HW-7-Subbu-Kandhaswamy-IST707

Subbu Kandhaswamy

11/21/2019

## Homework7 SVMs, kNN, and Random Forestfor handwriting recognition

In this we use SVMs, kNN, and Random Forest algorithms for handwriting recognition, and compare their performance with the naïve Bayes and decision tree models you built in previous week

# Load packages.  
  
# install.packages('randomForest')  
library(caret)

## Loading required package: lattice

## Loading required package: ggplot2

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(stringr)  
library(e1071)  
library(rpart)  
library(randomForest)

## randomForest 4.6-14

## Type rfNews() to see new features/changes/bug fixes.

##   
## Attaching package: 'randomForest'

## The following object is masked from 'package:dplyr':  
##   
## combine

## The following object is masked from 'package:ggplot2':  
##   
## margin

# Load Kggle-digit-train dataset.  
dataSet <- read.csv("C:/Users/Administrator/Documents/MSDatascience/IST707/Week7/Kaggle-digit-train.csv")  
  
# Explore the dimension of the dataset, and the structure and summary of the  
# first ten rows.  
  
dim(dataSet)

## [1] 42000 785

str(dataSet[, 1:10])

## 'data.frame': 42000 obs. of 10 variables:  
## $ label : int 1 0 1 4 0 0 7 3 5 3 ...  
## $ pixel0: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel1: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel2: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel3: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel4: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel5: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel6: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel7: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ pixel8: int 0 0 0 0 0 0 0 0 0 0 ...

summary(dataSet[, 1:10])

## label pixel0 pixel1 pixel2 pixel3   
## Min. :0.000 Min. :0 Min. :0 Min. :0 Min. :0   
## 1st Qu.:2.000 1st Qu.:0 1st Qu.:0 1st Qu.:0 1st Qu.:0   
## Median :4.000 Median :0 Median :0 Median :0 Median :0   
## Mean :4.457 Mean :0 Mean :0 Mean :0 Mean :0   
## 3rd Qu.:7.000 3rd Qu.:0 3rd Qu.:0 3rd Qu.:0 3rd Qu.:0   
## Max. :9.000 Max. :0 Max. :0 Max. :0 Max. :0   
## pixel4 pixel5 pixel6 pixel7 pixel8   
## Min. :0 Min. :0 Min. :0 Min. :0 Min. :0   
## 1st Qu.:0 1st Qu.:0 1st Qu.:0 1st Qu.:0 1st Qu.:0   
## Median :0 Median :0 Median :0 Median :0 Median :0   
## Mean :0 Mean :0 Mean :0 Mean :0 Mean :0   
## 3rd Qu.:0 3rd Qu.:0 3rd Qu.:0 3rd Qu.:0 3rd Qu.:0   
## Max. :0 Max. :0 Max. :0 Max. :0 Max. :0

# Let's also check what the first 5 rows of the dataset look like.  
head(dataSet[, 1:10], n = 5)

## label pixel0 pixel1 pixel2 pixel3 pixel4 pixel5 pixel6 pixel7 pixel8  
## 1 1 0 0 0 0 0 0 0 0 0  
## 2 0 0 0 0 0 0 0 0 0 0  
## 3 1 0 0 0 0 0 0 0 0 0  
## 4 4 0 0 0 0 0 0 0 0 0  
## 5 0 0 0 0 0 0 0 0 0 0

Now that we have checked what the dataset look like, it’s time to do some additional processing and then train the models.

#Datasets usage. In this case, we’ll work on two new models: kNN (k-Nearest Neighbours) , SVM (Support Vector Machine) and Random Forest - which is an expansion of the decision tree algorithm we have previously worked on. Also, due to the size of the training set, we’ll only work with 30 percent of the dataset, setting the seeds and selecting two-thirds of this subset to serve as our training set and the remaining third to serve as our testing set. All these subsets will be chosen at random.

