

Assignment40

John Bute

2024-11-19

R Markdown

7. Build bagging models for Wine datasets.

```
library(ipred)
Wine <- read.csv("C:/Users/johnb/Desktop/Machine Learning/data/wine.csv",
stringsAsFactors = TRUE)

set.seed(1)
Wine$Class = as.factor(Wine$Class)

index <- sample(1:nrow(Wine), 0.7 * nrow(Wine))
train_data <- Wine[index, ]
test_data <- Wine[-index, ]

bagging_model <- bagging(Class ~ ., data = train_data,
                        nbagg = 150, coob = TRUE,
                        controll = rpart::rpart.control(minsplit = 5, cp =
0),
                        importance = TRUE)

bagging_preds <- predict(bagging_model, newdata = test_data, type = "class")
conf_matrix_bagging <- table(Actual = test_data$Class, Predicted =
bagging_preds)
print(conf_matrix_bagging)

##          Predicted
## Actual   1   2   3
##      1 20   0   0
##      2   0 21   1
##      3   0   0 12

test_error <- 1 - sum(diag(conf_matrix_bagging)) / sum(conf_matrix_bagging)
cat("Test Error Rate:", test_error, "\n")

## Test Error Rate: 0.01851852
```

The model shows incredible performance as demonstrated by the confusion matrix and the test error.

8. Build random forest models for the Wine dataset

```
library(randomForest)

## Warning: package 'randomForest' was built under R version 4.4.2
```

```

## randomForest 4.7-1.2

## Type rfNews() to see new features/changes/bug fixes.

rf_model <- randomForest(Class ~ ., data = train_data,
                          ntree = 500)

rf_preds <- predict(rf_model, newdata = test_data, type = "class")
conf_matrix_rf <- table(Actual = test_data$Class, Predicted = rf_preds)
print(conf_matrix_rf)

##          Predicted
## Actual  1  2  3
##      1 20  0  0
##      2  0 21  1
##      3  0  0 12

test_error_rf <- 1 - sum(diag(conf_matrix_rf)) / sum(conf_matrix_rf)
cat("Test Error Rate:", test_error_rf, "\n")

## Test Error Rate: 0.01851852

```

It literally performed the exact same way as the bagging model.