

xgboost

```
library(mlr3)
library(mlr3learners)
library(mlr3filters)
library(mlr3pipelines)
library(tidyverse)
library(ggplot2)
```

xgboost Upload Rate Prediction

```
data_dir = "../datasets/"

dataset_ul = read_csv(
  str_c(data_dir, "dataset_ul.csv"),
  col_types = cols(scenario=col_factor())
)

dataset_ul_prediction = dataset_ul %>% select(
  scenario,
  velocity_mps,
  acceleration_mpss,
  rsrp_dbm,
  rsrq_db,
  rssnr_db,
  cqi,
  ta,
  payload_mb,
  f_mhz,
  throughput_mbits
)

# remove missing values
dataset_ul_prediction = dataset_ul_prediction %>% drop_na()
glimpse(dataset_ul_prediction)
```

```
## Rows: 6,168
## Columns: 11
## $ scenario      <fct> campus, campus, campus, campus, campus, campus, c...
## $ velocity_mps  <dbl> 11.80, 11.83, 11.70, 11.45, 11.49, 7.93, 8.15, 8....
## $ acceleration_mpss <dbl> 0.13, 0.03, 0.06, -0.32, -0.26, 0.23, 0.24, 0.32,...
## $ rsrp_dbm      <dbl> -99, -85, -121, -84, -97, -96, -74, -108, -111, -...
## $ rsrq_db       <dbl> -9, -5, -15, -6, -12, -12, -5, -9, -13, -11, -6, ...
## $ rssnr_db      <dbl> -1, 22, -8, 11, -2, 5, 29, 2, 6, 11, 13, 16, -3, ...
## $ cqi           <dbl> 8, 10, 4, 13, 9, 5, 15, 2, 6, 15, 12, 9, 6, 11, 1...
## $ ta            <dbl> 9, 7, 63, 4, 7, 7, 4, 21, 16, 7, 4, 4, 7, 16, 4, ...
## $ payload_mb    <dbl> 1.0, 4.0, 6.0, 2.0, 6.0, 5.0, 4.0, 10.0, 0.1, 7.0...
```

```

## $ f_mhz          <dbl> 1750, 1720, 1770, 1720, 1750, 1750, 1720, 1770, 1...
## $ throughput_mbits <dbl> 4.66, 24.52, 1.29, 14.86, 3.97, 6.52, 16.27, 3.18...

task = TaskRegr$new(
  id = "ul_prediction",
  backend = dataset_ul_prediction,
  target = "throughput_mbits"
)
task

## <TaskRegr:ul_prediction> (6168 x 11)
## * Target: throughput_mbits
## * Properties: -
## * Features (10):
##   - dbl (9): acceleration_mpss, cqi, f_mhz, payload_mb, rsrp_dbm,
##     rsrq_db, rssnr_db, ta, velocity_mps
##   - fct (1): scenario

factor_encoding = po("encode", method = "one-hot", affect_columns = selector_type("factor"))
xgboost = lrn("regr.xgboost", nrounds=10)
pipe = factor_encoding %>% xgboost
learner = GraphLearner$new(pipe)

resampling = rsmpl("holdout", ratio = 0.8)
result = resample(
  task = task,
  learner = learner,
  resampling = resampling
)

## INFO [14:01:57.035] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:01:57] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated

result

## <ResampleResult> of 1 iterations
## * Task: ul_prediction
## * Learner: encode.regr.xgboost
## * Warnings: 0 in 0 iterations
## * Errors: 0 in 0 iterations

# get  $r^2$ 
result$aggregate(msr("regr.rsq"))

## regr.rsq
## 0.8319859

# get MSE
result$aggregate(msr("regr.mse"))

## regr.mse
## 13.76918

# get MAE
result$aggregate(msr("regr.mae"))

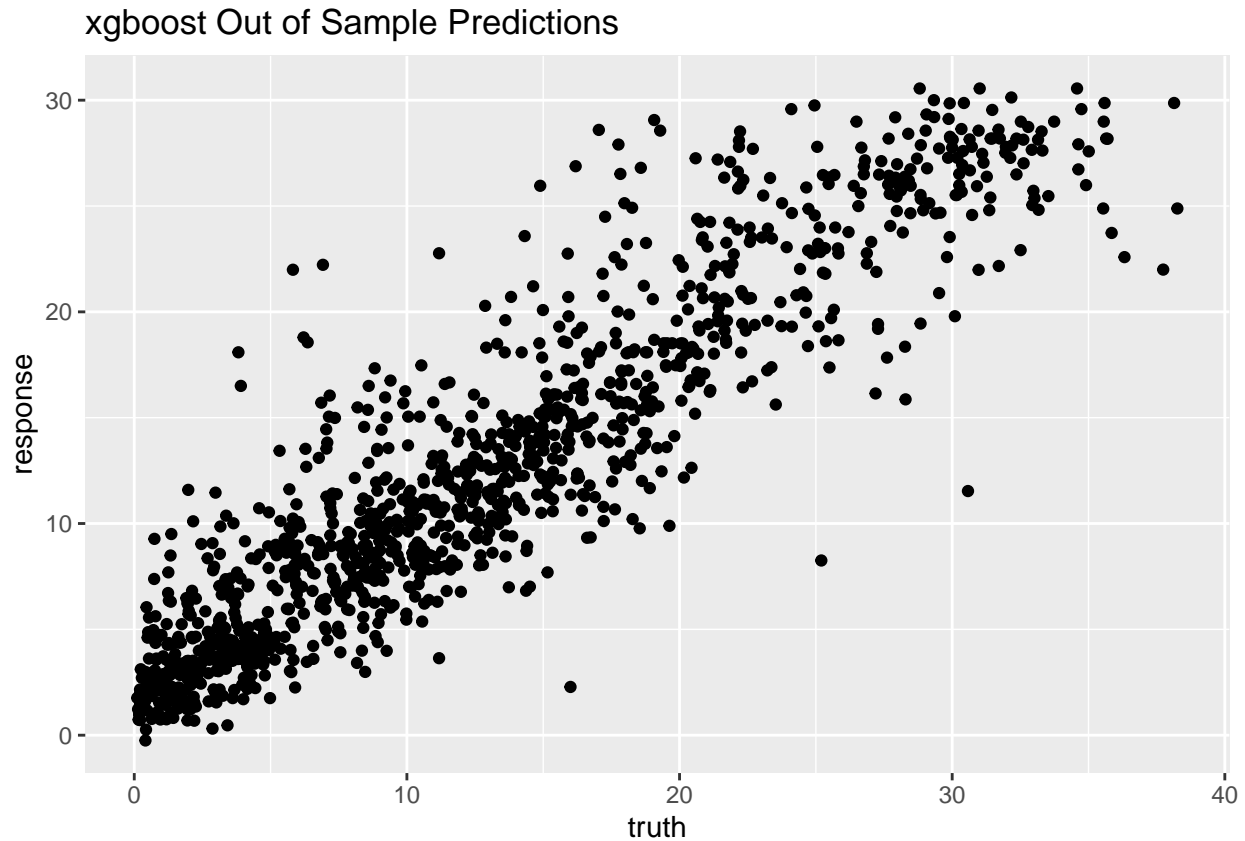
## regr.mae
## 2.691314