##Pre-Processing

Considering the number of variables, we’ll make sure that we work with a select number of objects,to ensure that we don’t eliminate unneccesary columns, only columns where there is no variance will be eliminated.

# Update/Replace all NAs with 0, count number of columns, and create a list  
# that will populate the columns that sum 0.  
  
dataSet[is.na(dataSet)] <- 0  
cols <- ncol(dataSet)  
vars <- list()  
  
# reviewing variables in dataset to find which present no variance.  
for (i in 2:cols) {  
 colVar <- var(dataSet[[i]])  
 if (colVar == 0) {  
 vars <- append(vars, i)  
 }  
}  
  
# Remove/Drop the columns that have no variance in both the test and  
# training set.  
colsToDrop <- unlist(vars)  
  
dataSet <- dataSet[, -colsToDrop]  
  
# With the columns now selected, we'll now choose 20 percent of both the  
# training and testing datasets.  
set.seed(1024)  
  
split <- sample(nrow(dataSet), nrow(dataSet) \* 0.3)  
dataSubset <- dataSet[split, ]  
  
# Now to select the training and testing sets from this subset.  
set.seed(760)  
  
trainSplit <- sample(nrow(dataSubset), nrow(dataSubset) \* 2/3)  
titanictrainSet <- dataSubset[trainSplit, ]  
titanictestSet <- dataSubset[-trainSplit, ]

After selecting the columns, we’ll end up working with 8,400 records to train the model and 4,200 records to test the model.

# Training and testing the models.

## k-Nearest Neighbours

Starting off with kNN, we’ll build the model using two-thirds of the training set and then use the remaining third of the same data to test the results. Just to avoid redoing this process, we’ll set the seed for the split and work from there.

# We'll first start by converting the label variable into a factor. Because  
# we are building a classifier that will determine the probability that a  
# observation belongs to a certain number, we need to convert the labels  
# into factors that can then be converted to variable names  
titanictrainSet$label <- factor(paste0("X", titanictrainSet$label), levels = c("X0",   
 "X1", "X2", "X3", "X4", "X5", "X6", "X7", "X8", "X9"))  
titanictestSet$label <- factor(paste0("X", titanictestSet$label), levels = c("X0",   
 "X1", "X2", "X3", "X4", "X5", "X6", "X7", "X8", "X9"))  
  
# Now to start creating the kNN. We'll do three-fold CV as in the previous  
# exercise.  
set.seed(239)  
x <- trainControl(method = "repeatedcv", repeats = 3, classProbs = TRUE)  
knnTrain <- train(label ~ ., data = titanictrainSet, method = "knn", preProcess = c("center",   
 "scale"), trControl = x, metric = "ROC", tuneLength = 3)

## Warning in train.default(x, y, weights = w, ...): The metric "ROC" was not  
## in the result set. Accuracy will be used instead.

