

SPRAWOZDANIE

Zajęcia: Analiza Procesów Ucznienia

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium 7

Data 30.05.2019r.

Temat: "Uczenie maszynowe z użyciem
drzew decyzyjnych"

Wariant 4

Artur Kozieł
Informatyka II stopień,
stacjonarne
1 semestr

1. Polecenie:

Celem ćwiczenia było nauczenie się obsługi biblioteki C5.0, a także stworzenie drzewa decyzyjnego do prognozowania oceny klientów. Obiektem oceny były lodówki.

Wariant 4 - Lodówka. Uwzględniamy następujące dane: pojemność użytkową chłodziarki, pojemność użytkową zamrażarki oraz jakieś inne parametry. Dane (10 lodówek) pobrać ze strony <http://www.euro.com.pl>

2. Wprowadzane dane:

Dane lodówek (pojemność lodówki, pojemność zamrażarki, wysokość, poziom hałasu oraz ocena użytkowników) zostały pobrane ze strony www.euro.com.pl.

3. Wykorzystane komendy:

```
library("C50") # dołączenie pakietu C.50
setwd("C:/Users/Bernadka/Desktop/APU Artur/4") # ustawienie katalogu
dane <- read.csv('fridges.csv') # załadowanie danych z pliku .csv
dane$Ocena <- factor(dane$Ocena) #zamienienie kolumny decyzyjnej w faktor
head(dane) # wypisanie kilku pierwszych danych
treeModel <- C5.0(x=dane[, -6], y=dane$Ocena) # zbudowanie drzewa
# decyzyjnego z użyciem kolumny Ocena jako danej wyjściowej
treeModel # wypisanie drzewa decyzyjnego do konsoli
summary(treeModel) # podsumowanie uczenia
plot(treeModel) # wyrysowanie drzewa do wykresu
```

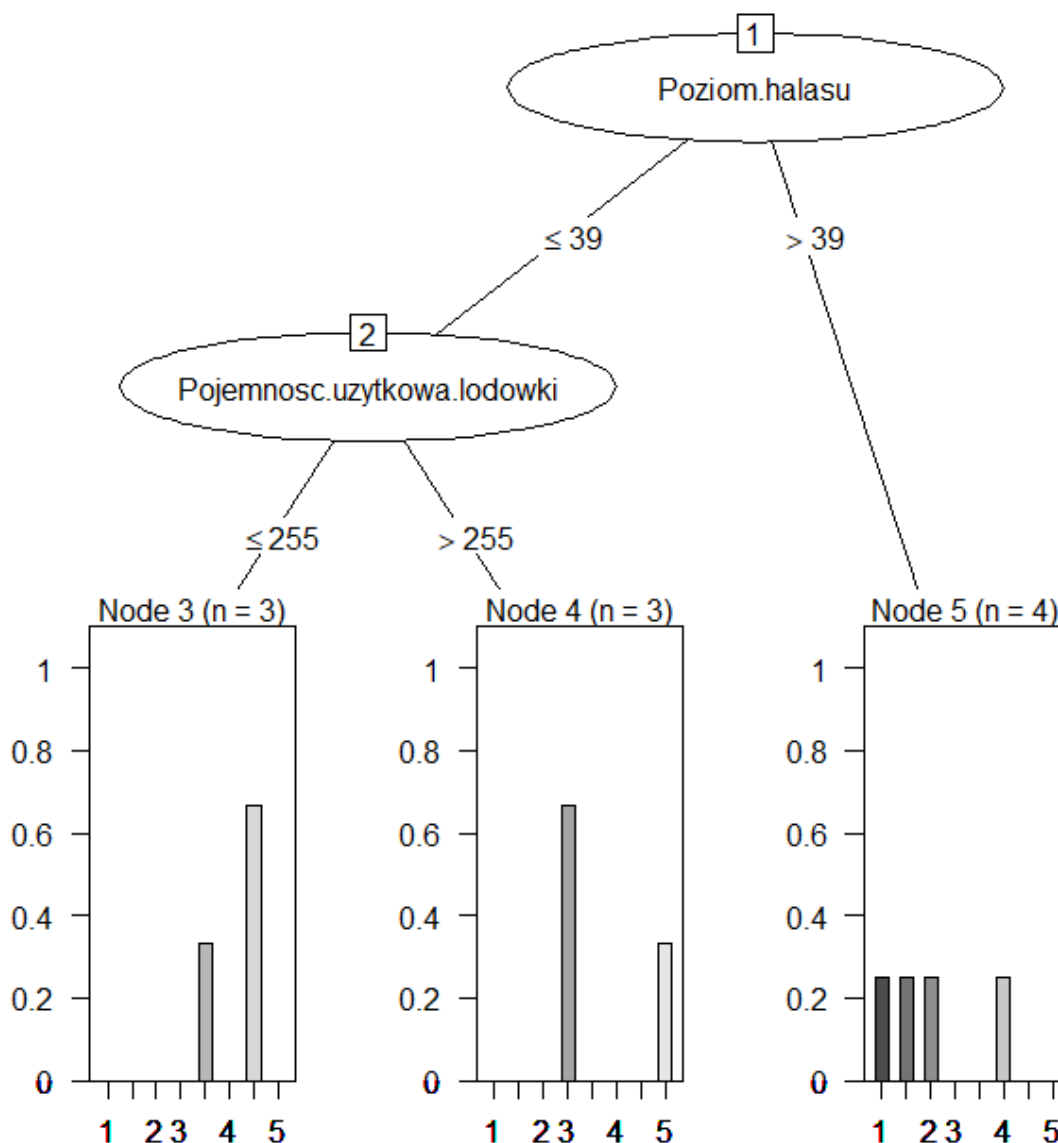
4. Wynik działania:

Kod programu dostępny w repozytorium: <https://github.com/arturkoziel/APU>

Przygotowane dane:

Model	Pojemnosc uzytkowa lodowki	Pojemnosc uzytkowa zamrazarki	Wysokosc	Poziom halasu	Ocena
LG GSL761PZU	405	196	179	39	3
Amica FD2325.4XI	160	44	144	40	2
Samsung RB37J501MB1	255	98	201	39	4.5
Samsung RB33N341MSS	217	98	185	38	3.5
Samsung RS50N3913BC	357	144	179	43	4
LG Premium GBB59PZRZS	225	75	190	37	4.5
Samsung Chef Collection RB41J7839S4	276	130	201	39	5
Bosch KGN39MB3A	279	87	203	39	3
Amica FK205.4	100	56	144	40	1
Amica FK2695.4FTHAA	161	70	170	42	1.5

Z przygotowanych danych utworzone zostało drzewo decyzyjne, które przedstawia poniższy schemat:



5. Wnioski:

Na otrzymanym wykresie można stwierdzić, że bazową wartością, która decyduje o wyborze łódki jest poziom hałasu. Jeżeli jest on wyższy niż 39 to łódki te otrzymują po równo oceny 1, 1,5, 2 i 4. Jeżeli poziom hałasu jest mniejszy niż 39 to ważnym czynnikiem staje się pojemność użytkowa łódki. Jeśli jest mniejsza niż 255 litrów to wybór otrzymuje ocenę 4,5 lub w mniejszym stopniu 3,5. Gdy pojemność jest wyższa od 255 litrów to najwyższą ocenę jest 3 oraz w mniejszym stopniu 5.