KNN(K-NEAREST NEIGHBOUR)

#KNN

getwd()

setwd("C:\\Users\\10ani\\Desktop\\DS\_classes")

prostate\_cancer=read.csv("Prostate\_Cancer.csv")

prostate\_cancer$id=NULL

str(prostate\_cancer)

table(prostate\_cancer$diagnosis\_result)

normalize=function(x){

return((x-min(x))/(max(x)-min(x)))}

lapply(prostate\_cancer[,c(2:9)],normalize)

pros\_cac=data.frame(lapply(prostate\_cancer[,c(2:9)],normalize))

prostate\_cac=cbind(prostate\_cancer$diagnosis\_result,pros\_cac)

colnames(prostate\_cac)

colnames(prostate\_cac)[1]=c("diag")

ind=sort(sample(nrow(prostate\_cac),nrow(prostate\_cac)\*0.8)) #sampling

train=prostate\_cac[ind,]

test=prostate\_cac[-ind,]

#install.packages("class")

library(class)

prc\_pred=knn(train = train[,-1], test[,-1], cl=train$diag, k=5)

View(prc\_pred)

prc\_pred

tbl=table(test$diag,prc\_pred)

sum(diag(tbl))/sum(tbl)