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel447,  
## pixel449, pixel477, pixel531, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel702, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel447,  
## pixel449, pixel477, pixel531, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel702, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel447,  
## pixel449, pixel477, pixel531, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel702, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel14, pixel15, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel642, pixel643, pixel670, pixel698, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel14, pixel15, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel642, pixel643, pixel670, pixel698, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel14, pixel15, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel642, pixel643, pixel670, pixel698, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel752,  
## pixel753, pixel761, pixel777, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel752,  
## pixel753, pixel761, pixel777, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel752,  
## pixel753, pixel761, pixel777, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel114, pixel167, pixel195,  
## pixel223, pixel224, pixel251, pixel336, pixel337, pixel363, pixel364,  
## pixel391, pixel419, pixel449, pixel477, pixel504, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel114, pixel167, pixel195,  
## pixel223, pixel224, pixel251, pixel336, pixel337, pixel363, pixel364,  
## pixel391, pixel419, pixel449, pixel477, pixel504, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel114, pixel167, pixel195,  
## pixel223, pixel224, pixel251, pixel336, pixel337, pixel363, pixel364,  
## pixel391, pixel419, pixel449, pixel477, pixel504, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel49, pixel50, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel724, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel49, pixel50, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel724, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel49, pixel50, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel724, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel34, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel34, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel34, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel169, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel762, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel169, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel762, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel169, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel762, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel365, pixel391, pixel393,  
## pixel419, pixel449, pixel477, pixel505, pixel533, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel365, pixel391, pixel393,  
## pixel419, pixel449, pixel477, pixel505, pixel533, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel365, pixel391, pixel393,  
## pixel419, pixel449, pixel477, pixel505, pixel533, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel587, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel587, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel587, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel365, pixel391, pixel393,  
## pixel419, pixel449, pixel477, pixel505, pixel533, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel752, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel365, pixel391, pixel393,  
## pixel419, pixel449, pixel477, pixel505, pixel533, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel752, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel365, pixel391, pixel393,  
## pixel419, pixel449, pixel477, pixel505, pixel533, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel752, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel777, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel777, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel777, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel762, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel762, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel762, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel49, pixel50, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel114, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel449, pixel477, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel49, pixel50, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel114, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel449, pixel477, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel49, pixel50, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel114, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel449, pixel477, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel702, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel702, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel702, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel12, pixel13, pixel14,  
## pixel15, pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113,  
## pixel167, pixel169, pixel224, pixel251, pixel336, pixel337, pixel363,  
## pixel364, pixel391, pixel419, pixel449, pixel477, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel12, pixel13, pixel14,  
## pixel15, pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113,  
## pixel167, pixel169, pixel224, pixel251, pixel336, pixel337, pixel363,  
## pixel364, pixel391, pixel419, pixel449, pixel477, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel12, pixel13, pixel14,  
## pixel15, pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113,  
## pixel167, pixel169, pixel224, pixel251, pixel336, pixel337, pixel363,  
## pixel364, pixel391, pixel419, pixel449, pixel477, pixel559, pixel561,  
## pixel588, pixel589, pixel615, pixel616, pixel617, pixel643, pixel698,  
## pixel725, pixel726, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel587, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel587, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel587, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel33, pixel34,  
## pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel447, pixel449, pixel477, pixel504, pixel531, pixel559,  
## pixel561, pixel588, pixel589, pixel615, pixel616, pixel617, pixel643,  
## pixel698, pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel33, pixel34,  
## pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel447, pixel449, pixel477, pixel504, pixel531, pixel559,  
## pixel561, pixel588, pixel589, pixel615, pixel616, pixel617, pixel643,  
## pixel698, pixel725, pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel33, pixel34,  
## pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel447, pixel449, pixel477, pixel504, pixel531, pixel559,  
## pixel561, pixel588, pixel589, pixel615, pixel616, pixel617, pixel643,  
## pixel698, pixel725, pixel726, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel195, pixel223,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel449, pixel477, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel195, pixel223,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel449, pixel477, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel195, pixel223,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel449, pixel477, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel169, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel169, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel32, pixel51, pixel58,  
## pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel169, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel505, pixel533, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel505, pixel533, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel505, pixel533, pixel559, pixel561, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel365, pixel391,  
## pixel393, pixel419, pixel449, pixel477, pixel559, pixel561, pixel588,  
## pixel589, pixel615, pixel616, pixel617, pixel643, pixel698, pixel702,  
## pixel725, pixel726, pixel752, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel365, pixel391,  
## pixel393, pixel419, pixel449, pixel477, pixel559, pixel561, pixel588,  
## pixel589, pixel615, pixel616, pixel617, pixel643, pixel698, pixel702,  
## pixel725, pixel726, pixel752, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel110, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel365, pixel391,  
## pixel393, pixel419, pixel449, pixel477, pixel559, pixel561, pixel588,  
## pixel589, pixel615, pixel616, pixel617, pixel643, pixel698, pixel702,  
## pixel725, pixel726, pixel752, pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel587, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel587, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel279, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel587, pixel588, pixel589,  
## pixel615, pixel616, pixel617, pixel643, pixel698, pixel725, pixel726,  
## pixel753, pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel14, pixel15, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel14, pixel15, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel14, pixel15, pixel51,  
## pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224,  
## pixel251, pixel336, pixel337, pixel363, pixel364, pixel391, pixel419,  
## pixel449, pixel477, pixel559, pixel561, pixel588, pixel589, pixel615,  
## pixel616, pixel617, pixel643, pixel698, pixel725, pixel726, pixel753,  
## pixel761, pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel80, pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel114, pixel167, pixel195,  
## pixel223, pixel224, pixel251, pixel336, pixel337, pixel363, pixel364,  
## pixel391, pixel419, pixel449, pixel477, pixel559, pixel561, pixel588,  
## pixel589, pixel615, pixel616, pixel617, pixel643, pixel698, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel114, pixel167, pixel195,  
## pixel223, pixel224, pixel251, pixel336, pixel337, pixel363, pixel364,  
## pixel391, pixel419, pixel449, pixel477, pixel559, pixel561, pixel588,  
## pixel589, pixel615, pixel616, pixel617, pixel643, pixel698, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel114, pixel167, pixel195,  
## pixel223, pixel224, pixel251, pixel336, pixel337, pixel363, pixel364,  
## pixel391, pixel419, pixel449, pixel477, pixel559, pixel561, pixel588,  
## pixel589, pixel615, pixel616, pixel617, pixel643, pixel698, pixel725,  
## pixel726, pixel753, pixel761, pixel778, pixel779

