

En la gráfica se muestran dos curvas de Sobreajuste/Sobre entrenamiento en aprendizaje supervisado. El error de entrenamiento se muestra en azul, mientras que el error de validación se muestra en rojo. Si el error de validación se incrementa mientras que el de entrenamiento decrece puede que se esté produciendo una situación de sobreajuste.



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

Sobreajuste es aquel donde un modelo coincide con los datos de entrenamiento casi a la perfección, pero no funciona bien en la validación y otros datos nuevos.

Adaptación insuficiente es cuando un modelo no logra capturar distinciones y patrones importantes en los datos, por lo que no se desempeña correctamente en los datos de entrenamiento

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

El árbol aleatorio no busca el ajuste ni sobreajuste, lo que hace es tener un mejor rendimiento en las decisiones por medio de muchos árboles, realizando una predicción promedio de predicciones de otros arboles, funcionando mejor con parámetros predeterminados

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

Que es más exacta la predicción comparada con otras porque evita que haya pocos valores en un árbol o bien muchos valores. Funciona de manera razonable incluso sin el ajuste.

¿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 en el Random Forest debido a que elimina las predicciones sobrantes, precisamente busca que no haya pocos datos, ni a la vez datos de más, para que sus resultados sean lo mas precisos .