Project 3 - Baseline(GBM) and Improved(XGBoost with Canny feature) models

Grp2

Load library

```
if(!require("EBImage")){
  source("https://bioconductor.org/biocLite.R")
  biocLite("EBImage")
}
```

```
## Loading required package: EBImage
```

```
## Warning: package 'EBImage' was built under R version 3.5.1
```

```
if(!require("gbm")){
  install.packages("gbm")
}
```

```
## Loading required package: gbm
```

```
## Loaded gbm 2.1.4
```

```
library("EBImage")
library("gbm")
```

Step 0: specify directories.

Set the working directory to the image folder. Specify the training and the testing set. For data without an independent test/validation set, you need to create your own testing data by random subsampling. In order to obain reproducible results, set.seed() whenever randomization is used.

```
set.seed(2018)
setwd("../")
```

Provide directories for training images. Low-resolution (LR) image set and High-resolution (HR) image set will be in different subfolders.

```
train_dir <- "../data/train_set/" # This will be modified for different data sets.
train_LR_dir <- paste(train_dir, "LR/", sep="")
train_HR_dir <- paste(train_dir, "HR/", sep="")
train_label_path <- paste(train_dir, "label.csv", sep="")
```

Step 1: set up controls for evaluation experiments.

In this chunk, we have a set of controls for the evaluation experiments.

- (T/F) cross-validation on the training set
- (number) K, the number of CV folds
- (T/F) process features for training set
- (T/F) run evaluation on an independent test set
- (T/F) process features for test set

```
run.cv=FALSE # run cross-validation on the training set
K <- 5  # number of CV folds
run.feature.train=TRUE # process features for training set
run.test=TRUE # run evaluation on an independent test set
run.feature.test=TRUE # process features for test set
run.train=TRUE
```

Using cross-validation or independent test set evaluation, we compare the performance of models with different specifications. In this example, we use GBM with different `depth`. In the following chunk, we list, in a vector, setups (in this case, `depth`) corresponding to models that we will compare. In your project, you might compare very different classifiers. You can assign them numerical IDs and labels specific to your project.

```
model_values <- seq(3, 11, 2)
model_labels = paste("GBM with depth =", model_values)
```

Step 2: import training images class labels.

We provide extra information of image label: car (0), flower (1), market (2). These labels are not necessary for your model.

```
#extra_label <- read.csv(train_label_path, colClasses=c("NULL", NA, NA))
```

Step 3: construct features and responses

`feature.R` should be the wrapper for all your feature engineering functions and options. The function `feature( )` should have options that correspond to different scenarios for your project and produces an R object that contains features and responses that are required by all the models you are going to evaluate later. + `feature.R` + Input: a path for low-resolution images. + Input: a path for high-resolution images. + Output: an RData file that contains extracted features and corresponding responses

```
source("../lib/feature2.R")
tm_feature_train <- NA
if(run.feature.train){
  tm_feature_train <- system.time(dat_train <- feature(train_LR_dir, train_HR_dir))
  feat_train <- dat_train$feature
  label_train <- dat_train$label
  save(dat_train, file="../output/feature_train.RData")
}
```

```
## Loading required package: magrittr
```

```
##
## Attaching package: 'imager'
```

```
## The following object is masked from 'package:magrittr':
##
##     add
```

```
## The following objects are masked from 'package:EBImage':
##
##     channel, dilate, display, erode, resize, watershed
```

```
## The following objects are masked from 'package:stats':
##
##     convolve, spectrum
```

```
## The following object is masked from 'package:graphics':
##
##     frame

## The following object is masked from 'package:base':
##
##     save.image

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour
```

```
## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour
```

```
## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour
```

```
## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour
```

```
## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour
```

```
## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour
```

```
## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour
```

```
## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel

## Warning in as.cimg.array(.): Assuming third dimension corresponds to colour

## Warning in class(X) <- c("cimg", "imager_array", "numeric"): Setting
## class(x) to multiple strings ("cimg", "imager_array", ...); result will no
## longer be an S4 object

## Warning in cannyEdges(.): Running Canny detector on luminance channel
```

Step 4: Train a classification model with training images

Call the train model and test model from library.

`train.R` and `test.R` should be wrappers for all your model training steps and your classification/prediction steps. + `train.R` + Input: a path that points to the training set features and responses. + Output: an RData file that contains trained classifiers in the forms of R objects: models/settings/links to external trained configurations. + `test.R` + Input: a path that points to the test set features. + Input: an R object that contains a trained classifier. + Output: an R object of response predictions on the test set. If there are multiple classifiers under evaluation, there should be multiple sets of label predictions.

```r
source("../lib/train.R")
source("../lib/test.R")
```

Model selection with cross-validation

- Do model selection by choosing among different values of training model parameters, that is, the interaction depth for GBM in this example.

```r
source("../lib/cross_validation.R")

if(run.cv){
  err_cv <- array(dim=c(length(model_values), 2))
```

```
  for(k in 1:length(model_values)){
    cat("k=", k, "\n")
    err_cv[k,] <- cv.function(feat_train, label_train, model_values[k], K)
  }
  save(err_cv, file="../output/err_cv.RData")
}
```

Visualize cross-validation results.

```
if(run.cv){
  load("../output/err_cv.RData")
  plot(model_values, err_cv[,1], xlab="Interaction Depth", ylab="CV Error",
       main="Cross Validation Error", type="n", ylim=c(0, 0.25))
  points(model_values, err_cv[,1], col="blue", pch=16)
  lines(model_values, err_cv[,1], col="blue")
  arrows(model_values, err_cv[,1]-err_cv[,2], model_values, err_cv[,1]+err_cv[,2],
         length=0.1, angle=90, code=3)
}
```