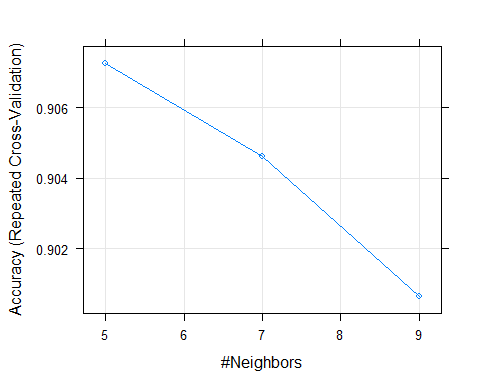
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel34, pixel49, pixel50,  
## pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel447, pixel449, pixel477, pixel504, pixel531, pixel559,  
## pixel561, pixel588, pixel589, pixel615, pixel616, pixel617, pixel643,  
## pixel698, pixel725, pixel726, pixel753, pixel761, pixel777, pixel778,  
## pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel34, pixel49, pixel50,  
## pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel447, pixel449, pixel477, pixel504, pixel531, pixel559,  
## pixel561, pixel588, pixel589, pixel615, pixel616, pixel617, pixel643,  
## pixel698, pixel725, pixel726, pixel753, pixel761, pixel777, pixel778,  
## pixel779  
  
## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel34, pixel49, pixel50,  
## pixel51, pixel58, pixel59, pixel81, pixel86, pixel87, pixel113, pixel167,  
## pixel224, pixel251, pixel336, pixel337, pixel363, pixel364, pixel391,  
## pixel419, pixel447, pixel449, pixel477, pixel504, pixel531, pixel559,  
## pixel561, pixel588, pixel589, pixel615, pixel616, pixel617, pixel643,  
## pixel698, pixel725, pixel726, pixel753, pixel761, pixel777, pixel778,  
## pixel779

## Warning in preProcess.default(thresh = 0.95, k = 5, freqCut = 19, uniqueCut  
## = 10, : These variables have zero variances: pixel51, pixel58, pixel59,  
## pixel81, pixel86, pixel87, pixel113, pixel167, pixel224, pixel251,  
## pixel336, pixel337, pixel363, pixel364, pixel391, pixel419, pixel449,  
## pixel477, pixel559, pixel561, pixel588, pixel589, pixel615, pixel616,  
## pixel617, pixel643, pixel698, pixel725, pixel726, pixel753, pixel761,  
## pixel778, pixel779

