msasnur_UniversalBank

msasnur@kent.edu

22/10/2019

```
#Assignment 1 - Machine Learning#
#----Email- msasnur@kent.edu----#
#-------Date - 10/13/2019-------#

library(readr)

ubank<-read.csv("UniversalBank.csv")
View(ubank)

library(caret)

## Loading required package: lattice

## Loading required package: ggplot2

library(ISLR)

head(ubank,5)

##   ID Age Experience Income ZIP.Code Family CCAvg Education Mortgage
## 1  1  25          1     49    91107      4   1.6         1        0
## 2  2  45         19     34    90089      3   1.5         1        0
## 3  3  39         15     11    94720      1   1.0         1        0
## 4  4  35          9    100    94112      1   2.7         2        0
## 5  5  35          8     45    91330      4   1.0         2        0
##   Personal.Loan Securities.Account CD.Account Online CreditCard
## 1             0                  1          0      0          0
## 2             0                  1          0      0          0
## 3             0                  0          0      0          0
## 4             0                  0          0      0          0
## 5             0                  0          0      0          1

summary(ubank)

##        ID            Age          Experience       Income
##  Min.   :   1   Min.   :23.00   Min.   :-3.0   Min.   :  8.00
##  1st Qu.:1251   1st Qu.:35.00   1st Qu.:10.0   1st Qu.: 39.00
##  Median :2500   Median :45.00   Median :20.0   Median : 64.00
##  Mean   :2500   Mean   :45.34   Mean   :20.1   Mean   : 73.77
##  3rd Qu.:3750   3rd Qu.:55.00   3rd Qu.:30.0   3rd Qu.: 98.00
##  Max.   :5000   Max.   :67.00   Max.   :43.0   Max.   :224.00
##     ZIP.Code         Family          CCAvg          Education
##  Min.   : 9307   Min.   :1.000   Min.   : 0.000   Min.   :1.000
```

```
##  1st Qu.:91911   1st Qu.:1.000   1st Qu.: 0.700   1st Qu.:1.000
##  Median :93437   Median :2.000   Median : 1.500   Median :2.000
##  Mean   :93153   Mean   :2.396   Mean   : 1.938   Mean   :1.881
##  3rd Qu.:94608   3rd Qu.:3.000   3rd Qu.: 2.500   3rd Qu.:3.000
##  Max.   :96651   Max.   :4.000   Max.   :10.000   Max.   :3.000
##     Mortgage     Personal.Loan   Securities.Account   CD.Account
##  Min.   :  0.0   Min.   :0.000   Min.   :0.0000     Min.   :0.0000
##  1st Qu.:  0.0   1st Qu.:0.000   1st Qu.:0.0000     1st Qu.:0.0000
##  Median :  0.0   Median :0.000   Median :0.0000     Median :0.0000
##  Mean   : 56.5   Mean   :0.096   Mean   :0.1044     Mean   :0.0604
##  3rd Qu.:101.0   3rd Qu.:0.000   3rd Qu.:0.0000     3rd Qu.:0.0000
##  Max.   :635.0   Max.   :1.000   Max.   :1.0000     Max.   :1.0000
##      Online         CreditCard
##  Min.   :0.0000   Min.   :0.000
##  1st Qu.:0.0000   1st Qu.:0.000
##  Median :1.0000   Median :0.000
##  Mean   :0.5968   Mean   :0.294
##  3rd Qu.:1.0000   3rd Qu.:1.000
##  Max.   :1.0000   Max.   :1.000
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
m_ubank <- ubank[,c(-1,-5)]
m_ubank$Personal.Loan<-factor(m_ubank$Personal.Loan)

#Data Partition - Training Data = 60% and Validation Data = 40%

set.seed(20)
Train_Index = createDataPartition(m_ubank$Age,p=0.6, list=FALSE)
Train_Data = m_ubank[Train_Index,]
Val_Data = m_ubank[-Train_Index,]

Test_Index = createDataPartition(m_ubank$Age,p=0.2, list=FALSE)
Test_Data = m_ubank[Test_Index,]
Traval_Data = m_ubank[-Test_Index,]
View(m_ubank)
summary(Train_Data)
```

```
##       Age          Experience        Income           Family
##  Min.   :23.00   Min.   :-3.00   Min.   :  8.00   Min.   :1.000
```

```
