MostSimple Prediction Task with mlr3

川田恵介

Simple RoadMap

0. Motivation!!!!
1. Define Data with Y and X
2. Define Learner, Task, Evaluation
3. SubSampling
4. Learning
5. Prediction
6. Evaluation

SetUp

- ranger パッケージをインストール

```
library(mlr3verse)
library(tidyverse)

set.seed(1)

Data <- read_csv("ExampleData/Example.csv")

Task <- as_task_regr(Data,
                     "Price") # Define Task

OLS <- lrn("regr.lm") # Define OLS
Tree <- lrn("regr.rpart") # Define Tree Learner
RandomForest <- lrn("regr.ranger") # Define Random Forest Learner
```

```
R2 <- msr("regr.rsq") # Define Evaluation with R2
```

Modefied Learner

```
Bagging <- lrn("regr.ranger") # Define Bagging
Bagging$param_set$values$mtry <- ncol(Data) - 1

LargeForest <- lrn("regr.ranger") # Define Larger Forest
LargeForest$param_set$values$num.trees <- 2000
```

SubSampling

```
Group <- partition(Task, ratio = 0.8) # 0.8 for tainning, 0.2 for testing
```

Learning

```
OLS$train(Task, Group$train)

Tree$train(Task, Group$train)

RandomForest$train(Task,Group$train)

Bagging$train(Task,Group$train)

LargeForest$train(Task,Group$train)
```

Prediction

```
PredOLS <- OLS$predict(Task, Group$test)

PredTree <- Tree$predict(Task, Group$test)

PredRandomForest <- RandomForest$predict(Task, Group$test)

PredBagging <- Bagging$predict(Task, Group$test)
```

```
PredLargeForest <- LargeForest$predict(Task, Group$test)
```

Evaluation

```
PredOLS$score(R2)
```

```
 regr.rsq
0.5103169
```

```
PredTree$score(R2)
```

```
 regr.rsq
0.4760191
```

```
PredRandomForest$score(R2)
```

```
 regr.rsq
0.6048222
```

```
PredBagging$score(R2)
```

```
 regr.rsq
0.5903706
```

```
PredLargeForest$score(R2)
```

```
 regr.rsq
0.6044596
```