Processing data “Iris” dengan metode klasifikasi menggunakan algoritma C5.0

1. **Pengaturan lokasi directory**

```
setwd("E:/Tugas_DM/Tugas_kelompok")
```

```
> setwd("E:/Tugas_DM/Tugas_kelompok")
```

2. **baca data**

```
dataset <- read.csv("iris.csv", sep = ";")
```

```
> dataset <- read.csv("iris.csv", sep = ";")
```

3. **Pembuatan model decision tree menggunakan algoritman C5.0**

```
model <- C5.0(Species ~., data=dataset)
```

```
> model <- C5.0(Species ~., data=dataset)
```

4. **Melihat model**

```
model
```

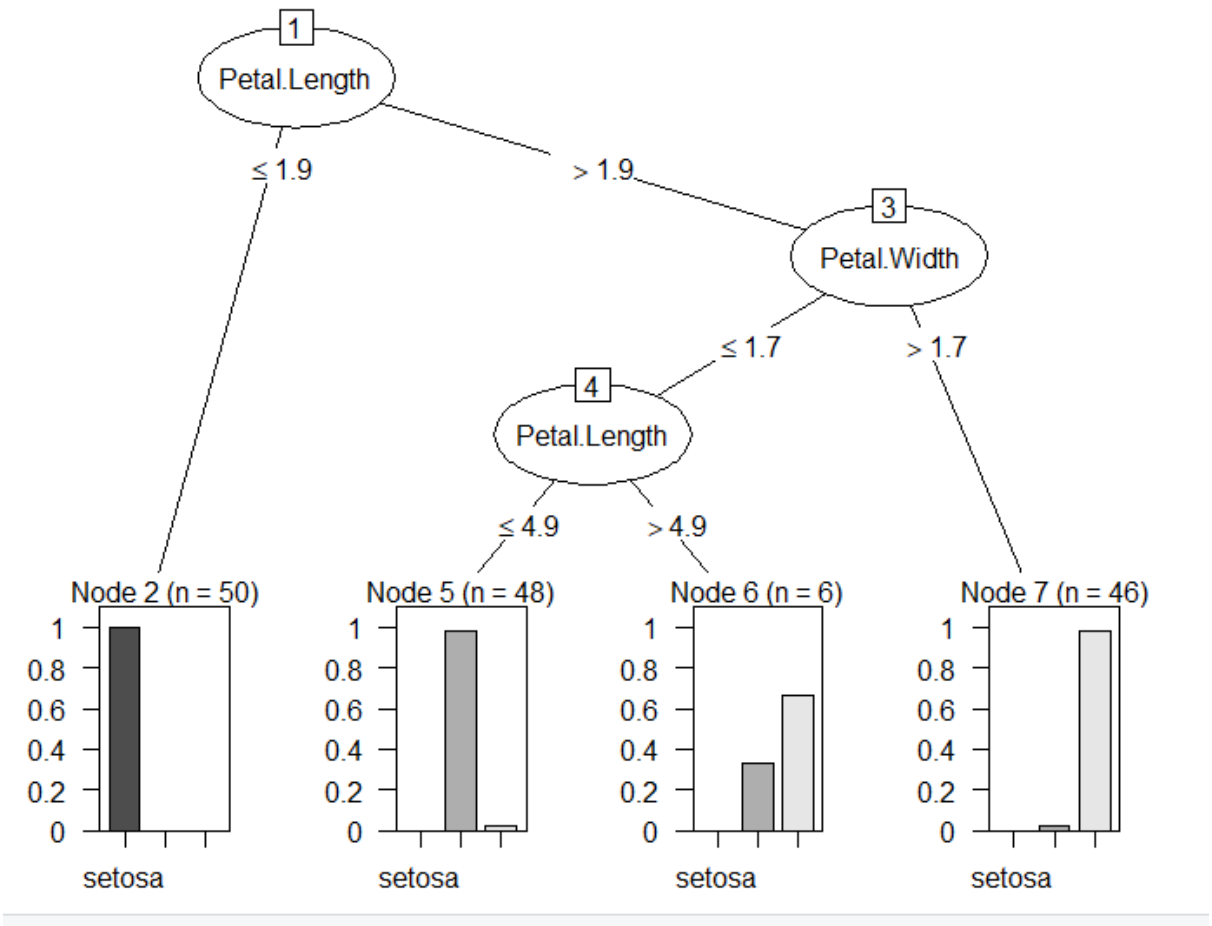
```
Call:
C5.0.formula(formula = Species ~ ., data = dataset)

Classification Tree
Number of samples: 150
Number of predictors: 4

Tree size: 4

Non-standard options: attempt to group attributes
```

```
summary(model)
```

```
> summary(model)

Call:
C5.0.formula(formula = Species ~ ., data = dataset)


C5.0 [Release 2.07 GPL Edition]         Sun Nov 24 15:29:31 2019
-------------------------------

Class specified by attribute `outcome'

Read 150 cases (5 attributes) from undefined.data

Decision tree:

Petal.Length <= 1.9: setosa (50)
Petal.Length > 1.9:
:...Petal.Width > 1.7: virginica (46/1)
    Petal.Width <= 1.7:
    :...Petal.Length <= 4.9: versicolor (48/1)
        Petal.Length > 4.9: virginica (6/2)


Evaluation on training data (150 cases):

           Decision Tree
         ----------------
         Size      Errors

            4    4( 2.7%)   <<


          (a)   (b)   (c)    <-classified as
         ----  ----  ----
           50                (a): class setosa
                 47     3    (b): class versicolor
                  1    49    (c): class virginica


        Attribute usage:

        100.00% Petal.Length
         66.67% Petal.Width
```

5. **Menampilkan pohon dari model yang sudah dibangun**

```
plot(model)
```



6. **Menjadikan dataset, sebagai data testing. Namun hanya kolom 1,2,3,4 saja, dan tanpa label**

```
datatesting <- dataset[,1:4]
```

```
datatesting <- dataset[,1:4]
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
1	5.1	3.5	1.4	0.2
2	4.9	3.0	1.4	0.2
3	4.7	3.2	1.3	0.2
4	4.6	3.1	1.5	0.2
5	5.0	3.6	1.4	0.2
6	5.4	3.9	1.7	0.4
7	4.6	3.4	1.4	0.3
8	5.0	3.4	1.5	0.2
9	4.4	2.9	1.4	0.2
10	4.9	3.1	1.5	0.1
11	5.4	3.7	1.5	0.2
12	4.8	3.4	1.6	0.2
13	4.8	3.0	1.4	0.1
14	4.3	3.0	1.1	0.1
15	5.8	4.0	1.2	0.2

7. **Membandingkan hasil prediksi dengan dataset**

```
table(predictions, dataset$Species)
```

```
> table(predictions, dataset$Species)

predictions  setosa versicolor virginica
  setosa         50          0         0
  versicolor      0         47         1
  virginica       0          3        49
>
```