##  1st Qu.:35.00   1st Qu.:10.00   1st Qu.: 39.00   1st Qu.:1.000
##  Median :45.00   Median :20.00   Median : 64.00   Median :2.000
##  Mean   :45.38   Mean   :20.18   Mean   : 74.75   Mean   :2.415
##  3rd Qu.:55.00   3rd Qu.:30.00   3rd Qu.:102.00   3rd Qu.:4.000
##  Max.   :67.00   Max.   :43.00   Max.   :224.00   Max.   :4.000
##      CCAvg         Education        Mortgage      Personal.Loan
##  Min.   : 0.00   Min.   :1.000   Min.   :  0.00   0:2700
##  1st Qu.: 0.70   1st Qu.:1.000   1st Qu.:  0.00   1: 301
##  Median : 1.50   Median :2.000   Median :  0.00
##  Mean   : 1.94   Mean   :1.866   Mean   : 57.78
##  3rd Qu.: 2.60   3rd Qu.:3.000   3rd Qu.:100.00
##  Max.   :10.00   Max.   :3.000   Max.   :635.00
##  Securities.Account   CD.Account          Online         CreditCard
##  Min.   :0.0000     Min.   :0.00000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000     1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000     Median :0.00000   Median :1.0000   Median :0.0000
##  Mean   :0.0993     Mean   :0.05598   Mean   :0.5978   Mean   :0.2982
##  3rd Qu.:0.0000     3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000     Max.   :1.00000   Max.   :1.0000   Max.   :1.0000
```

```
summary(Val_Data)
```

```
##       Age           Experience        Income           Family
##  Min.   :23.00   Min.   :-3.00   Min.   :  8.00   Min.   :1.000
##  1st Qu.:35.00   1st Qu.:10.00   1st Qu.: 39.00   1st Qu.:1.000
##  Median :45.00   Median :20.00   Median : 62.00   Median :2.000
##  Mean   :45.27   Mean   :19.99   Mean   : 72.31   Mean   :2.369
##  3rd Qu.:55.00   3rd Qu.:30.00   3rd Qu.: 93.00   3rd Qu.:3.000
##  Max.   :67.00   Max.   :42.00   Max.   :204.00   Max.   :4.000
##      CCAvg          Education        Mortgage      Personal.Loan
##  Min.   : 0.000   Min.   :1.000   Min.   :  0.00   0:1820
##  1st Qu.: 0.700   1st Qu.:1.000   1st Qu.:  0.00   1: 179
##  Median : 1.500   Median :2.000   Median :  0.00
##  Mean   : 1.935   Mean   :1.903   Mean   : 54.57
##  3rd Qu.: 2.500   3rd Qu.:3.000   3rd Qu.:102.00
##  Max.   :10.000   Max.   :3.000   Max.   :582.00
##  Securities.Account   CD.Account          Online         CreditCard
##  Min.   :0.0000     Min.   :0.00000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000     1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000     Median :0.00000   Median :1.0000   Median :0.0000
##  Mean   :0.1121     Mean   :0.06703   Mean   :0.5953   Mean   :0.2876
##  3rd Qu.:0.0000     3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000     Max.   :1.00000   Max.   :1.0000   Max.   :1.0000
```

```
summary(Test_Data)
```

```
##       Age           Experience        Income           Family
##  Min.   :23.00   Min.   :-2.00   Min.   :  8.00   Min.   :1.000
##  1st Qu.:35.00   1st Qu.:10.00   1st Qu.: 39.00   1st Qu.:1.000
##  Median :45.00   Median :20.00   Median : 63.00   Median :2.000
##  Mean   :45.47   Mean   :20.24   Mean   : 73.99   Mean   :2.379
```

```
##  3rd Qu.:55.00   3rd Qu.:30.00   3rd Qu.: 98.00   3rd Qu.:3.000
##  Max.   :67.00   Max.   :43.00   Max.   :205.00   Max.   :4.000
##      CCAvg          Education        Mortgage      Personal.Loan
##  Min.   : 0.000   Min.   :1.000   Min.   :  0.00   0:901
##  1st Qu.: 0.700   1st Qu.:1.000   1st Qu.:  0.00   1:100
##  Median : 1.500   Median :2.000   Median :  0.00
##  Mean   : 1.918   Mean   :1.884   Mean   : 58.33
##  3rd Qu.: 2.500   3rd Qu.:3.000   3rd Qu.:103.00
##  Max.   :10.000   Max.   :3.000   Max.   :589.00
##  Securities.Account   CD.Account         Online         CreditCard
##  Min.   :0.0000     Min.   :0.00000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000     1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000     Median :0.00000   Median :1.0000   Median :0.0000
##  Mean   :0.1059     Mean   :0.06993   Mean   :0.6184   Mean   :0.2927
##  3rd Qu.:0.0000     3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000     Max.   :1.00000   Max.   :1.0000   Max.   :1.0000
```

