MD2201: Data Science

Name of the student: Bhavin Patil **Roll No. 78**

Div: D **Batch: B-3**

Date of performance: 11-05-2022

Experiment No.5

Title: Classifier.

Aim: To apply Nearest Neighbor algorithm.

Software used: Programming language R.

Code Statement:

Consider 18 points data set referred in theory class. Consider test sample P (3,2). Apply following algorithms to find the class of this test point.

i. NN
ii. KNN with K=5 and K=7.
iii. MKNN with K=5
iv. R-NN Radius based algorithm with radius as 1.45 units.

Code:

```
cat("Nearest Neighbour\n")
dataset <- read.csv("knn1_csv.csv")
eucld <- sqrt(((3-dataset$x)**2)+((2-dataset$y)**2))
dataset <- cbind(dataset,eucld)
sorted_data <- dataset[order(dataset$eucld),]
cat("Class of P for NN is", sorted_data[1,4])

cat("\n----------------------------\n")
cat("\n\nK-Nearest Neighbour\n")

#k=5

df <- sorted_data[1:5,]
l1<-length(which(df$class==1))
l2<-length(which(df$class==2))
l3<-length(which(df$class==3))
if(l1>l2 & l1>l3){
  cat("class of P for KNN where k = 5 is : ",1)
  }
```

```
if(l2>l1 & l2 > l3){
  cat("class of P for KNN where k = 5 is : ",2)
}else{
  cat("class of P for KNN where k = 5 is : ",3)
}

#k=7

df1 <- sorted_data[1:7,]
l1<-length(which(df1$class==1))
l2<-length(which(df1$class==2))
l3<-length(which(df1$class==3))
if(l1>l2 & l1>l3){
  cat("\nclass of P for KNN where k = 7 is : ",1)
}
if(l2>l1 & l2 > l3){
  cat("\nclass of P for KNN where k = 7 is : ",2)
}else{
  cat("\nclass of P for KNN where k = 7 is : ",3)
}


cat("\n----------------------------\n")
cat("\n\nR-Nearest Neighbour\n")

#R = 1.45

df2 <- sorted_data[sorted_data$eucld<1.45,]
l1<-length(which(df2$class==1))
l2<-length(which(df2$class==2))
l3<-length(which(df2$class==3))
if(l1>l2 & l1>l3){
  cat("\nclass of P for RNN where R = 1.45 is : ",1)
}
if(l2>l1 & l2 > l3){
  cat("\nclass of P for RNN where R = 1.45 is : ",2)
}else{
  cat("\nclass of P for RNN where R = 1.45 is : ",3)
}

cat("\n----------------------------\n")
cat("\n\nModified K-Nearest Neighbour\n")

#K = 5

df3 <- sorted_data[1:5,]
dk <- max(df3$eucld)
d1 <- min(df3$eucld)
w <- (dk - df3$eucld)/(dk-d1)
df3 <- cbind(df3,w)
```

```
sum1 <- sum(df3$w[df3$class==1])
sum2 <- sum(df3$w[df3$class==2])
sum3 <- sum(df3$w[df3$class==3])
if(sum1>sum2 & sum1>sum3){
  cat("\nClass of P for MKNN where K = 5 is : ",1)
}
if(sum2>sum1 & sum2>sum3){
  cat("\nClass of P for MKNN where K = 5 is : ",2)
}else{
    cat("\nClass of P for MKNN where K = 5 is : ",3)
}
```

Results:

Nearest Neighbour

Class of P for NN is 3

K-Nearest Neighbour

class of P for KNN where k = 5 is : 3

class of P for KNN where k = 7 is : 2

R-Nearest Neighbour

class of P for RNN where R = 1.45 is : 2

Modified K-Nearest Neighbour

Class of P for MKNN where K = 5 is : 3

Conclusion: Implemented the different distance metric classifiers like NN, KNN, MKNN and RNN on 18 points dataset which was referred in the theory class we have successfully classified the class for test sample.