¿Qué es el boosting?

Es una técnica de ensamblaje de aprendizaje automático supervisado que combina modelos débiles en un modelo más fuerte y preciso. En lugar de construir un modelo grande y complejo.

¿Cuál es la diferencia entre el bagging y boosting?

El bagging entrena modelos independientes utilizando diferentes muestras aleatorias de datos y luego combina sus predicciones para obtener una prediccion final mas precisa.

El boosting se enfoca en mejorar la precisión de un modelo mediante la creacion de un conjunto de modelos secuenciales. Cada modelo se entrena para corregir los errores del modelo anterior.

¿En qué se diferencia una potenciación del gradiente de los algoritmos tradicionales de un árbol de decisión?

El enfoque, la potenciacion del gradiente busca mejorar iterativamente el modelo mediante la adición de modelos más simples, mientras que los árboles de decisión se basan en la construcción de un solo árbol que pueda predecir la variable objetivo.

¿Cuáles hiperparámetros se pueden afinar en una potenciación del gradiente además de los hiperparámetros de cada árbol?

Fraccion de muestras, subsample, controla la fraccion de muestras de entrenamiento que se utilizan para ajustar cada árbol.

Fraccion de columnas, colsample\_bytree, controla la fraccion de columnas que se utilizan para ajustar cada árbol.

Metodo de optimizacion, controla el metodo de optimizacion utilizado para ajustar los pesos de los arboles.

¿Cómo se puede reducir el sobreajuste cuando se realiza la potenciación del gradiente?

Regularizacion, regresión de L1 y L2.

Reduccion del numero de arboles, n\_stimators, esto puede hacer que el modelo sea menos complejo y por lo tanto menos propenso al sobreajuste.

¿Por qué pueden ser mejores varios árboles débiles aleatorios que un árbol largo?

Un árbol largo, con muchas capas y nodos, puede tener un alto poder predictivo, pero tambien es propenso a sufrir de sobreajuste. En contraste un conjunto de arboles de decision debiles y aleatorios, que se combinan mediante tecnicas de ensemble, puede ofrecer una mejor capacidad de generalizacion y reducir el riesgo de sobreajuste.

¿Qué es la potenciación del gradiente?

Es un metodo de aprendizaje automatico supervisado utilizado para la clasificacion y regresion de problemas. En esencia es un conjunto de modelos de árbol de decision debil, tambien conocido como weak learnersm que se combinan en un modelo mas fuerte mediante la aplicacion de un algoritmo de optimizacion por gradiente.

¿Qué es stacking?

Tecnica de aprendizaje automatico que combina multiples modelos de aprendizaje supervisado para mejorar la precision de las predicciones.

Este combina las predicciones de varios modelos para mejorar la precision final.

¿Cuáles son las ventajas y desventajas de bagging?

Mejora la precision: Al combinar multiples modelos en un ensamblado, puede mejorar la precision de las predicciones en comparacion con un solo modelo.

Reduce el sobreajuste: Al entrenar multiple modelos en conjuntos de datos de bootstrap aleatorios, puede reducir el sobreajuste y mejorar la generalizacion del modelo.

Escalable: Puede paralelizar para mejorar la eficiencia computacional, lo que lo hace adecuado para grandes conjuntos de datos y modelos complejos.

Desventajas.

Menos interpretable: Al usar multiples modelos en lugar de uno solo, el modelo resultante puede ser menos interpretable, lo que puede dificultar la explicacion de las predicciones.

Aumenta la complejidad: Al usar multiples modelos, el baggin puede aumentar la complejidad del modelo resultante, lo que puede dificultar la depuracion y el ajuste de los hiperparametros.

Sensible a datos ruidoso: Si se usan datos ruidosos, el bagging puede generar modelos con un sesgo aun mayor hacia los valores atipicos

¿Cuáles son las ventajas y desventajas de boosting?

Mejora la precision: Al combinar multiples modelos en un ensamblado, el boosting puede mejorar significativamente la precision de las predicciones en comparacion a un solo modelo.

Reduccion del sesgo: puede reducir el sesgo y mejorar la capacidad de generalizacion del modelo.

Trata con datos desequilibrados: Puede manejar eficazmente conjuntos de datos desequilibrados, lo que significa que es util para problemas de clasificacion binaria en los que una clase es mucho menos comun que la otra.

Puede trabajar con cualquier algoritmo:

Desventajas

Sensible al ruido y valores atipicos.

Riesgo de sobreajuste

Menos interpretable.