Homesite HOME INSURANCE

$20,000 • 1,534 teams

Homesite Quote Conversion

Merger and 1st Submission Deadlin

Mon 9 Nov 2015 Mon 8 Feb 2016 (16 days to go)

How to prevent a messed up leaderboard in the end?

Start Watching

0

After the end of the Rossmann Store Sales challenge it appears, that alot of people badly overfit the public LB. So there are changes from top 10 to 1200 or something. So how can one prevent this in this challenge where a change of 0.00001 can make a huge difference? Or will this challenge end totally random?

#1 | Posted 38 days ago

Sterby

5

I'm not as concerned about it for this competition. The thing that made Rossmann challenging was the way the data was split -- the first 3 weeks of August for the Public LB, and then the next 5 weeks or whatever it was for the Private LB. It was apparent from almost the very beginning that overfitting would be a problem (different random seeds producing very different LB results, for example). It would appear that this data is randomly split with no time component to worry about. For that reason, the data we're using to train is probably more representative of what we're testing on.

#2 | Posted 38 days ago

Branden Murray

1

I agree with Branden. Having said that, a good way to measure the quality of your model is to take a weighted average of your CV score and your public LB score. The weights should be based on the size of the training set and the size of the 30% of the test set that the public LB is based upon.

#3 | Posted 38 days ago

BreakfastPirate

0

BreakfastPirate wrote:

> I agree with Branden. Having said that, a good way to measure the quality of your model is to take a weighted average of your CV score and your public LB score. The weights should be based on the size of the

> training set and the size of the 30% of the test set that the public LB is based upon.

@BreakfastPirate is there any formula for calculating it? I saw this idea somewhere else, but I cannot find it. Thanks.

#4 | Posted 25 days ago

Bazinga!

4

Bazinga! wrote:

> **BreakfastPirate wrote:**
>
> > I agree with Branden. Having said that, a good way to measure the quality of your model is to take a weighted average of your CV score and your public LB score. The weights should be based on the size of the training set and the size of the 30% of the test set that the public LB is based upon.
>
> @BreakfastPirate is there any formula for calculating it? I saw this idea somewhere else, but I cannot find it. Thanks.

It's just

$$[CV score * \frac{n_{train}}{n_{train} + LB\% * n_{test}}] + [LB score * \frac{LB\% * n_{test}}{n_{train} + LB\% * n_{test}}]$$

isn't it? Where LB% is the percentage of the test set used on the public LB. Unless I've misunderstood.

[Edit] I suppose this assumes you've done k-fold CV and have out-of-fold predictions for your entire training set. If you have a single validation set then n_train would be n_validation instead.

#5 | Posted 25 days ago

Branden Murray

0

Branden Murray wrote:

> isn't it? Where LB% is the percentage of the test set used on the public LB. Unless I've misunderstood.
>
> [Edit] I suppose this assumes you've done k-fold CV and have out-of-fold predictions for your entire training set. If you have a single validation set then n_train would be n_validation instead.

Correct.

#6 | Posted 25 days ago

BreakfastPirate

Reply

+ Add attachment(s) ...

Post Reply Email me when someone replies

Start Watching « Back to forum