**

**Sprawozdanie z laboratorium nr 1**

**Student: Bartosz Ochędowski**

**Nr Albumu: 59300**

**Zadanie 1 - Z repozytorium baz danych dla uczenia maszynowego *http://archive.ics.uci.edu/ml/***

**nalezy wybrac:**

**a) 3 bazy danych, spełniajace nastepujace kryteria:**

\_ liczba atrybutów wieksza niz 4 i mniejsza niz 20,

\_ liczba przypadków wieksza niz 100 mniejsza niz 5000,

\_ wartosci – **symboliczne**,

\_ charakter danych – do klasyfikacji.

**b) 3 bazy danych, spełniajace nastepujace kryteria:**

\_ liczba atrybutów wieksza niz 4 i mniejsza niz 20,

\_ liczba przypadków wieksza niz 100 mniejsza niz 5000,

\_ wartosci – **numeryczne**,

\_ charakter danych – do klasyfikacji.

**Zadanie 2 - Sprawdz czy wybrane bazy mozna wczytac do R-Studio. Ustawic ostatnia kolumne**

**jako kolumne decyzyjna. Zapisac przygotowane bazy.**

**Zadanie 3 - Przeprowadz wstepna analize danych dla wybranych baz danych. W analizie**

**powinny sie znalezc informacje o (dla kazdej bazy):**

liczbie przypadków,

liczbie klas,

brakujacych wartosciach,

wartosciach poszczególnych atrybutów.

*Dodatkowo zapoznaj sie z:*

wykresem dla rozkładu poszczególnych wartosci,

wykresem gestosci,

rozproszeniem dla par atrybut wartosc.

**Numeryczne:**

**Baza 1:**

MyData <- read.table(file = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data", header = FALSE , sep = ",")

#Zmiana nazwy kolumny V5 na Decyzja

colnames(MyData)[5] <- c('Decyzja')

MyData[1:5]

Wyświetlone informacje:

V1 V2 V3 V4 Decyzja

1 5.1 3.5 1.4 0.2 Iris-setosa

2 4.9 3.0 1.4 0.2 Iris-setosa

3 4.7 3.2 1.3 0.2 Iris-setosa

4 4.6 3.1 1.5 0.2 Iris-setosa

\*

\*

\*

143 5.8 2.7 5.1 1.9 Iris-virginica

144 6.8 3.2 5.9 2.3 Iris-virginica

145 6.7 3.3 5.7 2.5 Iris-virginica

146 6.7 3.0 5.2 2.3 Iris-virginica

147 6.3 2.5 5.0 1.9 Iris-virginica

148 6.5 3.0 5.2 2.0 Iris-virginica

149 6.2 3.4 5.4 2.3 Iris-virginica

150 5.9 3.0 5.1 1.8 Iris-virginica

Zapis bazy danych:

write.csv(MyData, file ="results.csv")

*Liczba Przypadków:*

**150**

*Liczba klas:*

**3**

*Brakujących wartości:*

**0**

*Wartości poszczególnych atrybutów:*

MyData[,1] **ETC…**

*V1:*

[1] 5.1 4.9 4.7 4.6 5.0 5.4 4.6 5.0 4.4 4.9 5.4 4.8 4.8 4.3 5.8 5.7 5.4 5.1 5.7 5.1 5.4 5.1 4.6 5.1 4.8 5.0 5.0 5.2 5.2

[30] 4.7 4.8 5.4 5.2 5.5 4.9 5.0 5.5 4.9 4.4 5.1 5.0 4.5 4.4 5.0 5.1 4.8 5.1 4.6 5.3 5.0 7.0 6.4 6.9 5.5 6.5 5.7 6.3 4.9

[59] 6.6 5.2 5.0 5.9 6.0 6.1 5.6 6.7 5.6 5.8 6.2 5.6 5.9 6.1 6.3 6.1 6.4 6.6 6.8 6.7 6.0 5.7 5.5 5.5 5.8 6.0 5.4 6.0 6.7

[88] 6.3 5.6 5.5 5.5 6.1 5.8 5.0 5.6 5.7 5.7 6.2 5.1 5.7 6.3 5.8 7.1 6.3 6.5 7.6 4.9 7.3 6.7 7.2 6.5 6.4 6.8 5.7 5.8 6.4

[117] 6.5 7.7 7.7 6.0 6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.8 6.7

[146] 6.7 6.3 6.5 6.2 5.9

*V2:*

[1] 3.5 3.0 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 3.7 3.4 3.0 3.0 4.0 4.4 3.9 3.5 3.8 3.8 3.4 3.7 3.6 3.3 3.4 3.0 3.4 3.5 3.4