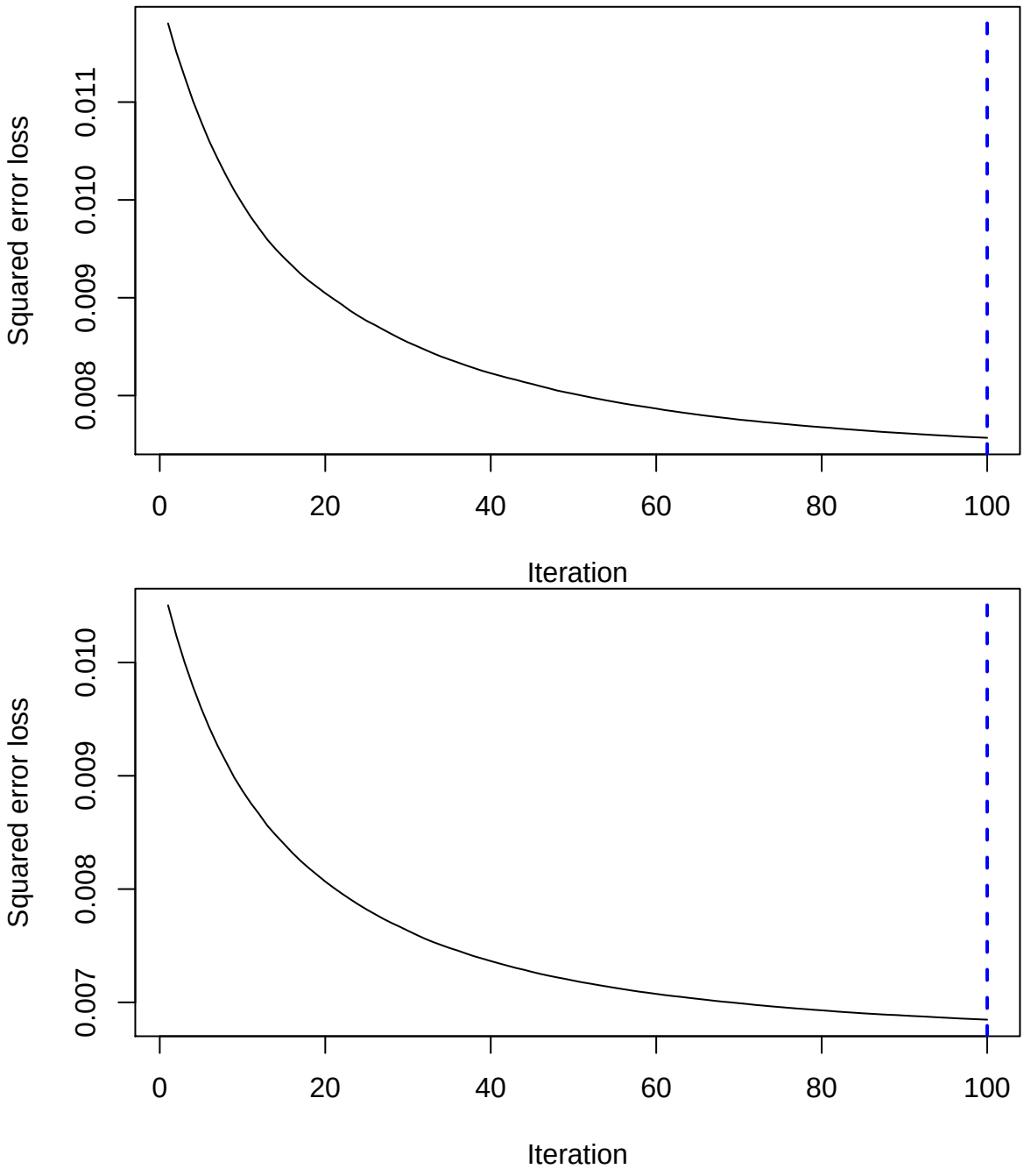
- Choose the "best"" parameter value

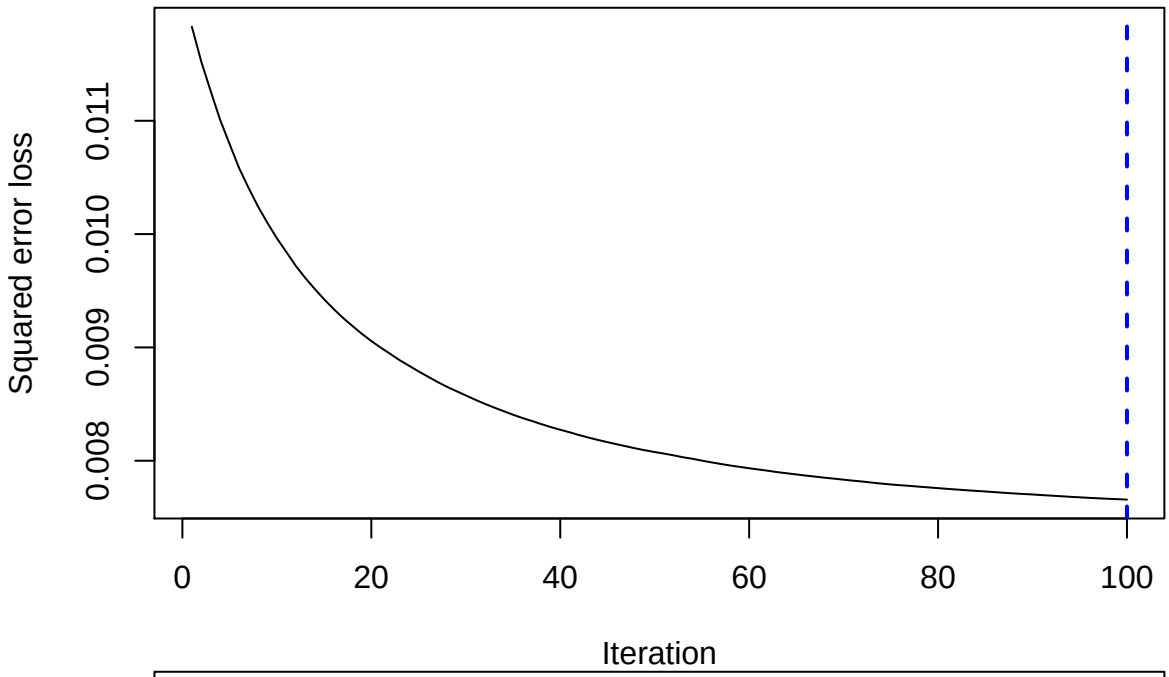
```
model_best=model_values[1]
if(run.cv){
  model_best <- model_values[which.min(err_cv[,1])]
}

par_best <- list(depth=model_best)
```

