Exercises 6

Brady Lamson

2022-03-14

11: Supervised Learning

11.1 Classifiers

Exercise 1

Based on the decision tree I would make the following predictions:

- A 9-inch, 10-pound pet would likely be a **cat** and be pretty confident.
- A 14-inch, 21-pound pet would likely be a **dog** and be very confident.

Exercise 2

```
type <- c("dog", "dog", "cat", "dog", "cat", "dog", "cat", "dog", "cat",
          "dog", "cat", "dog", "dog", "cat", "dog", "cat", "dog", "cat", "dog")

wt <- c(8, 17, 8, 18, 7, 22, 6, 16, 7, 20, 10, 15, 14, 11, 13, 13, 15, 17, 10)
ht <- c(7.5, 10, 8, 15, 7, 15, 7, 13, 11, 16, 7, 10.5, 9, 9.5, 9, 8, 9, 8, 12)

pets <- data.frame(Type = type, Ht = ht, Wt = wt)

my.tree <- rpart(
    Type ~ Wt + Ht, data = pets,
    control = rpart.control(minsplit = 7)
)

newPets <- data.frame(Ht = c(9, 14), Wt = c(10, 21))

predict(my.tree, newdata = newPets, type = "class")
```

```
##   1   2
## cat dog
## Levels: cat dog
```

First pet is predicted to be a cat and the second pet is predicted to be a dog. This prediction makes the same predictions I made in exercise 1.

Exercise 3

```
my.tree <- rpart(Species ~ Petal.Length + Petal.Width, data = iris)
my.tree
```

```
## n= 150
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
## 1) root 150 100 setosa (0.33333333 0.33333333 0.33333333)
##   2) Petal.Length< 2.45 50   0 setosa (1.00000000 0.00000000 0.00000000) *
##   3) Petal.Length>=2.45 100  50 versicolor (0.00000000 0.50000000 0.50000000)
##     6) Petal.Width< 1.75 54   5 versicolor (0.00000000 0.90740741 0.09259259) *
##     7) Petal.Width>=1.75 46   1 virginica (0.00000000 0.02173913 0.97826087) *
```

a) This tree has 3 different terminal nodes.

```
predSpecies <- predict(my.tree, type = "class")
species <- data.frame(Actual = iris$Species, Predicted = predSpecies)

confusion <- table(species)
confusion
```

```
##             Predicted
## Actual       setosa versicolor virginica
##   setosa         50          0         0
##   versicolor      0         49         1
##   virginica       0          5        45
```

b) The trees correct classification rate is **96%** and the mis-classification rate is, by extension, **4%**