

Colin Pham

Professor Brooks

Foundations of AI

13 November 2024

Assignment 5: Working with Uncertainty

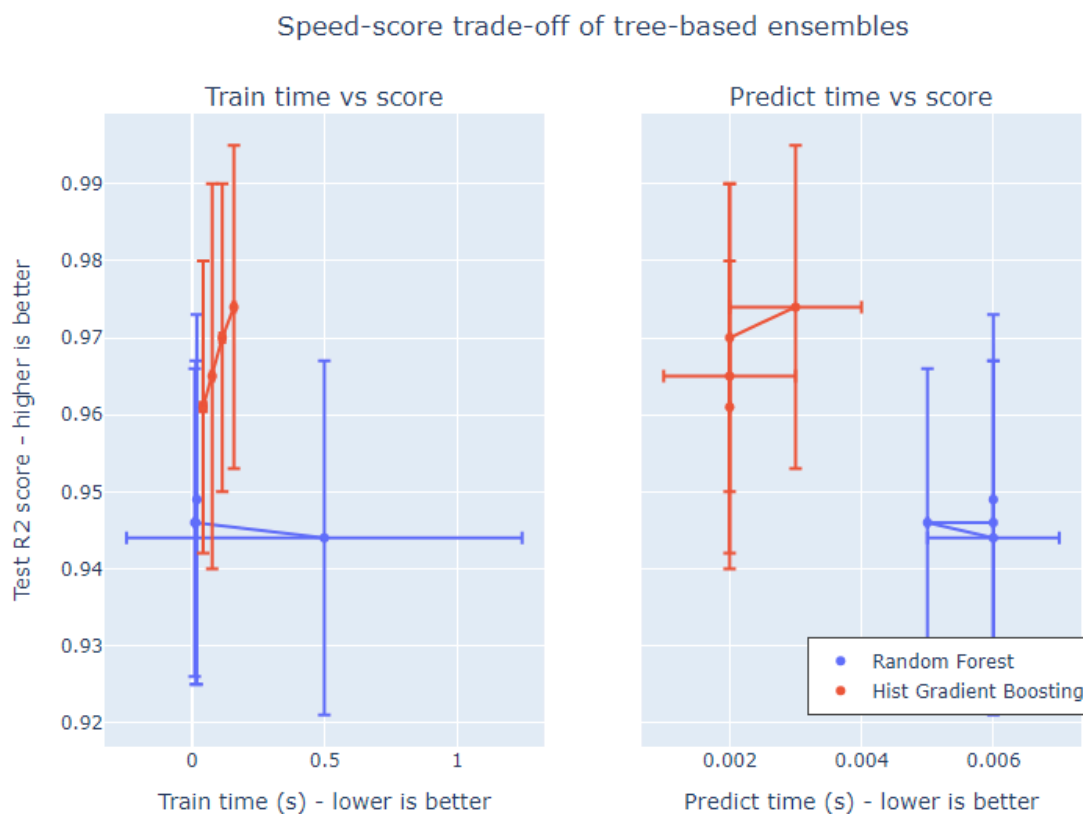
Problem 1: Decision Trees in scikit-learn

Run RandomForestClassifier on your dataset with 10, 25, and 50 estimators using both gini and entropy as separators

Legend: N = n_estimators. S = Separators. Can be G for Gino or E for Entropy

N:10 S:G	N:10 S:E	N:25 S:G	N:25 S:E	N:50 S:G	N:50 S:E
0.908333333 3333333	0.888888888 8888888	0.933333333 3333333	0.911111111 111111	0.930555555 5555556	0.95
0.841666666 6666667	0.908333333 3333333	0.902777777 7777778	0.883333333 3333333	0.911111111 111111	0.908333333 3333333
0.930362116 9916435	0.930362116 9916435	0.952646239 5543176	0.961002785 5153204	0.963788300 8356546	0.952646239 5543176
0.935933147 632312	0.933147632 3119777	0.958217270 194986	0.969359331 4763231	0.961002785 5153204	0.969359331 4763231
0.894150417 8272981	0.891364902 5069638	0.905292479 1086351	0.922005571 0306406	0.919220055 7103064	0.935933147 632312

Doing hyperparameter search by hand is annoying. Sklearn makes it easy for us with the GridSearchCV class. This lets us provide a list of models and hyperparameters, and it does cross-validation for each model and parameter combination and summarizes the result. Very handy!



Woah! Look at that forest go (you can't because it's so fast)!

Problem 4: Utility

min/km

Rocky 2 km/hr. X ... 30 min/km
 Sandy 3 km/hr. ~ ... 20 min/km
 Smooth 5 km/hr. ★ ... 12 min/km

Routes

1: 5 km
 20% sandy
 30% smooth
 50% rocky

2: 7 km
 40% Sandy
 20% smooth
 40% rocky

3: 6 km
 50% Sandy
 30% Smooth
 10% Rocky

Which Route?

1:

$$\begin{array}{r} (0.2 \times 100) \\ (0.3 \times 60) + \\ (0.5 \times 150) \\ \hline 20 + 18 + 75 = \\ \underline{113 \text{ min.}} \end{array}$$

2:

$$\begin{array}{r} (0.4 \times 140) \\ (0.2 \times 84) + \\ (0.4 \times 210) \\ \hline 56 + 16.8 + 84 = \\ \underline{156.8 \text{ min.}} \end{array}$$

3:

$$\begin{array}{r} (0.5 \times 120) \\ (0.3 \times 72) + \\ (0.1 \times 180) \\ \hline 60 + 21.6 + 18 = \\ \underline{99.6 \text{ min.}} \end{array}$$

Choose Route 3 w/ ≈ 99.6 min

Colin Pham

Which route should we pick?