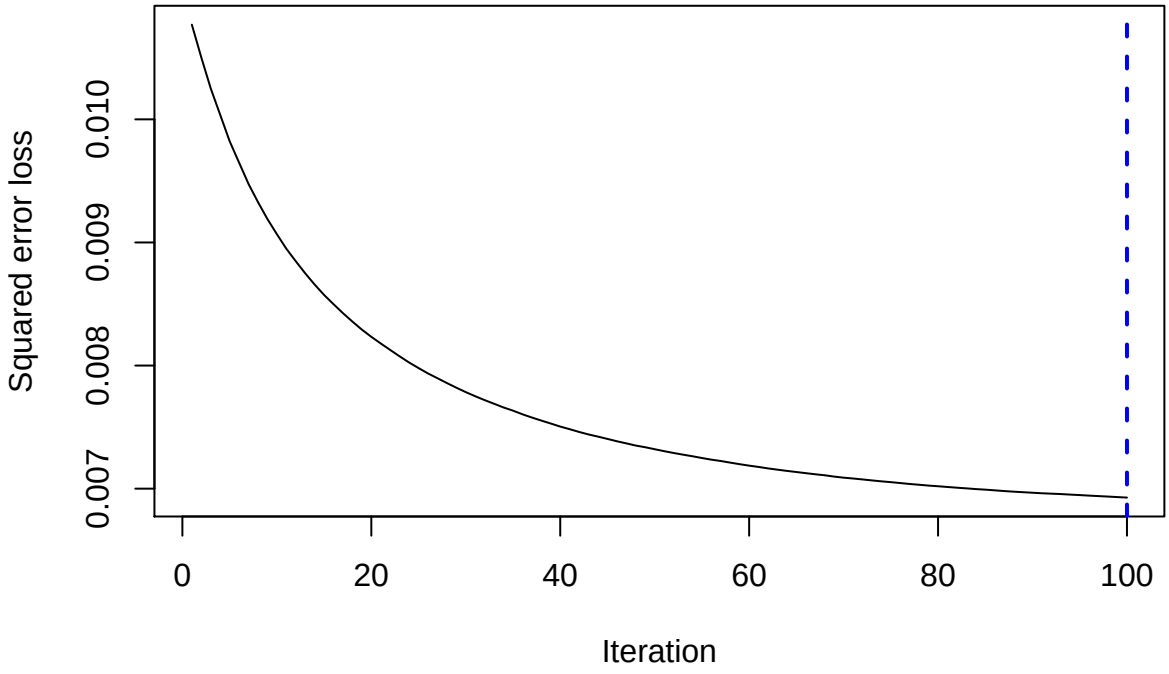
GBM

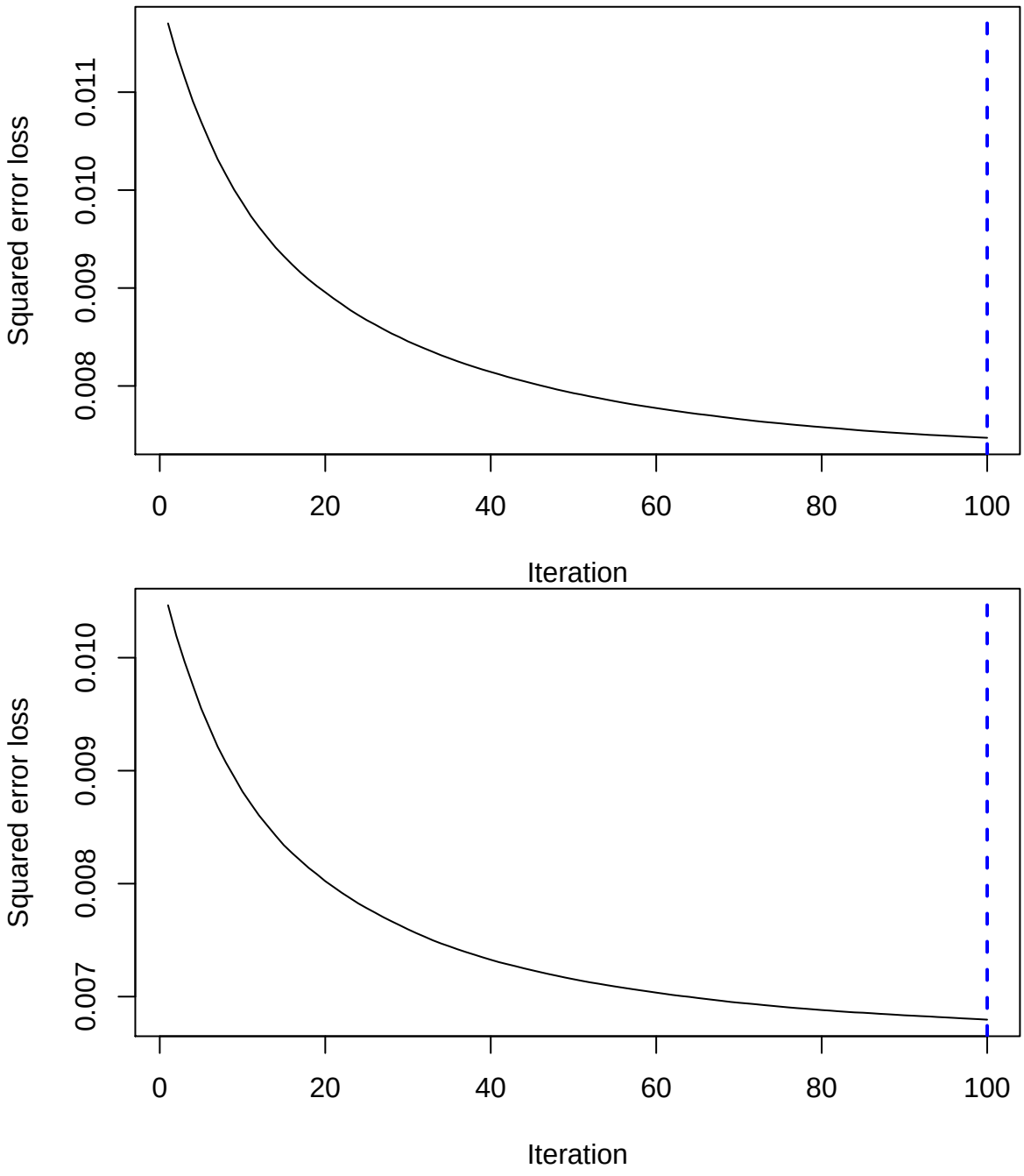
- Train the model with the entire training set using the selected model (model parameter) via cross-validation.

