

# **SPRAWOZDANIE**

Zajęcia: Analiza Procesów Ucznienia

Prowadzący: prof. dr hab. Vasyl Martsenyuk

## **Laboratorium 4**

05.01.2022

**Temat: Modelowanie procesów uczenia maszynowego w pakiecie mlr.  
Trenowanie, ocena i porównywanie modeli w pakiecie mlr**

**Wariant 7**

Smolec Bartłomiej

Informatyka II stopień  
niestacjonarne (zaoczne)

1 semestr

## 1 . Polecenie

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

1. iris
2. infert
3. mtcars
4. Aids2
5. bacteria
6. biopsy
7. cats
8. genotype
9. shuttle
10. Pima.tr2
11. OME
12. Melanoma

## 2. Rozwiązanie

Link do github: <https://github.com/barteksmolec12/APU4>

*Konfiguracja bibliotek*

```
> library(MASS)
> require(C50)
Ładowanie wymaganego pakietu: C50
Warning message:
pakiet 'C50' został zbudowany w wersji R 4.1.2
```

*Ładowanie danych (zbiór danych „cats”)*

```
> data(cats)
> str(cats)
'data.frame': 144 obs. of 3 variables:
 $ Sex: Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 1 ...
 $ Bwt: num 2 2 2 2.1 2.1 2.1 2.1 2.1 2.1 2.1 ...
 $ Hwt: num 7 7.4 9.5 7.2 7.3 7.6 8.1 8.2 8.3 8.5 ...
>
```

	Sex	Bwt	Hwt
1	F	2.0	7
2	F	2.0	7.4
3	F	2.0	9.5
4	F	2.1	7.2

## Operacje na danych

```
> cats[, 'Hwt'] <- factor(cats[, 'Hwt'])
> str(cats)
'data.frame': 144 obs. of 3 variables:
 $ Sex: Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 ...
 $ Bwt: num 2 2 2 2.1 2.1 2.1 2.1 2.1 2.1 2.1 ...
 $ Hwt: Factor w/ 73 levels "6.3","6.5","7",...: 3 7 24 5 6 8 12 13 14 16 ...
> table(cats$Hwt)

6.3 6.5 7 7.1 7.2 7.3 7.4 7.6 7.7 7.9 8 8.1 8.2 8.3 8.4 8.5 8.6 8.7 8.8 9 9.1 9.3 9.4
9.5 9.6 9.7
1 2 1 1 2 3 1 2 1 5 1 1 1 2 1 3 1 4 3 4 4 2 3
2 6 2
9.8 9.9 10 10.1 10.2 10.4 10.5 10.6 10.7 10.8 10.9 11 11.1 11.2 11.3 11.4 11.5 11.6 11.7 11.8 11.9 12 12.1 1
2.2 12.3 12.4
2 2 2 8 3 2 1 4 1 1 2 3 1 2 1 1 3 2 1 2 1 3 1
2 1 2
12.5 12.7 12.8 12.9 13 13.3 13.5 13.6 13.8 14.1 14.3 14.4 14.8 14.9 15 15.4 15.6 15.7 16.8 17.2 20.5
2 3 1 1 3 3 2 1 1 1 1 2 2 1 1 1 1 1 1 1 1
> m1 <- C5.0(cats[1:144, -3], cats[1:144, 3])
> summary(m1)
```