# Model summary  
knnTrain

## k-Nearest Neighbors   
##   
## 8400 samples  
## 708 predictor  
## 10 classes: 'X0', 'X1', 'X2', 'X3', 'X4', 'X5', 'X6', 'X7', 'X8', 'X9'   
##   
## Pre-processing: centered (708), scaled (708)   
## Resampling: Cross-Validated (10 fold, repeated 3 times)   
## Summary of sample sizes: 7561, 7561, 7558, 7561, 7560, 7561, ...   
## Resampling results across tuning parameters:  
##   
## k Accuracy Kappa   
## 5 0.9072603 0.8968522  
## 7 0.9046020 0.8938913  
## 9 0.9006326 0.8894715  
##   
## Accuracy was used to select the optimal model using the largest value.  
## The final value used for the model was k = 5.

plot(knnTrain)



# Given these numbers, it appears that the model is extremely good at  
# predicting the class label. Let's run the model on the test set.  
knnPred <- predict(knnTrain, titanictestSet, type = "prob")  
knnPred <- as.data.frame(knnPred)  
  
# Select maximum prediction. We subtract 1 because we are showing the class  
# label (i.e 0 is 1, 1 is 2, and so on.)  
knnPredictedValues <- data.frame(apply(knnPred, 1, which.max) - 1)  
colnames(knnPredictedValues) <- "prediction"  
  
# Compare to actual values  
knnResults <- titanictestSet %>% select(label) %>% mutate(real = str\_remove(label,   
 "X")) %>% bind\_cols(knnPredictedValues) %>% mutate(real = as.factor(real),   
 prediction = as.factor(prediction))  
  
confusionMatrix(knnResults$real, knnResults$prediction)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction 0 1 2 3 4 5 6 7 8 9  
## 0 395 0 0 0 0 2 3 0 2 0  
## 1 0 427 2 0 0 0 2 1 0 0  
## 2 8 10 401 10 5 1 4 7 7 1  
## 3 1 6 6 391 1 11 1 4 8 3  
## 4 1 11 2 0 343 2 0 4 0 20  
## 5 5 3 1 25 1 337 12 0 7 5  
## 6 10 1 1 0 2 4 392 0 0 0  
## 7 0 10 5 0 5 0 0 419 0 24  
## 8 6 6 6 15 7 23 4 1 323 8  
## 9 4 2 4 8 10 0 0 18 4 379  
##   
## Overall Statistics  
##   
## Accuracy : 0.9064   
## 95% CI : (0.8972, 0.9151)  
## No Information Rate : 0.1133   
## P-Value [Acc > NIR] : < 2.2e-16   
##   
## Kappa : 0.896   
##   
## Mcnemar's Test P-Value : NA   
##   
## Statistics by Class:  
##   
## Class: 0 Class: 1 Class: 2 Class: 3 Class: 4 Class: 5  
## Sensitivity 0.91860 0.8971 0.93692 0.8708 0.91711 0.88684  
## Specificity 0.99814 0.9987 0.98595 0.9891 0.98955 0.98455  
## Pos Pred Value 0.98259 0.9884 0.88326 0.9051 0.89556 0.85101  
## Neg Pred Value 0.99078 0.9870 0.99279 0.9846 0.99188 0.98870  
## Prevalence 0.10238 0.1133 0.10190 0.1069 0.08905 0.09048  
## Detection Rate 0.09405 0.1017 0.09548 0.0931 0.08167 0.08024  
## Detection Prevalence 0.09571 0.1029 0.10810 0.1029 0.09119 0.09429  
## Balanced Accuracy 0.95837 0.9479 0.96143 0.9299 0.95333 0.93570  
## Class: 6 Class: 7 Class: 8 Class: 9  
## Sensitivity 0.93780 0.92291 0.92023 0.86136  
## Specificity 0.99524 0.98825 0.98025 0.98670  
## Pos Pred Value 0.95610 0.90497 0.80952 0.88345  
## Neg Pred Value 0.99314 0.99063 0.99263 0.98382  
## Prevalence 0.09952 0.10810 0.08357 0.10476  
## Detection Rate 0.09333 0.09976 0.07690 0.09024  
## Detection Prevalence 0.09762 0.11024 0.09500 0.10214  
## Balanced Accuracy 0.96652 0.95558 0.95024 0.92403

