Tree_Visualizer_vignette

2023-06-07

Tree Visualizer

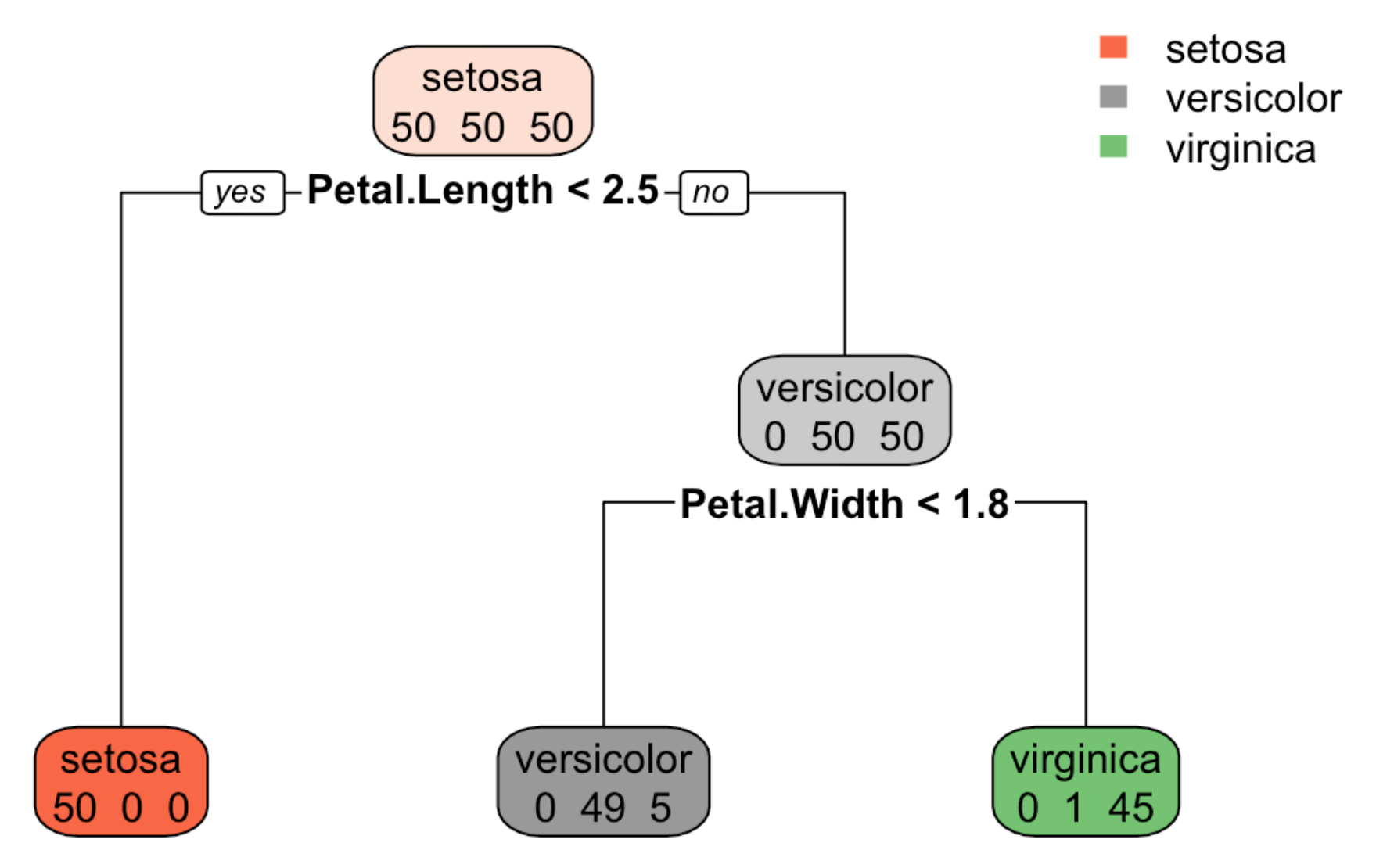
The package Rpart has a very high quality visualization suite, but it is only compatible with Rpart objects, which have a very complex structure. This limits interoperability with decision trees produced by other software or by hand.

By contrast, the structure used by the randomForest package is user-friendly and clear. Therefore, this package provides an adapter between the randomForest tree object and the rpart object.

