Assignment 3

Jacob Aylward

2025-06-13

```
library(caret)

library(readr)
Universal_Bank <- read_csv("C:/Users/jacob/Downloads/UniversalBank.csv")
View(Universal_Bank)

set.seed(123)
Index_Train <- createDataPartition(Universal_Bank$`Personal Loan`, p=0.6,
list = FALSE)
Train <- Universal_Bank[Index_Train,]
Test <- Universal_Bank[Index_Train,]

pivot_table <- table(Train$CreditCard,Train$`Personal Loan`,Train$Online) #
creating pivot table with CC and Personal Loan = Rows and Online = Column
pivot_table  #Pivot Table

## , ,  = 0
##
##
##        0    1
##   0  785   65
##   1  317   34
##
## , ,  = 1
##
##
##        0    1
##   0 1145  122
##   1  475   57

loan_accepted <- pivot_table["1","1","1"]
total_accepted <- sum(pivot_table["1" , ,"1"]) #Probability based on Credit
Card and Online
probability <- loan_accepted / total_accepted
probability

## [1] 0.1071429

pivottable_loan_online <- table(Train$`Personal Loan`,Train$Online) #Pivot
table with Personal Loan = Row and Online = Column
pivottable_loan_online

##
##        0    1
```

```
##   0 1102 1620
##   1   99  179
```

```
pivottable_loan_cc <- table(Train$`Personal Loan`,Train$CreditCard) #Pivot table with Personal Loan = Row and Credit Card = Column
pivottable_loan_cc
```

```
##
##        0    1
##   0 1930  792
##   1  187   91
```

```
p1 <- sum(pivottable_loan_cc["1","1"]) / sum(pivottable_loan_cc["1",]) # P(CC = 1 Loan = 1)
p1
```

```
## [1] 0.3273381
```

```
p2 <- sum(pivottable_loan_online["1","1"]) / sum(pivottable_loan_online["1",]) # P(Online = 1  Loan = 1)
p2
```

```
## [1] 0.6438849
```

```
p3 <- sum(pivottable_loan_online["1" ,]) / sum(pivottable_loan_online) # P(Loan = 1)
p3
```

```
## [1] 0.09266667
```

```
p4 <- sum(pivottable_loan_cc["0","1"]) / sum(pivottable_loan_cc["0",]) # P(CC = 1 Loan = 0)
p4
```

```
## [1] 0.2909625
```

```
p5 <- sum(pivottable_loan_online["0","1"]) / sum(pivottable_loan_online["0",]) # P(Online = 1  Loan = 0)
p5
```

```
## [1] 0.5951506
```

```
p6 <- sum(pivottable_loan_online["0" ,]) / sum(pivottable_loan_online)  # P(Loan = 0)
p6
```

```
## [1] 0.9073333
```

```
library(e1071)
```

```
nb_prob <- (p1 * p2) / p5 # Naive Bayes Probability P(Loan = 1  CC = 1 Online = 1)
nb_prob
```

```
## [1] 0.3541424
```

When comparing the probability from the Naive Bayes and the pivot table of part B the more accurate estimate is likely from the pivot table. Based on the video lecture "Naive Bayes Classifier" by Dr.Wu it discusses that with NB there is an assumption of conditional independence but in this scenario the probabilities are based on the dependency of the needed variables.

```
nb_model <-naiveBayes(`Personal Loan`~Online+CreditCard,data = Train)  # Naive Bayes model on training data
nb_model

##
## Naive Bayes Classifier for Discrete Predictors
##
## Call:
## naiveBayes.default(x = X, y = Y, laplace = laplace)
##
## A-priori probabilities:
## Y
##          0          1
## 0.90733333 0.09266667
##
## Conditional probabilities:
##    Online
## Y        [,1]      [,2]
##   0 0.5951506 0.4909531
##   1 0.6438849 0.4797134
##
##    CreditCard
## Y        [,1]      [,2]
##   0 0.2909625 0.4542897
##   1 0.3273381 0.4700881
```