INSTITUTO FEDERAL
Norte de Minas Gerais
Campus
Montes Claros

Árvore de Decisão

Equipe: Jaudir Aguiar, Jonas Diego, Matheus Henrique
Profa.: Luciana Cosme Balieiro

.: Sumário

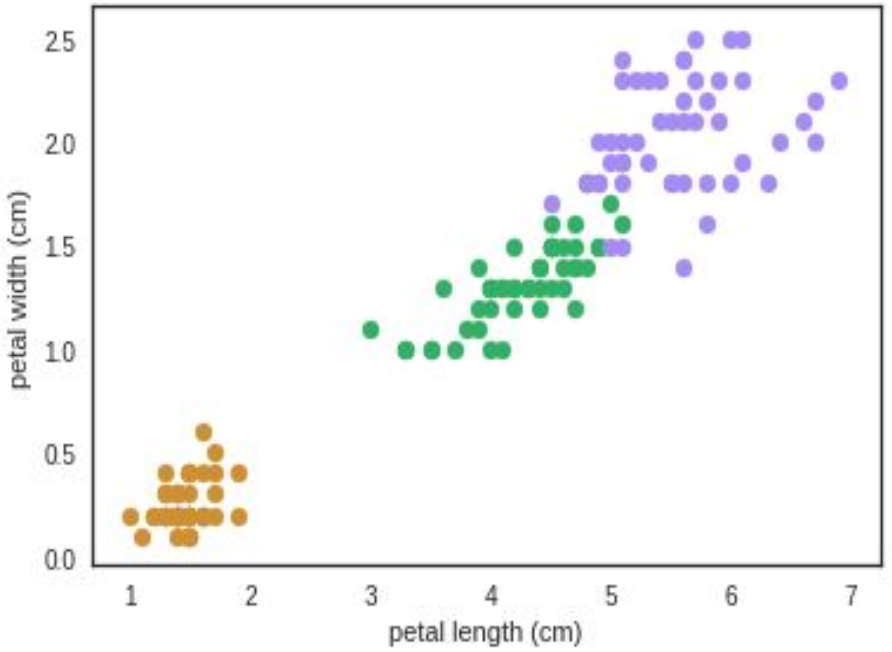
- Problema
- Método
- Construção da Árvore de Decisão
- Código
- Resultados
- Simulação

.: Árvores de decisão

“Árvores de decisão são **modelos estatísticos** que utilizam um **treinamento supervisionado** para a classificação e previsão de dados.”

.: Explicação do problema

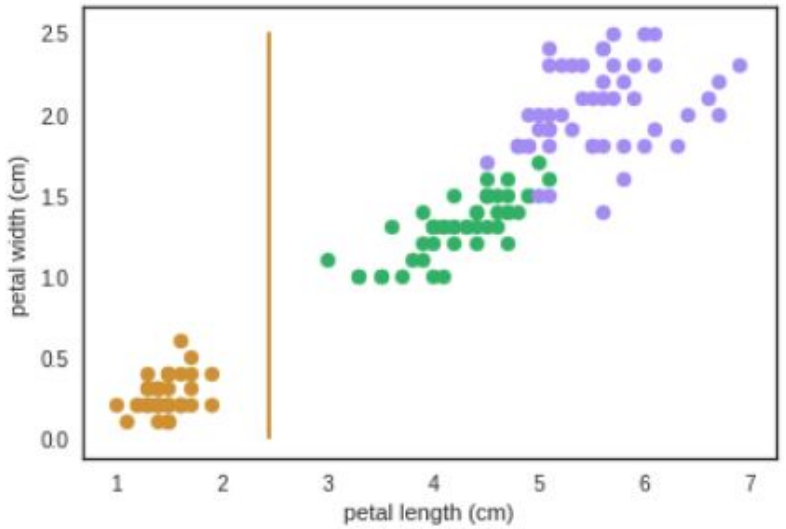
- Base de dados
  - Setosa (amarela)
  - Versicolor (verde)
  - Viriginica (roxa)
- Caracterítiscas
  - Comprimento sépala
  - Largura sépala
  - Comprimento pétala
  - Largura pétala



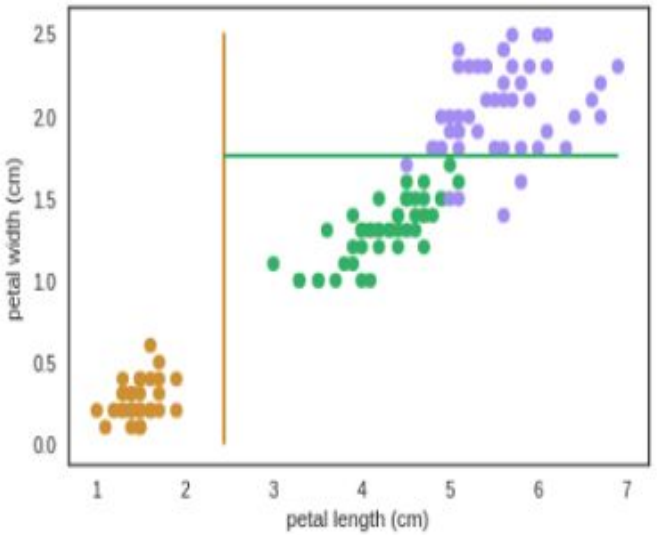
.: Construção da Árvores de decisão

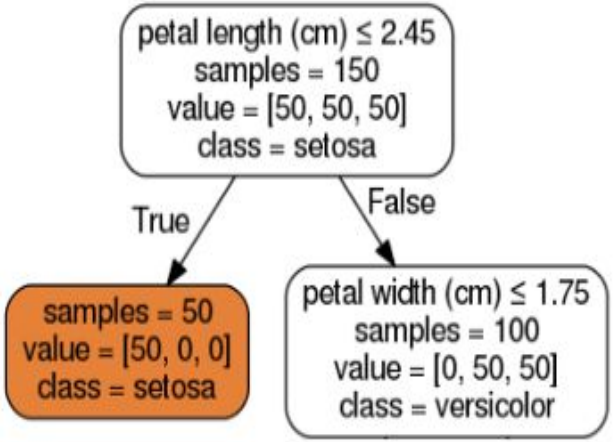
Divide o espaço recursivamente

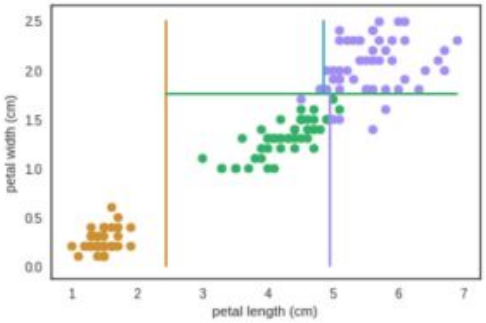
- Cortes ortogonais
- Uma variável por vez
  - Maximiza a pureza das sub-regiões resultantes
- Cada divisão é representada por um nó na árvore de decisão

