Assignment_3

Gowtham Chakri Mallepaka

2022-10-17

```
library(caret)
```

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

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

```
library(ISLR)
library(reshape2)
library(e1071)

ubank<-read.csv("c:/Users/reliance digital/Downloads/UniversalBank.csv")

#converting varibles to factors
ubank$Personal.Loan=factor(ubank$Personal.Loan)
ubank$Online=factor(ubank$Online)
ubank$CreditCard=factor(ubank$CreditCard)
set.seed(8)

#Splitting of data
set.seed(8)
Train_index=createDataPartition(ubank$Personal.Loan,p=0.6,list=FALSE)
Train.df=ubank[Train_index,]
validation.df=ubank[-Train_index,]
nrow(validation.df)
```

```
## [1] 2000
```

```
#taskA
table= xtabs(~CreditCard+Personal.Loan+Online,data=Train.df)
ftable(table)
```

```
##                          Online    0    1
## CreditCard Personal.Loan
## 0          0                     787 1146
##            1                      70  128
## 1          0                     319  460
##            1                      40   50
```

```r
#task b
#the proability that the person will accept the loan (with online= 1, cc= 1)
50/(50+460)
```

```
## [1] 0.09803922
```

```r
#task c
table(Personal.Loan=Train.df$Personal.Loan,CreditCard=Train.df$CreditCard)
```

```
##              CreditCard
## Personal.Loan    0    1
##             0 1933  779
##             1  198   90
```

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

```
##              Online
## Personal.Loan    0    1
##             0 1106 1606
##             1  110  178
```

```r
table(Personal.Loan=Train.df$Personal.Loan)
```

```
## Personal.Loan
##    0    1
## 2712  288
```

```r
#task d- finding the probabilities
#1 p(cc/l)
p1=90/(90+198)

#2 p(online/l)
p2=178/(178+110)

#3 p(l)
p3=288/(288+2712)

#4 p(cc=1/l=0)
p4=779/(779+1933)

#5 p(onine=1/l=0)
p5=1606/(1606+1106)

#6 p(l=0)
p6=2712/(2712+288)

#task E
(p1*p2*p3)/((p1*p2*p3)+(p4*p5*p6))
```

```
## [1] 0.1076053
```

##task f

Probability from Pivot table is 0.09803922 and probablity from naive bayes is 0.1076053 probablity from pivot table is more accurate than naive bayes. This is because in naive bayes we make an assumption that in attributes are independent of each other

```
#task g
model=naiveBayes(Personal.Loan~CreditCard+Online,data=Train.df)
test.set=data.frame(CreditCard=1,Online=1)
test.set$CreditCard=factor(test.set$CreditCard)
test.set$Online=factor(test.set$Online)
predict(model,test.set,type = "raw")
```

```
##              0         1
## [1,] 0.8923947 0.1076053
```

predicted probability is same as that of task E