- Route 3's odds outputted a calculated time of 99.6 minutes, which is less than the other routes of time 113 minutes and 156.8 minutes. **Route 3 is the best**

min/km

Rocky 2 km/hr. X ... 30 min/km
 Sandy 3 km/hr. ~ ... 20 min/km
 Smooth 5 km/hr. ★ ... 12 min/km

Routes

1: 5 km
 20% Sandy
 30% Smooth
 50% Rocky

2: 7 km
 40% Sandy
 20% Smooth
 40% Rocky

3: 6 km
 50% Sandy
 30% Smooth
 10% Rocky

Crater
 70% intact = -20 min
 30% damaged = +15 min

Bridge
 60% good = +40 min
 40% damaged = ± 0 min

1:
 (0.2×100)
 $(0.3 \times 60) +$
 (0.5×150)
 $20 + 18 + 75 =$
113 min.

Crater: $(0.7 \times -20) + (0.3 \times 15)$
 $= -14 + 4.5 = -9.5 \text{ min}$
 $113 - 9.5 = \underline{103.5 \text{ min.}}$

2:
 (0.4×140)
 $(0.2 \times 84) +$
 (0.4×210)
 $56 + 16.8 + 84 =$
156.8 min.

3:
 (0.5×120)
 $(0.3 \times 72) +$
 (0.1×180)
 $60 + 21.6 + 18 =$
99.6 min.

Bridge: $(0.4 \times 0) + (0.6 \times 40)$
 $= 24 \text{ min.}$
 $99.6 + 24 = \underline{123.6 \text{ min.}}$

Now Route 1 is the best choice w/ $\approx 103.5 \text{ min.}$

QNT

Now which route seems best?

- With the crater and the bridge cases for routes 1 and 3 respectively, Route 1's time decreased by 9.5 minutes and route 3's time increased by 24 minutes. **Route 1 became the more favorable route with a time of 103.5 minutes**

min/km

Rocky 2km/hr. X ... 30 min/km
 Sandy 3km/hr. ~ ... 20 min/km
 Smooth 5km/hr. ★ ... 12 min/km

Routes

1: 5km
 20% sandy
 30% smooth
 50% rocky

2: 7km
 40% Sandy
 20% smooth
 40% rocky

3: 6km
 50% Sandy
 30% Smooth
 10% Rocky

Crater
 70% intact = -20 min
 30% damaged = +15 min

1:
 (0.2×100)
 $(0.3 \times 60) +$
 (0.5×150)
 $20 + 18 + 75 =$
 113 min.

Crater: $(0.7 \times -20) + (0.3 \times 15)$
 $= -14 + 4.5 = -9.5 \text{ min}$
 $113 - 9.5 = 103.5 \text{ min.}$

Satellite
 If 2 NOT rocky,
 MIN travel time = $12 \times 7 = 84 \text{ min.}$
 MAX travel time = $20 \times 7 = 140 \text{ min.}$
 Uh Oh!

Satellite
 tells us IS rocky: 40%
 tells us IS NOT rocky: 60%

IF rocky, travel time = $30 \times 7 = 210 \text{ min.}$
 I would rather NOT wait
 for the satellite! $210 > 103.5 \text{ route 1}$
 $210 > 123.6 \text{ route 2}$

Rocky = NO WAIT (210 min)
 Sandy = NO WAIT (140 min)
 Smooth = WAIT (84 min)
 $103.5 - 84 = 19.5 \text{ min}$
 (Route 1) Wait time
 (no more than) ↗

Bridge
 60% damaged = +40 min
 40% good = ± 0 min

3:
 (0.5×120)
 $(0.3 \times 72) +$
 (0.1×180)
 $60 + 21.6 + 18 =$
 99.6 min.

Bridge: $(0.4 \times 0) + (0.6 \times 40)$
 $= 24 \text{ min.}$
 $99.6 + 24 = 123.6 \text{ min.}$

- John

First: If the satellite said that route 2 was not rocky, how long would we expect it to take?

- If route 2 is NOT rocky, it can either be smooth or sandy. A smooth path gives us an 84 minutes of travel while a sandy path gives us 140 minutes of travel

Second: What's the probability that the satellite will tell us this?

- A satellite telling us the path is NOT rocky would be the probability of smooth (20) plus the probability of sandy (40) equalling to 60%

Third: If the satellite tells us route 2 is in fact rocky, what do we do? How long will that take?

- If route 2 IS rocky, then the travel time will be **210 minutes**. How awful! Literally any other route would be a better choice at this point. **I would take route 1 at that point**

Last: given all of this, how long should we wait for the satellite?

- The satellite telling us it is NOT rocky and taking our chances to get a smooth path, our travel time would be 84 minutes. Taking that time and subtracting it from the next best path, route 1, of time 103.5, we would have to wait **no more than 19.5 minutes for a satellite response**. If the time exceeds, take route 1.
- However, there is STILL a possibility that the path can be sandy, making it still slower than route 1. Some more math has to be done here. After the 19.5 minute wait time: $(0.66666 \times 140) + (0.333333 \times 84) = 121.332$ minutes. ($\frac{2}{3}$ chance for sandy and $\frac{1}{3}$ chance for smooth) Wow! **Route 1 is still technically better**. Man. I would take route 1.