

SPRAWOZDANIE

Zajęcia: Analiza Procesów Ucznia

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium: 5.1

Data: 29.05.2020

Temat: „Modelowanie procesów uczenia maszynowego w pakiecie MLR
Trenowanie, ocena i porównanie modeli w pakiecie MLR”

Wariant: 1, zadanie 1

Jacek Adamczyk

Informatyka II stopień,

Stacjonarne,

1 semestr,

Gr. A

<https://github.com/jacekaGIT/ATH-1g>

1. Polecenia dla wariantu 1 (Zadanie 1):

Celem ćwiczenia jest zbudowanie drzewa decyzyjnego na podstawie danych ze zbioru IRIS który zawiera rozmiary płatków kwiatów dla 150 egzemplarzy 3 gatunków irysów. Uzyskane drzewo decyzyjne umożliwia rozpoznawanie poszczególnych gatunków na podstawie rozmiarów płatków.

Polecenie:

Zadanie 1. Zadanie dotyczy konstruowania drzew decyzyjnych oraz reguł klasyfikacyjnych na podstawie zbioru danych (library(MASS lub datasets)).
Warianty zadania

1. iris

2. Skrypt:

Po uruchomieniu skryptu zostaną wykonane kolejno wszystkie polecenia zadania. Instalacja pakietów „C5.0” i „MASS” została wyłączona (ustawiona jako komentarz) żeby niepotrzebnie nie instalować pakietów przy każdym uruchomieniu skryptu.

```
#..... Jacek Adamczyk, sem 1, II st .....
#..... Lab 5, gr A, wariant 1 .....
# Modelowanie procesów uczenia maszynowego ...
# w pakiecie MLR. Trenowanie, ocena .....
# i porównywanie modeli w pakiecie MLR .....
#.....

# Zadanie 1

#install.packages("C50")
#install.packages("MASS")
library("C50")
library("MASS")

View(iris)

data(iris)
head(iris)
```

3. Wyniki działania:

Poniżej przedstawiono wyniki działania skryptu. Pełne wydruki z konsoli można znaleźć w repozytorium GitHub (adres na stronie tytułowej).

- Podsumowanie z konstrukcji drzewa:

```
> summary(drzewo)
```

```
Call:
```

```
C5.0.default(x = iris[, -5], y = iris$Species)
```

```
C5.0 [Release 2.07 GPL Edition]      Fri May 29 09:36:54 2020
```

```
-----  
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
```

```
      47
```

```
      3
```

```
    (a): class setosa
```

```
    (b): class versicolor
```

```
    (c): class virginica
```

```
      1
```

```
      49
```

```
Attribute usage:
```

```
100.00%Petal.Length
```

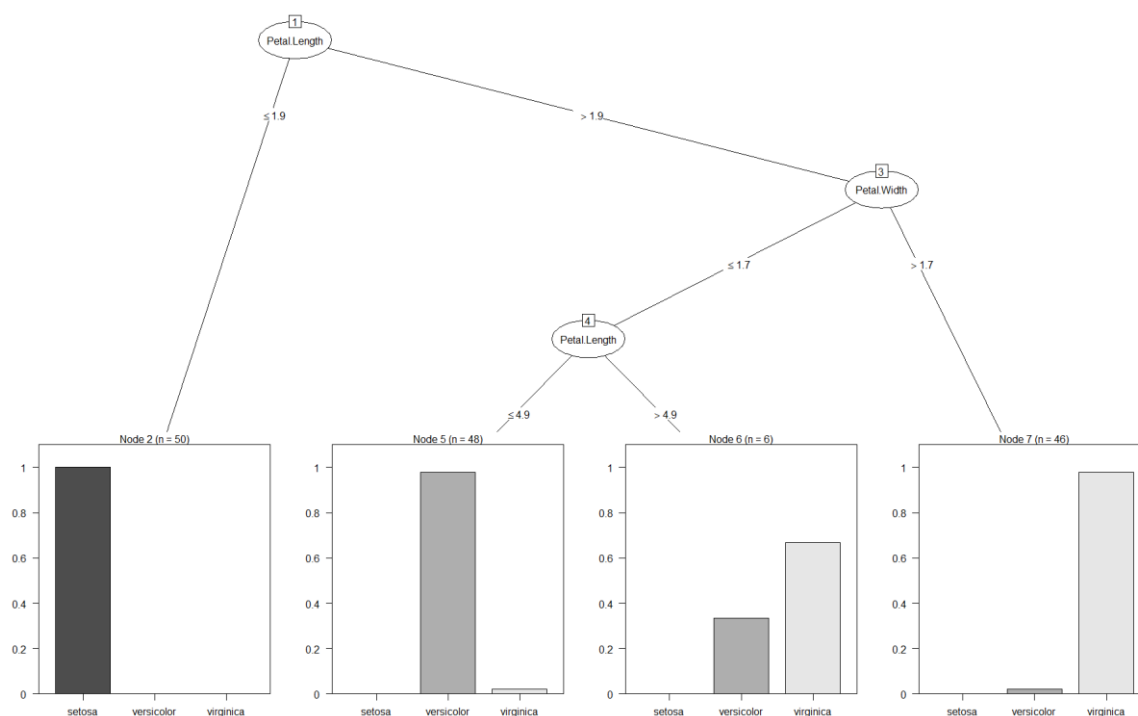
```
66.67%Petal.Width
```

```
Time: 0.0 secs
```

| | | | | | |
|----|-----|-----|-----|-----|------------|
| 47 | 5.1 | 3.8 | 1.6 | 0.2 | setosa |
| 48 | 4.6 | 3.2 | 1.4 | 0.2 | setosa |
| 49 | 5.3 | 3.7 | 1.5 | 0.2 | setosa |
| 50 | 5.0 | 3.3 | 1.4 | 0.2 | setosa |
| 51 | 7.0 | 3.2 | 4.7 | 1.4 | versicolor |
| 52 | 6.4 | 3.2 | 4.5 | 1.5 | versicolor |
| 53 | 6.9 | 3.1 | 4.9 | 1.5 | versicolor |

Rysunek 2. Fragment zbioru IRIS

- Uzyskane drzewo decyzyjne:



Rysunek 1. Drzewo decyzyjne

4. Wnioski:

Mając nieznaną egzemplarz irysa możemy na podstawie rozmiarów płatków i drzewa decyzyjnego określić z dużym prawdopodobieństwem jego gatunek. Kłopotliwe będzie jedynie określanie gatunku w węźle 6. Przy długości płatka $> 4,9$ istnieje duże prawdopodobieństwo pomyłki pomiędzy gatunkami Versicolor i Virginica.