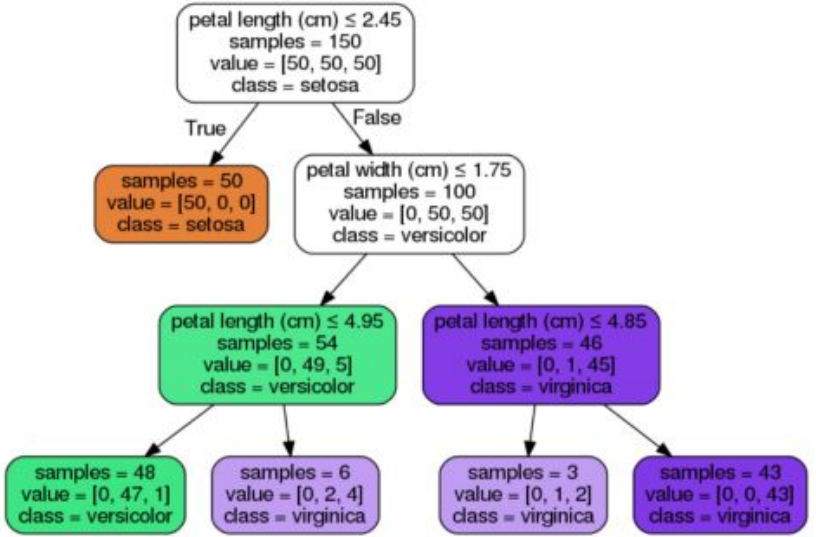


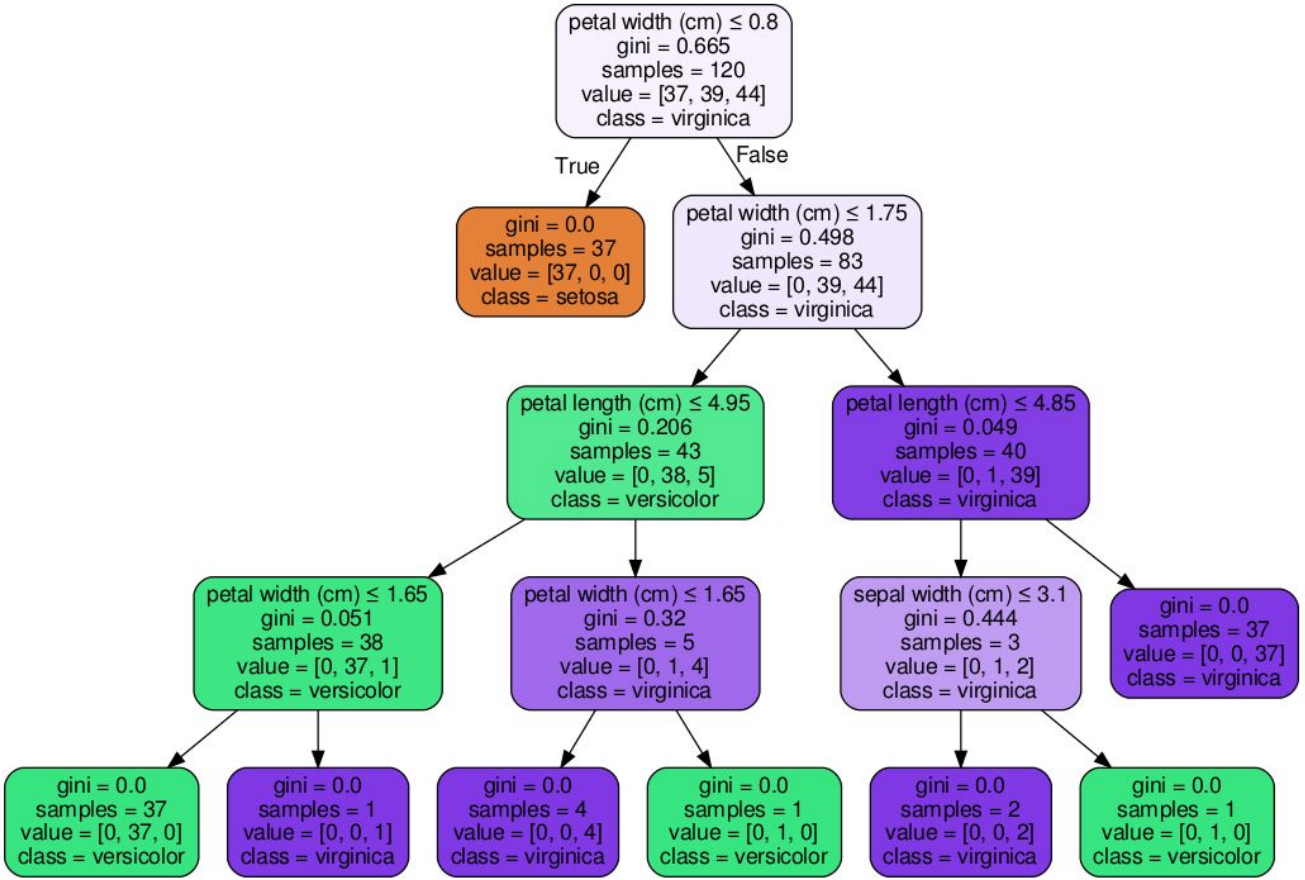
petal length (cm) ≤ 2.45
samples = 150
value = [50, 50, 50]
class = setosa


petal width (cm)
petal length (cm)
0.0
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
7


petal length (cm) ≤ 2.45
samples = 150
value = [50, 50, 50]
class = setosa
True
False
samples = 50
value = [50, 0, 0]
class = setosa
petal width (cm) ≤ 1.75
samples = 100
value = [0, 50, 50]
class = versicolor


petal width (cm)
petal length (cm)
0.0
0.5
1.0
1.5
2.0
2.5
1
2
3
4
5
6
7


petal length (cm) ≤ 2.45
samples = 150
value = [50, 50, 50]
class = setosa
True
False
samples = 50
value = [50, 0, 0]
class = setosa
petal width (cm) ≤ 1.75
samples = 100
value = [0, 50, 50]
class = versicolor
petal length (cm) ≤ 4.95
samples = 54
value = [0, 49, 5]
class = versicolor
petal length (cm) ≤ 4.85
samples = 46
value = [0, 1, 45]
class = virginica
samples = 48
value = [0, 47, 1]
class = versicolor
samples = 6
value = [0, 2, 4]
class = virginica
samples = 3
value = [0, 1, 2]
class = virginica
samples = 43
value = [0, 0, 43]
class = virginica

