Assignment 2

Vamshee Deepak Goud Katta

10/2/2021

1. Inserting Data and Libraries

Reading the UniversalBank csv file and inserting approproate libraries

```
library(class)
library(caret)
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
library(ISLR)
library(dummies)
```

```
## dummies-1.5.6 provided by Decision Patterns
```

```
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
```

```
library(tidyr)
library(ggplot2)

UBank_data <- read.csv("UniversalBank.csv")
```

2. Data Cleaning and Splitting

Removing the ID and Zipcode columns from the Bank_data dataset

Converting Education column into factor and creating a dummy database of Bank_data

Moving personal.loan column to the first column position

```
library(caret)
library(class)
library(ISLR)
library(tidyr)
library(dummies)
library(dplyr)
library(ggplot2)

UBank_data$Education = as.factor(UBank_data$Education)
UBank_data$Personal.Loan = as.factor(UBank_data$Personal.Loan)
UBank_data_d = dummy.data.frame(select(UBank_data, -c(ZIP.Code,ID)))
```

```
## Warning in model.matrix.default(~x - 1, model.frame(~x - 1), contrasts = FALSE):
## non-list contrasts argument ignored

## Warning in model.matrix.default(~x - 1, model.frame(~x - 1), contrasts = FALSE):
## non-list contrasts argument ignored
```

```
n <- which(colnames(UBank_data_d)=="Personal.Loan1")
Personal.Loan <- UBank_data_d$Personal.Loan1
UBank_data_d <- cbind(Personal.Loan, UBank_data_d[, -n])
UBank_data_d <- UBank_data_d[,-11]

head(UBank_data_d)
```

```
##   Personal.Loan Age Experience Income Family CCAvg Education1 Education2
## 1             0  25          1     49      4   1.6          1          0
## 2             0  45         19     34      3   1.5          1          0
## 3             0  39         15     11      1   1.0          1          0
## 4             0  35          9    100      1   2.7          0          1
## 5             0  35          8     45      4   1.0          0          1
## 6             0  37         13     29      4   0.4          0          1
##   Education3 Mortgage Securities.Account CD.Account Online CreditCard
## 1          0        0                  1          0      0          0
## 2          0        0                  1          0      0          0
## 3          0        0                  0          0      0          0
## 4          0        0                  0          0      0          0
## 5          0        0                  0          0      0          1
## 6          0      155                  0          0      1          0
```

3. Data Partition

Partitioning the UniversalBank dataset into Training and validation sets

```
set.seed(123)
Index_Train <- createDataPartition(UBank_data_d$Personal.Loan, p=0.6, list = FALSE)
# 60% of data is taken as Training Data

Train <- UBank_data_d[Index_Train,]
Validation <- UBank_data_d[-Index_Train,]
# Rest of the data is taken as Validation Data

summary(Train)
```

```
##  Personal.Loan          Age          Experience        Income
##  Min.   :0.00000   Min.   :23.00   Min.   :-3.00   Min.   :  8.0
##  1st Qu.:0.00000   1st Qu.:36.00   1st Qu.:10.00   1st Qu.: 39.0
##  Median :0.00000   Median :46.00   Median :20.00   Median : 64.0
##  Mean   :0.09267   Mean   :45.43   Mean   :20.21   Mean   : 73.4
##  3rd Qu.:0.00000   3rd Qu.:55.00   3rd Qu.:30.00   3rd Qu.: 95.0
##  Max.   :1.00000   Max.   :67.00   Max.   :43.00   Max.   :224.0
##      Family          CCAvg          Education1      Education2
##  Min.   :1.000   Min.   : 0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:1.000   1st Qu.: 0.700   1st Qu.:0.000   1st Qu.:0.000
##  Median :2.000   Median : 1.500   Median :0.000   Median :0.000
##  Mean   :2.399   Mean   : 1.924   Mean   :0.424   Mean   :0.285
##  3rd Qu.:4.000   3rd Qu.: 2.600   3rd Qu.:1.000   3rd Qu.:1.000
##  Max.   :4.000   Max.   :10.000   Max.   :1.000   Max.   :1.000
##    Education3       Mortgage      Securities.Account   CD.Account
##  Min.   :0.000   Min.   :  0.00   Min.   :0.000      Min.   :0.000
##  1st Qu.:0.000   1st Qu.:  0.00   1st Qu.:0.000      1st Qu.:0.000
##  Median :0.000   Median :  0.00   Median :0.000      Median :0.000
##  Mean   :0.291   Mean   : 55.29   Mean   :0.103      Mean   :0.059
##  3rd Qu.:1.000   3rd Qu.: 98.00   3rd Qu.:0.000      3rd Qu.:0.000
##  Max.   :1.000   Max.   :635.00   Max.   :1.000      Max.   :1.000
##      Online         CreditCard
##  Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000
##  Median :1.0000   Median :0.0000
##  Mean   :0.5997   Mean   :0.2943
##  3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.0000
```

```
summary(Validation)
```

```
##  Personal.Loan        Age          Experience        Income
##  Min.   :0.000   Min.   :23.0   Min.   :-3.00   Min.   :  8.00
##  1st Qu.:0.000   1st Qu.:35.0   1st Qu.:10.00   1st Qu.: 39.00
##  Median :0.000   Median :45.0   Median :20.00   Median : 63.00
```

```
##  Mean   :0.101   Mean   :45.2   Mean   :19.95   Mean   : 74.34
##  3rd Qu.:0.000   3rd Qu.:55.0   3rd Qu.:30.00   3rd Qu.:100.00
##  Max.   :1.000   Max.   :67.0   Max.   :43.00   Max.   :204.00
##      Family          CCAvg          Education1      Education2
##  Min.   :1.000   Min.   : 0.000   Min.   :0.000   Min.   :0.000
##  1st Qu.:1.000   1st Qu.: 0.700   1st Qu.:0.000   1st Qu.:0.000
##  Median :2.000   Median : 1.600   Median :0.000   Median :0.000
##  Mean   :2.393   Mean   : 1.959   Mean   :0.412   Mean   :0.274
##  3rd Qu.:3.000   3rd Qu.: 2.500   3rd Qu.:1.000   3rd Qu.:1.000
##  Max.   :4.000   Max.   :10.000   Max.   :1.000   Max.   :1.000
##    Education3       Mortgage      Securities.Account   CD.Account
##  Min.   :0.000   Min.   :  0.00   Min.   :0.0000     Min.   :0.0000
##  1st Qu.:0.000   1st Qu.:  0.00   1st Qu.:0.0000     1st Qu.:0.0000
##  Median :0.000   Median :  0.00   Median :0.0000     Median :0.0000
##  Mean   :0.314   Mean   : 58.32   Mean   :0.1065     Mean   :0.0625
##  3rd Qu.:1.000   3rd Qu.:103.00   3rd Qu.:0.0000     3rd Qu.:0.0000
##  Max.   :1.000   Max.   :601.00   Max.   :1.0000     Max.   :1.0000
##      Online         CreditCard
##  Min.   :0.0000   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:0.0000
##  Median :1.0000   Median :0.0000
##  Mean   :0.5925   Mean   :0.2935
##  3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.0000
```