[30] 3.2 3.1 3.4 4.1 4.2 3.1 3.2 3.5 3.1 3.0 3.4 3.5 2.3 3.2 3.5 3.8 3.0 3.8 3.2 3.7 3.3 3.2 3.2 3.1 2.3 2.8 2.8 3.3 2.4

[59] 2.9 2.7 2.0 3.0 2.2 2.9 2.9 3.1 3.0 2.7 2.2 2.5 3.2 2.8 2.5 2.8 2.9 3.0 2.8 3.0 2.9 2.6 2.4 2.4 2.7 2.7 3.0 3.4 3.1

[88] 2.3 3.0 2.5 2.6 3.0 2.6 2.3 2.7 3.0 2.9 2.9 2.5 2.8 3.3 2.7 3.0 2.9 3.0 3.0 2.5 2.9 2.5 3.6 3.2 2.7 3.0 2.5 2.8 3.2

[117] 3.0 3.8 2.6 2.2 3.2 2.8 2.8 2.7 3.3 3.2 2.8 3.0 2.8 3.0 2.8 3.8 2.8 2.8 2.6 3.0 3.4 3.1 3.0 3.1 3.1 3.1 2.7 3.2 3.3

[146] 3.0 2.5 3.0 3.4 3.0

*V3:*

[1] 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 1.5 1.6 1.4 1.1 1.2 1.5 1.3 1.4 1.7 1.5 1.7 1.5 1.0 1.7 1.9 1.6 1.6 1.5 1.4

[30] 1.6 1.6 1.5 1.5 1.4 1.5 1.2 1.3 1.5 1.3 1.5 1.3 1.3 1.3 1.6 1.9 1.4 1.6 1.4 1.5 1.4 4.7 4.5 4.9 4.0 4.6 4.5 4.7 3.3

[59] 4.6 3.9 3.5 4.2 4.0 4.7 3.6 4.4 4.5 4.1 4.5 3.9 4.8 4.0 4.9 4.7 4.3 4.4 4.8 5.0 4.5 3.5 3.8 3.7 3.9 5.1 4.5 4.5 4.7

[88] 4.4 4.1 4.0 4.4 4.6 4.0 3.3 4.2 4.2 4.2 4.3 3.0 4.1 6.0 5.1 5.9 5.6 5.8 6.6 4.5 6.3 5.8 6.1 5.1 5.3 5.5 5.0 5.1 5.3

[117] 5.5 6.7 6.9 5.0 5.7 4.9 6.7 4.9 5.7 6.0 4.8 4.9 5.6 5.8 6.1 6.4 5.6 5.1 5.6 6.1 5.6 5.5 4.8 5.4 5.6 5.1 5.1 5.9 5.7

[146] 5.2 5.0 5.2 5.4 5.1

*V4:*

[1] 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 0.2 0.2 0.1 0.1 0.2 0.4 0.4 0.3 0.3 0.3 0.2 0.4 0.2 0.5 0.2 0.2 0.4 0.2 0.2

[30] 0.2 0.2 0.4 0.1 0.2 0.1 0.2 0.2 0.1 0.2 0.2 0.3 0.3 0.2 0.6 0.4 0.3 0.2 0.2 0.2 0.2 1.4 1.5 1.5 1.3 1.5 1.3 1.6 1.0

[59] 1.3 1.4 1.0 1.5 1.0 1.4 1.3 1.4 1.5 1.0 1.5 1.1 1.8 1.3 1.5 1.2 1.3 1.4 1.4 1.7 1.5 1.0 1.1 1.0 1.2 1.6 1.5 1.6 1.5

[88] 1.3 1.3 1.3 1.2 1.4 1.2 1.0 1.3 1.2 1.3 1.3 1.1 1.3 2.5 1.9 2.1 1.8 2.2 2.1 1.7 1.8 1.8 2.5 2.0 1.9 2.1 2.0 2.4 2.3

[117] 1.8 2.2 2.3 1.5 2.3 2.0 2.0 1.8 2.1 1.8 1.8 1.8 2.1 1.6 1.9 2.0 2.2 1.5 1.4 2.3 2.4 1.8 1.8 2.1 2.4 2.3 1.9 2.3 2.5

[146] 2.3 1.9 2.0 2.3 1.8

*Decyzja:*

[1] Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa

[8] Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa

[15] Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa

[22] Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa

[29] Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa

[36] Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa

[43] Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa

[50] Iris-setosa Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor

[57] Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor

[64] Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor

[71] Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor

[78] Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor

[85] Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor

[92] Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor Iris-versicolor

[99] Iris-versicolor Iris-versicolor Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica

