué ventaja tiene la regresión Random Forest en comparación con la regresión Decission Tree?

Random Forest usa muchos árboles, y realiza una predicción promediando las predicciones de cada árbol componente. En general, tiene una precisión predictiva mucho mejor que un solo árbol de decisión y funciona bien con los parámetros predeterminados. Si sigue modelando, puede aprender más modelos con un rendimiento aún mejor, pero muchos de ellos son sensibles para obtener los parámetros correctos.

¿El error medio absoluto obtenido en la regresión Random Forest, es menor, mayor o igual que el error medio absoluto obtenido en la regresión DecissionTree?

La regresión es menor debido a que a que Random Forest toma la decisión de forma más rápida.