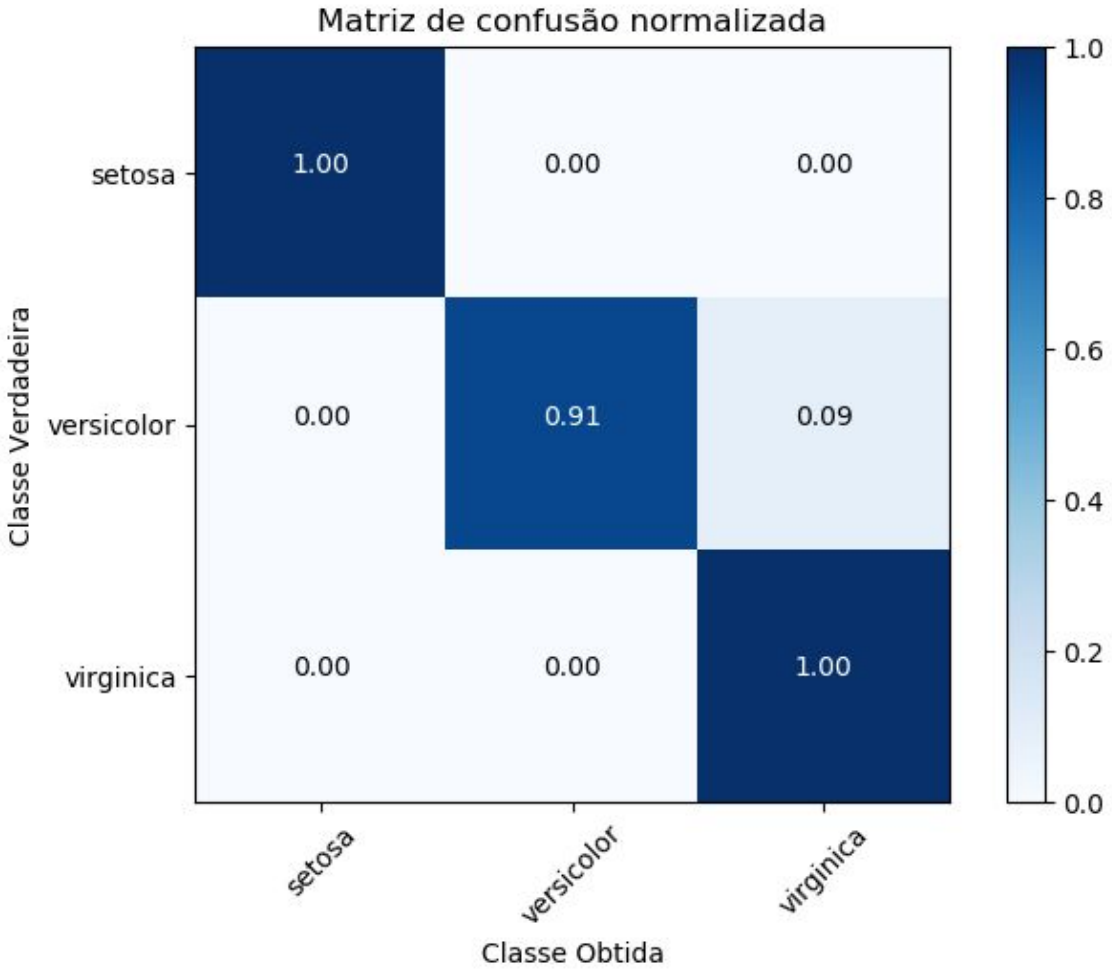

petal width (cm) ≤ 0.8
gini = 0.665
samples = 120
value = [37, 39, 44]
class = virginica
True
False
gini = 0.0
samples = 37
value = [37, 0, 0]
class = setosa
petal width (cm) ≤ 1.75
gini = 0.498
samples = 83
value = [0, 39, 44]
class = virginica
petal length (cm) ≤ 4.95
gini = 0.206
samples = 43
value = [0, 38, 5]
class = versicolor
petal length (cm) ≤ 4.85
gini = 0.049
samples = 40
value = [0, 1, 39]
class = virginica
petal width (cm) ≤ 1.65
gini = 0.051
samples = 38
value = [0, 37, 1]
class = versicolor
petal width (cm) ≤ 1.65
gini = 0.32
samples = 5
value = [0, 1, 4]
class = virginica
sepal width (cm) ≤ 3.1
gini = 0.444
samples = 3
value = [0, 1, 2]
class = virginica
gini = 0.0
samples = 37
value = [0, 0, 37]
class = virginica
gini = 0.0
samples = 37
value = [0, 37, 0]
class = versicolor
gini = 0.0
samples = 1
value = [0, 0, 1]
class = virginica
gini = 0.0
samples = 4
value = [0, 0, 4]
class = virginica
gini = 0.0
samples = 1
value = [0, 1, 0]
class = versicolor
gini = 0.0
samples = 2
value = [0, 0, 2]
class = virginica
gini = 0.0
samples = 1
value = [0, 1, 0]
class = versicolor

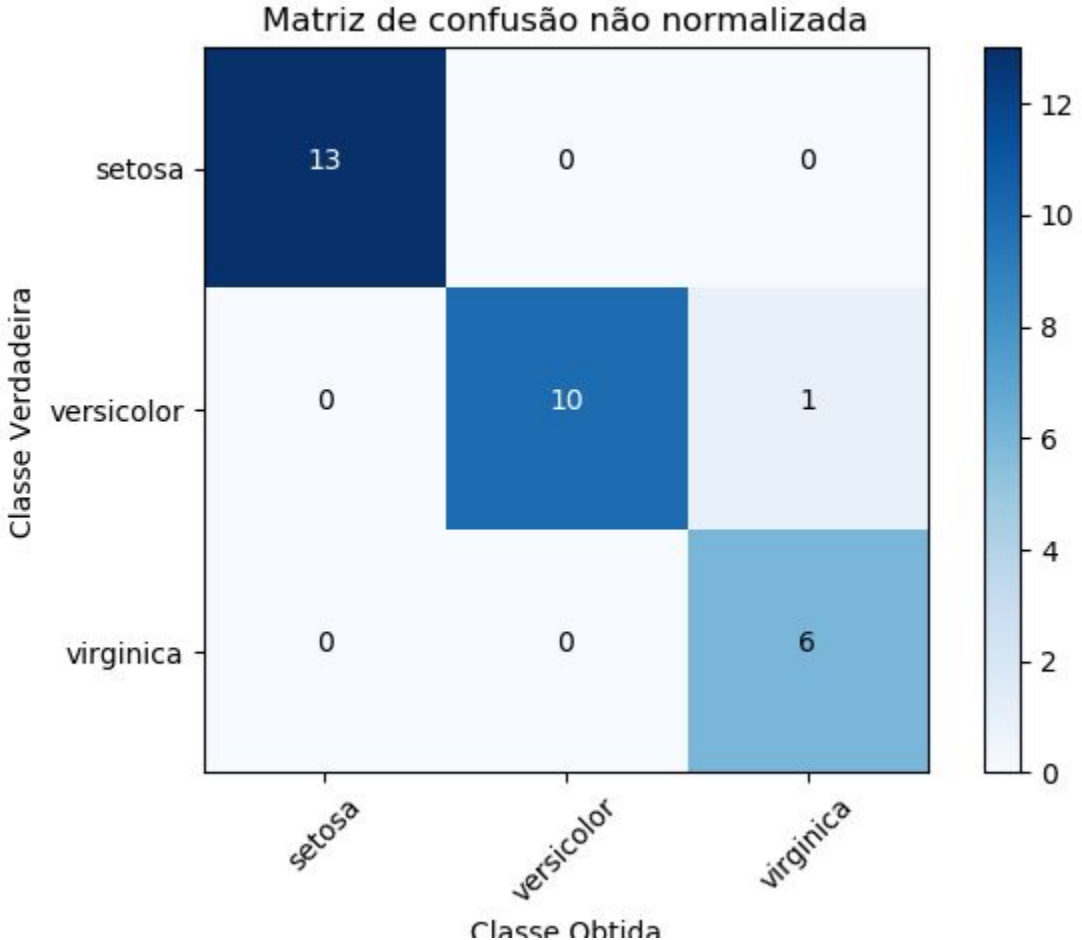
.: Árvores de decisão

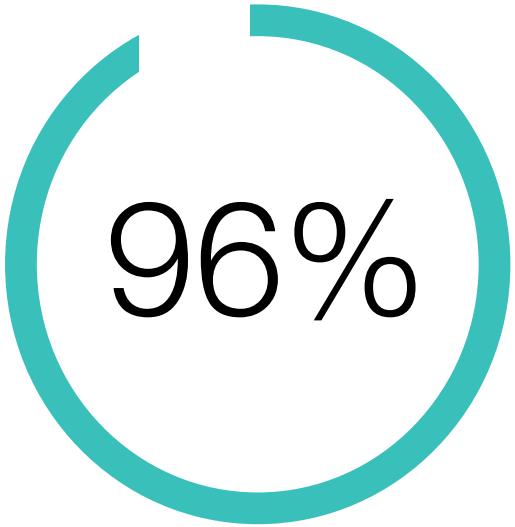
```
def DecisionTree(iris):
    X_train, X_test, y_train, y_test = train_test_split(iris.data, iris.target,\
    test_size=0.20,train_size = 0.80, random_state=80)

    clf = tree.DecisionTreeClassifier()
    clf = clf.fit(X_train, y_train)
    y_pred = clf.predict(X_test)
    return clf, y_test,y_pred,clf.score(X_test,y_test)
```