```

```

predictions = as.data.table(result$prediction())
ggplot(predictions) +
  geom_point(aes(x=truth, y=response)) +
  ggtitle("xgboost Out of Sample Predictions")

```



Feature Importance

```

filter_permutation = flt("permutation",
  learner = learner,
  resampling = rsmp("holdout", ratio=0.8),
  measure = msr("regr.mae"),
  standardize = TRUE,
  nmc=10
)
filter_permutation$calculate(task)

```

```

## INFO [14:01:58.676] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:01:58] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:01:59.078] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:01:59] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:01:59.520] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:01:59] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:01:59.963] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:00] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:00.413] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)

```


[illegible]

```
## [14:02:39] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:39.910] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:39] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:40.349] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:40] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:40.786] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:40] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:41.221] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:41] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:41.663] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:41] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:42.148] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:42] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:42.597] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:42] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:43.059] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:43] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:43.509] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:43] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:43.983] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:44] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:44.455] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:44] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:44.910] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:44] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:45.376] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:45] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:45.824] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:45] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
## INFO [14:02:46.288] Applying learner 'encode.regr.xgboost' on task 'ul_prediction' (iter 1/1)
## [14:02:46] WARNING: amalgamation/./src/objective/regression_obj.cu:174: reg:linear is now deprecated
```

```
filter_permutation_results = as.data.table(filter_permutation)
filter_permutation_results
```

```
##           feature      score
## 1:      payload_mb 1.8698401616
## 2:           f_mhz 0.9682512989
## 3:         rsrp_dbm 0.4047572336
## 4:              ta 0.0385780754
## 5:         rsrq_db 0.0121329598
## 6: acceleration_mpss 0.0058260855
## 7:         scenario 0.0001899041
## 8:         rssnr_db -0.0168326892
## 9:    velocity_mps -0.0219049419
## 10:              cqi -0.0343528891
```

```
ggplot(filter_permutation_results) +
  geom_bar(aes(x = reorder(feature, -score), y = score), stat="identity") +
  xlab("feature") +
  ylab("MAE difference") +
  scale_x_discrete(guide = guide_axis(angle = 20)) +
  ggtitle("Permutation Feature Importance")
```

Permutation Feature Importance

