SPRAWOZDANIE

Zajęcia: Analiza procesów uczenia

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium 4 14 czerwca 2020

Temat: Uczenie maszynowe z użyciem drzew decyzyjnych

Wariant: 1

Adres repozytorium: https://github.com/Konradbor/APU/tree/master/4

1. Polecenie:

Zadanie dotyczy prognozowania oceny klientów (w skali 5-punktowej, Error < 5%) urzadzeń RTV AGD. Używajac metody indukcji drzewa decyzji C5.0 opracować plik w jezyku R z wykorzystaniem paczki C50.

1. Smartfon Samsung z systemem Android 6, ośmiordzieniowym procesorem. Uwzglendniamy nastepijace dane: wyświetlacz, pamieć RAM, pamieć wbudowana, aparat foto. Komunikacja (wifi itp) nie ma znaczenia. Dane (>10 smartfonów) pobrać ze strony http://www.euro.com.pl

2. Wprowadzane dane:

Dane smartfonów z laboratorium 1.

_	nazwa [‡]	wyświetlacz 🍦	pamięć_RAM [‡]	pamieć_wbudowana ÷	aparat_foto	cena 🗦	liczba_opinii 🗦	ocena [‡]	status_opinii [‡]
1	Galaxy J2 Core (2020)	5.00	1	16	8	80	17	3	mniej niż 50 opinii
2	Galaxy Xcover FieldPro	5.10	4	64	12	1020	48	5	mniej niż 50 opinii
3	Galaxy A2 Core	5.00	1	8	5	120	36	5	mniej niż 50 opinii
4	Galaxy View2	17.30	3	64	0	660	50	4	50-100 opinii
5	Galaxy M30	6.40	3	32	13	300	316	4	więcej niż 100 opinii
6	Galaxy M20	6.30	3	32	13	300	358	4	więcej niż 100 opinii
7	Galaxy M10	6.22	2	16	13	135	107	4	więcej niż 100 opinii
8	Galaxy Tab Advanced2	10.10	3	32	8	200	8	5	mniej niż 50 opinii
9	Galaxy Tab A 8.0 (2018)	8.00	2	32	8	130	40	4.5	mniej niż 50 opinii
10	Galaxy A6s	6.00	6	64	12	300	86	5	50-100 opinii
11	Galaxy A9 (2018)	6.30	6	64	24	359	320	4.5	więcej niż 100 opinii
12	Galaxy A7 (2018)	6.00	4	64	12	309	223	4	więcej niż 100 opinii
13	Galaxy Note9	6.40	6	128	12	820	1243	5	więcej niż 100 opinii
14	Galaxy J6+	6.00	3	32	10	230	198	4	więcej niż 100 opinii
15	Galaxy J4 Core	6.00	1	16	5	150	87	3.5	50-100 opinii

3. Wykorzystane komendy:

a) kod źródłowy A

```
if (!require("C50")){ install.packages("C50"); library("C50")}
load("../1/ramka_smartfony")
input <- data.frame(subset(ramka, select = -c(status_opinii, nazwa, ocena)))
tree <- C5.0.default(x=input, y=ramka$ocena)
predict.C5.0(tree, input)
summary(tree)
plot(tree)
    4. Wynik działania:
> source('~/Dokumenty/APU/4/4.R', echo=TRUE)
> if (!require("C50")){ install.packages("C50"); library("C50")}
```

```
load("../1/ramka smartfony")
>
   input <- data.frame(subset(ramka, select = -c(status opinii,
>

¬ nazwa, ocena)))
  tree <- C5.0.default(x=input, y=ramka$ocena)
> predict.C5.0(tree, input)
 [1] 3 5 3 4 4 4 4 5 5 5 5 4 5 4 3
Levels: 3 4 4.5 5 3.5
   summary(tree)
Call:
C5.0.default(x = input, y = ramka$ocena)
C5.0 [Release 2.07 GPL Edition] Sun Jun 14 04:46:34 2020
_____
Class specified by attribute `outcome'
Read 15 cases (7 attributes) from undefined.data
Decision tree:
pamięć RAM <= 1: 3 (3/2)
pamięć RAM > 1:
:...pamięć RAM > 4: 5 (3/1)
   pamięć_RAM <= 4:</pre>
   :...liczba opinii <= 48: 5 (3/1)
       liczba opinii > 48: 4 (6)
Evaluation on training data (15 cases):
       Decision Tree
     _____
     Size Errors
        4 4(26.7%) <<
```

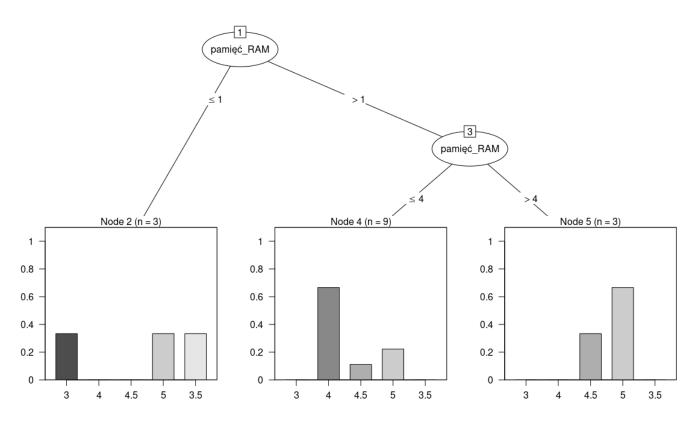
(a)	(b)	(c)	(d)	(e)	<-classified as
					() 7 0
1					(a): class 3
	6				(b): class 4
			2		(c): class 4.5
1			4		(d): class 5
1					(e): class 3.5

Attribute usage:

```
100.00% pamięć_RAM 60.00% liczba_opinii
```

Time: 0.0 secs

> plot(tree)



5. Wnioski:

Z drzewa decyzyjnego można wywnioskować, że najważniejszą cechą smartfona dla kupujących jest pamięć RAM. Im większa jej ilość, tym produkt ma lepszą ocenę.

Telefony z ilością pamięci RAM mniejszą niż 1 GB, jako jedyne dostały ocenę 3. Za to telefony z więcej niż 4 GB pamięci, otrzymały ocenę co najmniej 4,5.