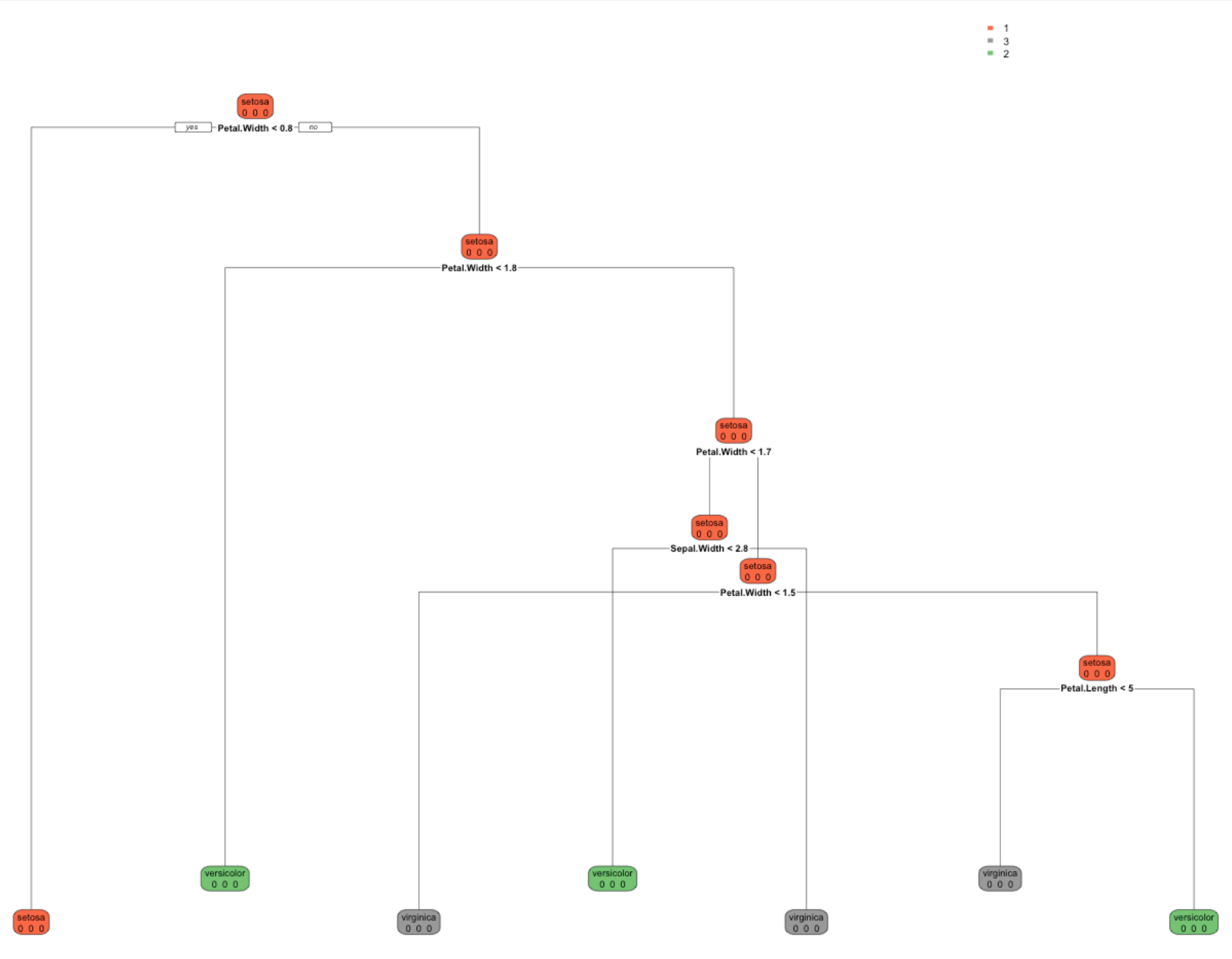
Let's demonstrate this with the iris dataset

```
data(iris)
iris_rpart = rpart(Species ~ .,data = iris)
iris_rf = randomForest(Species ~ .,data = iris)
iris_tree = iris_rf %>% getTree(k = 3,labelVar=T)  %>% as.data.frame()

iris_adapted = adapt_rf2rpart(iris_tree,add_data=F)
rpart.plot(iris_rpart,extra = 1,type = 2)
```



