

Chapter 4

Bijesh Mishra

2/17/2021

Machine Learning Chapter 4: Classification

This document contains codes related to STAT 5063 Machine Learning Chapter 4 from Statistics Department of Oklahoma State University. This RMD is stored in lecture folder. Lecture folder also contains RMD file and pdf generated after knitting RMD file. So, for each lecture, there must be three files inside this folder: 1) Lecture notes (Chapter 4), 2) RMD File (Chapter 4.Rmd) and 3) Pdf file (Chapter-4) after knitting with same name. Note that RMD file for homework is stored inside homework folder.

Lecture and Video Codes:

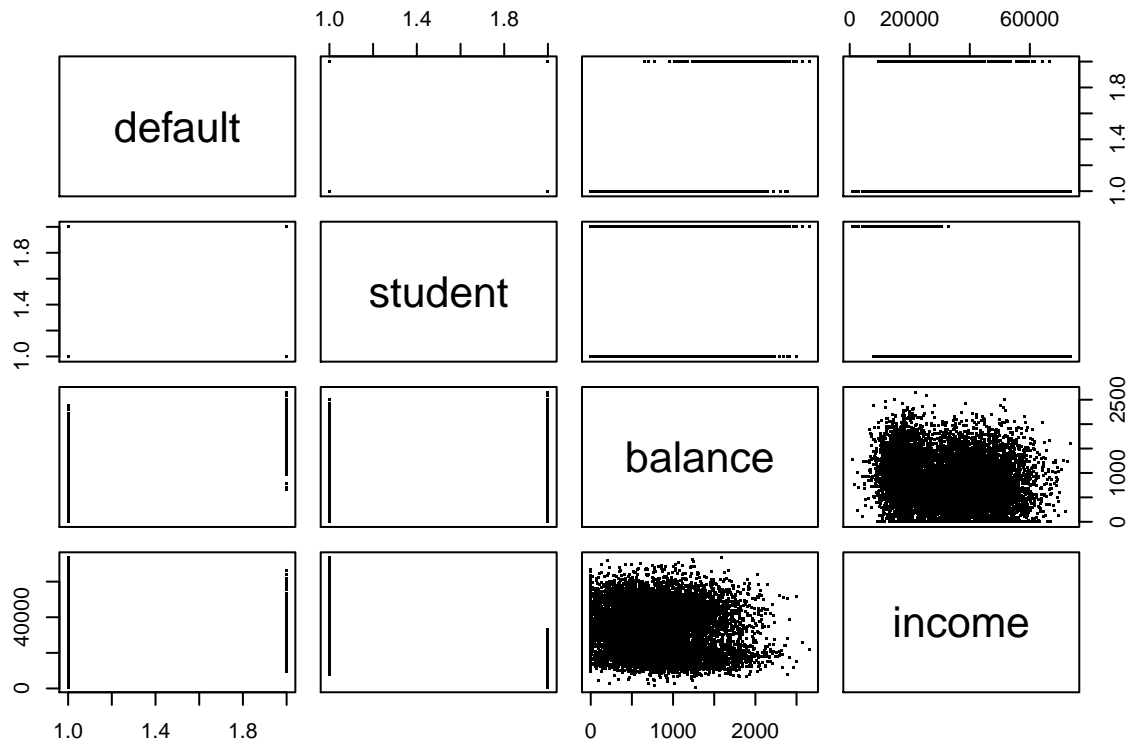
```
rm(list = ls()) #Clear environment. #Run manually.
```

Motivation and Overview

- Logistic regression
- Linear discriminant analysis
- Quadratic discriminant analysis
- K-nearest neighbors

4.1 Classification: Motivation and Overview

```
library(ISLR)  
help(Default)  
plot(Default, pch = ".")
```



```
dim(Default)
```

```
## [1] 10000      4
```

```
Default[1:3,]
```

```
##   default student  balance  income
## 1      No      No  729.5265 44361.63
## 2      No     Yes  817.1804 12106.13
## 3      No      No 1073.5492 31767.14
```

```
summary(Default)
```

```
##   default      student      balance      income
##   No :9667    No :7056   Min.   :  0.0   Min.   :  772
##   Yes: 333    Yes:2944  1st Qu.: 481.7  1st Qu.:21340
##                                     Median : 823.6  Median :34553
##                                     Mean   : 835.4  Mean   :33517
##                                     3rd Qu.:1166.3  3rd Qu.:43808
##                                     Max.   :2654.3  Max.   :73554
```