4. Q1. Classifying customer using given constraints

Creating Test data from Q1 in Assignment

Normalizing the Training Data

Applying KNN Model with k=1

```r
# copying data

Train_norm <- Train
Validation_norm <- Validation
UBD_norm <- UBank_data_d

#Creating test data from the question

Test1 <- data.frame(Age = 40, Experience = 10, Income = 84, Family = 2, CCAvg = 2, Education1 = 0, Educa

#Normalizing the data

Test_norm1 <- Test1

norm_values <- preProcess(Train, method = c("center", "scale"))
Train_norm <- predict(norm_values, Train)
Validation_norm <- predict(norm_values, Validation)
```

```
UBD_norm <- predict(norm_values, UBank_data_d)

#combine normalized values of train and validation and then preprocess and normalize test data#########

summary(Train_norm)
```

```
##  Personal.Loan          Age              Experience           Income
##  Min.   :-0.3195   Min.   :-1.96925   Min.   :-2.03259   Min.   :-1.4286
##  1st Qu.:-0.3195   1st Qu.:-0.82784   1st Qu.:-0.89392   1st Qu.:-0.7514
##  Median :-0.3195   Median : 0.05016   Median :-0.01801   Median :-0.2053
##  Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.0000
##  3rd Qu.:-0.3195   3rd Qu.: 0.84037   3rd Qu.: 0.85789   3rd Qu.: 0.4720
##  Max.   : 3.1286   Max.   : 1.89398   Max.   : 1.99656   Max.   : 3.2901
##      Family            CCAvg           Education1        Education2
##  Min.   :-1.2142   Min.   :-1.1157   Min.   :-0.8578   Min.   :-0.6312
##  1st Qu.:-1.2142   1st Qu.:-0.7098   1st Qu.:-0.8578   1st Qu.:-0.6312
##  Median :-0.3463   Median :-0.2459   Median :-0.8578   Median :-0.6312
##  Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000
##  3rd Qu.: 1.3895   3rd Qu.: 0.3921   3rd Qu.: 1.1653   3rd Qu.: 1.5836
##  Max.   : 1.3895   Max.   : 4.6835   Max.   : 1.1653   Max.   : 1.5836
##    Education3         Mortgage       Securities.Account   CD.Account
##  Min.   :-0.6405   Min.   :-0.5477   Min.   :-0.3388    Min.   :-0.2504
##  1st Qu.:-0.6405   1st Qu.:-0.5477   1st Qu.:-0.3388    1st Qu.:-0.2504
##  Median :-0.6405   Median :-0.5477   Median :-0.3388    Median :-0.2504
##  Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000    Mean   : 0.0000
##  3rd Qu.: 1.5606   3rd Qu.: 0.4231   3rd Qu.:-0.3388    3rd Qu.:-0.2504
##  Max.   : 1.5606   Max.   : 5.7429   Max.   : 2.9506    Max.   : 3.9930
##      Online          CreditCard
##  Min.   :-1.2237   Min.   :-0.6457
##  1st Qu.:-1.2237   1st Qu.:-0.6457
##  Median : 0.8169   Median :-0.6457
##  Mean   : 0.0000   Mean   : 0.0000
##  3rd Qu.: 0.8169   3rd Qu.: 1.5481
##  Max.   : 0.8169   Max.   : 1.5481
```

```
var(Train_norm)
```

```
##                    Personal.Loan          Age   Experience       Income
## Personal.Loan         1.00000000 -0.021822200 -0.021166260  0.488102867
## Age                  -0.02182220  1.000000000  0.994046117 -0.058886301
## Experience           -0.02116626  0.994046117  1.000000000 -0.049627190
## Income                0.48810287 -0.058886301 -0.049627190  1.000000000
## Family                0.05096949 -0.047518269 -0.052070243 -0.158593072
## CCAvg                 0.34988271 -0.064959563 -0.063361729  0.649073730
## Education1           -0.15555490 -0.014705142  0.010836619  0.228897878
## Education2            0.09363791 -0.025452176 -0.025283471 -0.124462179
## Education3            0.07618321  0.041293806  0.013336986 -0.125345941
## Mortgage              0.14814232 -0.007099632 -0.003875213  0.201359111
## Securities.Account    0.02407606  0.002844950  0.003405624  0.009194703
## CD.Account            0.34931209 -0.001475115  0.002056915  0.198656079
## Online                0.02884123  0.013192870  0.012993003  0.020872689
## CreditCard            0.02314386  0.013069757  0.014634477  0.028104163
```

```
##                         Family         CCAvg   Education1   Education2
## Personal.Loan       0.050969493  0.3498827085 -0.155554896  0.093637912
## Age                -0.047518269 -0.0649595631 -0.014705142 -0.025452176
## Experience         -0.052070243 -0.0633617290  0.010836619 -0.025283471
## Income             -0.158593072  0.6490737302  0.228897878 -0.124462179
## Family              1.000000000 -0.1046245559 -0.118580297  0.124257664
## CCAvg              -0.104624556  1.0000000000  0.158046007 -0.085770034
## Education1         -0.118580297  0.1580460070  1.000000000 -0.541677873
## Education2          0.124257664 -0.0857700337 -0.541677873  1.000000000
## Education3          0.005524861 -0.0867127350 -0.549661177 -0.404476027
## Mortgage           -0.020468810  0.1156690366  0.031288479 -0.016957841
## Securities.Account  0.009242326  0.0037544277  0.002183559  0.033852865
## CD.Account         -0.008133090  0.1560439055 -0.028763610  0.029944231
## Online             -0.009921811 -0.0006108599 -0.023094425  0.032078283
## CreditCard         -0.012264272  0.0098128271  0.009779915 -0.005922037
##                      Education3     Mortgage Securities.Account   CD.Account
## Personal.Loan       0.076183212  0.148142317        0.024076056  0.349312088
## Age                 0.041293806 -0.007099632        0.002844950 -0.001475115
## Experience          0.013336986 -0.003875213        0.003405624  0.002056915
## Income             -0.125345941  0.201359111        0.009194703  0.198656079
## Family              0.005524861 -0.020468810        0.009242326 -0.008133090
## CCAvg              -0.086712735  0.115669037        0.003754428  0.156043906
## Education1         -0.549661177  0.031288479        0.002183559 -0.028763610
## Education2         -0.404476027 -0.016957841        0.033852865  0.029944231
## Education3          1.000000000 -0.017188557       -0.036019196  0.001535448
## Mortgage           -0.017188557  1.000000000       -0.011946442  0.098948506
## Securities.Account -0.036019196 -0.011946442        1.000000000  0.287483990
## CD.Account          0.001535448  0.098948506        0.287483990  1.000000000
## Online             -0.006753417  0.010952781       -0.002902940  0.169943664
## CreditCard         -0.004755028  0.005400099       -0.031158792  0.285277874
##                           Online   CreditCard
## Personal.Loan       0.0288412325  0.023143864
## Age                 0.0131928696  0.013069757
## Experience          0.0129930031  0.014634477
## Income              0.0208726892  0.028104163
## Family             -0.0099218106 -0.012264272
## CCAvg              -0.0006108599  0.009812827
## Education1         -0.0230944254  0.009779915
## Education2          0.0320782833 -0.005922037
## Education3         -0.0067534166 -0.004755028
## Mortgage            0.0109527811  0.005400099
## Securities.Account -0.0029029400 -0.031158792
## CD.Account          0.1699436636  0.285277874
## Online              1.0000000000  0.003723470
## CreditCard          0.0037234702  1.000000000
```