[106] Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica

[113] Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica

[120] Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica

[127] Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica

[134] Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica

[141] Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica Iris-virginica

[148] Iris-virginica Iris-virginica Iris-virginica

Levels: Iris-setosa Iris-versicolor Iris-virginica

***//Reszta wartości poszczególnych atrybutów będzie przedstawiana w skróconej formie***

**Baza 2:**

MyData <- read.table(file ="https://archive.ics.uci.edu/ml/machine-learning-databases/blood-transfusion/transfusion.data", header = FALSE , sep = ",")

#Zmiana nazwy kolumny V5 na Decyzja

colnames(MyData)[5] <- c('Decyzja')

MyData[1:5]

Wyświetlone informacje:

V1 V2 V3 V4 Decyzja

1 Recency (months) Frequency (times) Monetary (c.c. blood) Time (months) whether he/she donated blood in March 2007

2 2 50 12500 98 1

3 0 13 3250 28 1

4 1 16 4000 35 1

5 2 20 5000 45 1

6 1 24 6000 77 0

\*

\*

\*

197 2 5 1250 47 1

198 9 8 2000 38 1

199 4 6 1500 47 0

200 11 7 1750 29 0

[ reached 'max' / getOption("max.print") -- omitted 549 rows ]

Zapis bazy danych:

write.csv(MyData, file ="Blood Transfusion Service Center.csv")

*Liczba Przypadków:*

**749**

*Liczba klas:*

**2 czyli ‘0’ oraz ‘1’**

*Brakujących wartości:*

**0**

*Wartości poszczególnych atrybutów:*

MyData[,1] **ETC…**

*V1:*

[1] Recency (months) 2 0 1 2 1

[7] 4 2 1 2 5 4

[13] 0 2 1 2 2 2

[19] 2 2 2 2 4 2

[25] 2 9 4 4 4 4

[31] 4 1 4 4 4 2

[37] 2 2 4 2 4 2

[43] 2 2 2 4 4 2

**Etc…**

*V2:*

[1] Frequency (times) 50 13 16 20 24

[7] 4 7 12 9 46 23

[13] 3 10 13 6 5 14

[19] 15 6 3 3 11 6

[25] 6 9 14 6 12 5

[31] 8 14 10 10 9 16

[37] 8 12 6 14 7 13

[43] 5 5 5 20 9 9

**Etc…**

*V3:*

[1] Monetary (c.c. blood) 12500 3250 4000 5000

[6] 6000 1000 1750 3000 2250

[11] 11500 5750 750 2500 3250

[16] 1500 1250 3500 3750 1500

[21] 750 750 2750 1500 1500

[26] 2250 3500 1500 3000 1250

[31] 2000 3500 2500 2500 2250

[36] 4000 2000 3000 1500 3500

[41] 1750 3250 1250 1250 1250

**Etc…**

*V4:*

[1] Time (months) 98 28 35 45 77 4 14

[9] 35 22 98 58 4 28 47 15

[17] 11 48 49 15 4 4 28 16

[25] 16 16 40 14 34 11 21 58

[33] 28 28 26 64 28 47 16 57

[41] 22 53 16 16 16 69 28 36

**Etc…**

*Decyzja:*

[1] whether he/she donated blood in March 2007 1

[3] 1 1

[5] 1 0

[7] 0 1

[9] 0 1

[11] 1 0

[13] 0 1

[15] 0 1

[17] 1 1

[19] 1 1

[21] 1 1

[23] 0 1

[25] 1 0

[27] 0 0

[29] 1 1

[31] 0 0

[33] 1 1

[35] 1 0

[37] 1 1

[39] 1 1

[41] 1 1

[43] 0 1

**Etc…**

**Baza 3:**

MyData <- read.table(file = "https://archive.ics.uci.edu/ml/machine-learning-databases/ecoli/ecoli.data", header = FALSE , sep = ".")

#Przeniesienie kolumny v1 na koniec tabeli oraz zmiana nazwy na Decyzja

MyData<-cbind(MyData[2:8],MyData[,1])

colnames(MyData)[8] <- c('Decyzja')

MyData[1:8]

Wyświetlone informacje:

V2 V3 V4 V5 V6 V7 V8 Decyzja

1 49 0 29 0 48 0 50 0 56 0 24 0 35 cp AAT\_ECOLI 0

2 07 0 40 0 48 0 50 0 54 0 35 0 44 cp ACEA\_ECOLI 0

3 56 0 40 0 48 0 50 0 49 0 37 0 46 cp ACEK\_ECOLI 0

\*

\*

\*

119 25 0 51 0 48 0 50 0 37 0 42 0 50 cp SYGB\_ECOLI 0

120 29 0 28 0 48 0 50 0 50 0 42 0 50 cp SYH\_ECOLI 0

121 25 0 26 0 48 0 50 0 39 0 32 0 42 cp SYI\_ECOLI 0

122 24 0 41 0 48 0 50 0 49 0 23 0 34 cp SYK1\_ECOLI 0

123 17 0 39 0 48 0 50 0 53 0 30 0 39 cp SYK2\_ECOLI 0

124 04 0 31 0 48 0 50 0 41 0 29 0 39 cp SYL\_ECOLI 0

125 61 0 36 0 48 0 50 0 49 0 35 0 44 cp SYM\_ECOLI 0

[ reached 'max' / getOption("max.print") -- omitted 211 rows ]