## Utworzone zostało drzewo decyzyjne

```
Call:
C5.0.default(x = cats[1:144, -3], y = cats[1:144, 3])

C5.0 [Release 2.07 GPL Edition] Tue Feb 08 20:54:19 2022

-----

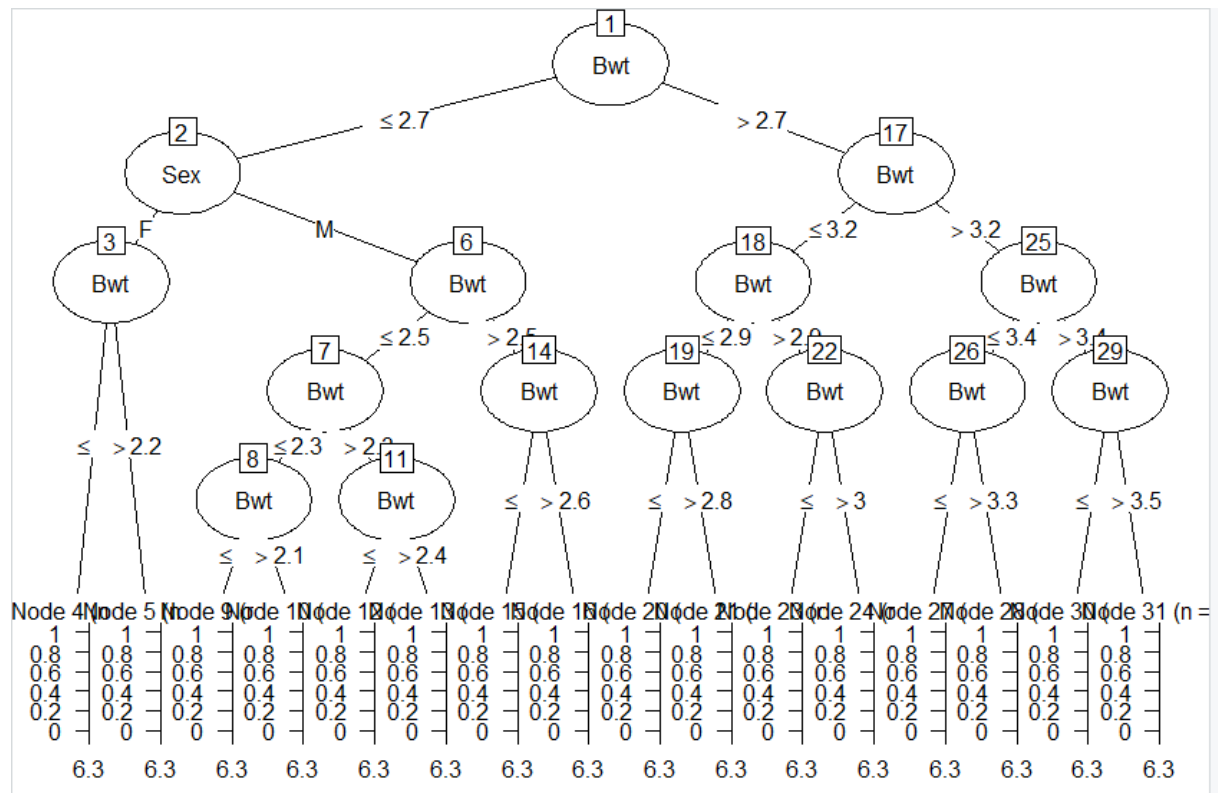
Class specified by attribute `outcome'

Read 144 cases (3 attributes) from undefined.data

Decision tree:

Bwt > 2.7:
...Bwt <= 3.2:
: ...Bwt <= 2.9:
: : ...Bwt <= 2.8: 9.1 (7/6)
: : : Bwt > 2.8: 10.1 (8/5)
: : Bwt > 2.9:
: : ...Bwt <= 3: 10.6 (11/9)
: : : Bwt > 3: 13 (12/10)
: Bwt > 3.2:
: ...Bwt <= 3.4:
: : ...Bwt <= 3.3: 11.5 (5/4)
: : : Bwt > 3.3: 11.2 (5/4)
: : Bwt > 3.4:
: : ...Bwt <= 3.5: 11.7 (5/4)
: : : Bwt > 3.5: 14.8 (9/7)
Bwt <= 2.7:
...Sex = F:
...Bwt <= 2.2: 8.7 (18/16)
: Bwt > 2.2: 10.1 (24/20)
```

### Wizualizacja w formie graficznej



### 3. Wnioski

Utworzone zadanie pozwoliło mi zapoznać się z tematyką tworzenia drzew decyzyjnych. Operowałem tutaj na zbiorze danych „cats”. Analizowanymi danymi były: płeć oraz waga ciała i serca.