```
summary(Validation_norm)
```

```
##  Personal.Loan           Age              Experience           Income
##  Min.   :-0.31953   Min.   :-1.96925   Min.   :-2.03259   Min.   :-1.42861
##  1st Qu.:-0.31953   1st Qu.:-0.91565   1st Qu.:-0.89392   1st Qu.:-0.75140
##  Median :-0.31953   Median :-0.03764   Median :-0.01801   Median :-0.22710
##  Mean   : 0.02873   Mean   :-0.01981   Mean   :-0.02213   Mean   : 0.02067
##  3rd Qu.:-0.31953   3rd Qu.: 0.84037   3rd Qu.: 0.85789   3rd Qu.: 0.58119
```

```
##  Max.   : 3.12859   Max.   : 1.89398   Max.   : 1.99656   Max.   : 2.85314
##      Family              CCAvg           Education1         Education2
##  Min.   :-1.214194   Min.   :-1.11573   Min.   :-0.85783   Min.   :-0.63124
##  1st Qu.:-1.214194   1st Qu.:-0.70979   1st Qu.:-0.85783   1st Qu.:-0.63124
##  Median :-0.346293   Median :-0.18786   Median :-0.85783   Median :-0.63124
##  Mean   :-0.005641   Mean   : 0.02029   Mean   :-0.02428   Mean   :-0.02436
##  3rd Qu.: 0.521609   3rd Qu.: 0.33407   3rd Qu.: 1.16535   3rd Qu.: 1.58365
##  Max.   : 1.389511   Max.   : 4.68347   Max.   : 1.16535   Max.   : 1.58365
##    Education3          Mortgage        Securities.Account   CD.Account
##  Min.   :-0.64055   Min.   :-0.54769   Min.   :-0.33881   Min.   :-0.25036
##  1st Qu.:-0.64055   1st Qu.:-0.54769   1st Qu.:-0.33881   1st Qu.:-0.25036
##  Median :-0.64055   Median :-0.54769   Median :-0.33881   Median :-0.25036
##  Mean   : 0.05063   Mean   : 0.03003   Mean   : 0.01151   Mean   : 0.01485
##  3rd Qu.: 1.56065   3rd Qu.: 0.47267   3rd Qu.:-0.33881   3rd Qu.:-0.25036
##  Max.   : 1.56065   Max.   : 5.40606   Max.   : 2.95057   Max.   : 3.99297
##      Online           CreditCard
##  Min.   :-1.22369   Min.   :-0.645725
##  1st Qu.:-1.22369   1st Qu.:-0.645725
##  Median : 0.81693   Median :-0.645725
##  Mean   :-0.01462   Mean   :-0.001828
##  3rd Qu.: 0.81693   3rd Qu.: 1.548131
##  Max.   : 0.81693   Max.   : 1.548131
```

```
var(Validation_norm)
```

```
##                    Personal.Loan           Age    Experience       Income
## Personal.Loan         1.08009777  0.0133266557  0.0132191744  0.551059837
## Age                   0.01332666  1.0327551325  1.0220182700 -0.051302177
## Experience            0.01321917  1.0220182700  1.0226757778 -0.042924214
## Income                0.55105984 -0.0513021773 -0.0429242139  1.028521084
## Family                0.07891768 -0.0451479570 -0.0534726737 -0.156530305
## CCAvg                 0.41939848 -0.0349824725 -0.0321607240  0.672463699
## Education1           -0.14386434 -0.0480099895 -0.0238708722  0.204299406
## Education2            0.04069611 -0.0028449500 -0.0078847351 -0.134250919
## Education3            0.11607841  0.0550617028  0.0338072229 -0.088854867
## Mortgage              0.14097190 -0.0207909520 -0.0205702104  0.221586939
## Securities.Account    0.01978492 -0.0052356992 -0.0080697325 -0.020556280
## CD.Account            0.28820198  0.0228544894  0.0234039913  0.133332209
## Online               -0.02705050  0.0145597473  0.0152625195  0.004632481
## CreditCard           -0.02757569 -0.0003045919  0.0005372617 -0.048136139
##                          Family        CCAvg  Education1   Education2
## Personal.Loan        0.07891768  0.419398479 -0.14386434  0.040696112
## Age                 -0.04514796 -0.034982473 -0.04800999 -0.002844950
## Experience          -0.05347267 -0.032160724 -0.02387087 -0.007884735
## Income              -0.15653031  0.672463699  0.20429941 -0.134250919
## Family               0.98080706 -0.118834966 -0.11719581  0.158583078
## CCAvg               -0.11883497  1.068246059  0.16041345 -0.098969863
## Education1          -0.11719581  0.160413452  0.99210723 -0.506116587
## Education2           0.15858308 -0.098969863 -0.50611659  0.976357112
## Education3          -0.03009458 -0.076170293 -0.57641529 -0.419669648
## Mortgage            -0.02050409  0.106779382  0.06127397 -0.053882485
## Securities.Account   0.03627116  0.032690069  0.01412891 -0.037765542
## CD.Account           0.04778511  0.115558998  0.00644197 -0.029385032
## Online               0.04065499 -0.008067683  0.04291686  0.002961914
```

```
## CreditCard          0.04724647 -0.031647341  0.02255025 -0.021483322
##                     Education3     Mortgage Securities.Account  CD.Account
## Personal.Loan       0.116078410  0.140971903        0.019784921  0.28820198
## Age                 0.055061703 -0.020790952       -0.005235699  0.02285449
## Experience          0.033807223 -0.020570210       -0.008069733  0.02340399
## Income             -0.088854867  0.221586939       -0.020556280  0.13333221
## Family             -0.030094584 -0.020504095        0.036271159  0.04778511
## CCAvg              -0.076170293  0.106779382        0.032690069  0.11555900
## Education1         -0.576415293  0.061273973        0.014128914  0.00644197
## Education2         -0.419669648 -0.053882485       -0.037765542 -0.02938503
## Education3          1.044207846 -0.013116159        0.022159882  0.02219451
## Mortgage           -0.013116159  1.038206516        0.003999834  0.07880069
## Securities.Account  0.022159882  0.003999834        1.030117842  0.37486936
## CD.Account          0.022194512  0.078800690        0.374869362  1.05555793
## Online             -0.049636676 -0.031290949        0.036256424  0.19032293
## CreditCard         -0.003183968 -0.026281792        0.008969063  0.27621566
##                           Online    CreditCard
## Personal.Loan      -0.027050497 -0.0275756944
## Age                 0.014559747 -0.0003045919
## Experience          0.015262519  0.0005372617
## Income              0.004632481 -0.0481361390
## Family              0.040654991  0.0472464673
## CCAvg              -0.008067683 -0.0316473407
## Education1          0.042916862  0.0225502527
## Education2          0.002961914 -0.0214833220
## Education3         -0.049636676 -0.0031839680
## Mortgage           -0.031290949 -0.0262817921
## Securities.Account  0.036256424  0.0089690627
## CD.Account          0.190322928  0.2762156637
## Online              1.005904426  0.0049325668
## CreditCard          0.004932567  0.9985127901
```

```
# Applying KNN Model on test data

Q1_KNN <- knn(train = Train_norm[,-10], test = Test1, cl = Train[,1], k=1, prob = 0.5)
Q1.attributes<- attributes(Q1_KNN)
row.names(Train_norm)[attr(Q1_KNN, "Q1_KNN.index")]
```

```
## character(0)
```

```
Q1.attributes[1]
```

```
## $levels
## [1] "0" "1"
```

```
Q1.attributes[3]
```

```
## $prob
## [1] 1
```