Zapis bazy danych:

write.csv(MyData, file ="Ecoli.csv")

*Liczba Przypadków:*

**125**

*Liczba klas:*

**336**

*Brakujących wartości:*

**0**

*Wartości poszczególnych atrybutów:*

MyData[,1] **ETC…**

*V1:*

[1] 49 0 07 0 56 0 59 0 23 0 67 0 29 0 21 0 20 0 42 0 42 0 25 0 39 0 51 0 22 0 25 0 34 0 44 0 23 0

[20] 41 0 40 0 31 0 51 0 30 0 36 0 29 0 25 0 21 0 43 0 43 0 53 0 34 0 56 0 40 0 24 0 36 0 29 0 65 0

[39] 32 0 38 0 33 0 30 0 40 0 28 0 61 0 17 0 44 0 43 0 42 0 23 0 37 0 29 0 22 0 23 0 47 0 54 0 51 0

**ETC…**

*V2:*

[1] 29 0 40 0 40 0 49 0 32 0 39 0 28 0 34 0 44 0 40 0 24 0 48 0 32 0 50 0 43 0 40 0 45 0 27 0 40 0

[20] 57 0 45 0 23 0 54 0 16 0 39 0 37 0 40 0 51 0 37 0 39 0 38 0 33 0 51 0 29 0 35 0 54 0 52 0 47 0

[39] 42 0 46 0 45 0 37 0 50 0 38 0 45 0 38 0 35 0 40 0 35 0 33 0 52 0 30 0 36 0 58 0 47 0 47 0 37 0

**ETC…**

*V3:*

[1] 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0

[20] 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0

[39] 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0 48 0

**ETC…**

*V4:*

[1] 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0

[20] 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0

[39] 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0 50 0

**ETC…**

*V5:*

[1] 56 0 54 0 49 0 52 0 55 0 36 0 44 0 51 0 46 0 56 0 57 0 44 0 46 0 46 0 48 0 46 0 38 0 55 0 39 0

[20] 39 0 38 0 73 0 41 0 56 0 48 0 48 0 47 0 50 0 53 0 47 0 44 0 38 0 34 0 42 0 31 0 41 0 42 0 59 0

[39] 35 0 48 0 52 0 59 0 45 0 50 0 48 0 45 0 55 0 39 0 58 0 43 0 42 0 45 0 35 0 37 0 22 0 28 0 35 0

**ETC…**

*V6:*

[1] 24 0 35 0 37 0 45 0 25 0 38 0 23 0 28 0 51 0 18 0 27 0 17 0 24 0 32 0 16 0 44 0 24 0 52 0 28 0

[20] 21 0 22 0 05 0 34 0 11 0 22 0 44 0 33 0 32 0 35 0 31 0 26 0 35 0 37 0 35 0 19 0 38 0 29 0 30 0

[39] 28 0 22 0 32 0 41 0 39 0 33 0 35 0 42 0 55 0 28 0 15 0 33 0 42 0 03 0 39 0 53 0 16 0 33 0 36 0

**ETC…**

*V7:*

[1] 35 cp 44 cp 46 cp 36 cp 35 cp 46 cp 34 cp 39 cp 57 cp 30 cp 37 cp 29 cp 35 cp 35 cp

[15] 28 cp 52 cp 35 cp 58 cp 38 cp 32 cp 00 cp 14 cp 43 cp 23 cp 23 cp 52 cp 42 cp 41 cp

[29] 44 cp 41 cp 36 cp 44 cp 46 cp 44 cp 31 cp 46 cp 39 cp 40 cp 38 cp 29 cp 41 cp 49 cp

[43] 47 cp 42 cp 41 cp 50 cp 61 cp 39 cp 27 cp 43 cp 36 cp 17 cp 47 cp 59 cp 26 cp 42 cp