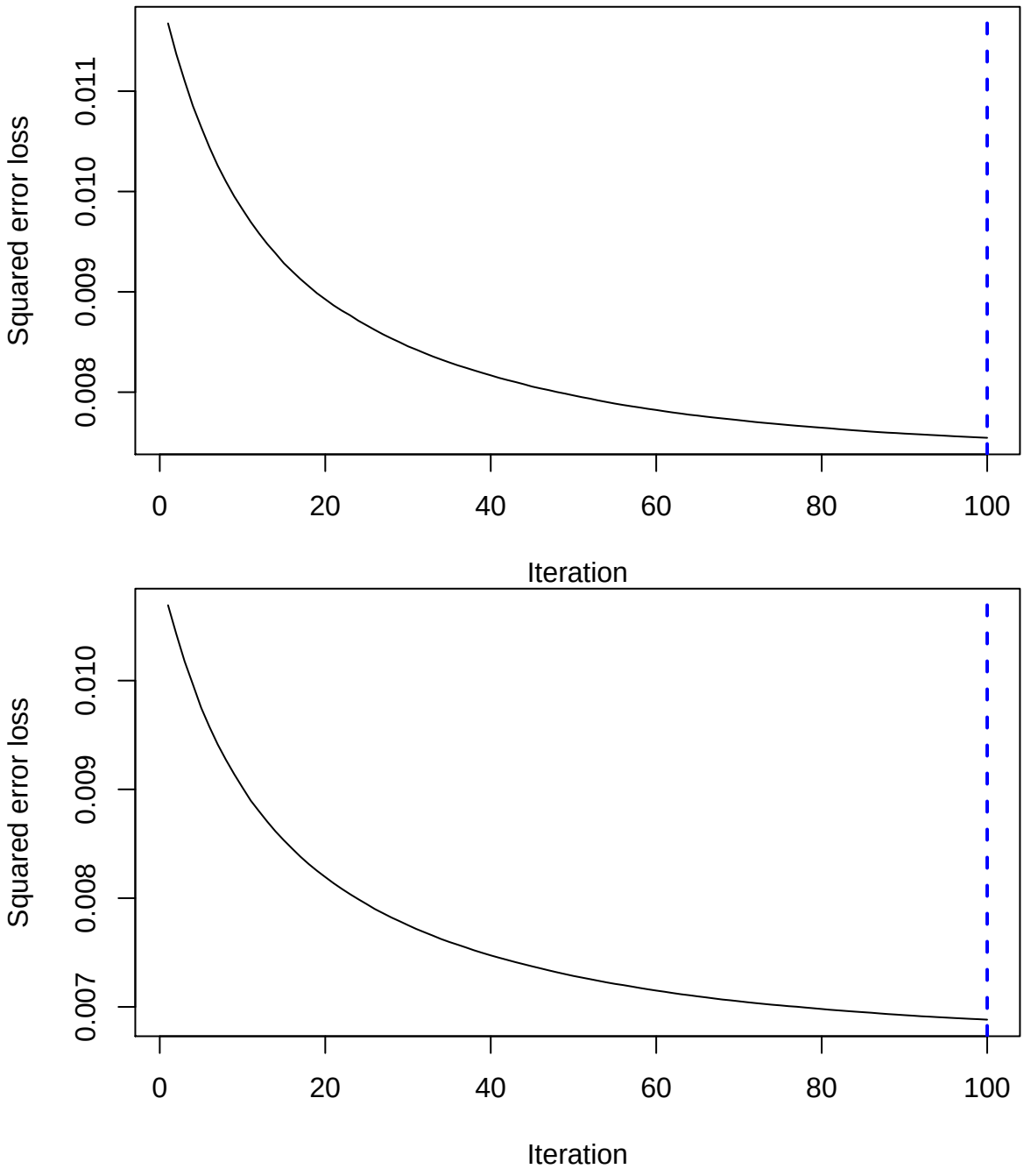
```
tm_train=NA
if(run.train){
  #tm_train_gbm <- system.time(fit_train_gbm <- train(feat_train, label_train, par_best))
  tm_train_gbm <- system.time(fit_train <- train(feat_train, label_train))
  #save(fit_train_gbm, file="../output/fit_train_gbm.RData")
}
```

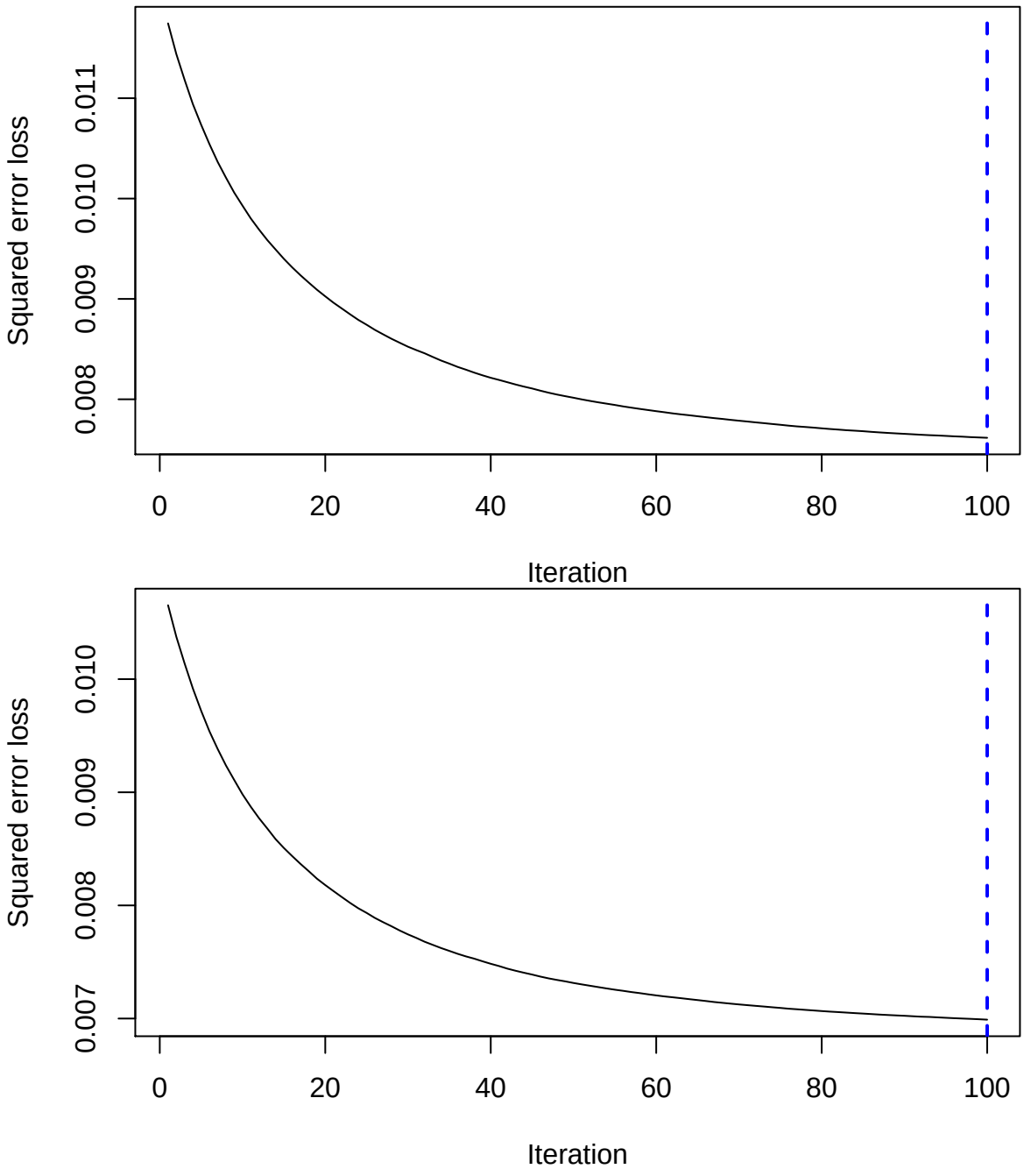

Squared error loss
0.008
0.009
0.010
0.011
0
20
40
60
80
100
Iteration
Squared error loss
0.007
0.008
0.009
0.010
0
20
40
60
80
100
Iteration

