Random Forest

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

Upload Rate Prediction

```
data_dir = "../datasets/"
dataset_ul = read_csv(str_c(data_dir, "dataset_ul.csv"), col_types = cols())
dataset_ul_prediction = dataset_ul %>% select(
  scenario,
  provider,
  velocity_mps,
  acceleration_mpss,
  rsrp_dbm,
  rsrq_db,
  rssnr_db,
  cqi,
  ta,
  throughput_mbits
)

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

```
## Rows: 6,168
## Columns: 10
## $ scenario          <chr> "campus", "campus", "campus", "campus", "campus",...
## $ provider          <chr> "o2", "tmobile", "vodafone", "tmobile", "o2", "o2...
## $ 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, ...
## $ 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 10)
## * Target: throughput_mbits
## * Properties: -
## * Features (9):
##   - dbl (7): acceleration_mpss, cqi, rsrp_dbm, rsrq_db, rssnr_db, ta,
##     velocity_mps
##   - chr (2): provider, scenario
```

```
learner_random_forest = mlr_learners$get("regr.ranger")
```

```
resampling = rsmp("holdout", ratio = 0.8)
result = resample(
  task = task,
  learner = learner_random_forest,
  resampling = resampling
)
```

```
## INFO  [19:59:05.112] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
```

```
result
```

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

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

```
##  regr.rsq
## 0.5084884
```

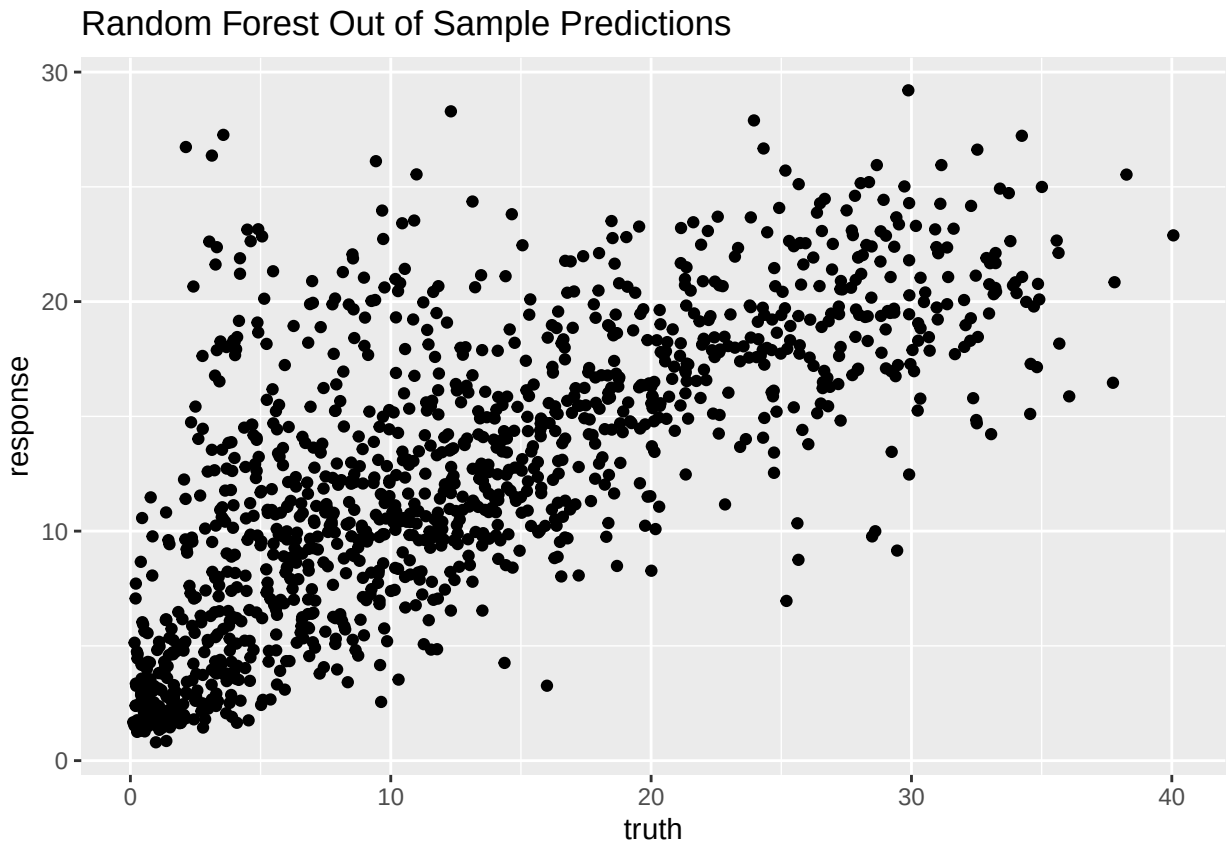
```
# get MSE
result$aggregate(msr("regr.mse"))
```

```
## regr.mse
## 43.11964
```

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

```
## regr.mae
## 4.926147
```

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



Feature Importance

