

**MA477: Data Science**  
**Lesson 20 Board Sheet — 03 March 2026**  
United States Military Academy, West Point  
Instructor: MAJ Patrick Kuiper

---

## 1 Decision Tree Bagging Classification Lesson Objectives

- Understand how bagging improves basic decision trees by reducing variance.
- Define and interpret the out-of-bag error.
- Use functions in Python's sklearn module to implement bagging.

### Discussion Questions:

1. Why do deep decision trees tend to have high variance?
2. Why does increasing  $B$  in bagging not increase overfitting?
3. What assumption about the individual trees is required in order for variance to decrease by a factor of  $B$ ?