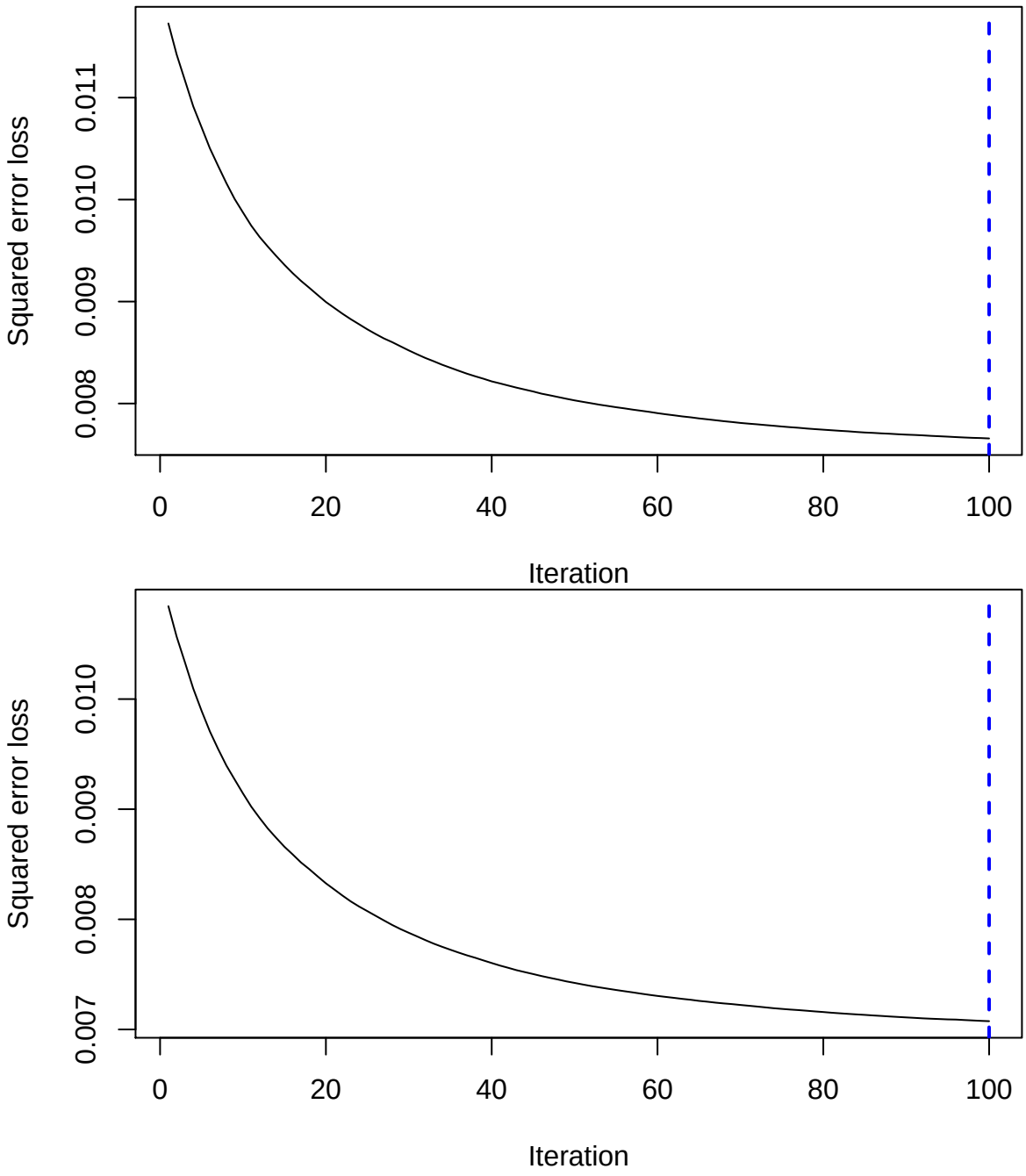

Squared error loss
0.008
0.009
0.010
0.011
0
20
40
60
80
100
Iteration


Squared error loss
0.007
0.008
0.009
0.010
0
20
40
60
80
100
Iteration


Squared error loss
0.008
0.009
0.010
0.011
0
20
40
60
80
100
Iteration
Squared error loss
0.007
0.008
0.009
0.010
0
20
40
60
80
100
Iteration


Squared error loss
0.008
0.009
0.010
0.011
0
20
40
60
80
100
Iteration
Squared error loss
0.007
0.008
0.009
0.010
0
20
40
60
80
100
Iteration


Squared error loss
0.008
0.009
0.010
0.011
0
20
40
60
80
100
Iteration
Squared error loss
0.007
0.008
0.009
0.010
0
20
40
60
80
100
Iteration