```
filter_permutation = flt("permutation",
  learner = learner_random_forest,
  resampling = rsmp("holdout", ratio=0.8),
  measure = msr("regr.mae"),
  standardize = TRUE,
  nmc=5
)
filter_permutation$calculate(task)
```

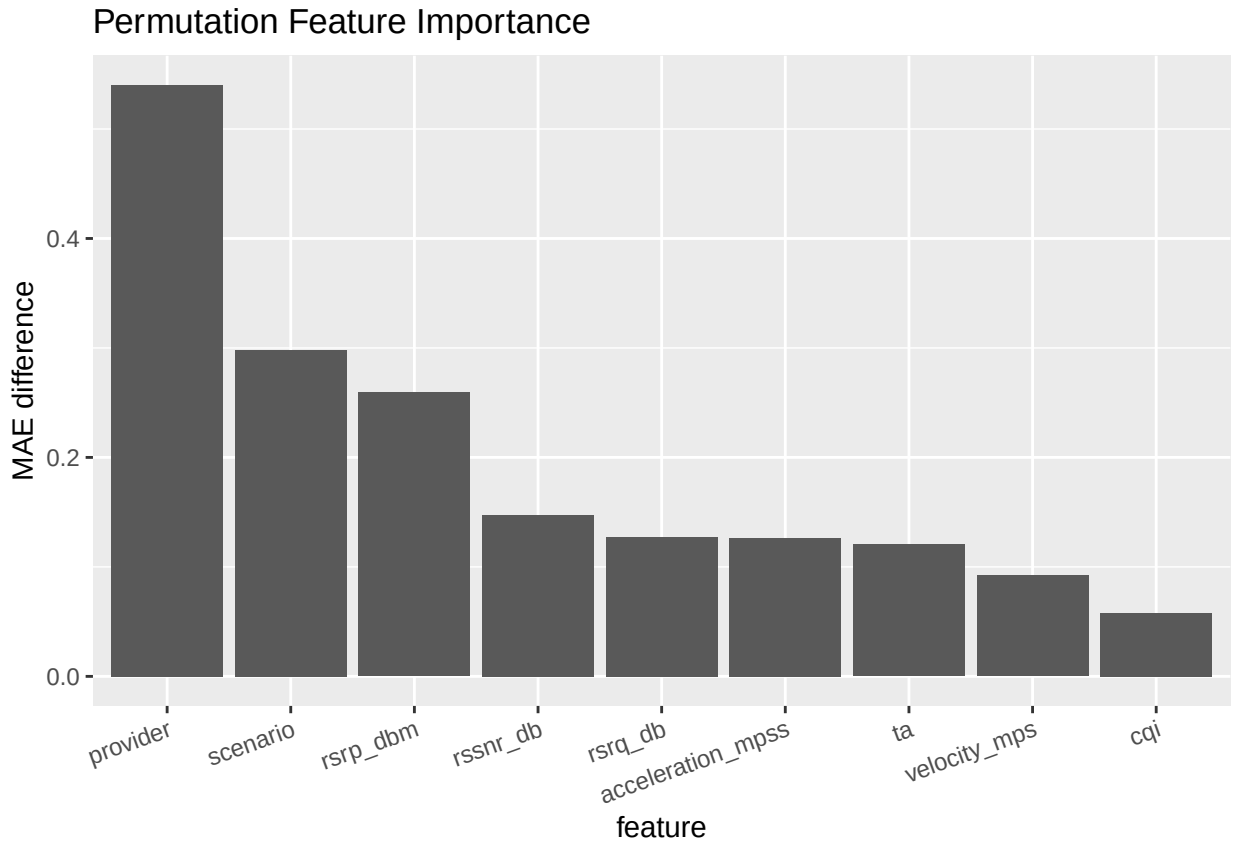
```
## INFO  [19:59:07.028] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:07.640] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:08.101] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:08.602] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:09.074] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:09.537] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:10.001] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:10.599] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:11.043] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:11.480] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:11.928] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:12.369] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:12.801] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:13.381] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:13.819] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
```

```
## INFO  [19:59:14.270] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:14.712] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:15.307] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:15.749] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:16.351] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:16.799] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:17.249] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:17.729] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:18.200] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:18.633] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:19.219] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:19.680] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:20.114] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:20.551] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:21.013] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:21.474] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:22.058] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:22.538] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:22.977] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:23.405] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:23.870] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:24.316] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:24.913] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:25.504] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:25.953] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:26.390] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:26.865] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:27.323] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:27.888] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:28.339] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
## INFO  [19:59:28.769] Applying learner 'regr.ranger' on task 'ul_prediction' (iter 1/1)
```

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

```
##              feature      score
## 1:          provider 0.53993203
## 2:          scenario 0.29827436
## 3:          rsrp_dbm 0.25943705
## 4:          rssnr_db 0.14753153
## 5:           rsrq_db 0.12764740
## 6: acceleration_mpss 0.12639543
## 7:                ta 0.12050201
## 8:      velocity_mps 0.09216798
## 9:               cqi 0.05810642
```

```
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
MAE difference
0.0
0.2
0.4
provider
scenario
rsrp_dbm
rssnr_db
rsrq_db
acceleration_mpss
ta
velocity_mps
cqi
feature