.: Resultados

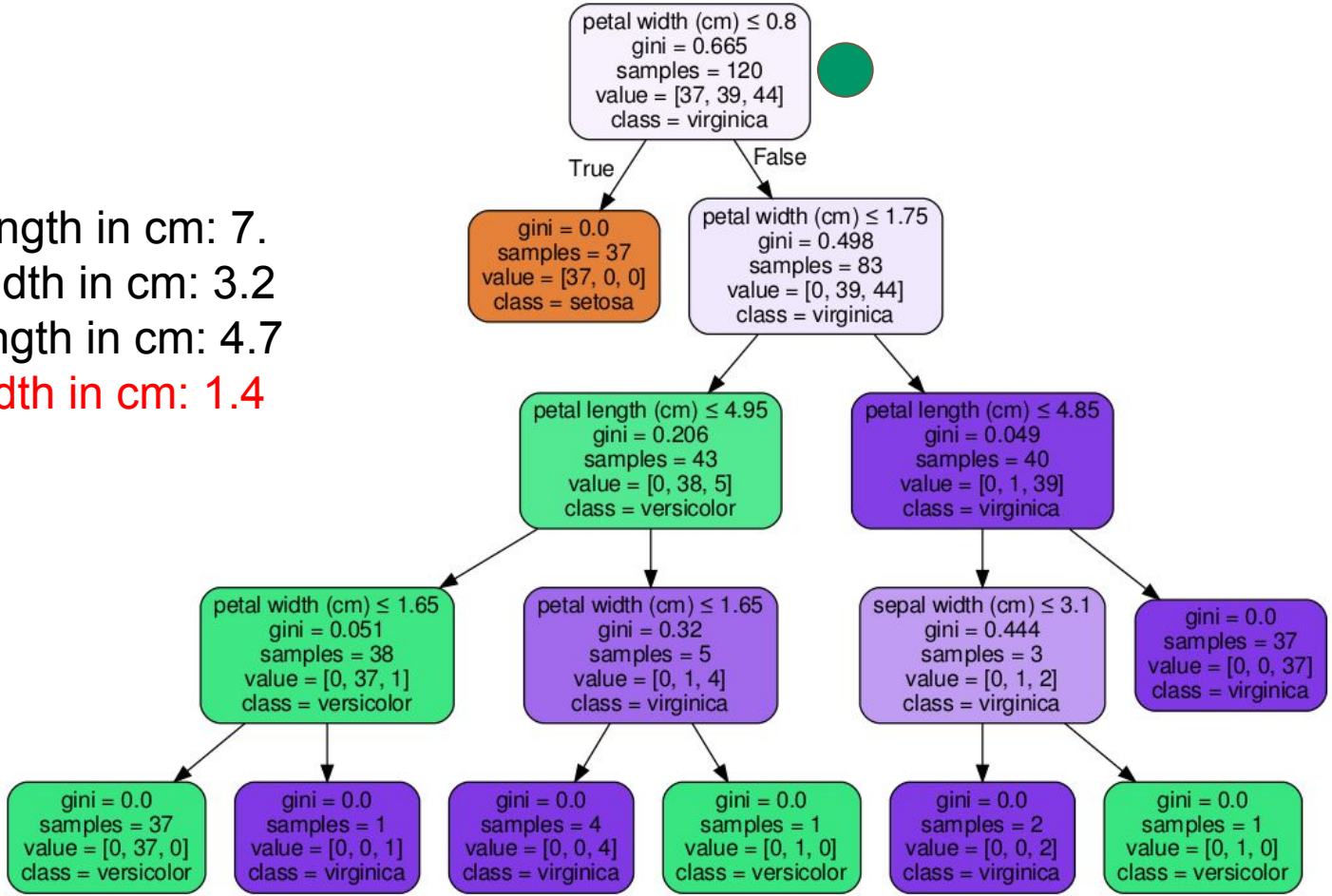

Matriz de confusão normalizada
setosa
versicolor
virginica
Classe Verdadeira
1.00
0.00
0.00
0.00
0.91
0.09
0.00
0.00
1.00
setosa
versicolor
virginica
Classe Obtida
1.0
0.8
0.6
0.4
0.2
0.0


Matriz de confusão não normalizada
setosa
versicolor
virginica
Classe Verdadeira
13
0
0
0
10
1
0
0
6
setosa
versicolor
virginica
Classe Obtida
12
10
8
6
4
2
0