5. Training the KNN Model

Combining training and validation datasets and renormalizing data

Expanding search grid to appropriate K value to gain highest possible accuracy

```
set.seed(123)
Search_grid <- expand.grid(k=c(2,3,5,7,14))
Model <- train(Personal.Loan~Income+Mortgage, data =UBD_norm, method = "knn", tuneGrid=Search_grid, pre
```

```
## Warning in train.default(x, y, weights = w, ...): You are trying to do
## regression and your outcome only has two possible values Are you trying to do
## classification? If so, use a 2 level factor as your outcome column.
```

```
head(Model)
```

```
## $method
## [1] "knn"
##
## $modelInfo
## $modelInfo$label
## [1] "k-Nearest Neighbors"
##
## $modelInfo$library
## NULL
##
## $modelInfo$loop
## NULL
##
## $modelInfo$type
## [1] "Classification" "Regression"
##
## $modelInfo$parameters
##   parameter   class      label
## 1         k numeric #Neighbors
##
## $modelInfo$grid
## function(x, y, len = NULL, search = "grid"){
##                     if(search == "grid") {
##                       out <- data.frame(k = (5:((2 * len)+4))[(5:((2 * len)+4))%%2 > 0])
##                     } else {
##                       by_val <- if(is.factor(y)) length(levels(y)) else 1
##                       out <- data.frame(k = sample(seq(1, floor(nrow(x)/3), by = by_val), size = len
##                     }
##                   }
##
## $modelInfo$fit
## function(x, y, wts, param, lev, last, classProbs, ...) {
##                     if(is.factor(y))
```

```
##                     {
##                       knn3(as.matrix(x), y, k = param$k, ...)
##                     } else {
##                       knnreg(as.matrix(x), y, k = param$k, ...)
##                     }
##                   }
## <bytecode: 0x00000000216a4a48>
##
## $modelInfo$predict
## function(modelFit, newdata, submodels = NULL) {
##                     if(modelFit$problemType == "Classification")
##                     {
##                       out <- predict(modelFit, newdata,  type = "class")
##                     } else {
##                       out <- predict(modelFit, newdata)
##                     }
##                     out
##                   }
## <bytecode: 0x0000000020ff7678>
##
## $modelInfo$predictors
## function(x, ...) colnames(x$learn$X)
##
## $modelInfo$tags
## [1] "Prototype Models"
##
## $modelInfo$prob
## function(modelFit, newdata, submodels = NULL)
##                     predict(modelFit, newdata, type = "prob")
##
## $modelInfo$levels
## function(x) levels(x$learn$y)
##
## $modelInfo$sort
## function(x) x[order(-x[,1]),]
##
##
## $modelType
## [1] "Regression"
##
## $results
##    k      RMSE  Rsquared       MAE     RMSESD RsquaredSD      MAESD
## 1  2 0.9776761 0.1832878 0.4094909 0.02788705 0.02042435 0.01639182
## 2  3 0.9573072 0.1961045 0.4103678 0.02761934 0.02132426 0.01654075
## 3  5 0.9319872 0.2144342 0.4126807 0.02757722 0.02326601 0.01576960
## 4  7 0.9139502 0.2302150 0.4121928 0.02407857 0.02375640 0.01511472
## 5 14 0.8831811 0.2614493 0.4100736 0.02058955 0.02211097 0.01283333
##
## $pred
## NULL
##
## $bestTune
##    k
## 5 14
```

7. Q3. Creating Confusion Matrix

```
Train_Predictors <- Train_norm
Test_Predictors <- Validation

Train_labels <-Train_norm[,1]
Test_labels  <-Validation_norm[,1]

Predicted_Test_labels <- knn(Train_Predictors, Test_Predictors, cl=Train_labels, k=1, prob = TRUE)
library(gmodels)
CrossTable(x=Test_labels, y=Predicted_Test_labels, prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Row Total |
## |           N / Col Total |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  2000
##
##
##                     | Predicted_Test_labels
##        Test_labels | -0.319525808959514 |   3.12859443161078 |          Row Total |
## -------------------|--------------------|--------------------|--------------------|
## -0.319525808959514 |                363 |               1435 |               1798 |
##                    |              0.202 |              0.798 |              0.899 |
##                    |              0.771 |              0.939 |                    |
##                    |              0.181 |              0.718 |                    |
## -------------------|--------------------|--------------------|--------------------|
##   3.12859443161078 |                108 |                 94 |                202 |
##                    |              0.535 |              0.465 |              0.101 |
##                    |              0.229 |              0.061 |                    |
##                    |              0.054 |              0.047 |                    |
## -------------------|--------------------|--------------------|--------------------|
##       Column Total |                471 |               1529 |               2000 |
##                    |              0.235 |              0.764 |                    |
## -------------------|--------------------|--------------------|--------------------|
##
##
```

