Handout



## Introduction:

The MLR package provides a way to easily use machine learning algorithmns in R. The features are designed to reduce the amount of time spent coding machine learning algorithmns from scratch to allow experimentation with machine learning and comparison between different models. This helps simplify the process of resampling models, optimizing hyperparameters, selecting features, and dealing with both pre and post processing of data. There are several supervisied algorithmns for classification, regression, and survival analysis. There are also a few un-supervised algorithms.

## Features and Basic Usage:

* Interfaces to R classification, regression, clustering and survival analysis methods
* Convenience methods and generic builidng blocks
* Resampling methods
  + Bootstrapping
  + Cross-validation
  + Subsampling
* Visualizations
* Simpler benchmarking of datasets and learners
* Hyperparameter tuning
* Variable selection with filters and wrappers
* Cost-sensitive learning, threshold tuning, and imbalance corrections
* OpenML connector
* Built-in parallelization
* Tutorial

## Other Information:

### Installation

install.packages(“mlr”)

library(mlr) #Load the library

### Example

library(mlr)

data(iris)

**Define the task**

task = makeClassifTask(id = “tutorial”,data= iris, target =“Species”)

**Define the learner**

lrn = makeLearner(“classif.lda”)

**Define the resampling strategy**

rdesc = makeResampleDesc(method = “CV”, stratify = TRUE)

**Do the resampling**

r = resample(learner = lrn, task = task, resampling = rdesc,show.info = FALSE)

**Get the mean misclassification error**

r$aggr#> mmce.test.mean

#> 0.02

### Popular Learners

|  |  |
| --- | --- |
| Package | Learner |
| survival | surv.coxph |
| e1071 | classif.naiveBayes, classif.svm, regr.svm |
| MASS | classif.lda, classif.qda |
| randomForest | classif.randomForest, regr.randomForest |
| kernlab | classif.gausspr, classif.ksvm, classif.lssvm, cluster.kkmeans, regr.gausspr, regr.ksvm, regr.rvm |
| glmnet | classif.cvglmnet, classif.glmnet, regr.cvglmnet, regr.glmnet, surv.cvglmnet, surv.glmnet |
| party | classif.cforest, classif.ctree, multilabel.cforest, regr.cforest, regr.ctree |
| party,modeltools | regr.mob |
| party,survival | surv.cforest |
| fpc | cluster.dbscan |
| rpart | classif.rpart, regr.rpart, surv.rpart |
| RWeka | classif.IBk, classif.J48, classif.JRip, classif.OneR, classif.PART, cluster.Cobweb, cluster.EM, cluster.FarthestFirst, cluster.SimpleKMeans, cluster.XMeans, regr.IBk |
| gbm | classif.gbm, regr.gbm, surv.gbm |
| nnet | classif.multinom, classif.nnet, regr.nnet |
| caret,pls | classif.plsdaCaret |
| pls | regr.pcr, regr.plsr |
| FNN | classif.fnn, regr.fnn |
| earth | regr.earth |
| neuralnet | classif.neuralnet |
| class | classif.knn, classif.lvq1 |

Additional Information (also citations)

<https://rpubs.com/ippromek/336804>

<https://arxiv.org/pdf/1609.06146.pdf>

<https://mlr.mlr-org.com/>

<https://www.r-bloggers.com/most-popular-learners-in-mlr/>