```r
#Normalization

train.norm.df <- Train_Data[, -8]
valid.norm.df <- Val_Data[, -8]
Test.norm.df <- Test_Data[, -8]
traval.norm.df <- Traval_Data[, -8]

norm.values <- preProcess(train.norm.df, method=c("center", "scale"))

train.norm.df <- predict(norm.values, train.norm.df)
valid.norm.df <- predict(norm.values, valid.norm.df)
traval.norm.df <- predict(norm.values, traval.norm.df)
Test.norm.df <- predict(norm.values, Test.norm.df)
View(train.norm.df)

library(FNN)
k_ubank<-knn(train.norm.df,Test.norm.df,cl=Train_Data$Personal.Loan,k=3,prob
= TRUE)

# Confusion Matrix

library(gmodels)
CrossTable(x=Test_Data$Personal.Loan,k_ubank,prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Row Total |
## |           N / Col Total |
## |         N / Table Total |
## |-------------------------|
```

```
##
##
## Total Observations in Table:  1001
##
##
##                            | k_ubank
## Test_Data$Personal.Loan |         0 |         1 | Row Total |
## -------------------------|-----------|-----------|-----------|
##                        0 |       896 |         5 |       901 |
##                          |     0.994 |     0.006 |     0.900 |
##                          |     0.976 |     0.060 |           |
##                          |     0.895 |     0.005 |           |
## -------------------------|-----------|-----------|-----------|
##                        1 |        22 |        78 |       100 |
##                          |     0.220 |     0.780 |     0.100 |
##                          |     0.024 |     0.940 |           |
##                          |     0.022 |     0.078 |           |
## -------------------------|-----------|-----------|-----------|
##            Column Total |       918 |        83 |      1001 |
##                          |     0.917 |     0.083 |           |
## -------------------------|-----------|-----------|-----------|
##
##
```

```
#Hypertuning

accuracy.df <- data.frame(k = seq(1, 55, 1), accuracy = rep(0, 55))
for(i in 1:55) {
  knn.pred <- knn(train.norm.df, valid.norm.df,
                  cl = Train_Data$`Personal.Loan`, k = i)
  accuracy.df[i, 2] <- confusionMatrix(knn.pred,
Val_Data$`Personal.Loan`)$overall[1]
}
accuracy.df
```

```
##     k  accuracy
## 1   1 0.9564782
## 2   2 0.9609805
## 3   3 0.9649825
## 4   4 0.9574787
## 5   5 0.9634817
## 6   6 0.9574787
## 7   7 0.9634817
## 8   8 0.9579790
## 9   9 0.9619810
## 10 10 0.9569785
## 11 11 0.9604802
## 12 12 0.9564782
## 13 13 0.9589795
## 14 14 0.9544772
```

```
## 15 15 0.9569785
## 16 16 0.9524762
## 17 17 0.9554777
## 18 18 0.9519760
## 19 19 0.9559780
## 20 20 0.9529765
## 21 21 0.9549775
## 22 22 0.9489745
## 23 23 0.9519760
## 24 24 0.9474737
## 25 25 0.9489745
## 26 26 0.9474737
## 27 27 0.9499750
## 28 28 0.9479740
## 29 29 0.9504752
## 30 30 0.9479740
## 31 31 0.9504752
## 32 32 0.9474737
## 33 33 0.9474737
## 34 34 0.9464732
## 35 35 0.9489745
## 36 36 0.9444722
## 37 37 0.9469735
## 38 38 0.9464732
## 39 39 0.9479740
## 40 40 0.9449725
## 41 41 0.9459730
## 42 42 0.9439720
## 43 43 0.9454727
## 44 44 0.9429715
## 45 45 0.9444722
## 46 46 0.9419710
## 47 47 0.9439720
## 48 48 0.9419710
## 49 49 0.9439720
## 50 50 0.9429715
## 51 51 0.9434717
## 52 52 0.9424712
## 53 53 0.9429715
## 54 54 0.9409705
## 55 55 0.9419710
```