**ETC…**

*Decyzja:*

[1] AAT\_ECOLI 0 ACEA\_ECOLI 0 ACEK\_ECOLI 0 ACKA\_ECOLI 0 ADI\_ECOLI 0 ALKH\_ECOLI 0 AMPD\_ECOLI 0 AMY2\_ECOLI 0

[9] APT\_ECOLI 0 ARAC\_ECOLI 0 ASG1\_ECOLI 0 BTUR\_ECOLI 0 CAFA\_ECOLI 0 CAIB\_ECOLI 0 CFA\_ECOLI 0 CHEA\_ECOLI 0

[17] CHEB\_ECOLI 0 CHEW\_ECOLI 0 CHEY\_ECOLI 0 CHEZ\_ECOLI 0 CRL\_ECOLI 0 CSPA\_ECOLI 0 CYNR\_ECOLI 0 CYPB\_ECOLI 0

[25] CYPC\_ECOLI 0 CYSB\_ECOLI 0 CYSE\_ECOLI 0 DAPD\_ECOLI 0 DCP\_ECOLI 0 DDLA\_ECOLI 0 DDLB\_ECOLI 0 DEOC\_ECOLI 0

[33] DLDH\_ECOLI 0 EFG\_ECOLI 0 EFTS\_ECOLI 0 EFTU\_ECOLI 0 ENO\_ECOLI 0 FABB\_ECOLI 0 FES\_ECOLI 0 G3P1\_ECOLI 0

[41] G3P2\_ECOLI 0 G6PI\_ECOLI 0 GCVA\_ECOLI 0 GLNA\_ECOLI 0 GLPD\_ECOLI 0 GLYA\_ECOLI 0 GSHR\_ECOLI 0 GT\_ECOLI 0

[49] HEM6\_ECOLI 0 HEMN\_ECOLI 0 HPRT\_ECOLI 0 IF1\_ECOLI 0 IF2\_ECOLI 0 ILVY\_ECOLI 0 IPYR\_ECOLI 0 KAD\_ECOLI 0

**ETC…**

**Symboliczne:**

**//Wiem że to nie są symboliczne bazy , niestety nie mogłem się ich do szukać nawet po przez //filtracje danych na stronie UCI .**

**//Udało mi się znaleźć tylko jedną taką jest to ostatnia baza**

**Baza 4:**

MyData <- read.table(file = "https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer/breast-cancer.data", header = FALSE , sep = ",")

#Zmiana nazwy kolumny V10 na Decyzja

colnames(MyData)[10] <- c('Decyzja')

MyData[1:10]

Wyświetlone informacje:

V1 V2 V3 V4 V5 V6 V7 V8 V9 Decyzja

1 no-recurrence-events 30-39 premeno 30-34 0-2 no 3 left left\_low no

2 no-recurrence-events 40-49 premeno 20-24 0-2 no 2 right right\_up no

3 no-recurrence-events 40-49 premeno 20-24 0-2 no 2 left left\_low no

4 no-recurrence-events 60-69 ge40 15-19 0-2 no 2 right left\_up no

\*

\*

\*

96 no-recurrence-events 40-49 premeno 30-34 0-2 no 2 right right\_low no

97 no-recurrence-events 60-69 ge40 30-34 0-2 no 2 left left\_up no

98 no-recurrence-events 60-69 ge40 15-19 0-2 no 2 right left\_up no

99 no-recurrence-events 40-49 premeno 30-34 0-2 no 1 left right\_up no

100 no-recurrence-events 30-39 premeno 25-29 0-2 no 2 left left\_low no

[ reached 'max' / getOption("max.print") -- omitted 186 rows ]

Zapis bazy danych:

write.csv(MyData, file ="Breast-Cancer.csv")

*Liczba Przypadków:*

**100**

*Liczba klas:*

**2 czyli ‘yas’ oraz ‘no’**

*Brakujących wartości:*

**0**

*Wartości poszczególnych atrybutów:*

MyData[,1] **ETC…**

*V1:*

[1] no-recurrence-events no-recurrence-events no-recurrence-events no-recurrence-events no-recurrence-events

[6] no-recurrence-events no-recurrence-events no-recurrence-events no-recurrence-events no-recurrence-events

[11] no-recurrence-events no-recurrence-events no-recurrence-events no-recurrence-events no-recurrence-events

**ETC…**

*V2:*

[1] 30-39 40-49 40-49 60-69 40-49 60-69 50-59 60-69 40-49 40-49 40-49 50-59 60-69 50-59 40-49 60-69 40-49 50-59 60-69

[20] 50-59 50-59 60-69 30-39 50-59 50-59 40-49 50-59 60-69 40-49 60-69 50-59 50-59 50-59 50-59 50-59 30-39 50-59 50-59

[39] 40-49 40-49 50-59 60-69 60-69 40-49 50-59 50-59 40-49 50-59 40-49 40-49 50-59 30-39 50-59 70-79 70-79 70-79 50-59

**ETC…**

*V3:*

[1] premeno premeno premeno ge40 premeno ge40 premeno ge40 premeno premeno premeno ge40 lt40 ge40

