

Assignment 5  
Eva DeThomas

Dataset with 10, 25, and 50 estimators using both gini and entropy as separators:

Results for model Random Forest:

	param_n_estimators	param_criterion	mean_test_score	mean_fit_time
0	10	gini	0.977528	0.121841
1	20	gini	0.977528	0.163288
2	50	gini	0.971910	0.345822
3	10	entropy	0.960674	0.084749
4	20	entropy	0.966292	0.164746
5	50	entropy	0.971910	0.339796

- Test 5, 10, 15, and 20 estimators for Random Forest
- test 25, 50, 75, and 100 iterations for Histogram Boosting
- do 5 splits.



4)

a)

Terrain time per km: Sandy = 20 (60 / 3), Smooth = 12 (60 / 5), Rocky = 30 (60 / 2)

$$1) (0.2 * 20) + (0.3 * 12) + (0.5 * 30) = 22.6 * 5 = 113$$

$$2) (0.4 * 20) + (0.2 * 12) + (0.4 * 30) = 22.4 * 7 = 156.8$$

$$3) (0.5 * 20) + (0.4 * 12) + (0.1 * 30) = 17.8 * 6 = 106.8$$

**Best route = Route 3! 106.8 minutes**

b)

$$\text{Route 1: } (0.3 * (15)) + (0.7 * -20) = -9.5$$
$$113 - 9.5 = 103.5$$

$$\text{Route 3: } (0.6 * 40) = 24$$
$$106.8 + 24 = 130.8$$

$$\text{Route 2} = 156.8$$

**New best route = Route 1!**

c)

Route 2:

(Not rocky):

Current probability:

$$20\% + 40\% = 60\%$$

Current probability:

$$20 / 60 + 40 / 60$$

Need:

$$X / 100 + y / 100$$

So...

If  $20 / 60 = x / 100$ ...

$$X = 20 * 100 / 60 = 33.333$$

$$Y = 40 * 100 / 60 = 66.667$$

New utility of route 2 w/balanced probabilities =

$$(0.666 * 20) + (0.333 * 12) = 17.316 * 7 = \mathbf{121.212 \text{ mins}}$$

**Route 2 without rocky = 121.212 mins**

d)

Probability of the satellite will find that it is rocky (assuming the satellite exists):

**0.6 or 60%** because that is the probability that it is not rocky is 100% - 40% (probability it is not rocky)

e)

$$7 * 30 = 210 \text{ minutes}$$

f)

At this point:

Route 1 is 103.5

Route 2 could be 210 or 121.212

Route 3 is 130.8

**In this case, no need to wait for the satellite because Route 1 has the best utility still.**