Question 17,18

Vyshnavi Pisupati

16/11/2020

Question 17

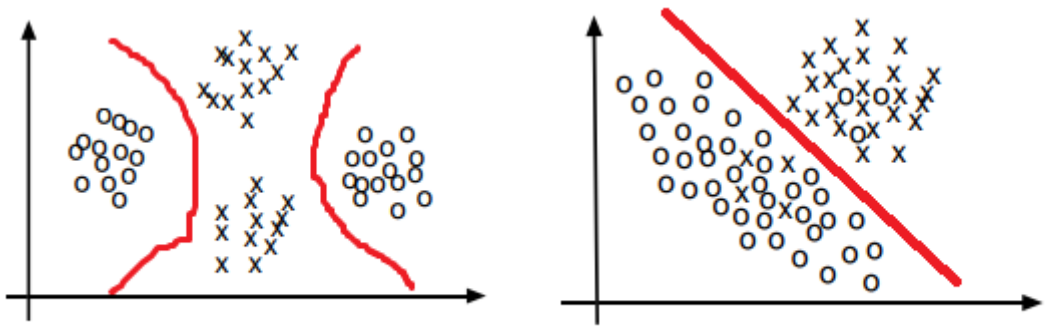
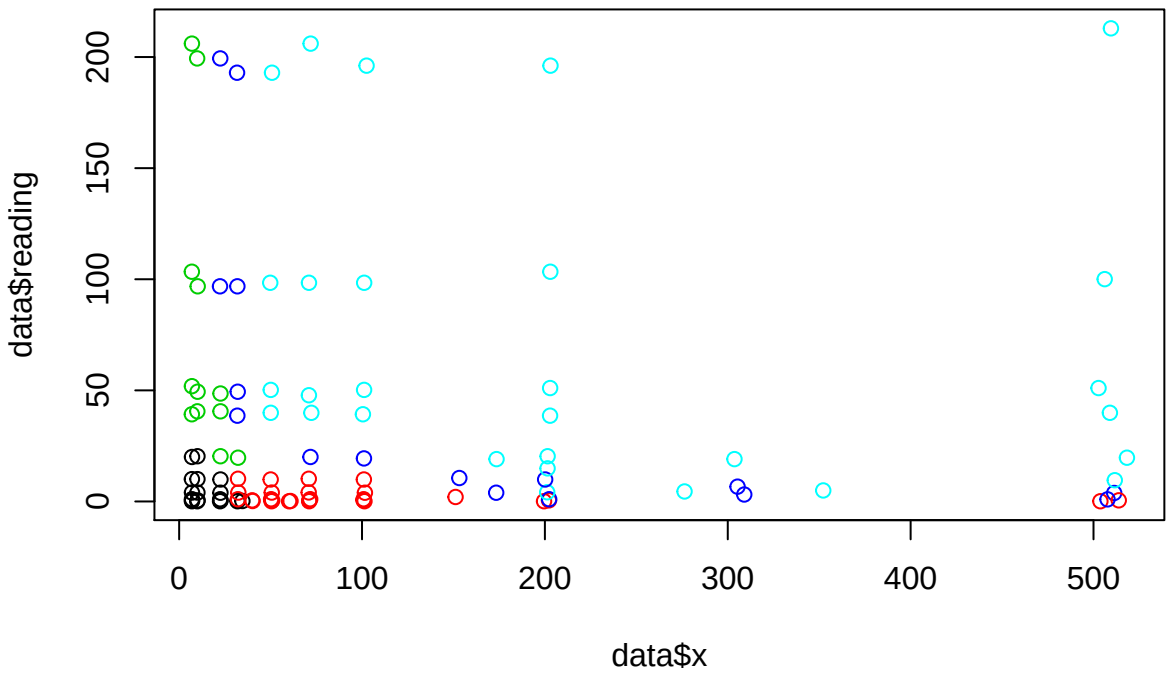


Figure 1: "SVM"

Question 18

```
library(kknn)
library(knitr)
r1 <- read.csv(file = "D:/Thesis/Exam/regime_I.csv",
               col.names = c("x","reading"))
r2 <- read.csv(file = "D:/Thesis/Exam/regime_II.csv",
               col.names = c("x","reading"))
r3 <- read.csv(file = "D:/Thesis/Exam/regime_III.csv",
               col.names = c("x","reading"))
r4 <- read.csv(file = "D:/Thesis/Exam/regime_IV.csv",
               col.names = c("x","reading"))
r5 <- read.csv(file = "D:/Thesis/Exam/regime_V.csv",
               col.names = c("x","reading"))

r1$type <- 1

r2$type <- 2
```

```
r3$type <- 3

r4$type <- 4

r5$type <- 5

n1 <- nrow(r1)
set.seed(12345)
id1<-sample(1:n1, floor(n1*0.5))

n2 <- nrow(r2)
set.seed(12345)
id2<-sample(1:n2, floor(n2*0.5))

n3 <- nrow(r3)
set.seed(12345)
id3<-sample(1:n3, floor(n3*0.5))

n4 <- nrow(r4)
set.seed(12345)
id4<-sample(1:n4, floor(n4*0.5))

n5 <- nrow(r5)
set.seed(12345)
id5<-sample(1:n5, floor(n5*0.5))

train <- rbind(r1[id1,],r2[id2,],r3[id3,],r4[id4,],r5[id5,])
test <- rbind(r1[-id1,],r2[-id2,],r3[-id3,],r4[-id4,],r5[-id5,])

data <- rbind(train,test)

plot(data$x,data$reading,col=data$type)
```




```
#Fitting a model using kknn model for K=5
output_test_30<-kknn(formula=type~.,train=train,test=test,k=5)
output_train_30<-kknn::kknn(formula=type~.,train=train,test=train,k=5)
fv_test<-output_test_30$fitted.values
fv_train<-output_train_30$fitted.values
#Classifying
fv_test<-ifelse(fv_test>0.5,1,0)
fv_train<-ifelse(fv_train>0.5,1,0)
#Confusion matrices and misclassification rates

cm_kknn_train_30<-table(train$type,fv_train)
mc_train_kknn_30<-1-(sum(diag(cm_kknn_train_30))/sum(cm_kknn_train_30))


cm_kknn_test_30<-table(test$type,fv_test)
mc_test_kknn_30<-1-(sum(diag(cm_kknn_test_30))/sum(cm_kknn_test_30))

cat("\nMissclassifiaction",mc_test_kknn_30)
```

Missclassifiaction 0.8113208

```
#Fitting model using kknn model with K=2
output_test_1<-kknn::kknn(formula = type~.,train=train,test = test,k=2)
output_train_1<-kknn::kknn(formula = type~.,train=train,test = train,k=2)
fv_test_1<-output_test_1$fitted.values
fv_train_1<-output_train_1$fitted.values
```

```r
#Classifying
fv_test_1<-ifelse(fv_test_1>0.5,1,0)
fv_train<-ifelse(fv_train>0.5,1,0)
#Confusion matrices and misclassification rates
cm_kknn_train_1<-table(train$type,fv_train_1)
mc_train_kknn_1<-1-(sum(diag(cm_kknn_train_1))/sum(cm_kknn_train_1))
cm_kknn_test_1<-table(test$type,fv_test_1)
mc_test_kknn_1<-1-(sum(diag(cm_kknn_test_1))/sum(cm_kknn_test_1))

cat("\nMissclassifiaction",mc_test_kknn_1)
```

Missclassifiaction 0.8113208

Looking at the misclassification rate, we can see that model performs moderately but still the performance of the model can be improved by tuning the parameters and and the threshold probability.