[15] premeno lt40 premeno premeno ge40 ge40 ge40 ge40 premeno premeno premeno premeno premeno ge40

[29] premeno ge40 ge40 premeno premeno ge40 ge40 premeno ge40 ge40 premeno premeno ge40 ge40

**ETC…**

*V4:*

[1] 30-34 20-24 20-24 15-19 0-4 15-19 25-29 20-24 50-54 20-24 0-4 25-29 10-14 25-29 30-34 30-34 15-19 30-34 30-34

[20] 30-34 40-44 15-19 25-29 40-44 35-39 25-29 20-24 25-29 40-44 30-34 40-44 15-19 10-14 10-14 10-14 30-34 0-4 15-19

[39] 10-14 30-34 20-24 25-29 5-9 10-14 50-54 30-34 25-29 25-29 20-24 20-24 15-19 20-24 15-19 20-24 40-44 40-44 0-4

**ETC…**

*V5:*

[1] 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2

[20] 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2

[39] 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2 0-2

**ETC…**

*V6:*

[1] no no no no no no no no no no no no no no no no no no no no no no no no no no no no

[29] no no no no no no no no no no no no no no no no no no no no no no no no no no no no

[57] no no no no no no no no no no no no no no no no no no no no no no no no no no no no

**ETC…**

*V7:*

[1] 3 2 2 2 2 2 2 1 2 2 3 2 1 3 3 1 2 3 3 1 2 2 2 2 2 2 1 3 2 2 3 2 3 1 1 2 2 1 2 1 1 2 1 2 1 1 2 1 1 1 2 2 1 3 1 1 1

[58] 2 1 1 2 1 1 2 1 1 1 3 3 2 1 2 2 2 2 1 2 1 2 2 2 1 2 2 1 3 2 1 3 1 2 3 2 2 1 2 2 2 1 2 3 3 2 2 2 1 2 2 3 1 1 1 2 1

[115] 2 2 1 3 1 1 1 2 3 1 1 2 2 2 2 2 2 3 2 3 2 2 3 1 2 2 2 2 1 2 3 2 2 2 3 3 3 3 3 3 3 2 3 1 1 1 3 2 2 1 1 2 2 2 2 2 2

**ETC…**

*V8:*

[1] left right left right right left left left left right left left left left left left left left left

[20] right left left right left right left left right right left right right left right left left left right

[39] left left right left left left right left right right right right left left left left right right right

**ETC…**

*V9:*

[1] left\_low right\_up left\_low left\_up right\_low left\_low left\_low left\_low left\_low left\_up central

[12] left\_low right\_up right\_up left\_up left\_low left\_low left\_low left\_low right\_up left\_low left\_low

[23] left\_low left\_up left\_up left\_up left\_low left\_up left\_low left\_low left\_up left\_low left\_low

**ETC…**

*Decyzja:*

[1] no no no no no no no no no no no no no no no no no no no no no no no no no no no no

[29] no no no no no no no no no no no no no no no no no no no no no no no no no no no no

[57] no no no no no no no no no no no no no no no no no no no no no no no no no no no no

**ETC…**

**Baza 5:**

MyData <- read.table(file = "https://archive.ics.uci.edu/ml/machine-learning-databases/hayes-roth/hayes-roth.data", header = FALSE , sep = ",")

#Zmiana nazwy kolumny V6 na Decyzja

colnames(MyData)[6] <- c('Decyzja')

MyData[1:6]

Wyświetlone informacje:

V1 V2 V3 V4 V5 Decyzja

1 92 2 1 1 2 1

2 10 2 1 3 2 2

3 83 3 1 4 1 3

4 61 2 4 2 2 3

5 107 1 1 3 4 3

6 113 1 1 3 2 2

\*

\*

\*

129 40 2 1 2 1 1

130 90 1 2 1 2 2

131 21 1 2 2 1 2

132 9 3 1 1 2 1

Zapis bazy danych:

write.csv(MyData, file ="Hayes-Roth.csv")

*Liczba Przypadków:*

**132**

*Liczba klas:*

**3 czyli ‘1’ , ‘2’ oraz ‘3’**

*Brakujących wartości:*

**0**

*Wartości poszczególnych atrybutów:*

MyData[,1] **ETC…**

*V1:*

[1] 92 10 83 61 107 113 80 125 36 105 81 122 94 60 8 20 85 50 68 89 52 19 118 16 91 79 23 25

[29] 30 57 3 114 37 66 110 116 88 77 82 64 84 86 6 74 106 115 130 54 33 67 69 39 53 127 96 121

[57] 70 123 42 78 11 129 128 5 4 95 73 26 48 104 102 2 41 119 75 47 93 46 132 108 18 62 120 35

