Decision Tree

Mr. Gulati

July 31, 2020

Importing libraries

```r
library(party)
```

```
## Loading required package: grid

## Loading required package: mvtnorm

## Loading required package: modeltools

## Loading required package: stats4

## Loading required package: strucchange

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Loading required package: sandwich
```

```r
library(plyr)
```

```
##
## Attaching package: 'plyr'

## The following object is masked from 'package:modeltools':
##
##     empty
```

```r
library(readr)
```

Importing dataset

```r
dataset <- read.csv("D:/Internship/Task 3/iris.csv")
#converting Species from character to factor
dataset$Species=as.factor(dataset$Species)
head(dataset)
```

```
##   Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm     Species
## 1  1           5.1          3.5           1.4          0.2 Iris-setosa
## 2  2           4.9          3.0           1.4          0.2 Iris-setosa
## 3  3           4.7          3.2           1.3          0.2 Iris-setosa
## 4  4           4.6          3.1           1.5          0.2 Iris-setosa
## 5  5           5.0          3.6           1.4          0.2 Iris-setosa
## 6  6           5.4          3.9           1.7          0.4 Iris-setosa
```

```r
#Removing Id column
dataset <- dataset[-1]
head(dataset)
```

```
##   SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm     Species
## 1           5.1          3.5           1.4          0.2 Iris-setosa
## 2           4.9          3.0           1.4          0.2 Iris-setosa
## 3           4.7          3.2           1.3          0.2 Iris-setosa
## 4           4.6          3.1           1.5          0.2 Iris-setosa
## 5           5.0          3.6           1.4          0.2 Iris-setosa
## 6           5.4          3.9           1.7          0.4 Iris-setosa
```
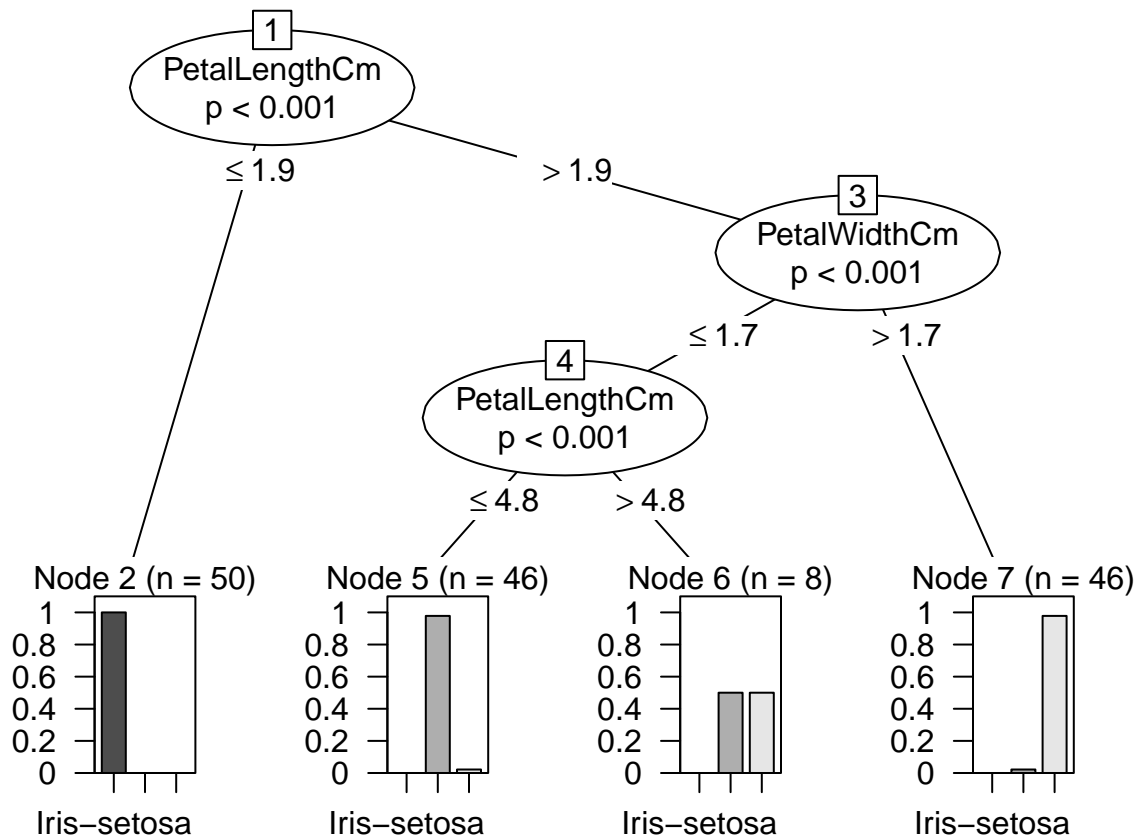
Defining The Decision Tree Algorithm

```r
model <- ctree(Species ~., data = dataset)
```

