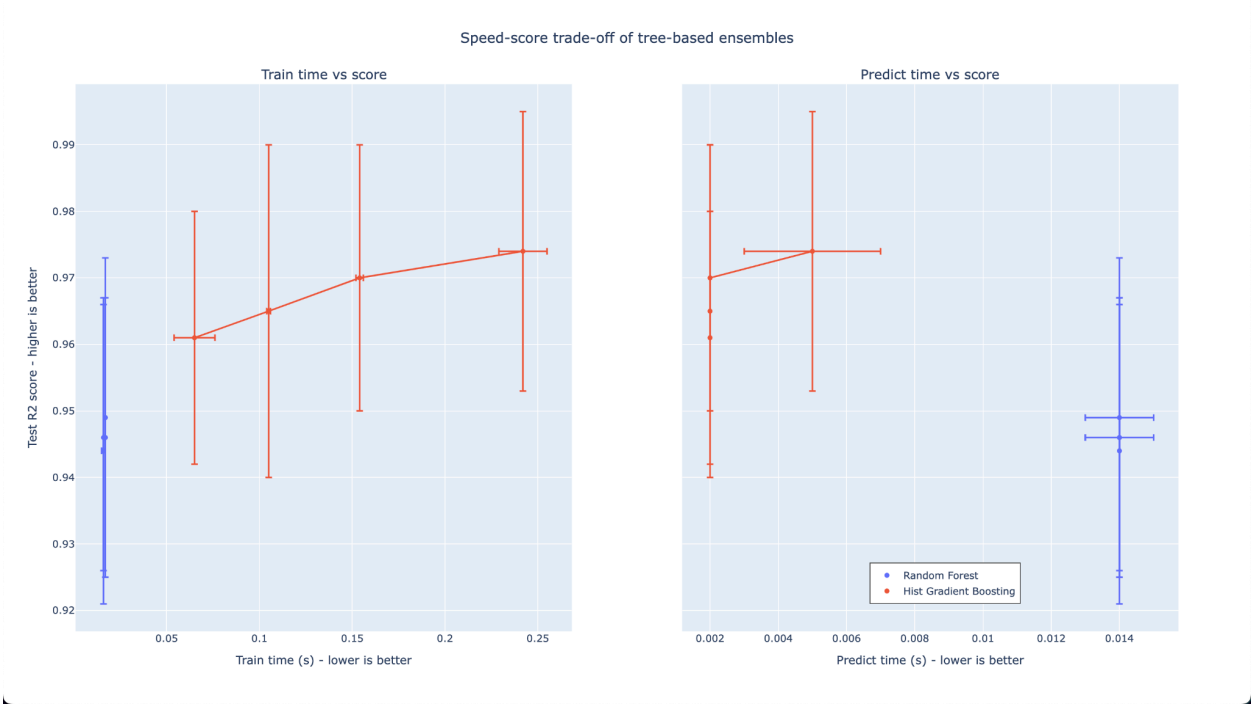
Theo Lecuyer
Foundations of AI
11/04/2024

Problem 1.

Using the digits dataset instead of the iris dataset

Estimators	Gini	Entropy
10	[0.91, 0.84, 0.93, 0.94, 0.89]	[0.90, 0.86, 0.91, 0.92, 0.89]
25	[0.91, 0.90, 0.94, 0.96, 0.91]	[0.93, 0.90, 0.95, 0.95, 0.91]
50	[0.94, 0.90, 0.95, 0.94, 0.89]	[0.93, 0.90, 0.95, 0.96, 0.92]

Graph for comparing the ensemble methods of Random forest and Histogram Gradient boosting



Random forest results in a much faster training time (more than 5x faster), but a Prediction time that is much slower and slightly less accurate. Histogram Gradient boosting provides the opposite, a slower upfront training time, but a much faster, more accurate prediction.

Question 4, Utility
(second page)

Prior info:
Over rocky terrain 30 min/km, over sandy 20 min/km, and over smooth 12 min/km.
Route 1 is 5 km long. There is a 20% chance it is sandy, 30% chance it is smooth, and a 50% chance it is rocky.
Route 2 is 7 km long. There is a 40% chance it is sandy, a 20% chance it is smooth, and a 40 % chance it is rocky.
Route 3 is 6 km long. There is a 50% chance it is sandy, a 40% chance it is smooth, and a 10% chance it is rocky.

1. Which route should we pick? Show your work. You may find it easier to convert km/hr to mins/km (how many minutes does it take to go 1km on each surface?)
(Rocky 30, sandy 20, smooth 12)
Route 1: (5km long) .5(rocky)*150min+.3(smooth)*60min+.2(sandy)*100min = 113 mins
Route 2: (7km long) .4(rocky)*210min+.2(smooth)*84min+.4(sandy)*140min = 156.8 mins
Route 3: (6km long) .1(rocky)*180min+.4(smooth)*72min+.5(sandy)*120min = 106.8 mins

The route we should pick with the shortest average time would be route 3.

2.
Route 1 contains a crater. If the wall of the crater is intact, we can take a shortcut through the crater, which will save 20 minutes. If the wall has been damaged, we will need to go around, which will add 15 minutes to our journey. There is a 30% chance that the wall is damaged.

Route 3 contains a bridge. If that bridge is damaged, we will need to repair it, which will add 40 minutes to our time. There is a 60% chance that the bridge is damaged.

Update your estimates for the travel time for each route. Now which route seems best?
Route 1: (5km long) .5(rocky)*150min+.3(smooth)*60min+.2(sandy)*100min = 113 mins
Adjusted utility for route 1:
.7*-20+.3*15 = -9.5 (Average time saved of 9.5 mins)
So route 1 113-9.5 = 103.5 mins
Route 2: (7km long) .4(rocky)*210min+.2(smooth)*84min+.4(sandy)*140min = 156.8 mins
Route 2 stays the same
Route 3: (6km long) .1(rocky)*180min+.4(smooth)*72min+.5(sandy)*120min = 106.8 mins
Adjusted utility for route 3:
(.60*40)+(0.4*0)=24 (Average added time of 24 mins)
So route 3 106.8+24 = 130.8 mins

Now adding the crater and bridge stipulations, the best route would be route 1.

There is an orbiting satellite that can tell us whether route 2 is rocky or not. If not, that would be great news, and would make it much more appealing! The only problem is that the satellite is not yet in position. How long should we wait for the satellite?

(2 points) First: If the satellite said that route 2 was not rocky, how long would we expect it to take?
If it said route 2 was not rocky, we could then expect:
(Redistribute the probabilities for rocky based on 20/60 and 40/60
Route 2: (7 km long) 0.6667 (sandy)×140 min+0.3333 (smooth)×84 min=121.3 mins

(3 points) Second: What's the probability that the satellite will tell us this?
Since we first estimated that the route was rocky with a probability of 40%, that means the probability the satellite tells us it isn't rocky is the complement.
So: 1-0.4 = 0.6
Therefore, there is a 60% probability the satellite will tell us it isn't rocky

(2 points) Third: If the satellite tells us route 2 is in fact rocky, what do we do? How long will that take?
If the satellite tells us that route 2 is rocky, then we would just calculate the distance and speed.
The speed for rocky terrain is 30 min/km, so for the 7 km route 2 if it were rocky we can expect
7 * 30 = **210 min**

(3 points) Last: given all of this, how long should we wait for the satellite?
The
Given this, we should not wait for the satellite. As the best outcome for route 2, even if we can guarantee it is not rocky, is still slower than the best result for route 1 which is 103.5 mins.