ica10_shuangyu_zhao

shuangyu_zhao

2023-02-09

```
library(ISLR2)
oj <- OJ
```

1.

```
split_pro <- 0.75
n <- length(oj$Purchase)*split_pro
row_samp <- sample(1:length(oj$Purchase), n, replace = FALSE)
train <- oj[row_samp, ]
test <- oj[-row_samp,]
```

2. LDA

```
library(MASS)
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:ISLR2':
##
##     Boston
```

```
oj_lda_mod <- lda(data = train, Purchase ~ PriceDiff  + LoyalCH)
summary(oj_lda_mod)
```

```
##         Length Class  Mode
## prior   2      -none- numeric
## counts  2      -none- numeric
## means   4      -none- numeric
## scaling 2      -none- numeric
## lev     2      -none- character
## svd     1      -none- numeric
## N       1      -none- numeric
## call    3      -none- call
## terms   3      terms  call
## xlevels 0      -none- list
```

```
prediction_test <- predict(oj_lda_mod, test)
library(caret)
```

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

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

```
confusionMatrix(data = prediction_test$class, reference = test$Purchase)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  CH  MM
##         CH 141  15
##         MM  25  87
##
##                Accuracy : 0.8507
##                  95% CI : (0.8024, 0.8912)
##     No Information Rate : 0.6194
##     P-Value [Acc > NIR] : <2e-16
##
##                   Kappa : 0.6893
##
##  Mcnemar's Test P-Value : 0.1547
##
##             Sensitivity : 0.8494
##             Specificity : 0.8529
##          Pos Pred Value : 0.9038
##          Neg Pred Value : 0.7768
##              Prevalence : 0.6194
##          Detection Rate : 0.5261
##    Detection Prevalence : 0.5821
##       Balanced Accuracy : 0.8512
##
##        'Positive' Class : CH
##
```

QDA

```
oj_qda_mod <- qda(data = train, Purchase ~ PriceDiff  + LoyalCH)

prediction_test_q <- predict(oj_qda_mod, test)
confusionMatrix(data = prediction_test_q$class, reference = test$Purchase)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  CH  MM
##         CH 143  16
##         MM  23  86
##
```

```
##                Accuracy : 0.8545
##                  95% CI : (0.8065, 0.8944)
##     No Information Rate : 0.6194
##     P-Value [Acc > NIR] : <2e-16
##
##                   Kappa : 0.6954
##
##  Mcnemar's Test P-Value : 0.3367
##
##             Sensitivity : 0.8614
##             Specificity : 0.8431
##          Pos Pred Value : 0.8994
##          Neg Pred Value : 0.7890
##              Prevalence : 0.6194
##          Detection Rate : 0.5336
##    Detection Prevalence : 0.5933
##       Balanced Accuracy : 0.8523
##
##        'Positive' Class : CH
##
```