8. Q4. Classifying customer using best K

```
Q4_KNN <- knn(Train_norm[,2:14], Test1, cl = Train[,1], k=3, prob = TRUE)
head(Q4_KNN)
```

```
## [1] 1
```

```
## Levels: 0 1
```

9. Q5. Repartitioning data into 5:3:2

```
set.seed(123)
Index_Train2 <- createDataPartition(UBank_data_d$Personal.Loan, p=0.5, list = FALSE)
# 50% of data is taken as Training Data

Train2 <- UBank_data_d[Index_Train2,]
Validation.test <- UBank_data_d[-Index_Train2,]
# Rest of the data is taken as Validation and Test Data
Index_Validation <- createDataPartition(Validation.test$Personal.Loan, p=0.6, list = FALSE)
# 30% of total or 60% of remaining data is taken as Validation Data
Validation2 <- Validation.test[Index_Validation,]
Test2 <- Validation.test[-Index_Validation,]

knn12 <- knn(train=Train2, test=Train2,cl=Train2[,1], k=3 ,prob=TRUE)
head(knn12)
```

```
## [1] 0 0 0 0 0 1
## Levels: 0 1
```

10. Normalizing the data

```
Train_norm2 <- Train2
Validation_norm2 <- Validation2
Test_norm2 <- Test2

norm_values2 <- preProcess(Train2, method = c("center", "scale"))
Train_norm2 <- predict(norm_values2, Train2)
Validation_norm2 <- predict(norm_values2, Validation2)

summary(Train_norm2)
```

```
##  Personal.Loan          Age              Experience            Income
##  Min.   :-0.3175   Min.   :-1.96620   Min.   :-2.028011   Min.   :-1.4337
##  1st Qu.:-0.3175   1st Qu.:-0.91026   1st Qu.:-0.887141   1st Qu.:-0.7553
##  Median :-0.3175   Median : 0.05769   Median :-0.009548   Median :-0.2081
##  Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.0000
##  3rd Qu.:-0.3175   3rd Qu.: 0.84965   3rd Qu.: 0.868044   3rd Qu.: 0.5360
##  Max.   : 3.1485   Max.   : 1.90559   Max.   : 2.008915   Max.   : 3.2937
##      Family            CCAvg           Education1        Education2
##  Min.   :-1.2072   Min.   :-1.1117   Min.   :-0.8627   Min.   :-0.6235
##  1st Qu.:-1.2072   1st Qu.:-0.7101   1st Qu.:-0.8627   1st Qu.:-0.6235
##  Median :-0.3442   Median :-0.2510   Median :-0.8627   Median :-0.6235
##  Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000
##  3rd Qu.: 1.3818   3rd Qu.: 0.3801   3rd Qu.: 1.1587   3rd Qu.: 1.6032
##  Max.   : 1.3818   Max.   : 4.6261   Max.   : 1.1587   Max.   : 1.6032
##    Education3         Mortgage       Securities.Account   CD.Account
```

```
##  Min.   :-0.6439   Min.   :-0.5478   Min.   :-0.3443    Min.   :-0.2472
##  1st Qu.:-0.6439   1st Qu.:-0.5478   1st Qu.:-0.3443    1st Qu.:-0.2472
##  Median :-0.6439   Median :-0.5478   Median :-0.3443    Median :-0.2472
##  Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000    Mean   : 0.0000
##  3rd Qu.: 1.5523   3rd Qu.: 0.4302   3rd Qu.:-0.3443    3rd Qu.:-0.2472
##  Max.   : 1.5523   Max.   : 5.7888   Max.   : 2.9035    Max.   : 4.0441
##      Online          CreditCard
##  Min.   :-1.2255   Min.   :-0.6508
##  1st Qu.:-1.2255   1st Qu.:-0.6508
##  Median : 0.8157   Median :-0.6508
##  Mean   : 0.0000   Mean   : 0.0000
##  3rd Qu.: 0.8157   3rd Qu.: 1.5360
##  Max.   : 0.8157   Max.   : 1.5360
```

```
var(Train_norm2)
```

```
##                    Personal.Loan          Age    Experience      Income
## Personal.Loan         1.00000000 -0.039892527 -0.0394261891  0.48831047
## Age                  -0.03989253  1.000000000  0.9938753927 -0.07315732
## Experience           -0.03942619  0.993875393  1.0000000000 -0.06362619
## Income                0.48831047 -0.073157322 -0.0636261874  1.00000000
## Family                0.05466992 -0.052886503 -0.0566271658 -0.15774567
## CCAvg                 0.34286680 -0.080084888 -0.0787215841  0.64293801
## Education1           -0.15345851 -0.028931723 -0.0032007434  0.23409981
## Education2            0.09845703 -0.002123290 -0.0007944896 -0.13699741
## Education3            0.06962520  0.033528932  0.0042612661 -0.11923015
## Mortgage              0.12344822 -0.009330027 -0.0062713370  0.18868170
## Securities.Account    0.02128853  0.007288790  0.0081171364  0.02023177
## CD.Account            0.33811664 -0.025475234 -0.0218012286  0.20688577
## Online                0.01559431  0.016175686  0.0169663649  0.02293378
## CreditCard            0.02077448  0.007297123  0.0086818520  0.02257521
##                          Family         CCAvg   Education1    Education2
## Personal.Loan        0.054669919  0.3428668037 -0.153458509  0.0984570322
## Age                 -0.052886503 -0.0800848877 -0.028931723 -0.0021232903
## Experience          -0.056627166 -0.0787215841 -0.003200743 -0.0007944896
## Income              -0.157745671  0.6429380113  0.234099807 -0.1369974100
## Family               1.000000000 -0.1073080770 -0.108562019  0.1121474674
## CCAvg               -0.107308077  1.0000000000  0.157720930 -0.0873376377
## Education1          -0.108562019  0.1577209302  1.000000000 -0.5381110149
## Education2           0.112147467 -0.0873376377 -0.538111015  1.0000000000
## Education3           0.007341507 -0.0852235891 -0.555767091 -0.4016488058
## Mortgage            -0.025484237  0.0882360290  0.024209999 -0.0107674476
## Securities.Account   0.020545392 -0.0019484187  0.010240351  0.0312548048
## CD.Account          -0.016934369  0.1440580262 -0.022420461  0.0293661537
## Online              -0.018749386  0.0001651959 -0.009290158  0.0267725326
## CreditCard          -0.023944582  0.0016695385  0.002583906 -0.0025720532
##                        Education3     Mortgage Securities.Account   CD.Account
## Personal.Loan        0.0696252018  0.123448217        0.021288532  0.338116639
## Age                  0.0335289318 -0.009330027        0.007288790 -0.025475234
## Experience           0.0042612661 -0.006271337        0.008117136 -0.021801229
## Income              -0.1192301463  0.188681704        0.020231773  0.206885769
## Family               0.0073415073 -0.025484237        0.020545392 -0.016934369
## CCAvg               -0.0852235891  0.088236029       -0.001948419  0.144058026
## Education1          -0.5557670911  0.024209999        0.010240351 -0.022420461
```

```
## Education2         -0.4016488058 -0.010767448        0.031254805  0.029366154
## Education3          1.0000000000 -0.015684406       -0.041953316 -0.004604107
## Mortgage           -0.0156844062  1.000000000       -0.011679779  0.086936391
## Securities.Account -0.0419533156 -0.011679779        1.000000000  0.271806163
## CD.Account         -0.0046041067  0.086936391        0.271806163  1.000000000
## Online             -0.0163122324  0.004979324       -0.013545219  0.166639952
## CreditCard         -0.0002705974  0.010598666       -0.045086060  0.278424713
##                           Online    CreditCard
## Personal.Loan       0.0155943070  0.0207744812
## Age                 0.0161756863  0.0072971226
## Experience          0.0169663649  0.0086818520
## Income              0.0229337773  0.0225752126
## Family             -0.0187493856 -0.0239445818
## CCAvg               0.0001651959  0.0016695385
## Education1         -0.0092901578  0.0025839062
## Education2          0.0267725326 -0.0025720532
## Education3         -0.0163122324 -0.0002705974
## Mortgage            0.0049793241  0.0105986657
## Securities.Account -0.0135452189 -0.0450860601
## CD.Account          0.1666399517  0.2784247125
## Online              1.0000000000  0.0041124343
## CreditCard          0.0041124343  1.0000000000
```