###XGBOOST

```
if(run.train){
  tm_train_xgboost <- system.time(fit_train_xgboost <- xgb_train(feat_train, label_train))
  save(fit_train_xgboost, file="../output/fit_train_xgboost.RData")
}
```

```
## [1]  train-rmse:0.274181
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.167082
## [3]  train-rmse:0.126893
## [4]  train-rmse:0.114665
```

```
## [5]  train-rmse:0.111399
## [6]  train-rmse:0.110567
## [7]  train-rmse:0.110358
## [8]  train-rmse:0.110306
## [9]  train-rmse:0.110293
## [10] train-rmse:0.110289
## [11] train-rmse:0.110288
## [12] train-rmse:0.110288
## [13] train-rmse:0.110288
## [14] train-rmse:0.110288
## [15] train-rmse:0.110288
## [16] train-rmse:0.110288
## [17] train-rmse:0.110288
## [18] train-rmse:0.110288
## [19] train-rmse:0.110288
## [20] train-rmse:0.110288
## [21] train-rmse:0.110288
## [22] train-rmse:0.110288
## [23] train-rmse:0.110288
## [24] train-rmse:0.110288
## [25] train-rmse:0.110288
## [26] train-rmse:0.110288
## [27] train-rmse:0.110288
## [28] train-rmse:0.110288
## [29] train-rmse:0.110288
## [30] train-rmse:0.110288
## [31] train-rmse:0.110288
## [32] train-rmse:0.110288
## [33] train-rmse:0.110288
## [34] train-rmse:0.110288
## [35] train-rmse:0.110288
## [36] train-rmse:0.110288
## [37] train-rmse:0.110288
## [38] train-rmse:0.110288
## [39] train-rmse:0.110288
## [40] train-rmse:0.110288
## [41] train-rmse:0.110288
## Stopping. Best iteration:
## [11] train-rmse:0.110288
##
## [1]  train-rmse:0.269741
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.162154
## [3]  train-rmse:0.121151
## [4]  train-rmse:0.108505
## [5]  train-rmse:0.105107
## [6]  train-rmse:0.104239
## [7]  train-rmse:0.104022
## [8]  train-rmse:0.103967
## [9]  train-rmse:0.103953
## [10] train-rmse:0.103950
## [11] train-rmse:0.103949
## [12] train-rmse:0.103949
```

```
## [13] train-rmse:0.103949
## [14] train-rmse:0.103949
## [15] train-rmse:0.103949
## [16] train-rmse:0.103949
## [17] train-rmse:0.103949
## [18] train-rmse:0.103949
## [19] train-rmse:0.103949
## [20] train-rmse:0.103949
## [21] train-rmse:0.103949
## [22] train-rmse:0.103949
## [23] train-rmse:0.103949
## [24] train-rmse:0.103949
## [25] train-rmse:0.103949
## [26] train-rmse:0.103949
## [27] train-rmse:0.103949
## [28] train-rmse:0.103949
## [29] train-rmse:0.103949
## [30] train-rmse:0.103949
## [31] train-rmse:0.103949
## [32] train-rmse:0.103949
## [33] train-rmse:0.103949
## [34] train-rmse:0.103949
## [35] train-rmse:0.103949
## [36] train-rmse:0.103949
## [37] train-rmse:0.103949
## [38] train-rmse:0.103949
## [39] train-rmse:0.103949
## [40] train-rmse:0.103949
## [41] train-rmse:0.103949
## Stopping. Best iteration:
## [11] train-rmse:0.103949
##
## [1]  train-rmse:0.274066
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.167121
## [3]  train-rmse:0.127020
## [4]  train-rmse:0.114826
## [5]  train-rmse:0.111570
## [6]  train-rmse:0.110741
## [7]  train-rmse:0.110533
## [8]  train-rmse:0.110481
## [9]  train-rmse:0.110468
## [10] train-rmse:0.110464
## [11] train-rmse:0.110463
## [12] train-rmse:0.110463
## [13] train-rmse:0.110463
## [14] train-rmse:0.110463
## [15] train-rmse:0.110463
## [16] train-rmse:0.110463
## [17] train-rmse:0.110463
## [18] train-rmse:0.110463
## [19] train-rmse:0.110463
## [20] train-rmse:0.110463
```

```
## [21] train-rmse:0.110463
## [22] train-rmse:0.110463
## [23] train-rmse:0.110463
## [24] train-rmse:0.110463
## [25] train-rmse:0.110463
## [26] train-rmse:0.110463
## [27] train-rmse:0.110463
## [28] train-rmse:0.110463
## [29] train-rmse:0.110463
## [30] train-rmse:0.110463
## [31] train-rmse:0.110463
## [32] train-rmse:0.110463
## [33] train-rmse:0.110463
## [34] train-rmse:0.110463
## [35] train-rmse:0.110463
## [36] train-rmse:0.110463
## [37] train-rmse:0.110463
## [38] train-rmse:0.110463
## [39] train-rmse:0.110463
## [40] train-rmse:0.110463
## [41] train-rmse:0.110463
## Stopping. Best iteration:
## [11] train-rmse:0.110463
##
## [1]  train-rmse:0.269887
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.162844
## [3]  train-rmse:0.122221
## [4]  train-rmse:0.109740
## [5]  train-rmse:0.106391
## [6]  train-rmse:0.105538
## [7]  train-rmse:0.105323
## [8]  train-rmse:0.105269
## [9]  train-rmse:0.105256
## [10] train-rmse:0.105253
## [11] train-rmse:0.105252
## [12] train-rmse:0.105252
## [13] train-rmse:0.105252
## [14] train-rmse:0.105252
## [15] train-rmse:0.105252
## [16] train-rmse:0.105252
## [17] train-rmse:0.105252
## [18] train-rmse:0.105252
## [19] train-rmse:0.105252
## [20] train-rmse:0.105252
## [21] train-rmse:0.105252
## [22] train-rmse:0.105252
## [23] train-rmse:0.105252
## [24] train-rmse:0.105252
## [25] train-rmse:0.105252
## [26] train-rmse:0.105252
## [27] train-rmse:0.105252
## [28] train-rmse:0.105252
```

```
## [29] train-rmse:0.105252
## [30] train-rmse:0.105252
## [31] train-rmse:0.105252
## [32] train-rmse:0.105252
## [33] train-rmse:0.105252
## [34] train-rmse:0.105252
## [35] train-rmse:0.105252
## [36] train-rmse:0.105252
## [37] train-rmse:0.105252
## [38] train-rmse:0.105252
## [39] train-rmse:0.105252
## [40] train-rmse:0.105252
## [41] train-rmse:0.105252
## Stopping. Best iteration:
## [11] train-rmse:0.105252
##
## [1]  train-rmse:0.274221
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.166856
## [3]  train-rmse:0.126498
## [4]  train-rmse:0.114202
## [5]  train-rmse:0.110915
## [6]  train-rmse:0.110078
## [7]  train-rmse:0.109867
## [8]  train-rmse:0.109815
## [9]  train-rmse:0.109802
## [10] train-rmse:0.109798
## [11] train-rmse:0.109798
## [12] train-rmse:0.109797
## [13] train-rmse:0.109797
## [14] train-rmse:0.109797
## [15] train-rmse:0.109797
## [16] train-rmse:0.109797
## [17] train-rmse:0.109797
## [18] train-rmse:0.109797
## [19] train-rmse:0.109797
## [20] train-rmse:0.109797
## [21] train-rmse:0.109797
## [22] train-rmse:0.109797
## [23] train-rmse:0.109797
## [24] train-rmse:0.109797
## [25] train-rmse:0.109797
## [26] train-rmse:0.109797
## [27] train-rmse:0.109797
## [28] train-rmse:0.109797
## [29] train-rmse:0.109797
## [30] train-rmse:0.109797
## [31] train-rmse:0.109797
## [32] train-rmse:0.109797
## [33] train-rmse:0.109797
## [34] train-rmse:0.109797
## [35] train-rmse:0.109797
## [36] train-rmse:0.109797
```

```
## [37] train-rmse:0.109797
## [38] train-rmse:0.109797
## [39] train-rmse:0.109797
## [40] train-rmse:0.109797
## [41] train-rmse:0.109797
## [42] train-rmse:0.109797
## Stopping. Best iteration:
## [12] train-rmse:0.109797
##
## [1]  train-rmse:0.269828
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.162087
## [3]  train-rmse:0.120990
## [4]  train-rmse:0.108306
## [5]  train-rmse:0.104895
## [6]  train-rmse:0.104025
## [7]  train-rmse:0.103807
## [8]  train-rmse:0.103752
## [9]  train-rmse:0.103738
## [10] train-rmse:0.103735
## [11] train-rmse:0.103734
## [12] train-rmse:0.103734
## [13] train-rmse:0.103734
## [14] train-rmse:0.103734
## [15] train-rmse:0.103734
## [16] train-rmse:0.103734
## [17] train-rmse:0.103734
## [18] train-rmse:0.103734
## [19] train-rmse:0.103734
## [20] train-rmse:0.103734
## [21] train-rmse:0.103734
## [22] train-rmse:0.103734
## [23] train-rmse:0.103734
## [24] train-rmse:0.103734
## [25] train-rmse:0.103734
## [26] train-rmse:0.103734
## [27] train-rmse:0.103734
## [28] train-rmse:0.103734
## [29] train-rmse:0.103734
## [30] train-rmse:0.103734
## [31] train-rmse:0.103734
## [32] train-rmse:0.103734
## [33] train-rmse:0.103734
## [34] train-rmse:0.103734
## [35] train-rmse:0.103734
## [36] train-rmse:0.103734
## [37] train-rmse:0.103734
## [38] train-rmse:0.103734
## [39] train-rmse:0.103734
## [40] train-rmse:0.103734
## [41] train-rmse:0.103734
## Stopping. Best iteration:
## [11] train-rmse:0.103734
```

