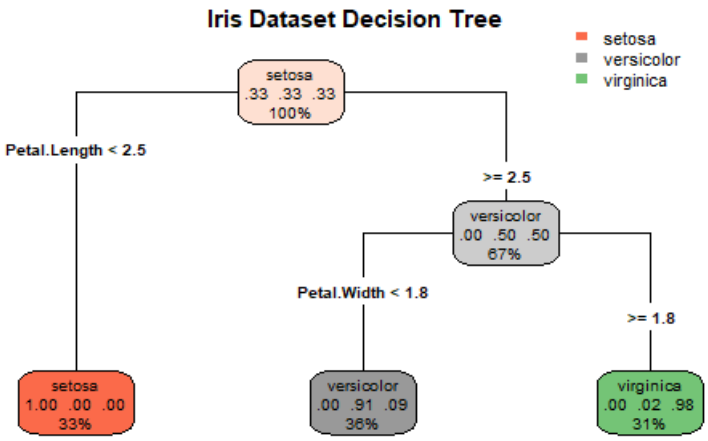
Javier-EchavarrenSuarez-HW3

We present the plot of the decision tree in order to discuss the results.



(a) How many levels are there in the decision tree?

In the decision tree there are two levels.

(b) What is the default class label associated with each vertex?

The upper vertex has as default class label “setosa”, and the second upper vertex has as default class label “versicolor”.

The final vertices have as default class labels “setosa”,”versicolor” and “virginica” from left to right.

Level 0, Vertex 1: Default class label is <setosa>

Level 1, Vertex 2: Default class label is <setosa>

Level 1, Vertex 3: Default class label is <versicolor>

Level 2, Vertex 6: Default class label is <versicolor>

Level 2, Vertex 7: Default class label is <virginica>

(c) Starting from the root note, what is the name of the first attribute used for a decision, and what are the split points?

Level 0, split on attribute: <Petal.Length>.

Split points: < 2.5 left subtree, >= 2.5 right subtree

(d) Each vertex has three lines.

(i) At each vertex, what do the three numbers in the middle line signify?

They signify the proportion of the samples in each vertex containing each of the class labels, what in other words is the probability of each sample in that vertex of having each of the class labels.

(ii) At each vertex, what does the last line signify?

The percentage of number of samples in that vertex in relation to the total number of samples.