```
summary(Validation_norm2)
```

```
##  Personal.Loan           Age               Experience           Income
##  Min.   :-0.31748   Min.   :-1.966201   Min.   :-1.940252   Min.   :-1.43373
##  1st Qu.:-0.31748   1st Qu.:-0.910258   1st Qu.:-0.887141   1st Qu.:-0.75526
##  Median :-0.31748   Median :-0.030306   Median :-0.009548   Median :-0.22999
##  Mean   : 0.01294   Mean   : 0.005479   Mean   : 0.008589   Mean   : 0.00752
##  3rd Qu.:-0.31748   3rd Qu.: 0.849647   3rd Qu.: 0.868044   3rd Qu.: 0.54150
##  Max.   : 3.14850   Max.   : 1.905589   Max.   : 1.833396   Max.   : 2.87786
##      Family             CCAvg           Education1          Education2
##  Min.   :-1.20717   Min.   :-1.1117   Min.   :-0.862725   Min.   :-0.62348
##  1st Qu.:-1.20717   1st Qu.:-0.7101   1st Qu.:-0.862725   1st Qu.:-0.62348
##  Median :-0.34416   Median :-0.2510   Median :-0.862725   Median :-0.62348
##  Mean   :-0.02946   Mean   :-0.0057   Mean   :-0.009703   Mean   :-0.02078
##  3rd Qu.: 0.51884   3rd Qu.: 0.3227   3rd Qu.: 1.158655   3rd Qu.: 1.60325
##  Max.   : 1.38184   Max.   : 4.0524   Max.   : 1.158655   Max.   : 1.60325
##    Education3          Mortgage        Securities.Account   CD.Account
##  Min.   :-0.64394   Min.   :-0.54777   Min.   :-0.34427   Min.   :-0.2472
##  1st Qu.:-0.64394   1st Qu.:-0.54777   1st Qu.:-0.34427   1st Qu.:-0.2472
##  Median :-0.64394   Median :-0.54777   Median :-0.34427   Median :-0.2472
##  Mean   : 0.03104   Mean   : 0.04763   Mean   :-0.00433   Mean   : 0.0246
##  3rd Qu.: 1.55231   3rd Qu.: 0.48006   3rd Qu.:-0.34427   3rd Qu.:-0.2472
##  Max.   : 1.55231   Max.   : 5.44953   Max.   : 2.90355   Max.   : 4.0441
##      Online          CreditCard
##  Min.   :-1.22552   Min.   :-0.650785
##  1st Qu.:-1.22552   1st Qu.:-0.650785
##  Median : 0.81565   Median :-0.650785
##  Mean   :-0.02803   Mean   : 0.009622
##  3rd Qu.: 0.81565   3rd Qu.: 1.535992
##  Max.   : 0.81565   Max.   : 1.535992
```

```
var(Validation_norm2)
```

```
##                    Personal.Loan          Age   Experience       Income
## Personal.Loan       1.0367562006  0.034760883  0.033977749  0.524482873
## Age                 0.0347608835  1.034390729  1.020937362 -0.021521017
## Experience          0.0339777489  1.020937362  1.018204680 -0.016387163
## Income              0.5244828729 -0.021521017 -0.016387163  1.022101427
## Family              0.0735369006 -0.082295131 -0.091855996 -0.142636126
## CCAvg               0.4090315990 -0.009366446 -0.007441807  0.644671543
## Education1         -0.1465059112 -0.038495762 -0.015126481  0.194247333
## Education2          0.0581522428 -0.037267710 -0.042877732 -0.104391821
## Education3          0.1018242731  0.078583805  0.058726012 -0.108089239
## Mortgage            0.1884327805 -0.006476600 -0.005513505  0.235021855
## Securities.Account  0.0002453135 -0.015795157 -0.019099334 -0.034320896
## CD.Account          0.2673380394  0.008657233  0.010392644  0.137092495
## Online              0.0099426249 -0.011606783 -0.013009669  0.007398131
## CreditCard         -0.0211655142  0.014862622  0.015667732 -0.037290749
##                          Family        CCAvg   Education1   Education2
## Personal.Loan        0.07353690  0.409031599 -0.146505911  0.058152243
## Age                 -0.08229513 -0.009366446 -0.038495762 -0.037267710
## Experience          -0.09185600 -0.007441807 -0.015126481 -0.042877732
## Income              -0.14263613  0.644671543  0.194247333 -0.104391821
## Family               0.96165678 -0.108245622 -0.105707510  0.147356325
## CCAvg               -0.10824562  1.006053439  0.143798447 -0.078782727
## Education1          -0.10570751  0.143798447  0.997299429 -0.514461096
## Education2           0.14735633 -0.078782727 -0.514461096  0.979458824
## Education3          -0.03048693 -0.078534451 -0.576159121 -0.407084716
## Mortgage            -0.01098045  0.122849332  0.035517010 -0.066948308
## Securities.Account   0.01448616  0.035852491 -0.005492059 -0.007211108
## CD.Account           0.02558667  0.130553167 -0.006307495 -0.030045444
## Online               0.01303622 -0.020178598 -0.023010842  0.038851554
## CreditCard           0.05892716 -0.018939831  0.002459329 -0.008484843
##                       Education3     Mortgage Securities.Account   CD.Account
## Personal.Loan       0.101824273  0.188432781       0.0002453135  0.267338039
## Age                 0.078583805 -0.006476600      -0.0157951568  0.008657233
## Experience          0.058726012 -0.005513505      -0.0190993344  0.010392644
## Income             -0.108089239  0.235021855      -0.0343208962  0.137092495
## Family             -0.030486929 -0.010980453       0.0144861591  0.025586673
## CCAvg              -0.078534451  0.122849332       0.0358524914  0.130553167
## Education1         -0.576159121  0.035517010      -0.0054920586 -0.006307495
## Education2         -0.407084716 -0.066948308      -0.0072111081 -0.030045444
## Education3          1.027517723  0.027442322       0.0130796085  0.036487407
## Mortgage            0.027442322  1.103049365      -0.0050065488  0.107889682
## Securities.Account  0.013079609 -0.005006549       0.9891579289  0.363135926
## CD.Account          0.036487407  0.107889682       0.3631359263  1.093139341
## Online             -0.013318229 -0.062805436       0.0525985088  0.200232474
## CreditCard          0.005696622 -0.026417305       0.0169904348  0.289909068
##                         Online   CreditCard
## Personal.Loan       0.009942625 -0.021165514
## Age                -0.011606783  0.014862622
## Experience         -0.013009669  0.015667732
## Income              0.007398131 -0.037290749
## Family              0.013036220  0.058927161
```

```
## CCAvg              -0.020178598 -0.018939831
## Education1         -0.023010842  0.002459329
## Education2          0.038851554 -0.008484843
## Education3         -0.013318229  0.005696622
## Mortgage           -0.062805436 -0.026417305
## Securities.Account  0.052598509  0.016990435
## CD.Account          0.200232474  0.289909068
## Online              1.010977649  0.009647790
## CreditCard          0.009647790  1.008697186
```

