

## Guía de lectura

### Introduction to Statistical Learning. Springer, 2013

**Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani**

#### Capítulo 8, sec. 8.2 (sesión 7)

1. ¿Cuáles son las similitudes y las diferencias entre que usen árboles de decisión como modelos de base para **ensamble** como bagging, random forest y boosting?
2. ¿Cómo se diferencia el proceso de entrenamiento entre Bagging y Boosting en términos de cómo se ajustan los árboles de decisión?
3. ¿Cuáles son las principales diferencias entre Random Forest y XGBoost en cuanto a la forma en que se manejan los errores y se mejoran los árboles de decisión?
4. ¿Qué métodos utilizan Bagging, Boosting, Random Forest y XGBoost para reducir el sobreajuste de los árboles de decisión?
5. ¿Cómo afecta la dependencia entre los árboles de decisión al desempeño del modelo en Bagging, Boosting, Random Forest y XGBoost?
6. ¿Cuáles son las ventajas y desventajas de cada uno de estos modelos de ensamble en términos de precisión, tiempo de entrenamiento y capacidad para manejar diferentes tipos de datos?
7. ¿Cómo manejan Bagging, Boosting, Random Forest y XGBoost los datos desbalanceados o con ruidos?
8. ¿En qué escenarios específicos sería más apropiado utilizar Bagging en lugar de Boosting, o viceversa?
9. ¿Cómo se pueden interpretar las predicciones de los modelos de ensamble en comparación con los árboles de decisión individuales?
10. ¿Cuál es el efecto del tamaño del conjunto de datos en la eficacia de cada uno de estos modelos de ensamble?