The model takes a long time to run - training this model took two hours! - but the end results show that it is highly accurate when predicting class labels. We can see that for both the training and test sets, the accuracy is slightly over 90 percent which, so far, are the best results we have seen.

## Support Vector Machines

Now, our attention will turn to SVMs. Since the data has already been processed, we will train the model.

# Run the model using Linear Kernel to classify the data. with CV = 3 (3  
# k-fold cross validatiion)  
  
svmTrain <- svm(label ~ ., data = titanictrainSet, type = "C", kernel = "linear",   
 cross = 3, probability = TRUE)

## Warning in svm.default(x, y, scale = scale, ..., na.action = na.action):  
## Variable(s) 'pixel51' and 'pixel58' and 'pixel59' and 'pixel81' and  
## 'pixel86' and 'pixel87' and 'pixel113' and 'pixel167' and 'pixel224' and  
## 'pixel251' and 'pixel336' and 'pixel337' and 'pixel363' and 'pixel364' and  
## 'pixel391' and 'pixel419' and 'pixel449' and 'pixel477' and 'pixel559' and  
## 'pixel561' and 'pixel588' and 'pixel589' and 'pixel615' and 'pixel616' and  
## 'pixel617' and 'pixel643' and 'pixel698' and 'pixel725' and 'pixel726' and  
## 'pixel753' and 'pixel761' and 'pixel778' and 'pixel779' constant. Cannot  
## scale data.

summary(svmTrain)

##   
## Call:  
## svm(formula = label ~ ., data = titanictrainSet, type = "C",   
## kernel = "linear", cross = 3, probability = TRUE)  
##   
##   
## Parameters:  
## SVM-Type: C-classification   
## SVM-Kernel: linear   
## cost: 1   
##   
## Number of Support Vectors: 2391  
##   
## ( 307 232 151 207 285 269 314 228 139 259 )  
##   
##   
## Number of Classes: 10   
##   
## Levels:   
## X0 X1 X2 X3 X4 X5 X6 X7 X8 X9  
##   
## 3-fold cross-validation on training data:  
##   
## Total Accuracy: 90.85714   
## Single Accuracies:  
## 90.64286 91.03571 90.89286

As we can see the SVM results has a higher degree of accuracy within the training set than what kNN showed originally. Now to run the test set on the model.

svmPred <- predict(svmTrain, titanictestSet, type = "prob")  
svmPred <- as.data.frame(svmPred)  
colnames(svmPred) <- "results"  
  
# Now, build the data frame to compare the results.  
svmResults <- titanictestSet %>% select(label) %>% bind\_cols(svmPred) %>% mutate(real = factor(as.character(str\_remove(label,   
 "X"))), prediction = factor(as.character(str\_remove(results, "X"))))  
  