```
##
## [1]  train-rmse:0.274093
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.166788
## [3]  train-rmse:0.126456
## [4]  train-rmse:0.114168
## [5]  train-rmse:0.110883
## [6]  train-rmse:0.110047
## [7]  train-rmse:0.109837
## [8]  train-rmse:0.109784
## [9]  train-rmse:0.109771
## [10] train-rmse:0.109768
## [11] train-rmse:0.109767
## [12] train-rmse:0.109767
## [13] train-rmse:0.109767
## [14] train-rmse:0.109767
## [15] train-rmse:0.109767
## [16] train-rmse:0.109767
## [17] train-rmse:0.109767
## [18] train-rmse:0.109767
## [19] train-rmse:0.109767
## [20] train-rmse:0.109767
## [21] train-rmse:0.109767
## [22] train-rmse:0.109767
## [23] train-rmse:0.109767
## [24] train-rmse:0.109767
## [25] train-rmse:0.109767
## [26] train-rmse:0.109767
## [27] train-rmse:0.109767
## [28] train-rmse:0.109767
## [29] train-rmse:0.109767
## [30] train-rmse:0.109767
## [31] train-rmse:0.109767
## [32] train-rmse:0.109767
## [33] train-rmse:0.109767
## [34] train-rmse:0.109767
## [35] train-rmse:0.109767
## [36] train-rmse:0.109767
## [37] train-rmse:0.109767
## [38] train-rmse:0.109767
## [39] train-rmse:0.109767
## [40] train-rmse:0.109767
## [41] train-rmse:0.109767
## Stopping. Best iteration:
## [11] train-rmse:0.109767
##
## [1]  train-rmse:0.270028
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.162731
## [3]  train-rmse:0.121954
## [4]  train-rmse:0.109411
## [5]  train-rmse:0.106043
```

```
## [6]  train-rmse:0.105185
## [7]  train-rmse:0.104969
## [8]  train-rmse:0.104915
## [9]  train-rmse:0.104901
## [10] train-rmse:0.104898
## [11] train-rmse:0.104897
## [12] train-rmse:0.104897
## [13] train-rmse:0.104897
## [14] train-rmse:0.104897
## [15] train-rmse:0.104897
## [16] train-rmse:0.104897
## [17] train-rmse:0.104897
## [18] train-rmse:0.104897
## [19] train-rmse:0.104897
## [20] train-rmse:0.104897
## [21] train-rmse:0.104897
## [22] train-rmse:0.104897
## [23] train-rmse:0.104897
## [24] train-rmse:0.104897
## [25] train-rmse:0.104897
## [26] train-rmse:0.104897
## [27] train-rmse:0.104897
## [28] train-rmse:0.104897
## [29] train-rmse:0.104897
## [30] train-rmse:0.104897
## [31] train-rmse:0.104897
## [32] train-rmse:0.104897
## [33] train-rmse:0.104897
## [34] train-rmse:0.104897
## [35] train-rmse:0.104897
## [36] train-rmse:0.104897
## [37] train-rmse:0.104897
## [38] train-rmse:0.104897
## [39] train-rmse:0.104897
## [40] train-rmse:0.104897
## [41] train-rmse:0.104897
## Stopping. Best iteration:
## [11] train-rmse:0.104897
##
## [1]  train-rmse:0.274330
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.167002
## [3]  train-rmse:0.126681
## [4]  train-rmse:0.114401
## [5]  train-rmse:0.111120
## [6]  train-rmse:0.110284
## [7]  train-rmse:0.110074
## [8]  train-rmse:0.110022
## [9]  train-rmse:0.110008
## [10] train-rmse:0.110005
## [11] train-rmse:0.110004
## [12] train-rmse:0.110004
## [13] train-rmse:0.110004
```

```
## [14] train-rmse:0.110004
## [15] train-rmse:0.110004
## [16] train-rmse:0.110004
## [17] train-rmse:0.110004
## [18] train-rmse:0.110004
## [19] train-rmse:0.110004
## [20] train-rmse:0.110004
## [21] train-rmse:0.110004
## [22] train-rmse:0.110004
## [23] train-rmse:0.110004
## [24] train-rmse:0.110004
## [25] train-rmse:0.110004
## [26] train-rmse:0.110004
## [27] train-rmse:0.110004
## [28] train-rmse:0.110004
## [29] train-rmse:0.110004
## [30] train-rmse:0.110004
## [31] train-rmse:0.110004
## [32] train-rmse:0.110004
## [33] train-rmse:0.110004
## [34] train-rmse:0.110004
## [35] train-rmse:0.110004
## [36] train-rmse:0.110004
## [37] train-rmse:0.110004
## [38] train-rmse:0.110004
## [39] train-rmse:0.110004
## [40] train-rmse:0.110004
## [41] train-rmse:0.110004
## Stopping. Best iteration:
## [11] train-rmse:0.110004
##
## [1]  train-rmse:0.270424
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.162766
## [3]  train-rmse:0.121794
## [4]  train-rmse:0.109174
## [5]  train-rmse:0.105784
## [6]  train-rmse:0.104919
## [7]  train-rmse:0.104702
## [8]  train-rmse:0.104648
## [9]  train-rmse:0.104634
## [10] train-rmse:0.104631
## [11] train-rmse:0.104630
## [12] train-rmse:0.104630
## [13] train-rmse:0.104630
## [14] train-rmse:0.104630
## [15] train-rmse:0.104630
## [16] train-rmse:0.104630
## [17] train-rmse:0.104630
## [18] train-rmse:0.104630
## [19] train-rmse:0.104630
## [20] train-rmse:0.104630
## [21] train-rmse:0.104630
```

```
## [22] train-rmse:0.104630
## [23] train-rmse:0.104630
## [24] train-rmse:0.104630
## [25] train-rmse:0.104630
## [26] train-rmse:0.104630
## [27] train-rmse:0.104630
## [28] train-rmse:0.104630
## [29] train-rmse:0.104630
## [30] train-rmse:0.104630
## [31] train-rmse:0.104630
## [32] train-rmse:0.104630
## [33] train-rmse:0.104630
## [34] train-rmse:0.104630
## [35] train-rmse:0.104630
## [36] train-rmse:0.104630
## [37] train-rmse:0.104630
## [38] train-rmse:0.104630
## [39] train-rmse:0.104630
## [40] train-rmse:0.104630
## [41] train-rmse:0.104630
## Stopping. Best iteration:
## [11] train-rmse:0.104630
##
## [1]  train-rmse:0.274078
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.166884
## [3]  train-rmse:0.126623
## [4]  train-rmse:0.114363
## [5]  train-rmse:0.111087
## [6]  train-rmse:0.110253
## [7]  train-rmse:0.110044
## [8]  train-rmse:0.109991
## [9]  train-rmse:0.109978
## [10] train-rmse:0.109975
## [11] train-rmse:0.109974
## [12] train-rmse:0.109974
## [13] train-rmse:0.109974
## [14] train-rmse:0.109974
## [15] train-rmse:0.109974
## [16] train-rmse:0.109974
## [17] train-rmse:0.109974
## [18] train-rmse:0.109974
## [19] train-rmse:0.109974
## [20] train-rmse:0.109974
## [21] train-rmse:0.109974
## [22] train-rmse:0.109974
## [23] train-rmse:0.109974
## [24] train-rmse:0.109974
## [25] train-rmse:0.109974
## [26] train-rmse:0.109974
## [27] train-rmse:0.109974
## [28] train-rmse:0.109974
## [29] train-rmse:0.109974
```

```
## [30] train-rmse:0.109974
## [31] train-rmse:0.109974
## [32] train-rmse:0.109974
## [33] train-rmse:0.109974
## [34] train-rmse:0.109974
## [35] train-rmse:0.109974
## [36] train-rmse:0.109974
## [37] train-rmse:0.109974
## [38] train-rmse:0.109974
## [39] train-rmse:0.109974
## [40] train-rmse:0.109974
## [41] train-rmse:0.109974
## Stopping. Best iteration:
## [11] train-rmse:0.109974
##
## [1]  train-rmse:0.270430
## Will train until train_rmse hasn't improved in 30 rounds.
##
## [2]  train-rmse:0.163239
## [3]  train-rmse:0.122578
## [4]  train-rmse:0.110091
## [5]  train-rmse:0.106742
## [6]  train-rmse:0.105888
## [7]  train-rmse:0.105673
## [8]  train-rmse:0.105619
## [9]  train-rmse:0.105606
## [10] train-rmse:0.105603
## [11] train-rmse:0.105602
## [12] train-rmse:0.105602
## [13] train-rmse:0.105602
## [14] train-rmse:0.105602
## [15] train-rmse:0.105602
## [16] train-rmse:0.105602
## [17] train-rmse:0.105602
## [18] train-rmse:0.105602
## [19] train-rmse:0.105602
## [20] train-rmse:0.105602
## [21] train-rmse:0.105602
## [22] train-rmse:0.105602
## [23] train-rmse:0.105602
## [24] train-rmse:0.105602
## [25] train-rmse:0.105602
## [26] train-rmse:0.105602
## [27] train-rmse:0.105602
## [28] train-rmse:0.105602
## [29] train-rmse:0.105602
## [30] train-rmse:0.105602
## [31] train-rmse:0.105602
## [32] train-rmse:0.105602
## [33] train-rmse:0.105602
## [34] train-rmse:0.105602
## [35] train-rmse:0.105602
## [36] train-rmse:0.105602
## [37] train-rmse:0.105602
```

```
## [38] train-rmse:0.105602
## [39] train-rmse:0.105602
## [40] train-rmse:0.105602
## [41] train-rmse:0.105602
## Stopping. Best iteration:
## [11] train-rmse:0.105602
```

