

Modelos Vía Ensamble

XGBoost o LightGBM para especificar los parámetros para el modelo. Algunos de los parámetros importantes son:

- La función de pérdida que se va a minimizar. Para la clasificación binaria, podemos usar ``binary:logistic`` para XGBoost o ``binary`` para LightGBM.
- La métrica que se utilizará para la evaluación. Para la clasificación binaria, podemos usar "auc" para el área bajo la curva ROC o "logloss" para la pérdida logarítmica.
- El número de árboles a construir. Podemos usar un número grande y confiar en detenernos temprano para encontrar el número óptimo.
- La profundidad máxima de cada árbol. Un valor mayor puede capturar patrones más complejos, pero también aumentar el riesgo de sobreajuste.
- El factor de contracción para cada árbol. Un valor menor puede reducir la varianza, pero también aumentar el sesgo y el tiempo de entrenamiento.
- La fracción de muestras que se utilizarán para cada árbol. Un valor menor puede reducir la correlación entre los árboles, pero también aumentar la varianza.
- La fracción de características que se utilizarán para cada árbol. Un valor menor puede reducir la correlación entre las características, pero también aumentar el sesgo.
- El término de regularización L1 para los pesos. Un valor mayor puede reducir la complejidad y el sobreajuste.
- El término de regularización L2 para los pesos. Un valor mayor puede reducir la complejidad y el sobreajuste.