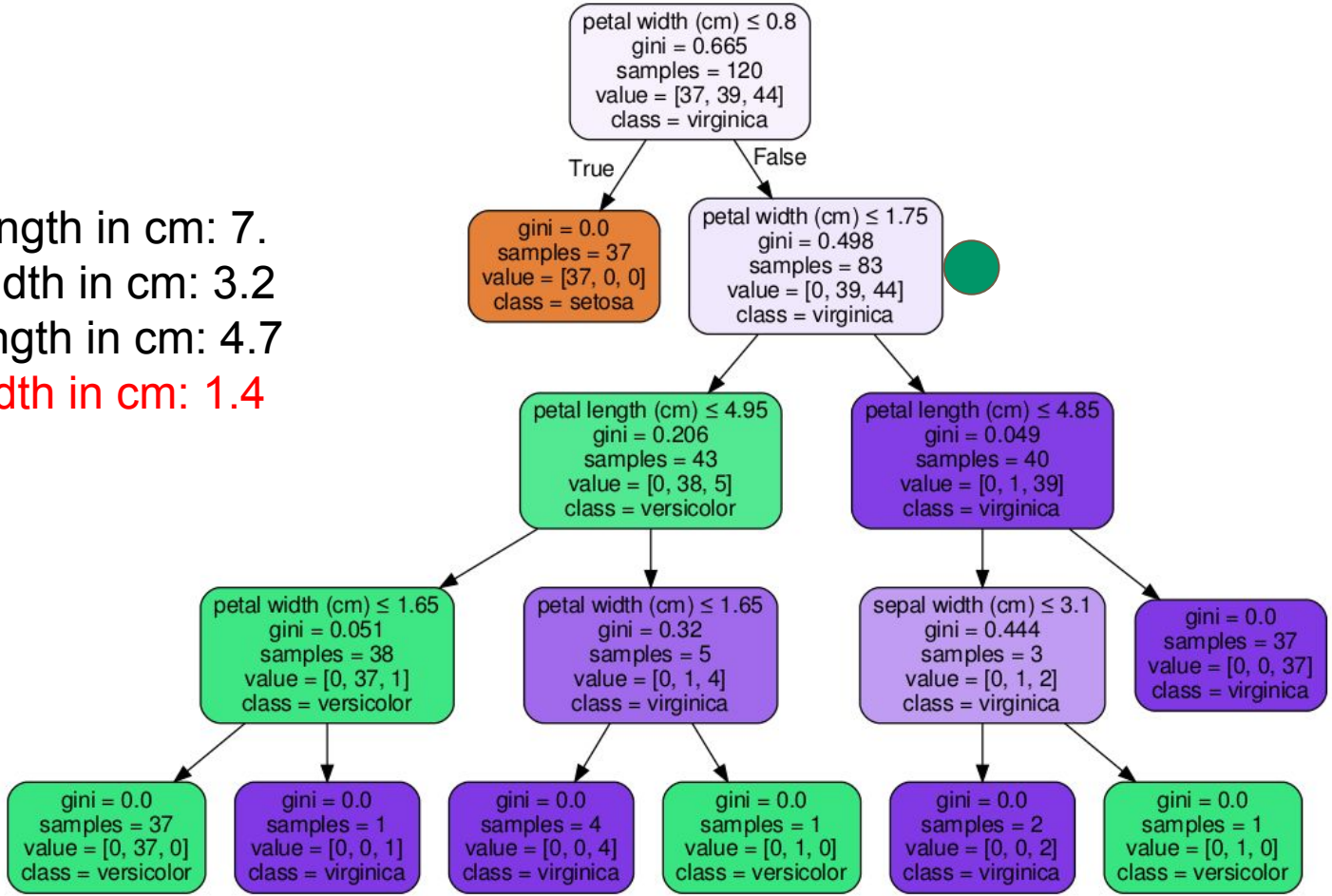

96%

.: Simulação

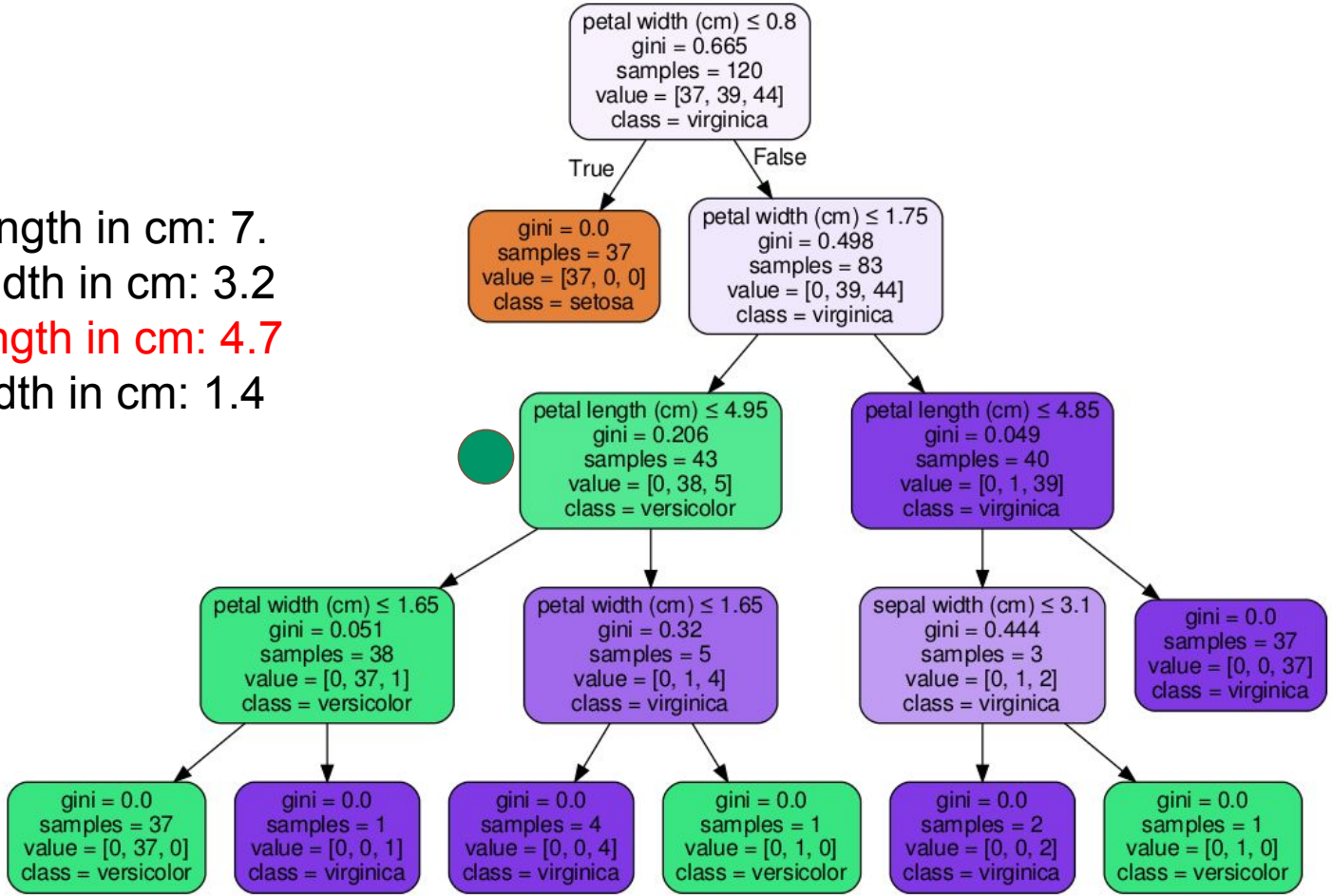
sepal length in cm: 7.
sepal width in cm: 3.2
petal length in cm: 4.7
petal width in cm: 1.4



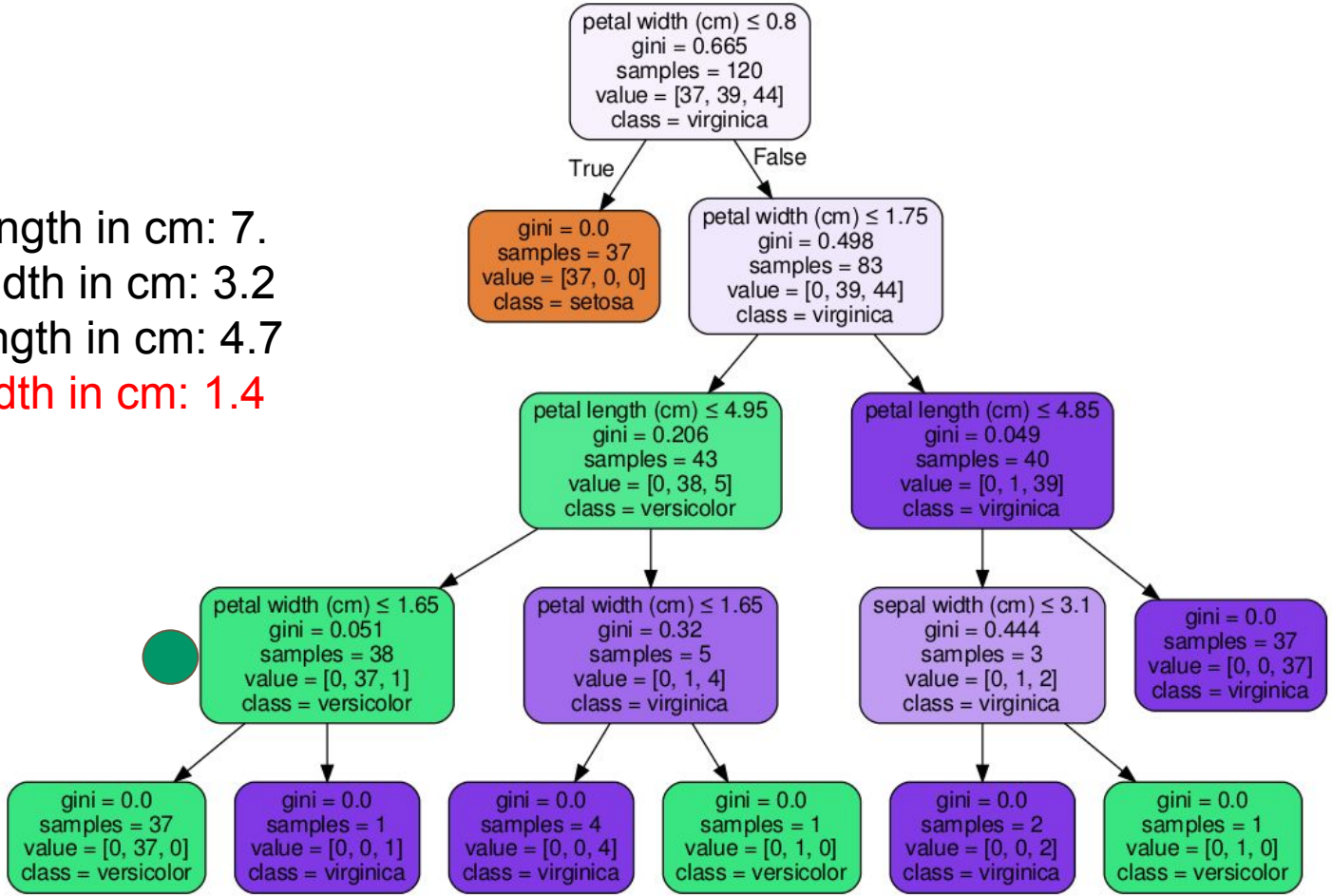
sepal length in cm: 7.
sepal width in cm: 3.2
petal length in cm: 4.7
petal width in cm: 1.4