Visualizing Decision Tree

```r
plot(model)
```

PetalLengthCm
p < 0.001

≤ 1.9   > 1.9

PetalWidthCm
p < 0.001

≤ 1.7   > 1.7

PetalLengthCm
p < 0.001

≤ 4.8   > 4.8

Node 2 (n = 50)   Node 5 (n = 46)   Node 6 (n = 8)   Node 7 (n = 46)

1   1   1   1
0.8   0.8   0.8   0.8
0.6   0.6   0.6   0.6
0.4   0.4   0.4   0.4
0.2   0.2   0.2   0.2
0   0   0   0

Iris−setosa   Iris−setosa   Iris−setosa   Iris−setosa

Prediction

```
table(predict(model, dataset[,1:4]), dataset$Species)
```

```
##
##                   Iris-setosa Iris-versicolor Iris-virginica
##   Iris-setosa              50               0              0
##   Iris-versicolor           0              49              5
##   Iris-virginica            0               1             45
```
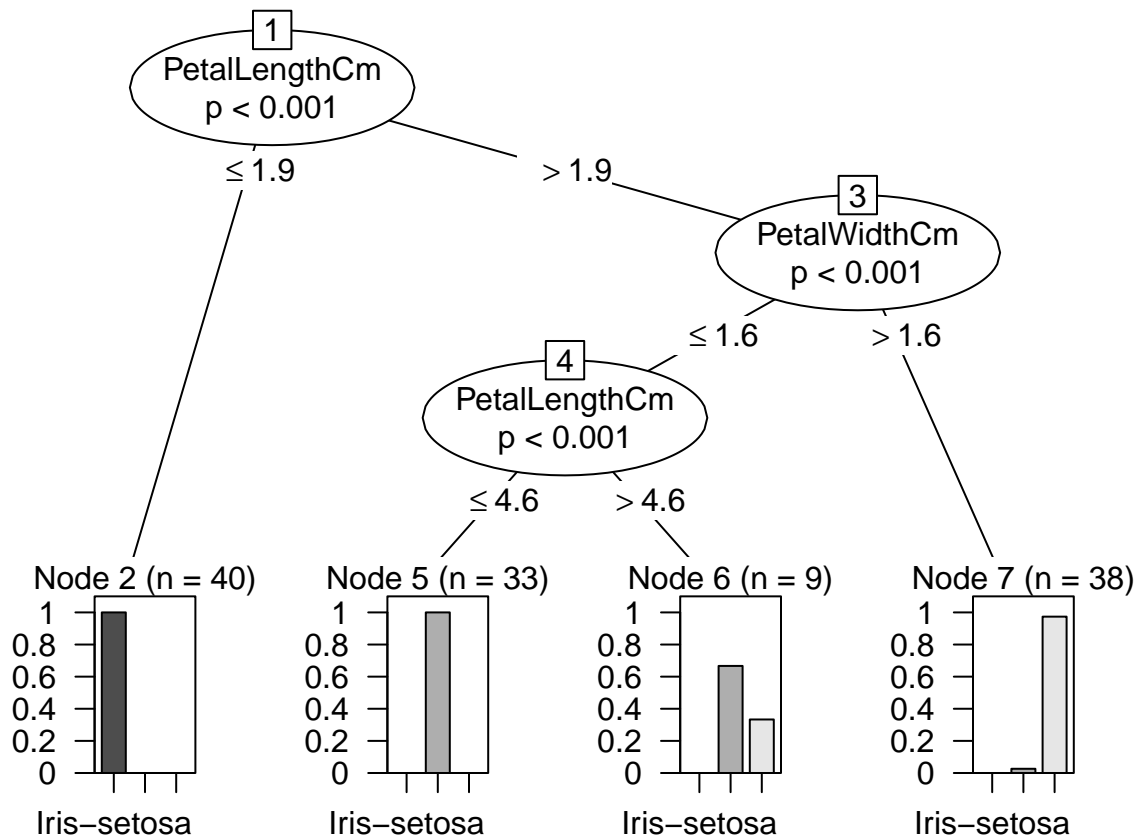
Decision Tree Classifier: 70:30 Train & Test samples

```
library("caTools")
set.seed(100)
splitm <-sample.split(dataset$Species, SplitRatio = 0.8)
train=subset(dataset,splitm==TRUE)
test=subset(dataset,splitm==FALSE)
model1 <- ctree(Species ~ .,data=train)
```

Visualising

```
plot(model1)
```

Prediction

```r
table(predict(model1,test[,1:4]),test$Species)
```

```
##
##                   Iris-setosa Iris-versicolor Iris-virginica
##   Iris-setosa              10               0              0
##   Iris-versicolor           0               9              1
##   Iris-virginica            0               1              9
```