```
accuracy.df[which.max(accuracy.df$accuracy),]
```

```
##   k  accuracy
## 3 3 0.9649825
```

```
#Accuracy obtained from train and test data (Confusion Matrix) = 97.3 when k
= 3
#Accuracy obtained from train and validation data = 96.4
```

```r
#Best K value is 3 with accuracy of 97.3

#------------------------------------------------------------
# Part 4 - Classifying customer using best K value

x <- data.frame("Age" = 40, "Experience" = 10, "Income" = 84, "Family" = 2,
"CCAvg" = 2, "Education"= 1, "Mortgage" = 0, "Personal Loan"=
"Deny","Securities Account" = 0, "CD Account" = 0, "Online" = 1,"Credit Card"
= 1)
y <- data.frame("Age" = 40, "Experience" = 10, "Income" = 84, "Family" = 2,
"CCAvg" = 2, "Education"= 2, "Mortgage" = 0, "Personal Loan"=
"Accept","Securities Account" = 0, "CD Account" = 0, "Online" = 1,"Credit
Card" = 1)
z <- data.frame("Age" = 40, "Experience" = 10, "Income" = 84, "Family" = 2,
"CCAvg" = 2, "Education"= 3, "Mortgage" = 0, "Personal Loan"=
"Deny","Securities Account" = 0, "CD Account" = 0, "Online" = 1,"Credit Card"
= 1)
test_pre<-as.data.frame(rbind(x,y,z))

test_pre.norm<-test_pre[,-8]
norm.values<- preProcess(test_pre.norm,method = c("center","scale"))

## Warning in preProcess.default(test_pre.norm, method = c("center",
## "scale")): These variables have zero variances: Age, Experience,
## Income, Family, CCAvg, Mortgage, Securities.Account, CD.Account, Online,
## Credit.Card

test_pre.norm<-predict(norm.values,test_pre.norm)
nn3<-
knn(train.norm.df,test=test_pre.norm,cl=Train_Data$`Personal.Loan`,k=4,prob =
TRUE)
CrossTable(x=test_pre$`Personal.Loan`,y=nn3,prop.chisq = FALSE)

##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  3
##
##
##                          | nn3
## test_pre$Personal.Loan |         0 | Row Total |
## -----------------------|-----------|-----------|
##                   Deny |         2 |         2 |
##                        |     0.667 |           |
```

```
## ----------------------|-----------|-----------|
##                Accept |         1 |         1 |
##                       |     0.333 |           |
## ----------------------|-----------|-----------|
##          Column Total |         3 |         3 |
## ----------------------|-----------|-----------|
##
##
```

```
#Customer's personal loan is rejected

#-----------------------------------------------------------------
#Part -5 Repartitioning the data, this time into training, validation, and
test sets (50% : 30% : 20%)

library(dplyr)

m_ubank1 <- ubank[,c(-1,-5)]
m_ubank1$Personal.Loan<-factor(m_ubank1$Personal.Loan)

set.seed(20)
Train_Index1 = createDataPartition(m_ubank1$Age,p=0.5, list=FALSE)
Train_Data1 = m_ubank1[Train_Index1,]
Val_Data1 = m_ubank1[-Train_Index1,]

Test_Index1 = createDataPartition(Val_Data1$Age,p=0.2, list=FALSE)
Test_Data1 = Val_Data1[Test_Index1,]
Val_Data1 = Val_Data1[-Test_Index1,]
View(m_ubank1)
summary(Train_Data1)
```

```