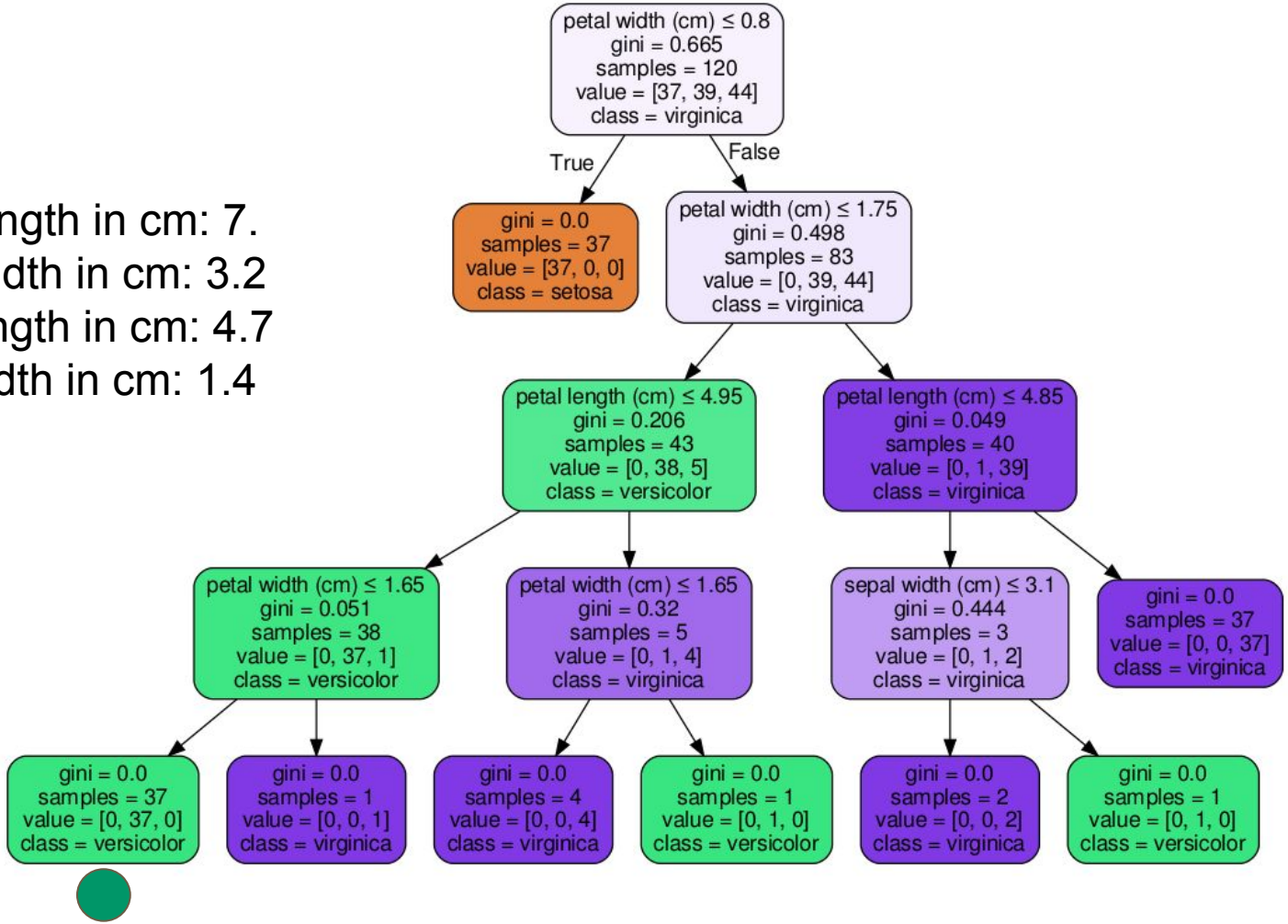
sepal length in cm: 7.
sepal width in cm: 3.2
petal length in cm: 4.7
petal width in cm: 1.4



sepal length in cm: 7.
sepal width in cm: 3.2
petal length in cm: 4.7
petal width in cm: 1.4

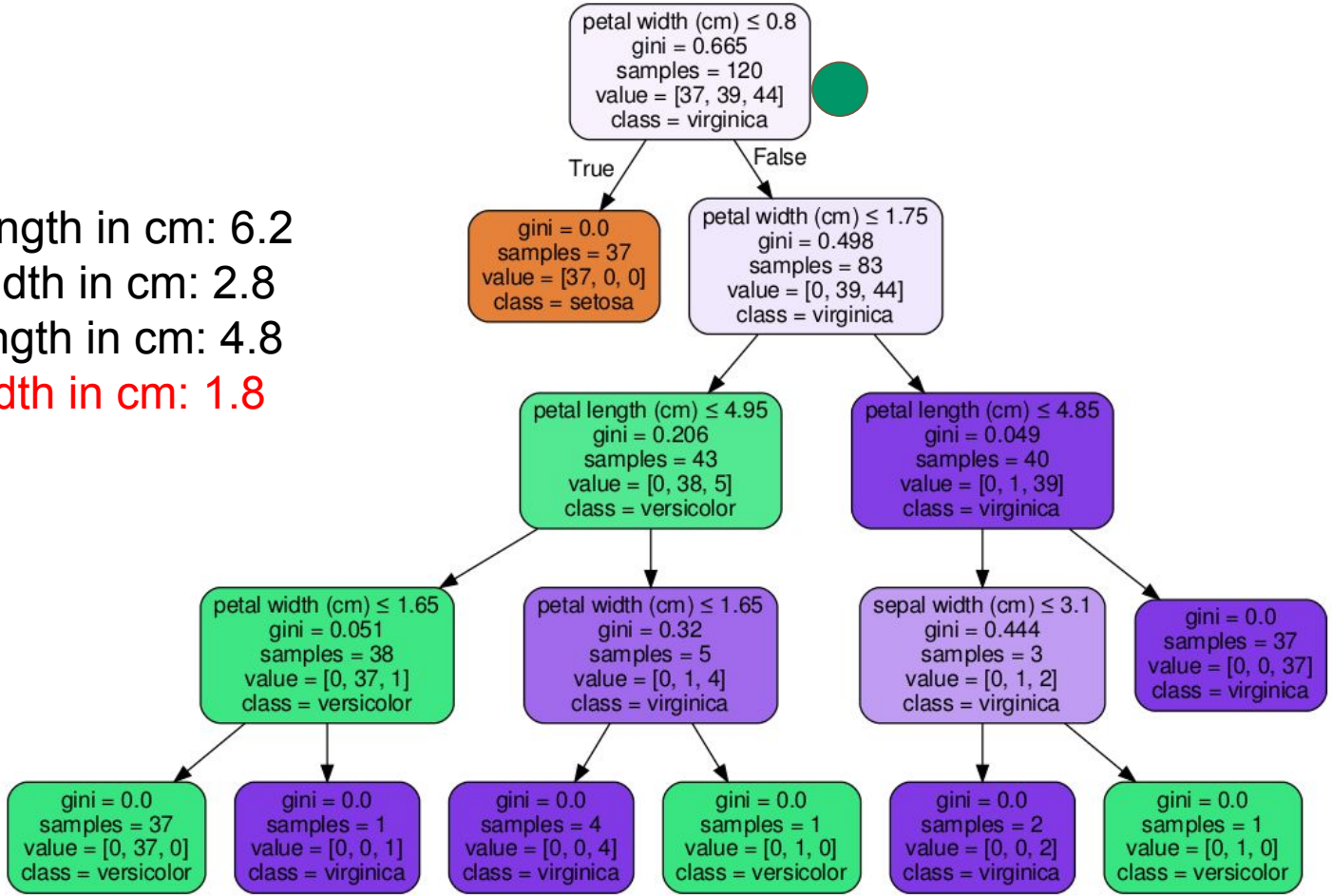


sepal length in cm: 7.
sepal width in cm: 3.2
petal length in cm: 4.7
petal width in cm: 1.4

