mlr3 入門

川田恵介

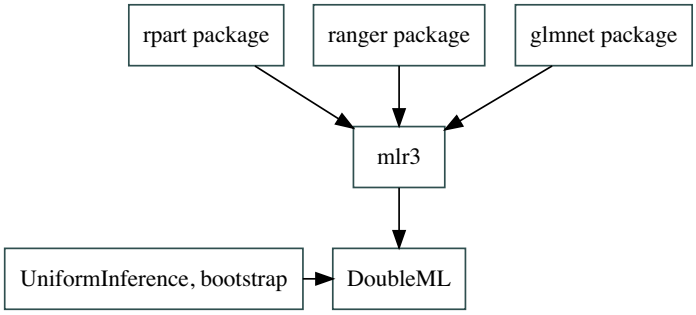
おすすめ

- R 入門
  - R for Data Science: https://r4ds.had.co.nz/
  - Advanced R: https://adv-r.hadley.nz/
- mlr3 入門
  - Package Page: https://mlr3.mlr-org.com/
- 講義に合わせたページ: https://github.com/tetokawata/BookEmiricalSocialML)

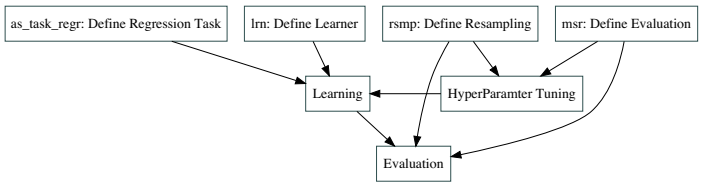
SetUp

```
library(mlr3verse) # Machine Learning
library(tidyverse) # PreProcess

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

mlr3 EcoSystem



RoadMap



Define

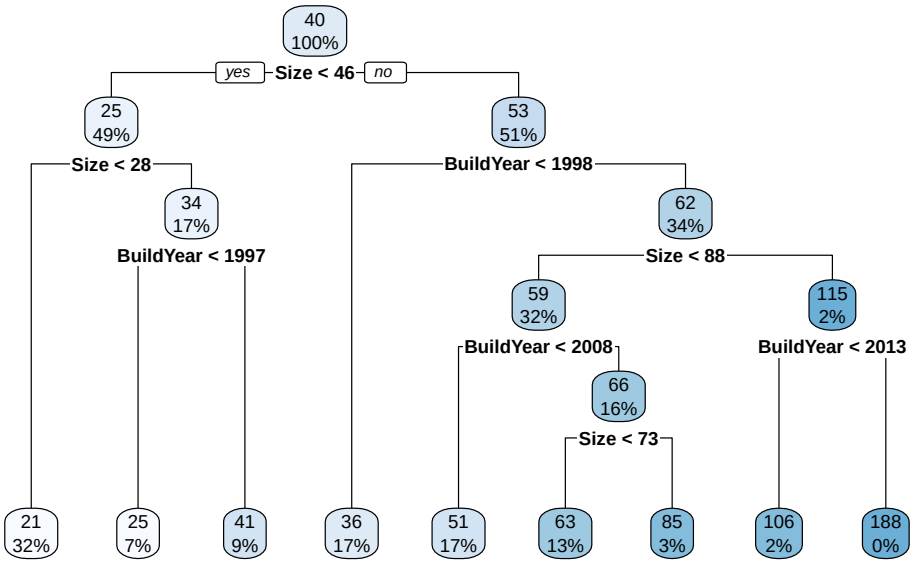
```
Tree <- lrn("regr.rpart") # Define Tree Learner

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

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

Learning

```
Tree$train(Task)
```

Visualization

```
rpart.plot::rpart.plot(Tree$model)
```



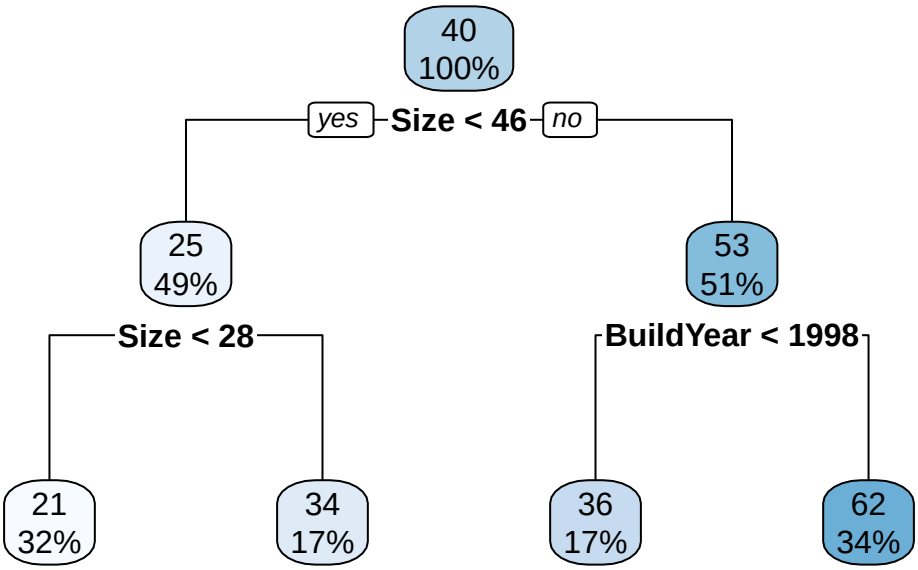
HyperParameter

```
Tree$param_set # Check default hyper parameters
```

```
<ParamSet>
                id    class lower upper nlevels        default value
 1:             cp ParamDbl     0     1     Inf           0.01
 2:     keep_model ParamLgl    NA    NA       2          FALSE
 3:     maxcompete ParamInt     0   Inf     Inf              4
 4:       maxdepth ParamInt     1    30      30             30
 5:   maxsurrogate ParamInt     0   Inf     Inf              5
 6:      minbucket ParamInt     1   Inf     Inf <NoDefault[3]>
 7:       minsplit ParamInt     1   Inf     Inf             20
 8: surrogatestyle ParamInt     0     1       2              0
 9:   usesurrogate ParamInt     0     2       3              2
10:           xval ParamInt     0   Inf     Inf             10     0
```

Set NewAlgorithm

```
ShallowTree <- Tree$clone(deep = TRUE) # Deep copy

ShallowTree$param_set$values$maxdepth <- 2 # Set MaxDepth

ShallowTree$train(Task)

rpart.plot::rpart.plot(ShallowTree$model)
```



Prediction

```
Tree$predict(Task) # 推定に用いたデータへの予想
```

```
<PredictionRegr> for 14793 observations:
    row_ids truth response
          1    21 25.00514
          2    98 51.46344
          3    30 21.05457
---
```

```
14791    34 36.39636
14792    33 36.39636
14793    51 62.58268
```

Evaluation

```
Tree$predict(Task)$score(R2)
```

```
 regr.rsq
0.4092182
```