

Anjali Assignment-3

2023-10-18

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.3      v tibble    3.2.1
## v lubridate  1.9.2      v tidyr     1.3.0
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(reshape)
```

```
##
## Attaching package: 'reshape'
##
## The following object is masked from 'package:lubridate':
##
##   stamp
##
## The following object is masked from 'package:dplyr':
##
##   rename
##
## The following objects are masked from 'package:tidyr':
##
##   expand, smiths
```

```
library(caret)
```

```
## Loading required package: lattice
##
## Attaching package: 'caret'
##
## The following object is masked from 'package:purrr':
##
##   lift
```

```
library(e1071)
```

```
UniB<- read_csv("C:/Users/Anjali/Desktop/Anjali FML Assignment 3/UniversalBank.csv")
```

```
## Rows: 5000 Columns: 14
## -- Column specification -----
## Delimiter: ","
## dbl (14): ID, Age, Experience, Income, ZIP Code, Family, CCAvg, Education, M...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(UniB)
```

```
## # A tibble: 6 x 14
##       ID Age Experience Income 'ZIP Code' Family CCAvg Education Mortgage
##   <dbl> <dbl>      <dbl>  <dbl>      <dbl>  <dbl> <dbl>      <dbl>      <dbl>
## 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         1         0
## 4     4   35         9    100      94112     1   2.7         2         0
## 5     5   35         8     45      91330     4   1         2         0
## 6     6   37        13     29      92121     4   0.4         2        155
## # i 5 more variables: Personal_Loan <dbl>, 'Securities Account' <dbl>,
## #   'CD Account' <dbl>, Online <dbl>, CreditCard <dbl>
```

```
colnames(UniB)
```

```
## [1] "ID"           "Age"           "Experience"
## [4] "Income"       "ZIP Code"      "Family"
## [7] "CCAvg"        "Education"     "Mortgage"
## [10] "Personal_Loan" "Securities Account" "CD Account"
## [13] "Online"       "CreditCard"
```

```
UniB$Personal_Loan = as.factor(UniB$Personal_Loan)
UniB$Online = as.factor(UniB$Online)
UniB$CreditCard = as.factor(UniB$CreditCard)
```

```
set.seed(456)
UniB_traindata <- sample(row.names(UniB), 0.6*dim(UniB)[1])
UniB_validdata <- setdiff(row.names(UniB), UniB_traindata)
UniB_train <- UniB[UniB_traindata, ]
UniB_valid <- UniB[UniB_validdata, ]
train <- UniB[UniB_traindata,]
valid <- UniB[UniB_validdata,]
```

```
#a
```

```
library(reshape2)
```

```
##
## Attaching package: 'reshape2'

## The following objects are masked from 'package:reshape':
##
##   colsplit, melt, recast

## The following object is masked from 'package:tidyr':
##
##   smiths

melt = melt(train,id=c("CreditCard","Personal_Loan"),variable= "Online")

## Warning: attributes are not identical across measure variables; they will be
## dropped

cast = dcast(melt,CreditCard + Personal_Loan ~ Online)

## Aggregation function missing: defaulting to length

cast[,c(1,2,3,14)]

##   CreditCard Personal_Loan   ID Online
## 1          0             0 1917   1917
## 2          0             1  200    200
## 3          1             0  794    794
## 4          1             1   89     89

#b

UniB.Loan.CC1 <- 89/3000
UniB.Loan.CC1

## [1] 0.02966667

#c

melt1 = melt(train,id=c("Personal_Loan"),variable = "Online")

## Warning: attributes are not identical across measure variables; they will be
## dropped

melt2 = melt(train,id=c("CreditCard"),variable = "Online")

## Warning: attributes are not identical across measure variables; they will be
## dropped
```

```

cast1 =dcast(melt1,`Personal_Loan`~Online)

## Aggregation function missing: defaulting to length

cast2=dcast(melt2,CreditCard~Online)

## Aggregation function missing: defaulting to length

UniB.Loanonline=cast1[,c(1,13)]
UniB.LoanCC = cast2[,c(1,14)]
UniB.Loanonline

##      Personal_Loan Online
## 1              0    2711
## 2              1     289

UniB.LoanCC

##      CreditCard Online
## 1              0    2117
## 2              1     883

#d

table(train[,c(14,10)])

##           Personal_Loan
## CreditCard    0      1
##           0 1917  200
##           1  794   89

table(train[,c(13,10)])

##           Personal_Loan
## Online       0      1
##           0 1046  112
##           1 1665  177

table(train[,c(10)])

## Personal_Loan
##      0      1
## 2711  289

#1.  $P(CC = 1 \mid Loan = 1)$ 
UniB.CCUB.Loan1 = 89/(89+200)
UniB.CCUB.Loan1

## [1] 0.3079585

```

```
#2.  $P(Online=1|Loan=1)$ 
UniB.ONUB.Loan1 = 177/(177+112)
UniB.ONUB.Loan1
```

```
## [1] 0.6124567
```

```
#3.  $P(Loan = 1)$ 
UniB.Loan1 = 289/(289+2711)
UniB.Loan1
```

```
## [1] 0.09633333
```

```
#4.  $P(CC=1|Loan=0)$ 
UniB.CCLoan.01 = 794/(794+1917)
UniB.CCLoan.01
```

```
## [1] 0.2928809
```

```
#5.  $P(Online=1|Loan=0)$ 
UniB.ON1.L0 = 1665/(1665+1046)
UniB.ON1.L0
```

```
## [1] 0.6141645
```

```
#6.  $P(Loan=0)$ 
UniB.Loan0 = 2711/(2711+289)
UniB.Loan0
```

```
## [1] 0.9036667
```

```
#e
```

```
UniB_Naivebayes = ((89/(89+200))*(177/(177+112))*(289/(289+2711)))/(((89/(89+200))*(177/(177+112))*(289/(289+2711)))+(2711/(2711+289)))
UniB_Naivebayes
```

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

```
#f
```

```
library(caret)
library(e1071)
UniB_nb_train = UniB_train[,c(10,13,14)]
UniB_naivebayes_1 = naiveBayes(`Personal_Loan`~.,data=UniB_nb_train)
UniB_naivebayes_1
```

```
##
## Naive Bayes Classifier for Discrete Predictors
##
## Call:
```

```

## naiveBayes.default(x = X, y = Y, laplace = laplace)
##
## A-priori probabilities:
## Y
##      0      1
## 0.90366667 0.09633333
##
## Conditional probabilities:
##      Online
## Y      0      1
## 0 0.3858355 0.6141645
## 1 0.3875433 0.6124567
##
##      CreditCard
## Y      0      1
## 0 0.7071191 0.2928809
## 1 0.6920415 0.3079585

```