##       Age          Experience        Income           Family
##  Min.   :23.00   Min.   :-3.00   Min.   :  8.00   Min.   :1.000
##  1st Qu.:35.00   1st Qu.:10.00   1st Qu.: 38.00   1st Qu.:1.000
##  Median :45.00   Median :20.00   Median : 63.00   Median :2.000
##  Mean   :45.32   Mean   :20.11   Mean   : 73.86   Mean   :2.395
##  3rd Qu.:55.00   3rd Qu.:30.00   3rd Qu.:100.00   3rd Qu.:3.000
##  Max.   :67.00   Max.   :42.00   Max.   :205.00   Max.   :4.000
##      CCAvg          Education        Mortgage      Personal.Loan
##  Min.   : 0.000   Min.   :1.000   Min.   :  0.00   0:2264
##  1st Qu.: 0.700   1st Qu.:1.000   1st Qu.:  0.00   1: 237
##  Median : 1.500   Median :2.000   Median :  0.00
##  Mean   : 1.933   Mean   :1.864   Mean   : 55.78
##  3rd Qu.: 2.500   3rd Qu.:3.000   3rd Qu.: 98.00
##  Max.   :10.000   Max.   :3.000   Max.   :617.00
##  Securities.Account   CD.Account          Online        CreditCard
##  Min.   :0.0000     Min.   :0.00000   Min.   :0.000   Min.   :0.0000
##  1st Qu.:0.0000     1st Qu.:0.00000   1st Qu.:0.000   1st Qu.:0.0000
##  Median :0.0000     Median :0.00000   Median :1.000   Median :0.0000
##  Mean   :0.1008     Mean   :0.05398   Mean   :0.595   Mean   :0.2939
```

```
##  3rd Qu.:0.0000     3rd Qu.:0.00000   3rd Qu.:1.000   3rd Qu.:1.0000
##  Max.   :1.0000     Max.   :1.00000   Max.   :1.000   Max.   :1.0000
```

```
summary(Val_Data1)
```

```
##       Age           Experience        Income           Family
##  Min.   :23.00   Min.   :-3.00   Min.   :  8.00   Min.   :1.000
##  1st Qu.:35.00   1st Qu.:10.00   1st Qu.: 40.00   1st Qu.:1.000
##  Median :45.00   Median :20.00   Median : 63.00   Median :2.000
##  Mean   :45.38   Mean   :20.11   Mean   : 73.03   Mean   :2.404
##  3rd Qu.:55.00   3rd Qu.:30.00   3rd Qu.: 95.00   3rd Qu.:3.000
##  Max.   :67.00   Max.   :43.00   Max.   :204.00   Max.   :4.000
##      CCAvg         Education        Mortgage      Personal.Loan
##  Min.   : 0.00   Min.   :1.000   Min.   :  0.00   0:1809
##  1st Qu.: 0.70   1st Qu.:1.000   1st Qu.:  0.00   1: 188
##  Median : 1.50   Median :2.000   Median :  0.00
##  Mean   : 1.93   Mean   :1.902   Mean   : 56.84
##  3rd Qu.: 2.50   3rd Qu.:3.000   3rd Qu.:102.00
##  Max.   :10.00   Max.   :3.000   Max.   :635.00
##  Securities.Account   CD.Account         Online         CreditCard
##  Min.   :0.0000     Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000     1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.0000     Median :0.0000   Median :1.0000   Median :0.0000
##  Mean   :0.1112     Mean   :0.0661   Mean   :0.5949   Mean   :0.2864
##  3rd Qu.:0.0000     3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000     Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
```

```
summary(Test_Data1)
```

```
##       Age           Experience        Income           Family
##  Min.   :23.00   Min.   :-2.00   Min.   :  8.00   Min.   :1.000
##  1st Qu.:35.00   1st Qu.:10.25   1st Qu.: 39.00   1st Qu.:1.000
##  Median :45.00   Median :20.00   Median : 64.50   Median :2.000
##  Mean   :45.26   Mean   :20.04   Mean   : 76.33   Mean   :2.373
##  3rd Qu.:55.00   3rd Qu.:30.00   3rd Qu.:102.00   3rd Qu.:3.000
##  Max.   :67.00   Max.   :42.00   Max.   :224.00   Max.   :4.000
##      CCAvg          Education       Mortgage      Personal.Loan
##  Min.   :0.0000   Min.   :1.00   Min.   :  0.00   0:447
##  1st Qu.:0.7125   1st Qu.:1.00   1st Qu.:  0.00   1: 55
##  Median :1.6700   Median :2.00   Median :  0.00
##  Mean   :1.9974   Mean   :1.88   Mean   : 58.74
##  3rd Qu.:2.6700   3rd Qu.:3.00   3rd Qu.:114.75
##  Max.   :8.8000   Max.   :3.00   Max.   :587.00
##  Securities.Account   CD.Account          Online         CreditCard
##  Min.   :0.00000    Min.   :0.00000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.00000    1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.0000
##  Median :0.00000    Median :0.00000   Median :1.0000   Median :0.0000
##  Mean   :0.09562    Mean   :0.06972   Mean   :0.6135   Mean   :0.3247
##  3rd Qu.:0.00000    3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.00000    Max.   :1.00000   Max.   :1.0000   Max.   :1.0000
```