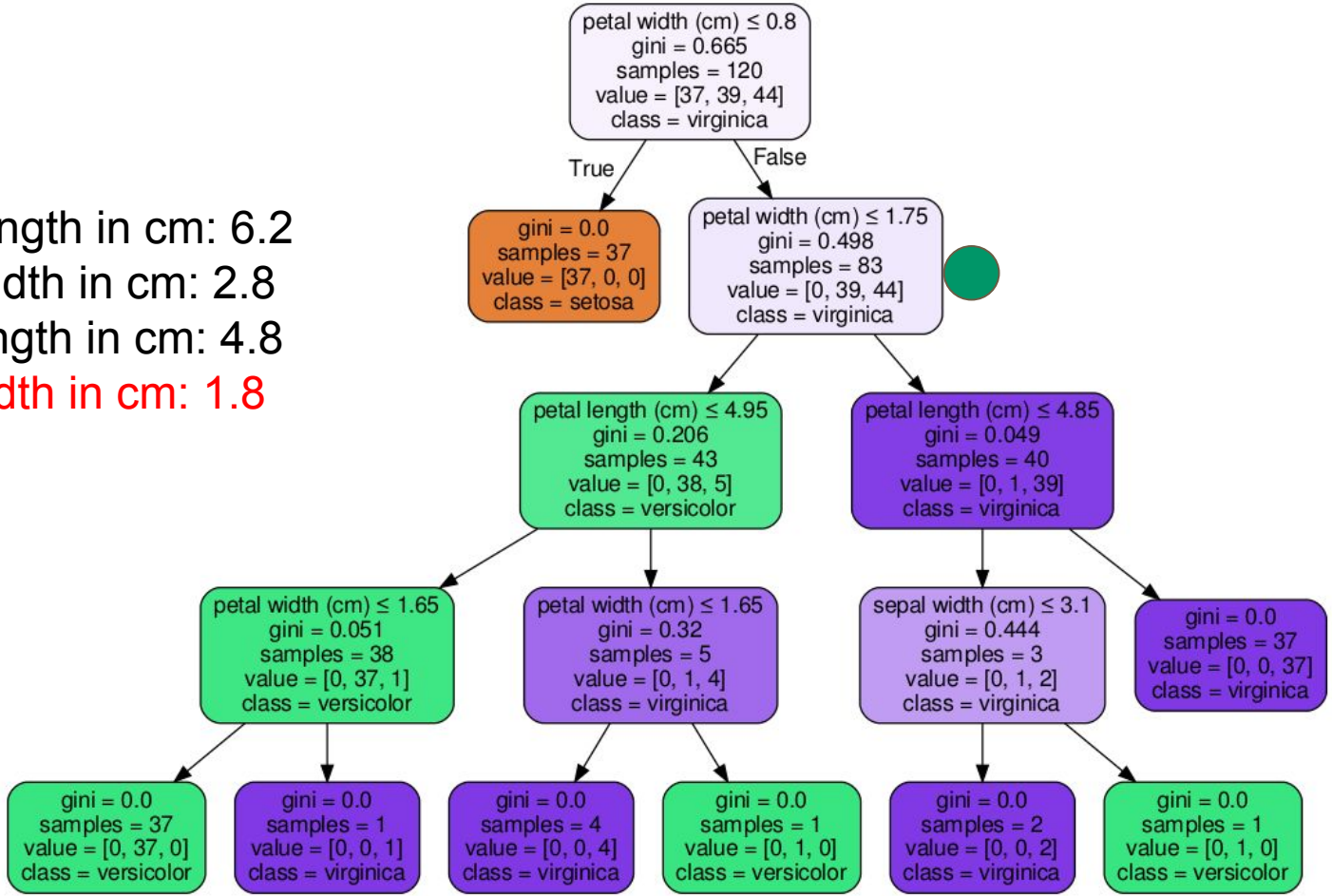


.: Simulação

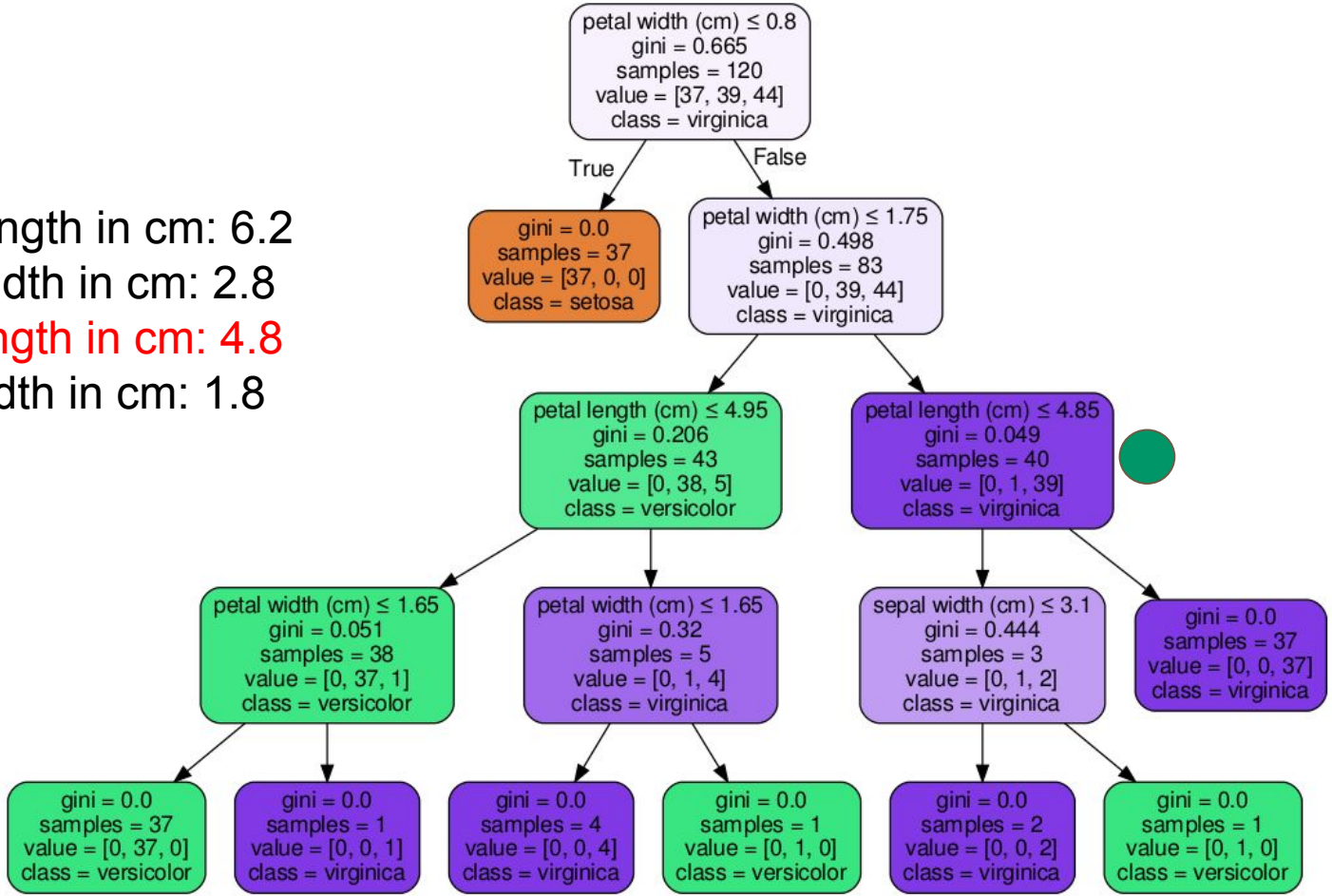
sepal length in cm: 6.2
sepal width in cm: 2.8
petal length in cm: 4.8
petal width in cm: 1.8



