**Lab 6**

1. **Precision increased** (0.4486 → **0.5191**) → Fewer false positives, meaning the model is making more reliable "Yes" predictions.

**Recall slightly decreased** (0.9289 → **0.8103**) → The model is catching fewer "Yes" cases than before, but it's a more balanced trade-off.

**F1 Score improved** (0.6051 → **0.6328**) → Better balance between precision and recall.

**Accuracy improved** (0.8634 → **0.8941**) → Overall classification performance increased.

SMOTE helped by **reducing bias towards the majority class**, improving **precision and F1-score**, while keeping recall high. This is a **better trade-off for a more balanced model**

**2)**

* n\_trees = range (1, 200, 20): Best 181 trees, Test RMSE = 0.5048
* n\_trees = range (1, 501, 20): Best 441 trees, Test RMSE = 0.5022

**RMSE Decreases Initially:**

* Increasing trees reduces variance and improves predictions (from 181 → 441 trees).
* However, the RMSE improvement is marginal (from **0.5048 → 0.5022**, just **0.5% better**).

**Adding more** trees to Random Forest can improve accuracy, but the returns diminish after a certain number of trees.

A very large number of trees leads to **increased computational cost** and **minimal performance gains** beyond a point.

For practical applications, a moderate number of trees (e.g., 100 to 300 trees) is often sufficient, and beyond that, the increase in model size and training time may not be justified by the small improvement in accuracy.