AWS Certified Machine Learning Specialty MLS-C01 [2023]

Back to All Questions

DB

Different method signature (defaults?) for XGBRegressor

0

David · Lecture 71 · 35 minutes ago

Hey Chandra. When I inspected `regressor` (defined in the lecture's notebook with `regressor = xgb.XGBRegressor()`), I got more parameters as well as parameters with different values that were shown on screen. Most intriguing to me was the regression.booster; mine was None. Does that mean my tree was not boosted?

Here's what was in my Jupyter notebook

```
In [n]:
--------------------
# Default Options #DWB# shown by inspecting regressor
#DWB# I don't like the format from just feeding in the name
#regressor
print(str(regressor)) # more than were there in Chandra's lecture
```

```
print(  (f"\nFor my default,\n  regression.booster = {regressor.booster},\n"
         "whereas in Chandra's lecture\n"
         "(\"Lab - Training Simple Regression\",\n"
         "at 2:08 in the video, as seen 20230720T185200-0600),\n"
         "I saw\n  regression.booster = 'gbtree'."
        )
     )
--------------------
```

[Output]

```
XGBRegressor(base_score=None, booster=None, callbacks=None,
             colsample_bylevel=None, colsample_bynode=None,
             colsample_bytree=None, early_stopping_rounds=None,
             enable_categorical=False, eval_metric=None, feature_types=None,
             gamma=None, gpu_id=None, grow_policy=None, importance_type=None,
             interaction_constraints=None, learning_rate=None, max_bin=None,
             max_cat_threshold=None, max_cat_to_onehot=None,
             max_delta_step=None, max_depth=None, max_leaves=None,
             min_child_weight=None, missing=nan, monotone_constraints=None,
             n_estimators=100, n_jobs=None, num_parallel_tree=None,
             predictor=None, random_state=None, ...)

For my default,
  regression.booster = None,
whereas in Chandra's lecture
("Lab - Training Simple Regression",
at 2:08 in the video, as seen 20230720T185200-0600),
we saw
  regression.booster = 'gbtree'.
```

Is this something to be concerned about?

1 reply **Following replies**

DB

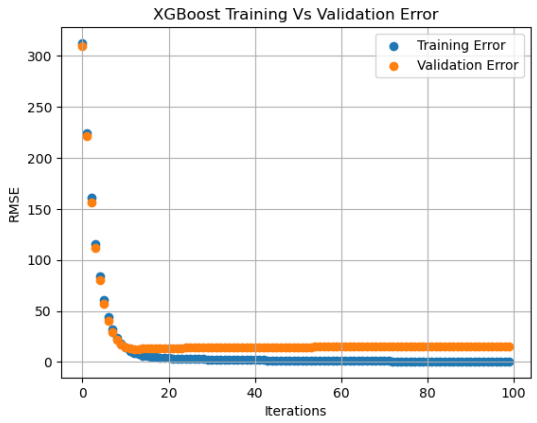
David

A few seconds ago

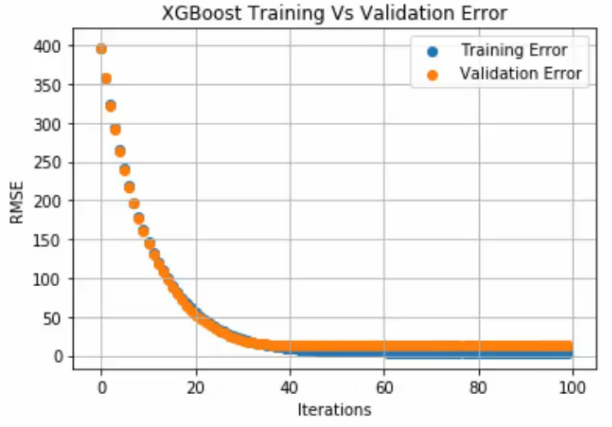
0

(Just for completeness, without making one post too long...)

The plots (mine and Chandra's) for XGBoost Training Vs Validation Error:



Mine



Chandra's

Corresponding RMSE Values at the end

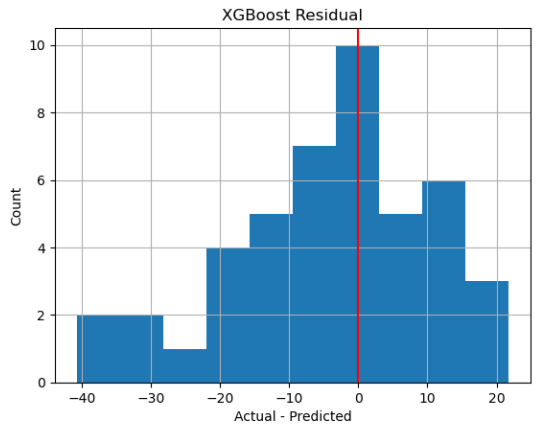
```
[95]    validation_0-rmse:0.43824       validation_1-rmse:15.05033
[96]    validation_0-rmse:0.43266       validation_1-rmse:15.05148
[97]    validation_0-rmse:0.41700       validation_1-rmse:15.05335
[98]    validation_0-rmse:0.40495       validation_1-rmse:15.05614
[99]    validation_0-rmse:0.39714       validation_1-rmse:15.06363
```

```
[95]    validation_0-rmse:3.5712        validation_1-rmse:13.6678
[96]    validation_0-rmse:3.56877       validation_1-rmse:13.6706
[97]    validation_0-rmse:3.56182       validation_1-rmse:13.6723
[98]    validation_0-rmse:3.54583       validation_1-rmse:13.6715
[99]    validation_0-rmse:3.50536       validation_1-rmse:13.6938
```

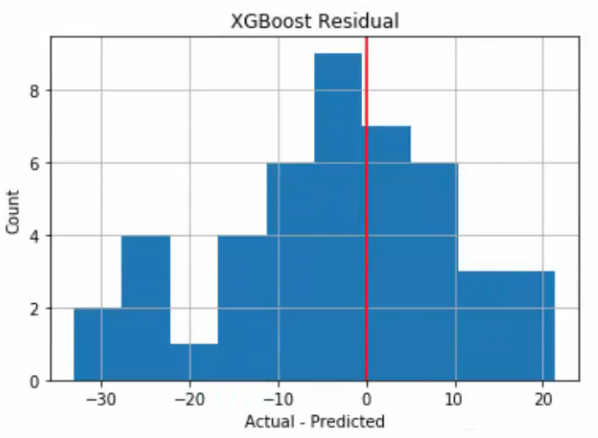
The plots for XGBoost Residual

(As before, my stuff is on the left, Chandra's stuff is on the right.)

```
XGBoost Algorithm Metrics
 Mean Squared Error: 226.91
 Root Mean Square Error: 15.06
```




```
XGBoost Algorithm Metrics
 Mean Squared Error: 187.52
 Root Mean Square Error: 13.69
```



These are non-trivially different, even when taking into account the differences in x-range and y-range. Does it matter?

DB

Add reply

Teach on Udemy

English