[85] 27 98 109 31 112 34 63 65 117 56 59 76 1 28 22 29 111 97 49 51 87 58 32 72 55 103 7 99

[113] 15 126 45 101 100 24 124 13 14 38 71 43 131 17 12 44 40 90 21 9

*V2:*

[1] 2 2 3 2 1 1 3 3 2 3 1 2 1 2 2 1 3 1 3 3 1 3 2 3 2 3 3 2 1 3 1 2 1 1 2 3 1 3 1 3 2 2 3 3 3 1 2 1 1 3 3 3 3 3 1 2 2

[58] 2 2 2 1 2 1 1 2 2 3 1 1 1 3 2 1 3 1 1 2 3 2 1 2 3 1 1 1 3 2 3 1 2 2 2 1 2 1 3 3 1 3 3 2 2 1 3 2 1 2 2 1 2 1 2 1 3

[115] 3 3 2 1 3 3 1 2 3 3 2 2 3 1 2 1 1 3

*V3:*

[1] 1 1 1 4 1 1 1 4 2 2 2 2 1 1 4 1 2 2 3 1 2 2 1 2 3 2 2 1 1 2 4 2 2 1 4 1 1 2 2 4 2 2 2 2 1 2 1 1 2 3 3 2 2 1 1 1 2

[58] 1 2 1 2 2 1 3 4 3 1 1 3 1 1 1 1 1 2 4 1 4 2 1 2 1 1 2 4 3 2 3 1 2 2 3 3 2 1 2 2 1 1 3 3 1 2 1 2 2 3 2 4 2 2 2 3 1

[115] 1 3 3 2 3 3 2 1 1 2 3 1 4 1 1 2 2 1

*V4:*

[1] 1 3 4 2 3 3 3 2 1 1 1 3 2 2 1 3 1 1 2 3 2 1 2 1 2 2 1 2 3 1 1 1 1 1 3 2 2 2 1 3 2 1 1 1 2 1 1 1 2 1 3 1 1 2 1 3 2

[58] 2 1 2 4 1 2 2 4 2 2 2 2 2 4 3 3 3 4 2 2 1 1 2 4 2 3 1 4 3 1 2 1 1 2 2 2 1 1 2 1 2 4 2 2 3 1 1 4 2 2 1 2 1 1 3 2 2

[115] 1 1 4 3 2 4 2 1 2 2 1 1 1 4 2 1 2 1

*V5:*

[1] 2 2 1 2 4 2 2 4 1 1 1 4 1 2 4 3 2 1 1 2 1 3 1 3 1 1 3 2 2 1 1 3 3 2 1 2 2 1 2 2 1 2 3 1 1 3 2 2 3 1 1 2 2 1 2 2 1

[58] 1 3 2 2 2 4 1 2 1 2 2 4 2 2 2 2 2 4 1 1 2 1 1 3 2 2 3 1 2 3 1 3 2 1 3 1 2 2 1 1 1 4 1 1 1 2 2 1 1 1 4 3 1 1 2 1 1

[115] 2 4 1 3 2 2 1 4 2 4 3 2 3 3 1 2 1 2

*Decyzja:*

[1] 1 2 3 3 3 2 2 3 1 1 1 3 1 2 3 1 2 1 1 1 2 1 1 1 1 2 2 2 1 1 3 1 2 1 3 2 2 2 2 3 2 2 2 1 1 2 1 1 2 1 1 2 2 1 1 1 2

[58] 1 1 2 3 2 3 1 3 1 2 2 3 2 3 2 2 1 3 3 1 3 1 1 3 2 1 1 3 2 2 2 1 2 2 2 2 2 1 2 1 1 3 2 2 1 2 1 3 2 2 3 3 1 1 2 1 1

[115] 1 3 3 2 2 3 2 3 2 3 1 1 3 3 1 2 2 1

**Baza 6:**

MyData <- read.table(file = "https://archive.ics.uci.edu/ml/machine-learning-databases/tic-tac-toe/tic-tac-toe.data", header = FALSE , sep = ",")#Zmiana nazwy kolumny V10 na Decyzja

colnames(MyData)[10] <- c('Decyzja')

MyData[1:10]

Wyświetlone informacje:

V1 V2 V3 V4 V5 V6 V7 V8 V9 Decyzja

1 x x x x o o x o o positive

2 x x x x o o o x o positive

3 x x x x o o o o x positive

4 x x x x o o o b b positive

\*

\*

\*

93 x x o o x b o x b positive

94 x x o o x b o b x positive

95 x x o o x b b x o positive

96 x x o o x b b o x positive

97 x x o b x o o x b positive

98 x x o b x o o b x positive

99 x x o b x o b o x positive

100 x x o b x b o x o positive

[ reached 'max' / getOption("max.print") -- omitted 858 rows ]

