Assignment 3

Madhukar Heerekar

2022-03-06

```
unibank = read.csv("C:\\Users\\heere\\OneDrive\\Desktop\\FMA ASSIGNMENT 3\\UniversalBank.csv")
library(gmodels)
library(caret)
```

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

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

```
library(ISLR)
library(e1071)
unibank$Personal.Loan<-factor(unibank$Personal.Loan)
unibank$Online<-factor(unibank$Online)
unibank$CreditCard<-factor(unibank$CreditCard)
df=unibank

#task1
set.seed(64060)
train_index <-createDataPartition(df$Personal.Loan, p =0.6, list = FALSE)
train.df = df[train_index,]

mytable <- xtabs(~ CreditCard + Personal.Loan , data = train.df)
ftable(mytable)
```

```
##            Personal.Loan    0    1
## CreditCard
## 0                        1924  195
## 1                         788   93
```

```
#task2
#Probablity for peronsl loan acceptance(1)conditional on having a bank
#credit_card(cc=1)and being an active user of online banking services
#(online=1)
#probability of loan acceptance given having a credit card and user
probability = (93/(93+788))
probability
```

```
## [1] 0.1055619
```

```r
#task3
#Create two separate pivot tables for the training data. One will have Loan (rows) as a
#function of Online (columns) and the other will have Loan (rows) as a function of CC.


table(Personal.Loan = train.df$Personal.Loan, Online= train.df$Online)
```

```
##              Online
## Personal.Loan    0    1
##             0 1081 1631
##             1  109  179
```

```r
table(Personal.Loan = train.df$Personal.Loan, CreditCard = train.df$CreditCard)
```

```
##              CreditCard
## Personal.Loan    0    1
##             0 1924  788
##             1  195   93
```

```r
table(personal. = train.df$Personal.Loan)
```

```
## personal.
##    0    1
## 2712  288
```

```r
#task4
#i.P(CC = 1 | Loan = 1) (the proportion of credit card holders among the loan acceptors)
probability1 <- 93/(93+195)
probability1
```

```
## [1] 0.3229167
```

```r
#2.p(online=1 | Loan=1)
probability2 <- 179/(179+109)
probability2
```

```
## [1] 0.6215278
```

```r
#iii. P(Loan=)(the proportion of loan acceptors)
probability3 <- 288/(288+2712)
probability3
```

```
## [1] 0.096
```

```r
#iv. P(CC=1 | Loan=0)
probability4 <- 788/(788+1924)


#v. P(Online=1 | Loan=0)
probability5 <- 1631/(1631+1081)
probability5
```

```
## [1] 0.6014012
```

```
#vi.P(Loan=0)
probability6 <- 2712/(2712+288)
probability6
```

```
## [1] 0.904
```

```
#Task6
#Compare this value with the one obtained from the pivot table in (B). Which is a more
#accurate estimate?
#Let a=
task5probability <- (0.28125 * 0.59375 * 0.096)/((0.28125 * 0.59375 * 0.096)) + (0.2971976 * 0.604351 *
task5probability
```

```
## [1] 1.162369
```

```
#Task7
#Run naive Bayes on the data
#Examine the model output on training data, and find the entry
#that corresponds to P(Loan = 1 | CC = 1, Online = 1
# Compare this to the number you
#obtained in (E).



nb.model <- naiveBayes(Personal.Loan~ Online + CreditCard, data=train.df)
To_predict=data.frame(online=1, creditcard=1)
predict(nb.model,To_predict,type='raw')
```

```
##          0     1
## [1,] 0.904 0.096
```