```r
train.norm.df1 <- Train_Data1[, -8]
valid.norm.df1 <- Val_Data1[, -8]
Test.norm.df1 <- Test_Data1[, -8]

norm.values1 <- preProcess(train.norm.df1, method=c("center", "scale"))

train.norm.df1 <- predict(norm.values1, train.norm.df1)
valid.norm.df1 <- predict(norm.values1, valid.norm.df1)
Test.norm.df1 <- predict(norm.values1, Test.norm.df1)
View(train.norm.df1)

library(FNN)
k_ubank1<-
knn(train.norm.df1,Test.norm.df1,cl=Train_Data1$Personal.Loan,k=3,prob =
TRUE)

#Confusion Matrix

library(gmodels)
CrossTable(x=Test_Data1$Personal.Loan,k_ubank1,prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Row Total |
## |           N / Col Total |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  502
##
##
##                            | k_ubank1
## Test_Data1$Personal.Loan |         0 |         1 | Row Total |
## -------------------------|-----------|-----------|-----------|
##                        0 |       443 |         4 |       447 |
##                          |     0.991 |     0.009 |     0.890 |
##                          |     0.957 |     0.103 |           |
##                          |     0.882 |     0.008 |           |
## -------------------------|-----------|-----------|-----------|
##                        1 |        20 |        35 |        55 |
##                          |     0.364 |     0.636 |     0.110 |
##                          |     0.043 |     0.897 |           |
##                          |     0.040 |     0.070 |           |
## -------------------------|-----------|-----------|-----------|
##             Column Total |       463 |        39 |       502 |
##                          |     0.922 |     0.078 |           |
```

```
## ------------------------|----------|----------|----------|
##
##

#Hypertuning

accuracy.df <- data.frame(k = seq(1, 55, 1), accuracy = rep(0, 55))
for(i in 1:55) {
  knn.pred1 <- knn(train.norm.df1, valid.norm.df1,
                   cl = Train_Data1$`Personal.Loan`, k = i)
  accuracy.df[i, 2] <- confusionMatrix(knn.pred1,
Val_Data1$`Personal.Loan`)$overall[1]
}
accuracy.df

##     k  accuracy
## 1   1 0.9554331
## 2   2 0.9519279
## 3   3 0.9584377
## 4   4 0.9529294
## 5   5 0.9584377
## 6   6 0.9479219
## 7   7 0.9524286
## 8   8 0.9474211
## 9   9 0.9489234
## 10 10 0.9459189
## 11 11 0.9474211
## 12 12 0.9454181
## 13 13 0.9489234
## 14 14 0.9454181
## 15 15 0.9449174
## 16 16 0.9424136
## 17 17 0.9434151
## 18 18 0.9414121
## 19 19 0.9439159
## 20 20 0.9384076
## 21 21 0.9404106
## 22 22 0.9364046
## 23 23 0.9394091
## 24 24 0.9369054
## 25 25 0.9394091
## 26 26 0.9369054
## 27 27 0.9389084
## 28 28 0.9349024
## 29 29 0.9379069
## 30 30 0.9354031
## 31 31 0.9379069
## 32 32 0.9349024
## 33 33 0.9379069
## 34 34 0.9359039
```

```
## 35 35 0.9384076
## 36 36 0.9354031
## 37 37 0.9364046
## 38 38 0.9339009
## 39 39 0.9359039
## 40 40 0.9318978
## 41 41 0.9328993
## 42 42 0.9308963
## 43 43 0.9334001
## 44 44 0.9313971
## 45 45 0.9328993
## 46 46 0.9303956
## 47 47 0.9318978
## 48 48 0.9283926
## 49 49 0.9288933
## 50 50 0.9283926
## 51 51 0.9293941
## 52 52 0.9273911
## 53 53 0.9283926
## 54 54 0.9263896
## 55 55 0.9278918
```

```
accuracy.df[which.max(accuracy.df$accuracy),]
```

```
##   k  accuracy
## 3 3 0.9584377
```

```
#Comparing the accuracies from confusion matrix test set with test and
validation set, we found accuracy of 95.2% from confusion matrix and 95.8%
for training & validation data
```