

Iris Data LDA

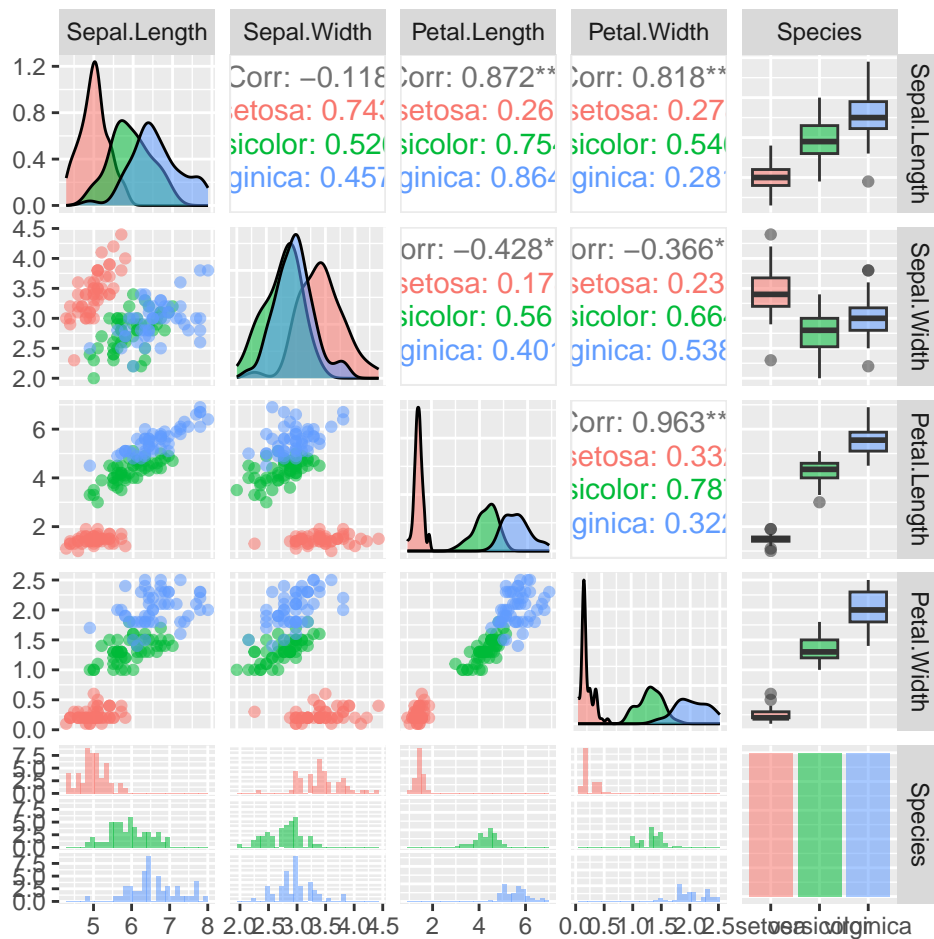
Carson Slater

2024-07-15

Colophon

```
library("tidymodels")  
library("GGally")  
library("MASS")  
library("caret")
```

Loading the Data, Exploring the data



Creating Cross Validation Folds

```
set.seed(613)

n_folds <- 10

# create 10 folds
folds <- createFolds(iris$Species, k = n_folds, list = TRUE, returnTrain = FALSE)
```

Performing the Linear Discriminant Analysis

```
apparent_error_rates <- numeric(n_folds)
preds <- vector("list", length = n_folds)
posteriors <- vector("list", length = n_folds)

for(i in 1:10) {

  test_indices <- folds[[i]]
  train_indices <- setdiff(1:nrow(iris), test_indices)

  train_data <- iris[train_indices, ]
  test_data <- iris[test_indices, ]

  trained_lda <- MASS::lda(Species ~ ., train_data)
  predictions <- predict(trained_lda, test_data)

  preds[[i]] <- predictions$class
  posteriors[[i]] <- predictions$posterior

  apparent_error_rate <- mean(preds[[i]] != test_data$Species)
  apparent_error_rates[i] <- apparent_error_rate

}

# apparent error rate
mean(apparent_error_rates)
```

```
## [1] 0.02
```

The apparent error rate is 0.02.