11. Combining Train2 and Validation2 datasets to renormalize

Re-normalizing the combined data

```
ForTest_norm2 <- rbind.data.frame(Train_norm2, Validation_norm2, deparse.level = 1, make.row.names = T,

norm_values3 <- preProcess(ForTest_norm2, method = c("center", "scale"))
Test_norm2 <- predict(norm_values3, ForTest_norm2)

head(Test_norm2)
```

```
##    Personal.Loan         Age Experience      Income     Family      CCAvg
## 1     -0.3201721 -1.78103674 -1.6746848 -0.53706088  1.4030355 -0.1913336
## 2     -0.3201721 -0.03215752 -0.1001986 -0.86404190  0.5337446 -0.2486542
## 5     -0.3201721 -0.90659713 -1.0623846 -0.62425581  1.4030355 -0.5352576
## 7     -0.3201721  0.66739417  0.5995731 -0.03568998 -0.3355463 -0.2486542
## 11    -0.3201721  1.71672171  1.6492306  0.68366827  1.4030355  0.2672319
## 13    -0.3201721  0.23017437  0.2496873  0.87985688 -0.3355463  1.0697213
##    Education1 Education2 Education3   Mortgage Securities.Account CD.Account
## 1   1.1630146 -0.6181217 -0.6522349 -0.5549367          2.9114521 -0.2520536
## 2   1.1630146 -0.6181217 -0.6522349 -0.5549367          2.9114521 -0.2520536
## 5  -0.8596195  1.6173999 -0.6522349 -0.5549367         -0.3433854 -0.2520536
## 7  -0.8596195  1.6173999 -0.6522349 -0.5549367         -0.3433854 -0.2520536
## 11 -0.8596195 -0.6181217  1.5328067 -0.5549367         -0.3433854 -0.2520536
## 13 -0.8596195 -0.6181217  1.5328067 -0.5549367          2.9114521 -0.2520536
##        Online CreditCard
## 1  -1.2125565 -0.6534032
## 2  -1.2125565 -0.6534032
## 5  -1.2125565  1.5300661
## 7   0.8244977 -0.6534032
## 11 -1.2125565 -0.6534032
## 13 -1.2125565 -0.6534032
```

```
var(Test_norm2)
```

```
##                    Personal.Loan          Age   Experience        Income
## Personal.Loan        1.000000000 -0.011728834 -0.011756844  0.4964071722
## Age                 -0.011728834  1.000000000  0.994222365 -0.0532232125
## Experience          -0.011756844  0.994222365  1.000000000 -0.0455544621
## Income               0.496407172 -0.053223212 -0.045554462  1.0000000000
```

```
## Family               0.061671383 -0.063997298 -0.070156107 -0.1525902184
## CCAvg                0.364727538 -0.053174403 -0.051771655  0.6401950805
## Education1          -0.149923480 -0.032338292 -0.007669307  0.2183448689
## Education2           0.083028718 -0.015287234 -0.016621659 -0.1247693027
## Education3           0.080807736  0.049876931  0.024531442 -0.1139249768
## Mortgage             0.144149945 -0.007991582 -0.005759501  0.2013840413
## Securities.Account   0.013321026 -0.001365822 -0.002093179 -0.0002292213
## CD.Account           0.304234410 -0.012350625 -0.009483211  0.1769401938
## Online               0.013270140  0.005674230  0.005639355  0.0169526801
## CreditCard           0.005035806  0.010064745  0.011263492  0.0001453466
##                          Family        CCAvg   Education1   Education2
## Personal.Loan       0.061671383  0.364727538 -0.149923480  0.083028718
## Age                -0.063997298 -0.053174403 -0.032338292 -0.015287234
## Experience         -0.070156107 -0.051771655 -0.007669307 -0.016621659
## Income             -0.152590218  0.640195080  0.218344869 -0.124769303
## Family              1.000000000 -0.108267865 -0.108247803  0.126874064
## CCAvg              -0.108267865  1.000000000  0.152415975 -0.084327778
## Education1         -0.108247803  0.152415975  1.000000000 -0.531482345
## Education2          0.126874064 -0.084327778 -0.531482345  1.000000000
## Education3         -0.007069571 -0.082230640 -0.560814034 -0.403261378
## Mortgage           -0.020130712  0.099112783  0.027815575 -0.031574008
## Securities.Account  0.018472115  0.012241324  0.004362557  0.016952465
## CD.Account         -0.001149917  0.136432425 -0.016162317  0.006877721
## Online             -0.006671358 -0.007400654 -0.014347086  0.031490543
## CreditCard          0.007099607 -0.006053718  0.002512432 -0.004846405
##                       Education3      Mortgage Securities.Account   CD.Account
## Personal.Loan       0.0808077360  0.1441499446       0.0133210257  0.304234410
## Age                 0.0498769309 -0.0079915820      -0.0013658217 -0.012350625
## Experience          0.0245314415 -0.0057595014      -0.0020931789 -0.009483211
## Income             -0.1139249768  0.2013840413      -0.0002292213  0.176940194
## Family             -0.0070695714 -0.0201307121       0.0184721153 -0.001149917
## CCAvg              -0.0822306397  0.0991127826       0.0122413238  0.136432425
## Education1         -0.5608140336  0.0278155745       0.0043625568 -0.016162317
## Education2         -0.4032613783 -0.0315740084       0.0169524654  0.006877721
## Education3          1.0000000000  0.0008120123      -0.0212825122  0.010737656
## Mortgage            0.0008120123  1.0000000000      -0.0090689856  0.091665279
## Securities.Account -0.0212825122 -0.0090689856       1.0000000000  0.301408689
## CD.Account          0.0107376558  0.0916652791       0.3014086888  1.000000000
## Online             -0.0152803732 -0.0203103990       0.0112817558  0.175636853
## CreditCard          0.0020228010 -0.0031070805      -0.0218296666  0.277500011
##                         Online    CreditCard
## Personal.Loan       0.013270140  0.0050358058
## Age                 0.005674230  0.0100647454
## Experience          0.005639355  0.0112634917
## Income              0.016952680  0.0001453466
## Family             -0.006671358  0.0070996073
## CCAvg              -0.007400654 -0.0060537178
## Education1         -0.014347086  0.0025124324
## Education2          0.031490543 -0.0048464054
## Education3         -0.015280373  0.0020228010
## Mortgage           -0.020310399 -0.0031070805
## Securities.Account  0.011281756 -0.0218296666
## CD.Account          0.175636853  0.2775000115
## Online              1.000000000  0.0061014669
```

```
## CreditCard          0.006101467  1.0000000000
```

