hw2.R

Muhammad

2021-09-01

#Question2.2  
  
library(kernlab)  
library(kknn)  
  
  
setwd("C:/Users/Muhammad/ISYE")  
  
data<-read.table("credit\_card\_data-headers.txt", header = TRUE)  
  
#ksvm model  
model <- ksvm(as.matrix(data[,1:10]),as.factor(data[,11]), C=100, scaled =TRUE, kernel="vanilladot", type = "C-svc")

## Setting default kernel parameters

#coeeficients of our svm  
a <- colSums(model@xmatrix[[1]] \* model@coef[[1]])  
a

## A1 A2 A3 A8 A9   
## -0.0010065348 -0.0011729048 -0.0016261967 0.0030064203 1.0049405641   
## A10 A11 A12 A14 A15   
## -0.0028259432 0.0002600295 -0.0005349551 -0.0012283758 0.1063633995

#intercept  
a0<- model@b  
a0

## [1] -0.08158492

#Assigning number of correct predictions to pred  
pred<- predict(model,data[,1:10])  
Accuracy\_of\_model <- sum(pred==data[,11])/nrow(data)  
Accuracy\_of\_model

## [1] 0.8639144

# Our model Accuracy is 86.39%  
  
#Q2.2.3  
   
accuracy\_function = function(x){  
   
 prediction<- matrix(nrow=nrow(data),ncol=1)   
 for (i in 1:nrow(data)){  
 knn\_model=kknn(R1~A1+A2+A3+A8+A9+A10+A11+A12+A14+A15,data[-i,],data[i,],k=x, scale = T)  
 # Rounding off by 50% threshold   
 prediction[i,1] = as.integer(fitted(knn\_model)+0.5)   
 }  
 accuracy = sum(prediction[,1] == data[,11]) / nrow(data)  
 return(accuracy)  
}  
  
K\_test <- matrix(ncol=1, nrow=27)  
#Usually value of K is root of number of sample size, hence K would be 25 or 26, to confirm lets loop all the number till 27  
for (i in 1:27){  
 K\_test[i,1] = accuracy\_function(i)   
}  
ind <- which.max(K\_test)  
ind

## [1] 12

K\_test[ind,1]

## [1] 0.853211

#I expected value of K to be approximately around 25 but that isn't the case. K is 12 in this case, as accuracy maximizes to 85.32%  
  
index\_vector<- matrix(1:27,ncol=1,nrow=27)  
plot(index\_vector[,1],K\_test[,1])

