**Assignment-KNN**

1. Prepare a model for glass classification using KNN

#Glass equation

glass<-read.csv("F:/Data Scientist/KNN/glass.csv")

View(glass)

#table of diagnosis

table(glass$Type)

#table of prop with more informative labels

round(prop.table(table(glass$Type))\*100,digits=1)

#create normalixation function

normalize<-function(x)

{

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

}

#normalize glass data

glass\_n<-as.data.frame(lapply(glass[1:9],normalize))

glass\_nl<-cbind(glass\_n,glass$Type)

#create trainig and test data

glass\_train<-glass\_n[1:154,]

glass\_test<-glass\_n[155:214,]

#create labels for train and test

glass\_train\_labels<-glass[1:154,1]

glass\_test\_labels<-glass[155:214,1]

#training a model on data

#load library class

library(class)

glass\_test\_pred<-knn(train=glass\_train,test=glass\_test,cl=glass\_train\_labels,k=5)

#Evaluating performance

#library load "gmodels

library(gmodels)

CrossTable(x=glass\_test\_labels,y=glass\_test\_pred,prop.chisq=FALSE,prop.c=FALSE

,prop.r=FALSE)

**ITERATIONS IN R**

> glass<-glass[,-1]

> #table of diagnosis

> table(glass$Type)

1 2 3 5 6 7

70 76 17 13 9 29

> #table of prop with more informative labels

> round(prop.table(table(glass$Type))\*100,digits=1)

1 2 3 5 6 7

32.7 35.5 7.9 6.1 4.2 13.6

#normalize glass data

> glass\_n<-as.data.frame(lapply(glass[1:9],normalize))

> glass\_nl<-cbind(glass\_n,glass$Type)

> View(glass\_nl)

> View(glass\_n)

> glass\_nl<-cbind(glass\_n,glass$Type)

> #create trainig and test data

> glass\_train<-glass\_n[1:154,]

> glass\_test<-glass\_n[155:214,]

> #create labels for train and test

> glass\_train\_labels<-glass[1:154,1]

> glass\_test\_labels<-wbcd[155:214,1]

Cell Contents

|-------------------------|

| N |

| N / Table Total |

|-------------------------|

Total Observations in Table: 60

| glass\_test\_pred

glass\_test\_labels | 11.02 | 11.23 | 11.45 | 12.16 | 12.3 | 12.35 | 12.64 | 12.85 | 12.96 | 13.14 | 13.24 | 13.33 | 13.36 | 13.41 | 13.42 | 13.55 | 13.7 | 13.75 | 13.78 | 13.98 | 14.21 | 14.43 | Row Total

1. Zoo Data

#zoo data

zoo<-read.csv("F:/Data Scientist/KNN/zoo.csv")

View(zoo)

#table of diagnosis

table(zoo$type)

#table of prop with more informative labels

round(prop.table(table(zoo$type))\*100,digits=1)

#create normalixation function

normalize<-function(x)

{

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

}

#normalize zoo data

zoo\_n<-as.data.frame(lapply(zoo[2:18],normalize))

zoo\_nl<-cbind(zoo\_n,zoo$type)

#create trainig and test data

zoo\_train<-zoo\_n[1:71,]

zoo\_test<-zoo\_n[72:101,]

#create labels for train and test

zoo\_train\_labels<-zoo[1:71,1]

zoo\_test\_labels<-zoo[72:101,1]

#training a model on data

#load library class

library(class)

zoo\_test\_pred<-knn(train=zoo\_train,test=zoo\_test,cl=zoo\_train\_labels,k=5)

#Evaluating performance

#library load "gmodels

library(gmodels)

CrossTable(x=zoo\_test\_labels,y=zoo\_test\_pred,prop.chisq=FALSE,prop.c=FALSE

,prop.r=FALSE)

**ITERATIONS in R**

#zoo data

> zoo<-read.csv("F:/Data Scientist/KNN/zoo.csv")

> View(zoo)

> #table of diagnosis

> table(zoo$type)

1 2 3 4 5 6 7

41 20 5 13 4 8 10

> #table of prop with more informative labels

> round(prop.table(table(zoo$type))\*100,digits=1)

1 2 3 4 5 6 7

40.6 19.8 5.0 12.9 4.0 7.9 9.9

> #create normalixation function

> normalize<-function(x)

+ {

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

+ }

> #normalize zoo data

> zoo\_n<-as.data.frame(lapply(zoo[2:18],normalize))

> zoo\_nl<-cbind(zoo\_n,zoo$type)

> #create trainig and test data

> zoo\_train<-zoo\_n[1:71,]

> zoo\_test<-zoo\_n[72:101,]

> #create labels for train and test

> zoo\_train\_labels<-zoo[1:71,1]

> zoo\_test\_labels<-zoo[72:101,1]

> #training a model on data

> #load library class

> library(class)

> zoo\_test\_pred<-knn(train=zoo\_train,test=zoo\_test,cl=zoo\_train\_labels,k=5)

> #Evaluating performance

> #library load "gmodels

> library(gmodels)

> CrossTable(x=zoo\_test\_labels,y=zoo\_test\_pred,prop.chisq=FALSE,prop.c=FALSE

+ ,prop.r=FALSE)

Cell Contents

|-------------------------|

| N |

| N / Table Total |

|-------------------------|

Total Observations in Table: 30

| zoo\_test\_pred

zoo\_test\_labels | antelope | bass | catfish | chub | clam | crab | crayfish | dove | duck | flamingo | flea | frog | giraffe | gull | haddock | honeybee | ladybird | lark | leopard | lion | mink | moth | opossum | oryx | ostrich | penguin | pheasant | piranha | Row Total |