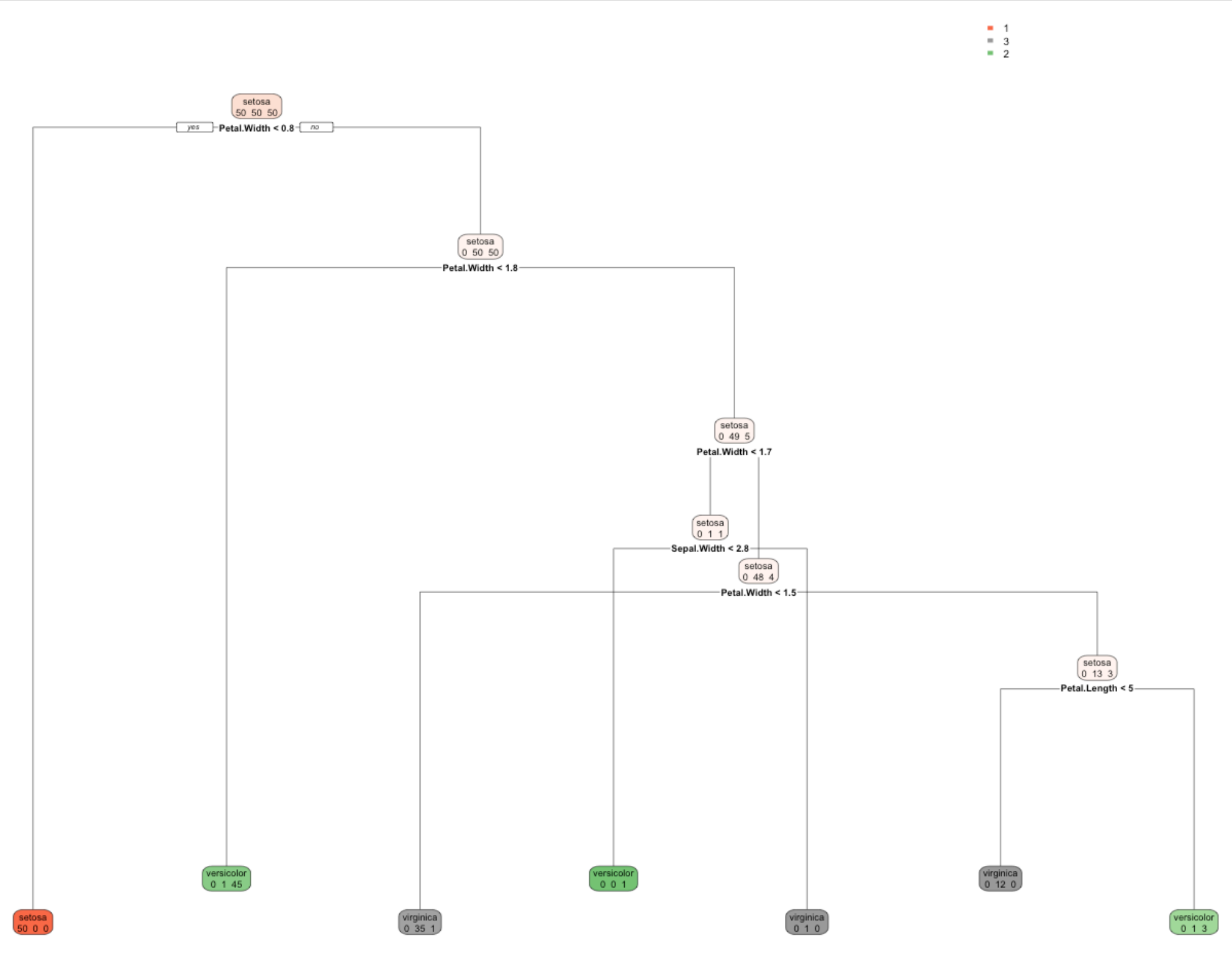
```
rpart.plot(iris_adapted,extra = 1,type = 2)
```



By default, the trees returned by adapt_rf2rpart are empty. It is worth noting that all intermediate nodes default to the first factor level, provided it is a classification tree

The tree can be populated by any dataset by setting add_data = T and supplying data

```
iris_adapted = adapt_rf2rpart(iris_tree,add_data=T,X = iris[,-5],y = iris[,5])
rpart.plot(iris_adapted,extra = 1,type = 2)
```



For now, intermediate nodes still display the first option, but the nodes are nevertheless fileld out, providing insight into model performance.