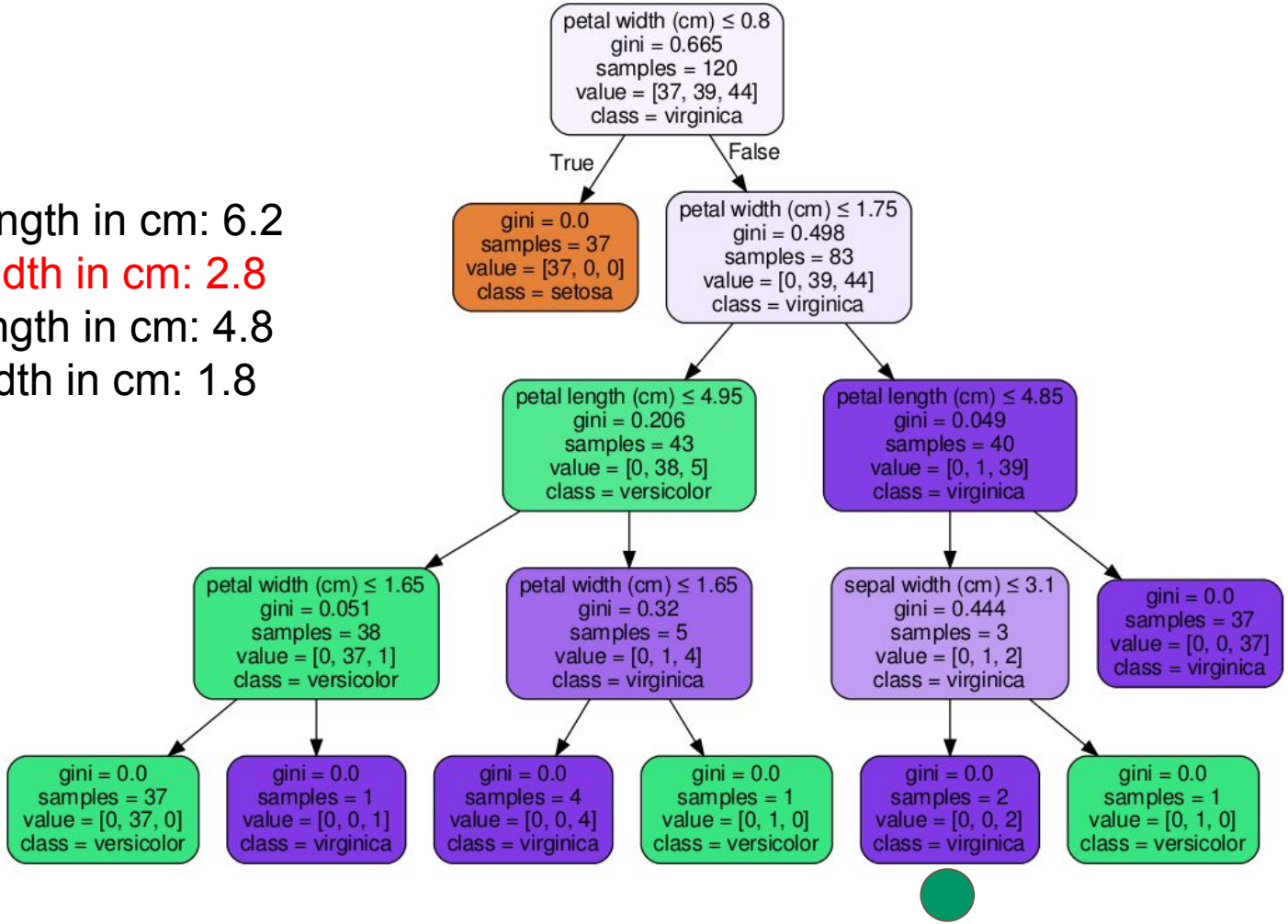
sepal length in cm: 6.2
sepal width in cm: 2.8
petal length in cm: 4.8
petal width in cm: 1.8



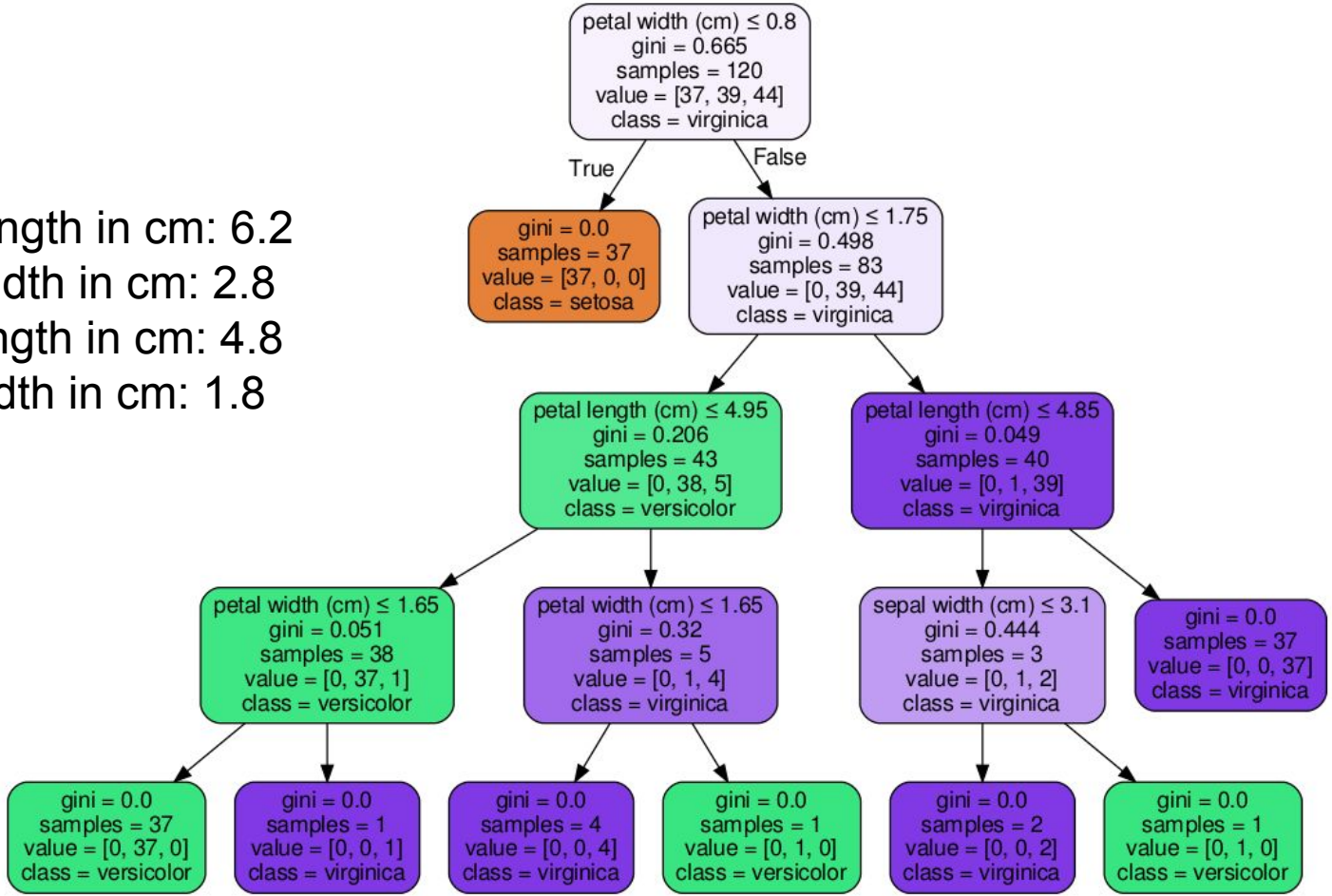
sepal length in cm: 6.2
sepal width in cm: 2.8
petal length in cm: 4.8
petal width in cm: 1.8



sepal length in cm: 6.2
sepal width in cm: 2.8
petal length in cm: 4.8
petal width in cm: 1.8



sepal length in cm: 6.2
sepal width in cm: 2.8
petal length in cm: 4.8
petal width in cm: 1.8



Dúvidas ?

.: Referências

https://medium.com/machine-learning-beyond-deep-learning/%C3%A1rvores-de-decis%C3%A3o-3f52f6420b69