confusionMatrix(svmResults$real, svmResults$prediction)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction 0 1 2 3 4 5 6 7 8 9  
## 0 385 0 1 0 1 3 5 0 4 3  
## 1 0 423 1 1 0 0 1 1 5 0  
## 2 2 3 402 5 9 4 10 3 9 7  
## 3 0 2 10 371 1 15 3 3 18 9  
## 4 1 1 5 0 354 1 2 2 3 14  
## 5 4 2 0 20 3 334 13 0 15 5  
## 6 3 0 4 0 3 6 392 0 1 1  
## 7 1 2 4 0 10 1 0 425 1 19  
## 8 2 2 6 23 4 12 6 0 341 3  
## 9 1 0 2 8 22 2 0 17 10 367  
##   
## Overall Statistics  
##   
## Accuracy : 0.9033   
## 95% CI : (0.894, 0.9121)  
## No Information Rate : 0.1074   
## P-Value [Acc > NIR] : < 2.2e-16   
##   
## Kappa : 0.8926   
##   
## Mcnemar's Test P-Value : NA   
##   
## Statistics by Class:  
##   
## Class: 0 Class: 1 Class: 2 Class: 3 Class: 4 Class: 5  
## Sensitivity 0.96491 0.9724 0.92414 0.86682 0.86978 0.88360  
## Specificity 0.99553 0.9976 0.98619 0.98383 0.99235 0.98378  
## Pos Pred Value 0.95771 0.9792 0.88546 0.85880 0.92428 0.84343  
## Neg Pred Value 0.99631 0.9968 0.99119 0.98487 0.98611 0.98843  
## Prevalence 0.09500 0.1036 0.10357 0.10190 0.09690 0.09000  
## Detection Rate 0.09167 0.1007 0.09571 0.08833 0.08429 0.07952  
## Detection Prevalence 0.09571 0.1029 0.10810 0.10286 0.09119 0.09429  
## Balanced Accuracy 0.98022 0.9850 0.95516 0.92533 0.93107 0.93369  
## Class: 6 Class: 7 Class: 8 Class: 9  
## Sensitivity 0.90741 0.9424 0.83784 0.85748  
## Specificity 0.99522 0.9899 0.98471 0.98356  
## Pos Pred Value 0.95610 0.9179 0.85464 0.85548  
## Neg Pred Value 0.98945 0.9930 0.98264 0.98382  
## Prevalence 0.10286 0.1074 0.09690 0.10190  
## Detection Rate 0.09333 0.1012 0.08119 0.08738  
## Detection Prevalence 0.09762 0.1102 0.09500 0.10214  
## Balanced Accuracy 0.95132 0.9661 0.91127 0.92052

The model’s consistency is very impressive. It manages to also predict at a high accuracy.

## Random Forest

Let’s turn to the final model that we will work on: random forest.

x <- trainControl(method = "repeatedcv", number = 3, repeats = 3)  
rfTrain <- train(label ~ ., data = titanictrainSet, method = "rf", metric = "Accuracy",   
 trControl = x, type = "C")  
  
rfTrain

## Random Forest   
##   
## 8400 samples  
## 708 predictor  
## 10 classes: 'X0', 'X1', 'X2', 'X3', 'X4', 'X5', 'X6', 'X7', 'X8', 'X9'   
##   
## No pre-processing  
## Resampling: Cross-Validated (3 fold, repeated 3 times)   
## Summary of sample sizes: 5598, 5602, 5600, 5599, 5601, 5600, ...   
## Resampling results across tuning parameters:  
##   
## mtry Accuracy Kappa   
## 2 0.8709923 0.8563668  
## 37 0.9444041 0.9381774  
## 707 0.9178953 0.9087060  
##   
## Accuracy was used to select the optimal model using the largest value.  
## The final value used for the model was mtry = 37.

# With the given results, the train has a minimum accuracy on two variables  
# sampled, but may this increases once we reach 37.

Accuracy of 94.6 percent accuracy, is by far the superior model when comparing the three models we’ve built and the results according to the training set. But the most important thing is to compare the results of the testing set.

