

Machine Learning

Importação e Leitura dos dados

	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

Breve análise para compreender o conjunto de dados

INFORMAÇÕES DO DATASET:

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 150 entries, 0 to 149
```

```
Data columns (total 5 columns):
```

#	Column	Non-Null Count	Dtype
0	SepalLengthCm	150 non-null	float64
1	SepalWidthCm	150 non-null	float64
2	PetalLengthCm	150 non-null	float64
3	PetalWidthCm	150 non-null	float64
4	Class	150 non-null	object

```
dtypes: float64(4), object(1)
```

```
memory usage: 6.0+ KB
```

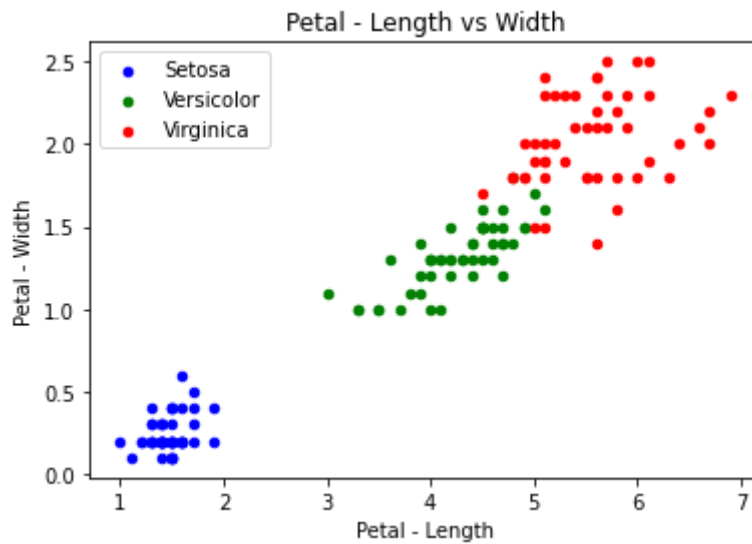
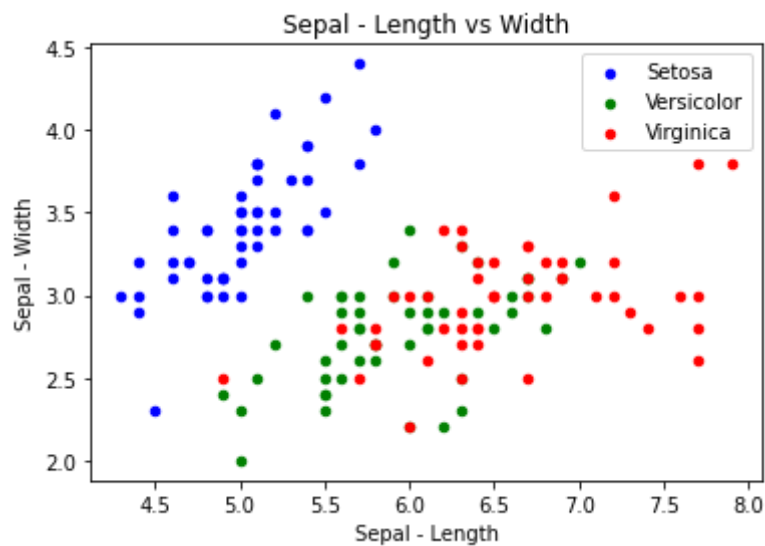
DESCRIÇÃO DO DATASET:

	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.054000	3.758667	1.198667
std	0.828066	0.433594	1.764420	0.763161
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

DISTRIBUIÇÃO DAS ESPÉCIES DE IRIS:

```
Class
Iris-setosa      50
Iris-versicolor  50
Iris-virginica   50
dtype: int64
```

Análise visual dos dados



Classificador - Modelo 1 - Todo o Conjunto de dados

Separando os dados e aplicando o algoritmo (DecisionTreeClassifier)

São 120 dados para treinamento e 30 dados para teste.