Step 5: Super-resolution for test images

Feed the final training model with the completely holdout testing data. + `superResolution.R` + Input: a path that points to the folder of low-resolution test images. + Input: a path that points to the folder (empty) of high-resolution test images. + Input: an R object that contains tuned predictors. + Output: construct high-resolution versions for each low-resolution test image.

```
source("../lib/superResolution.R")
test_dir <- "../data/test_set/" # This will be modified for different data sets.
test_LR_dir <- paste(test_dir, "LR/", sep="")
test_HR_dir <- paste(test_dir, "HR/", sep="")
###GBM test
if(run.test){
  load(file="../output/fit_train_gbm.RData")
  tm_test_gbm <- system.time(PSNR_gbm <- superResolution_gbm(test_LR_dir, test_HR_dir, fit_train_gbm))
}
#XGBoost test
if(run.test){
  load(file="../output/fit_train_gbm.RData")
  tm_test_xgboost <- system.time(PSNR_xgboost <- superResolution_xgboost(test_LR_dir, test_HR_dir, fit_
}
```

Summarize Running Time

Prediction performance matters, so does the running times for constructing features and for training the model, especially when the computation resource is limited.

```
cat("Time for constructing training features=", tm_feature_train[1], "s \n")
```

```
## Time for constructing training features= 13.436 s
```

```
cat("Time for training gbm model=", tm_train_gbm[1], "s \n")
```

```
## Time for training gbm model= 12.114 s
```

```
cat("Time for gbm super-resolution=", tm_test_gbm[1], "s \n")
```

```
## Time for gbm super-resolution= 61.554 s
```

```
cat("Time for training xgboost model=", tm_train_xgboost[1], "s \n")
```

```
## Time for training xgboost model= 1.49 s
```

```
cat("Time for xgboost super-resolution=", tm_test_xgboost[1], "s \n")
```

```
## Time for xgboost super-resolution= 16.235 s
```

```
print(paste("PSNR_gbm:", PSNR_gbm, sep=" "))
```

```
## [1] "PSNR_gbm: 19.0742734786922"
```

```r
print(paste("PSNR_xgboost:", PSNR_xgboost, sep=" "))
```

```
## [1] "PSNR_xgboost: 25.2325535569037"
```