Zapis bazy danych:

write.csv(MyData, file ="Tic-Tac-Toe Endgame.csv")

*Liczba Przypadków:*

**100**

*Liczba klas:*

**2 czyli ‘positive’ oraz ‘negative’**

*Brakujących wartości:*

**0**

*Wartości poszczególnych atrybutów:*

MyData[,1] **ETC…**

*V1:*

[1] x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x

[58] x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x

[115] x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x

**ETC…**

*V2:*

[1] x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x

[58] x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x

[115] x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x x

**ETC…**

*V3:*

[1] x x x x x x x x x x x x o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o o b b b b b b b

[58] b b b b b b b b b b b b b b b b b b b b b x x x x x x x x o o o o o o o o o o b b b b b x x x x o o o o o o b b x

[115] x x x x x x x o o o o o o o o o o o o o o o b b b b b b b b x x x x x x x x x x x x x x x x x x x x x x x x x x o

**ETC…**

*V4:*

[1] o o o o o o o o o b b b x x x x x x x x x o o o o o o o o o o o o o b b b b b b b b b b b b b b b b x x x o o o o

[58] o o o o o o o o o o o o b b b b b b b b b x o o o o o b b x x x x x x x x x x x x x x x o o o b x x x x x x x x x

[115] x o o o b b b x x x x x x x x x x x o o b b x x x x x x o b x x x x x x x x x x x x x x o o o o o b b b b b b b x

**ETC…**

*V5:*

[1] o o o o o o b b b o o o o o o o o o b b b x x x x x x b b b b b b b x x x o o o o o o o b b b b b b o o o x x x o

[58] o o o o o o b b b b b b o o o o o o b b b o x o o b b o b x x o o o o b b b b o o o b b o o b o o o o o b b o o o

[115] o o o b o o b x x o o o o o b b b b x x x x o o o o b b x x x x x x x x o o o o b b b b x x o b b x x o o b b b x

**ETC…**

*V6:*

[1] x o o o b b o o b o o b x o o o b b o o b x o o o b b x x o o b b b o o b x x o o b b b x o o o b b o o b o o b x

[58] x o o b b b x o o o b b x o o o b b o o b o x x x x x x x o o x o b b o o b b o o b o o x x x x o o b b o o o o x

[115] o x x x x x x x o x x x o b x x o b o b o b x x o b x o o o x o o o b b x x o b x x o b x x x x x x x x x x x x x

**ETC…**

*V7:*

[1] o x o b o b o b o o b o o x o b o b o b o o x o b o b o b x b x o b o b o o b x b x o b o x o b o b o b o o b o o

[58] b x b x o b o x o b o b o x o b o b o b o o o o b o b o o x o o x x b x b x o x b o x o o b o o x b x o x o x o o

[115] o x b b o b o o o x o b x b o b b o x b b o o b b o o o b o o x o b o b o b x b o b b o x b x x b o b x b x o b o

*V8:*

[1] o x o b o b o b o o b o o x o b o b o b o o x o b o b o b x b x o b o b o o b x b x o b o x o b o b o b o o b o o

[58] b x b x o b o x o b o b o x o b o b o b o o o o b o b o o x o o x x b x b x o x b o x o o b o o x b x o x o x o o

**ETC…**

[115] o x b b o b o o o x o b x b o b b o x b b o o b b o o o b o o x o b o b o b x b o b b o x b x x b o b x b x o b o

*V9:*

[1] o o x b b o b o o b o o o o x b b o b o o o o x b b o b o b x o x b b o o b o b x o x b o o x b b o b o o b o o b

[58] o b x o x b o o x b b o o o x b b o b o o x o x b b o b o o x x x b x b x o x b x x o x b o o o b x o x o x o x o

[115] x o b o b o o o x o x b x x b o x x x x x x b o x x o x x x o o x b b o x b x x b o x x o b x b x b o b x o x b x

*Decyzja:*

[1] positive positive positive positive positive positive positive positive positive positive positive positive

[13] positive positive positive positive positive positive positive positive positive positive positive positive

[25] positive positive positive positive positive positive positive positive positive positive positive positive

***LINKI:***

<https://archive.ics.uci.edu/ml/datasets/BuddyMove+Data+Set>

<https://archive.ics.uci.edu/ml/datasets/Blood+Transfusion+Service+Center>

<https://archive.ics.uci.edu/ml/datasets/Ecoli>

<https://archive.ics.uci.edu/ml/datasets/Breast+Cancer>

<https://archive.ics.uci.edu/ml/datasets/Hayes-Roth>

<https://archive.ics.uci.edu/ml/datasets/Tic-Tac-Toe+Endgame>