```
library(ElemStatLearn)
data(SAheart)
head(SAheart)

# 1. QDA

library(MASS)

#a)
set.seed(54)
train <- sample(1:nrow(SAheart),300)
train.data <- SAheart[train,]
qda.m <- qda(chd~., data=train.data)
qda.pred <- predict(qda.m, newdat=SAheart[-train,])
tab.qda <- table(SAheart[-train,"chd"], qda.pred$class)
missclass <- 1-sum(diag(tab.qda))/sum(tab.qda)

#b)
mis.qda <- rep(0,100)
mis.lda <- rep(0,100)
for (i in 1:100){
  train <- sample(1:nrow(SAheart),300)
  train.data <- SAheart[train,]
  #qda
  qda.m <- qda(chd~., data=train.data)
  qda.pred <- predict(qda.m, newdat=SAheart[-train,])
  tab.qda <- table(SAheart[-train,"chd"], qda.pred$class)
  mis.qda[i] <- 1-sum(diag(tab.qda))/sum(tab.qda)

  #lda-teil
  lda.m <- lda(chd~., data=train.data)
  lda.pred <- predict(lda.m, newdat=SAheart[-train,])
  tab.lda <- table(SAheart[-train,"chd"], lda.pred$class)
  mis.lda[i] <- 1-sum(diag(tab.lda))/sum(tab.lda)

}

boxplot(mis.qda, mis.lda, names=c("qda","lda"))

# 2. RDA
library(klaR)

#a)
set.seed(652)
train <- sample(1:nrow(SAheart),300)
train.data <- SAheart[train,]
rda.m <- rda(chd~., data=train.data)
rda.pred <- predict(rda.m, newdat=SAheart[-train,])
tab.rda <- table(SAheart[-train,"chd"],rda.pred$class)
missclass<- 1-sum(diag(tab.rda))/sum(tab.rda)

#b,c)

mis.rda <- rep(0,20)
mis.rda00 <- rep(0,20)
mis.rda01 <- rep(0,20)
for (i in 1:20){
  train <- sample(1:nrow(SAheart),300)
  train.data <- SAheart[train,]
```

```
  #rda
  rda.m <- rda(chd~., data=train.data)
  rda.pred <- predict(rda.m, newdat=SAheart[-train,])
  tab.rda<-table(SAheart[-train,"chd"], rda.pred$class)
  mis.rda[i]<- 1-sum(diag(tab.rda))/sum(tab.rda)

  #rda: gamma = 0 , lambda = 0 (qda)
  rda00.m <- rda(chd~., data=train.data, gamma=0, lambda=0)
  rda00.pred <- predict(rda00.m, newdat=SAheart[-train,])
  tab.rda00 <- table(SAheart[-train,"chd"],rda00.pred$class)
  mis.rda00[i] <- 1-sum(diag(tab.rda00))/sum(tab.rda00)

  #rda: gamma = 0 , lambda = 1 (lda)
  rda01.m <- rda(chd~.,data=train.data, gamma=0, lambda=1)
  rda01.pred <- predict(rda01.m, newdat=SAheart[-train,])
  tab.rda01 <- table(SAheart[-train,"chd"], rda01.pred$class)
  mis.rda01[i] <- 1-sum(diag(tab.rda01))/sum(tab.rda01)
}
boxplot(mis.qda[1:20],mis.lda[1:20],mis.rda[1:20],mis.rda00[1:20],mis.rda01[1:20],
        names=c("qda","lda","rda","rda00","rda01"))

# LDA gives the lowest rate of misclassifications and therefore offers
# a better solution. RDA with optimal parameter choice is better than the
# ones with fixed parameters.

# 3. LR
l
# a)
set.seed(33)
train <- sample(1:nrow(SAheart), 300)
glm.m <- glm(chd~., data=SAheart[train, ], family=binomial)
glm.pred <- predict(glm.m, newdat=SAheart[-train,])
plot(glm.pred, col=as.numeric(SAheart$chd)+2)
abline(h=0)
tab.glm <- table(SAheart[-train,"chd"], ifelse(glm.pred<0, 0, 1))
missclass <- 1-sum(diag(tab.glm))/sum(tab.glm)

# b)
mis.qda <- rep(0, 100)
mis.lda <- rep(0, 100)
mis.rda <- rep(0, 100)
mis.glm <- rep(0, 100)

for (i in 1:100){
  train <- sample(1:nrow(SAheart), 300)
  train.data <- SAheart[train,]

  # qda
  qda.m <- qda(chd~., data=train.data)
  qda.pred <- predict(qda.m, newdat=SAheart[-train,])
  tab.qda <- table(SAheart[-train, "chd"], qda.pred$class)
  mis.qda[i] <- 1-sum(diag(tab.qda))/sum(tab.qda)

  # lda
  lda.m <- lda(chd~., data=train.data)
  lda.pred <- predict(lda.m, newdat=SAheart[-train,])
  tab.lda <- table(SAheart[-train, "chd"], lda.pred$class)
  mis.lda[i] <- 1-sum(diag(tab.lda))/sum(tab.lda)
```

```
  # rda
  rda.m <- rda(chd~., data=train.data, gamma=0)
  rda.pred <- predict(rda.m, newdat=SAheart[-train,])
  tab.rda <- table(SAheart[-train, "chd"], rda.pred$class)
  mis.rda[i] <- 1-sum(diag(tab.rda))/sum(tab.rda)

  # glm
  glm.m <- glm(chd~., data=train.data, family=binomial)
  glm.pred <- predict(glm.m, newdat=SAheart[-train,])
  tab.glm <- table(SAheart[-train,"chd"], ifelse(glm.pred<0, 0, 1))
  mis.glm[i] <- 1-sum(diag(tab.glm))/sum(tab.glm)
}
boxplot(mis.lda, mis.qda, mis.rda, mis.glm,
        names=c("lda", "qda", "rda", "glm"))

#lda is better (glm shows very similar results)
```