12. Applying KNN Method with the best K

```
Q5_KNN <- knn(Train_norm2, Validation_norm2, cl = Train2[,1], k=3, prob = TRUE)
head(Q5_KNN)
```

```
## [1] 0 0 0 0 0 0
## Levels: 0 1
```

```
row.names(Train_norm2)[attr(Q5_KNN, "Q5_KNN.index")]
```

```
## character(0)
```

13. Creating Confusion Matrix 2

```
Train_Predictors <- Train_norm2
Test_Predictors <- Validation_norm2

Train_labels2 <-Train_norm2[,1]
Test_labels2  <-Validation_norm2[,1]

Predicted_Test_labels2 <- knn(Train_Predictors, Test_Predictors, cl=Train_labels2, k=1, prob = TRUE)
library(gmodels)
CrossTable(x=Test_labels2, y=Predicted_Test_labels2, prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Row Total |
## |           N / Col Total |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  1500
##
##
##                    | Predicted_Test_labels2
##       Test_labels2 | -0.317484332194368 |   3.14850182713279 |          Row Total |
## -------------------|--------------------|--------------------|--------------------|
## -0.317484332194368 |               1357 |                  0 |               1357 |
##                    |              1.000 |              0.000 |              0.905 |
##                    |              1.000 |              0.000 |                    |
##                    |              0.905 |              0.000 |                    |
## -------------------|--------------------|--------------------|--------------------|
```

```
##   3.14850182713279 |                   0 |                 143 |                 143 |
##                    |               0.000 |               1.000 |               0.095 |
##                    |               0.000 |               1.000 |                     |
##                    |               0.000 |               0.095 |                     |
## -------------------|---------------------|---------------------|---------------------|
##       Column Total |                1357 |                 143 |                1500 |
##                    |               0.905 |               0.095 |                     |
## -------------------|---------------------|---------------------|---------------------|
##
##
```

```
class_prob<-attr(Predicted_Test_labels, 'prob')
head(class_prob)
```

```
## [1] 1 1 1 1 1 1
```

14. Creating Confusion Matrix 3

```
Train_Predictors <- ForTest_norm2
Test_Predictors <- Test_norm2

Train_labels3 <-ForTest_norm2[,1]
Test_labels3  <-Test_norm2[,1]

Predicted_Test_labels3 <- knn(Train_Predictors, Test_Predictors, cl=Train_labels3, k=1, prob = TRUE)
library(gmodels)
CrossTable(x=Test_labels3, y=Predicted_Test_labels3, prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Row Total |
## |           N / Col Total |
## |         N / Table Total |
## |-------------------------|
##
##
## Total Observations in Table:  4000
##
##
##                    | Predicted_Test_labels3
##       Test_labels3 | -0.317484332194368 |   3.14850182713279 |          Row Total |
## -------------------|--------------------|--------------------|--------------------|
## -0.320172138788574 |               3628 |                  0 |               3628 |
##                    |              1.000 |              0.000 |              0.907 |
##                    |              1.000 |              0.000 |                    |
##                    |              0.907 |              0.000 |                    |
## -------------------|--------------------|--------------------|--------------------|
##   3.12253903098104 |                  0 |                372 |                372 |
```

```
##                    |              0.000 |              1.000 |              0.093 |
##                    |              0.000 |              1.000 |                    |
##                    |              0.000 |              0.093 |                    |
## -------------------|--------------------|--------------------|--------------------|
##       Column Total |               3628 |                372 |               4000 |
##                    |              0.907 |              0.093 |                    |
## -------------------|--------------------|--------------------|--------------------|
##
##
```

```
class_prob<-attr(Predicted_Test_labels3, 'prob')
head(class_prob)
```

```
## [1] 1 1 1 1 1 1
```