# Testing random forest model.  
rfPred <- predict(rfTrain, titanictestSet, type = "prob")  
rfPred <- as.data.frame(rfPred)  
  
rfPredictedValues <- data.frame(apply(rfPred, 1, which.max) - 1)  
colnames(rfPredictedValues) <- "results"  
  
# Random forest actual and predicted labels, and confusion matrix to  
# comparte the model's prediciton capabilities.  
rfResults <- titanictestSet %>% select(label) %>% bind\_cols(rfPredictedValues) %>%   
 mutate(real = factor(as.character(str\_remove(label, "X"))), prediction = factor(results))  
  
confusionMatrix(rfResults$real, rfResults$prediction)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction 0 1 2 3 4 5 6 7 8 9  
## 0 392 0 0 1 1 0 4 0 4 0  
## 1 0 425 1 1 0 1 1 1 2 0  
## 2 2 0 426 3 6 1 4 5 5 2  
## 3 0 1 9 402 2 5 0 4 4 5  
## 4 0 2 2 0 355 0 5 2 3 14  
## 5 3 1 0 8 1 373 7 1 0 2  
## 6 3 0 0 0 1 2 400 1 3 0  
## 7 0 2 9 0 3 0 0 434 1 14  
## 8 1 4 2 3 5 6 7 0 365 6  
## 9 3 2 3 8 8 3 0 6 7 389  
##   
## Overall Statistics  
##   
## Accuracy : 0.9431   
## 95% CI : (0.9357, 0.9499)  
## No Information Rate : 0.1081   
## P-Value [Acc > NIR] : < 2.2e-16   
##   
## Kappa : 0.9367   
##   
## Mcnemar's Test P-Value : NA   
##   
## Statistics by Class:  
##   
## Class: 0 Class: 1 Class: 2 Class: 3 Class: 4 Class: 5  
## Sensitivity 0.97030 0.9725 0.9425 0.94366 0.92932 0.95396  
## Specificity 0.99737 0.9981 0.9925 0.99205 0.99267 0.99396  
## Pos Pred Value 0.97512 0.9838 0.9383 0.93056 0.92689 0.94192  
## Neg Pred Value 0.99684 0.9968 0.9931 0.99363 0.99293 0.99527  
## Prevalence 0.09619 0.1040 0.1076 0.10143 0.09095 0.09310  
## Detection Rate 0.09333 0.1012 0.1014 0.09571 0.08452 0.08881  
## Detection Prevalence 0.09571 0.1029 0.1081 0.10286 0.09119 0.09429  
## Balanced Accuracy 0.98383 0.9853 0.9675 0.96786 0.96099 0.97396  
## Class: 6 Class: 7 Class: 8 Class: 9  
## Sensitivity 0.93458 0.9559 0.92640 0.90046  
## Specificity 0.99735 0.9923 0.99107 0.98938  
## Pos Pred Value 0.97561 0.9374 0.91479 0.90676  
## Neg Pred Value 0.99261 0.9946 0.99237 0.98860  
## Prevalence 0.10190 0.1081 0.09381 0.10286  
## Detection Rate 0.09524 0.1033 0.08690 0.09262  
## Detection Prevalence 0.09762 0.1102 0.09500 0.10214  
## Balanced Accuracy 0.96596 0.9741 0.95873 0.94492

So results are on par with what was expected: 94.7 percent accuracy, by far the best out of these three models.

# Conclusion based on the above results :

The Models tested presented very high degrees of accuracy, (over 90 percent of the time). However, points to remember is a)The size of the data is immense, we had to cut corners. b) We dismissed using the original test set as it did not have class labels, which we would’ve only been able to actually test in Kaggle. 3) In consideration to this, training dataset was used. And additionally, considering the size of this dataset(was over 40,000 records long), we were working with 30 percent of this - with all observations chosen at random. Since this was done prior to any splitting into a new training and testing set, there is no chance for observations to appear in both sets. Also, what’s concerning is that maybe this would probably cause the models to overfit because of the limited size of the training sample. This could be ammended by using a larger set and adding more computing power. However to conclude, the results we got ’re satisfying and merit additional investigation towards finer parameter tuning and slowly add more observations to the training set to generate the best model.