In addition to the plotting adapter, the package provides a function for evaluating tree performance. This can be used to either evaluate a decision tree, or identify which trees within a random forest may be best to plot, for example

```
## $tree_info
##    left daughter right daughter    split var split point status prediction
## 1              2              3  Petal.Width        0.80      1       <NA>
## 2              0              0         <NA>        0.00     -1     setosa
## 3              4              5  Petal.Width        1.75      1       <NA>
## 4              6              7  Petal.Width        1.65      1       <NA>
## 5              0              0         <NA>        0.00     -1  virginica
## 6              8              9  Petal.Width        1.45      1       <NA>
## 7             10             11  Sepal.Width        2.75      1       <NA>
## 8              0              0         <NA>        0.00     -1 versicolor
## 9             12             13 Petal.Length        4.95      1       <NA>
## 10             0              0         <NA>        0.00     -1  virginica
## 11             0              0         <NA>        0.00     -1 versicolor
## 12             0              0         <NA>        0.00     -1 versicolor
## 13             0              0         <NA>        0.00     -1  virginica
##    setosa virginica versicolor bin    purity
## 1       0         0          0   0 0.0000000
## 2      50         0          0  50 1.0000000
## 3       0         0          0   0 0.0000000
## 4       0         0          0   0 0.0000000
## 5       0        45          1  46 0.9782609
## 6       0         0          0   0 0.0000000
## 7       0         0          0   0 0.0000000
## 8       0         1         35  36 0.9722222
## 9       0         0          0   0 0.0000000
## 10      0         1          0   1 1.0000000
## 11      0         0          1   1 1.0000000
## 12      0         0         12  12 1.0000000
## 13      0         3          1   4 0.7500000
## 
## $conf_mat
##        predicts true_val
##   [1,]        0        0
##   [2,]        0        0
##   [3,]        0        0
##   [4,]        0        0
##   [5,]        0        0
##   [6,]        0        0
##   [7,]        0        0
##   [8,]        0        0
##   [9,]        0        0
##  [10,]        0        0
##  [11,]        0        0
##  [12,]        0        0
##  [13,]        0        0
##  [14,]        0        0
##  [15,]        0        0
##  [16,]        0        0
##  [17,]        0        0
##  [18,]        0        0
##  [19,]        0        0
##  [20,]        0        0
##  [21,]        0        0
##  [22,]        0        0
##  [23,]        0        0
##  [24,]        0        0
##  [25,]        0        0
##  [26,]        0        0
##  [27,]        0        0
##  [28,]        0        0
##  [29,]        0        0
##  [30,]        0        0
##  [31,]        0        0
##  [32,]        0        0
##  [33,]        0        0
##  [34,]        0        0
##  [35,]        0        0
##  [36,]        0        0
##  [37,]        0        0
##  [38,]        0        0
##  [39,]        0        0
##  [40,]        0        0
##  [41,]        0        0
##  [42,]        0        0
##  [43,]        0        0
##  [44,]        0        0
##  [45,]        0        0
##  [46,]        0        0
##  [47,]        0        0
##  [48,]        0        0
##  [49,]        0        0
##  [50,]        0        0
##  [51,]        2        2
##  [52,]        2        2
##  [53,]        2        2
##  [54,]        2        2
##  [55,]        2        2
##  [56,]        2        2
##  [57,]        2        2
##  [58,]        2        2
##  [59,]        2        2
##  [60,]        2        2
##  [61,]        2        2
##  [62,]        2        2
##  [63,]        2        2
##  [64,]        2        2
##  [65,]        2        2
##  [66,]        2        2
##  [67,]        2        2
##  [68,]        2        2
##  [69,]        2        2
##  [70,]        2        2
##  [71,]        2        2
##  [72,]        2        2
##  [73,]        2        2
##  [74,]        2        2
##  [75,]        2        2
##  [76,]        2        2
##  [77,]        2        2
##  [78,]        2        2
##  [79,]        2        2
##  [80,]        2        2
##  [81,]        2        2
##  [82,]        2        2
##  [83,]        2        2
##  [84,]        2        2
##  [85,]        2        2
##  [86,]        2        2
##  [87,]        2        2
##  [88,]        2        2
##  [89,]        2        2
##  [90,]        2        2
##  [91,]        2        2
##  [92,]        2        2
##  [93,]        2        2
##  [94,]        2        2
##  [95,]        2        2
##  [96,]        2        1
##  [97,]        1        2
##  [98,]        1        1
##  [99,]        1        1
## [100,]        1        1
## [101,]        1        1
## [102,]        1        1
## [103,]        1        1
## [104,]        1        1
## [105,]        1        1
## [106,]        1        1
## [107,]        1        1
## [108,]        1        1
## [109,]        1        1
## [110,]        1        1
## [111,]        1        1
## [112,]        1        1
## [113,]        1        1
## [114,]        1        1
## [115,]        1        1
## [116,]        1        1
## [117,]        1        1
## [118,]        1        1
## [119,]        1        1
## [120,]        1        1
## [121,]        1        1
## [122,]        1        1
## [123,]        1        1
## [124,]        1        1
## [125,]        1        1
## [126,]        1        1
## [127,]        1        1
## [128,]        1        1
## [129,]        1        1
## [130,]        1        1
## [131,]        1        1
## [132,]        1        1
## [133,]        2        2
## [134,]        1        1
## [135,]        1        1
## [136,]        1        1
## [137,]        1        1
## [138,]        1        1
## [139,]        1        1
## [140,]        1        1
## [141,]        1        1
## [142,]        1        1
## [143,]        1        1
## [144,]        1        1
## [145,]        1        1
## [146,]        1        1
## [147,]        2        2
## [148,]        2        2
## [149,]        2        2
## [150,]        2        1
## 
## $sensitivity
## [1] 0.96
## 
## $specificity
## [1] 1
## 
## $precision
## [1] 1
## 
## $npv
## [1] 1
## 
## $fpr
## [1] 0
## 
## $fdr
## [1] 0
## 
## $fnr
## [1] 0
## 
## $f1
## [1] 0.9795918
## 
## $accuracy
## [1] 0.6533333
```