```
[ 'Iris-setosa' 'Iris-versicolor' 'Iris-setosa' 'Iris-virginica'
  'Iris-setosa' 'Iris-setosa' 'Iris-setosa' 'Iris-versicolor'
  'Iris-virginica' 'Iris-versicolor' 'Iris-versicolor' 'Iris-virgini
ca'
  'Iris-setosa' 'Iris-setosa' 'Iris-versicolor' 'Iris-virginica'
  'Iris-virginica' 'Iris-virginica' 'Iris-virginica' 'Iris-virginic
a'
  'Iris-versicolor' 'Iris-versicolor' 'Iris-virginica' 'Iris-virgini
ca'
  'Iris-setosa' 'Iris-setosa' 'Iris-versicolor' 'Iris-setosa'
  'Iris-versicolor' 'Iris-setosa']
```

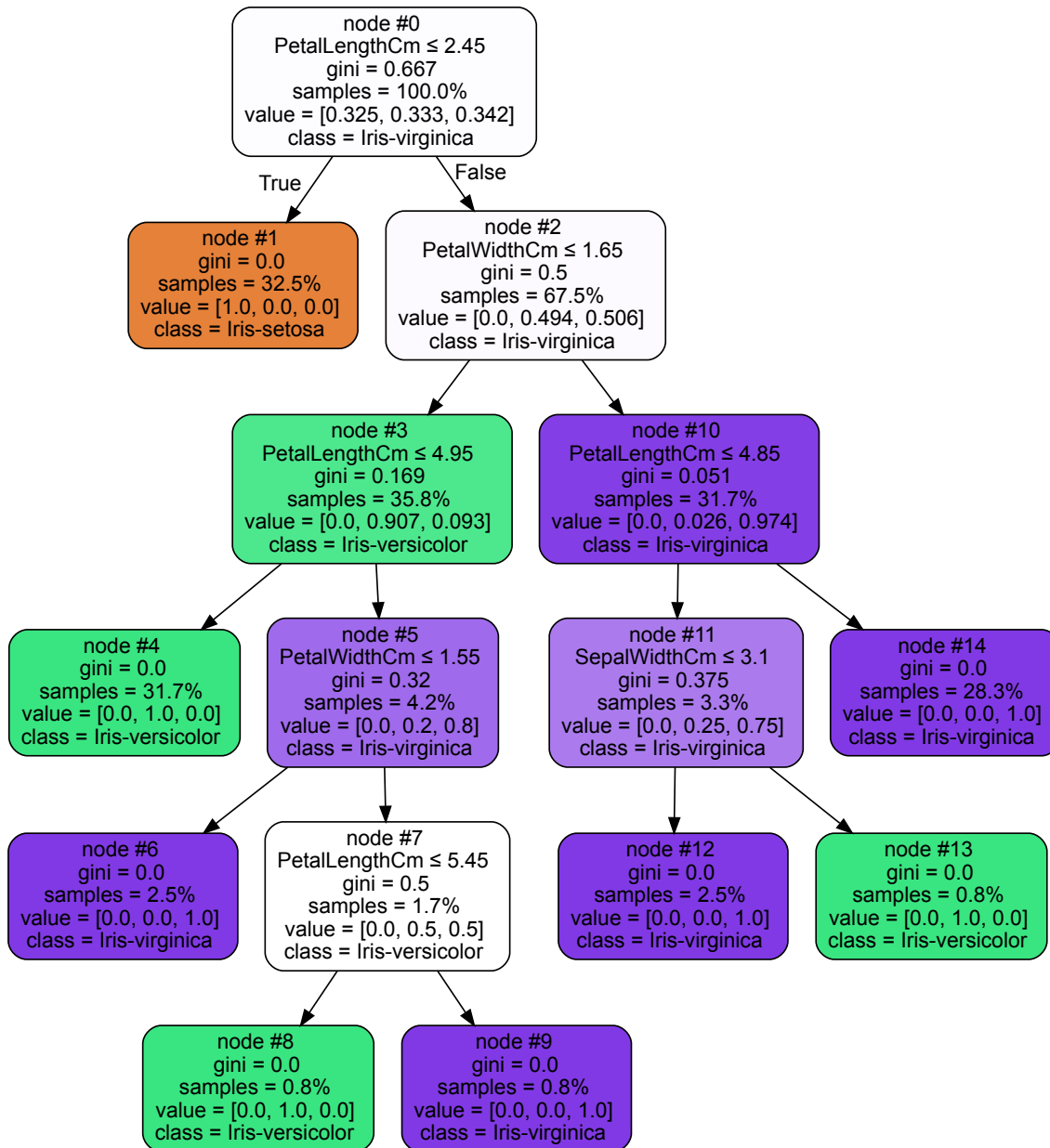
Análisis de performance do algoritmo

Precisão do modelo: 100.0

SepalLengthCm: 0.0
 SepalWidthCm: 0.01875390706397166
 PetalLengthCm: 0.5825299672521005
 PetalWidthCm: 0.3987161256839277

	precision	recall	f1-score	support
Iris-setosa	1.00	1.00	1.00	11
Iris-versicolor	1.00	0.90	0.95	10
Iris-virginica	0.90	1.00	0.95	9
accuracy			0.97	30
macro avg	0.97	0.97	0.96	30
weighted avg	0.97	0.97	0.97	30

Renderizando a árvore graficamente



Navegando na estrutura da árvore de decisão

Número de nós: 9

A árvore binária tem 9 nós e a seguinte estrutura:

```
nó = 0 (nó teste): vai para o nó 1 se o valor do atributo PetalWidthCm <= 0.800000011920929
se não, vai para o nó 2.
    nó = 1 (nó folha).
    nó = 2 (nó teste): vai para o nó 3 se o valor do atributo PetalWidthCm <= 1.75
se não, vai para o nó 6.
    nó = 3 (nó teste): vai para o nó 4 se o valor do atributo PetalLengthCm <= 4.950000047683716
se não, vai para o nó 5.
        nó = 4 (nó folha).
        nó = 5 (nó folha).
    nó = 6 (nó teste): vai para o nó 7 se o valor do atributo PetalLengthCm <= 4.8500001430511475
se não, vai para o nó 8.
        nó = 7 (nó folha).
        nó = 8 (nó folha).
```

Classificador - Modelo 2 - Dados correspondentes às sépalas

Preparação dos dados

São 120 dados das sépalas para treinamento e 30 dados para teste.

Aplicando o algoritmo

```
['Iris-virginica' 'Iris-virginica' 'Iris-setosa' 'Iris-virginica'
'Iris-virginica' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor'
'Iris-versicolor' 'Iris-versicolor' 'Iris-virginica' 'Iris-setosa'
'Iris-versicolor' 'Iris-virginica' 'Iris-virginica' 'Iris-versicolor'
'Iris-virginica' 'Iris-versicolor' 'Iris-versicolor' 'Iris-versicolor'
'Iris-versicolor' 'Iris-setosa' 'Iris-versicolor' 'Iris-setosa'
'Iris-versicolor' 'Iris-versicolor' 'Iris-setosa' 'Iris-virginica'
'Iris-setosa' 'Iris-setosa']
```

Precisão do modelo: 94.16666666666667

Classificador - Modelo 3 - Dados correspondentes às pétalas

Preparação dos dados

São 120 dados das pétalas para treinamento e 30 dados para teste.

Aplicando o algoritmo

```
['Iris-virginica' 'Iris-versicolor' 'Iris-virginica' 'Iris-versicol  
or'  
 'Iris-versicolor' 'Iris-virginica' 'Iris-versicolor' 'Iris-setosa'  
 'Iris-versicolor' 'Iris-setosa' 'Iris-virginica' 'Iris-virginica'  
 'Iris-versicolor' 'Iris-setosa' 'Iris-versicolor' 'Iris-versicolo  
r'  
 'Iris-setosa' 'Iris-virginica' 'Iris-setosa' 'Iris-setosa' 'Iris-s  
etosa'  
 'Iris-setosa' 'Iris-versicolor' 'Iris-virginica' 'Iris-versicolor'  
 'Iris-setosa' 'Iris-virginica' 'Iris-virginica' 'Iris-virginica'  
 'Iris-setosa']
```

Precisão do modelo: 99.16666666666667