RML\_060.R

darki

2020-04-03

# Rozdział 6 - Naiwny klasufikator Bayesa  
  
library(caret)

## Loading required package: lattice

## Loading required package: ggplot2

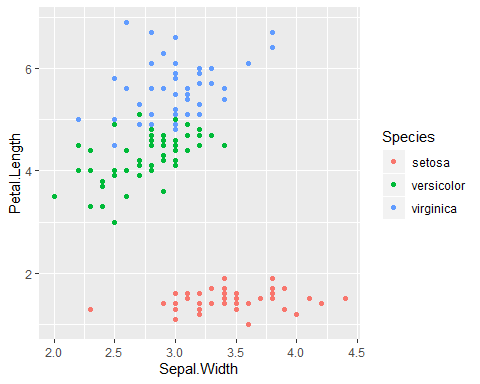
library(e1071)  
library(class)  
  
rm(list=ls())  
  
#odczyt danych  
iris\_dataset <- iris  
  
#losowanie danych  
iris\_dataset <- iris\_dataset[sample.int(nrow(iris\_dataset)),]  
iris\_dataset

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 118 7.7 3.8 6.7 2.2 virginica  
## 131 7.4 2.8 6.1 1.9 virginica  
## 80 5.7 2.6 3.5 1.0 versicolor  
## 113 6.8 3.0 5.5 2.1 virginica  
## 55 6.5 2.8 4.6 1.5 versicolor  
## 143 5.8 2.7 5.1 1.9 virginica  
## 123 7.7 2.8 6.7 2.0 virginica  
## 20 5.1 3.8 1.5 0.3 setosa  
## 66 6.7 3.1 4.4 1.4 versicolor  
## 57 6.3 3.3 4.7 1.6 versicolor  
## 111 6.5 3.2 5.1 2.0 virginica  
## 50 5.0 3.3 1.4 0.2 setosa  
## 138 6.4 3.1 5.5 1.8 virginica  
## 9 4.4 2.9 1.4 0.2 setosa  
## 90 5.5 2.5 4.0 1.3 versicolor  
## 122 5.6 2.8 4.9 2.0 virginica  
## 30 4.7 3.2 1.6 0.2 setosa  
## 128 6.1 3.0 4.9 1.8 virginica  
## 41 5.0 3.5 1.3 0.3 setosa  
## 99 5.1 2.5 3.0 1.1 versicolor  
## 98 6.2 2.9 4.3 1.3 versicolor  
## 65 5.6 2.9 3.6 1.3 versicolor  
## 75 6.4 2.9 4.3 1.3 versicolor  
## 23 4.6 3.6 1.0 0.2 setosa  
## 147 6.3 2.5 5.0 1.9 virginica  
## 27 5.0 3.4 1.6 0.4 setosa  
## 13 4.8 3.0 1.4 0.1 setosa  
## 3 4.7 3.2 1.3 0.2 setosa  
## 134 6.3 2.8 5.1 1.5 virginica  
## 59 6.6 2.9 4.6 1.3 versicolor  
## 121 6.9 3.2 5.7 2.3 virginica  
## 144 6.8 3.2 5.9 2.3 virginica  
## 120 6.0 2.2 5.0 1.5 virginica  
## 54 5.5 2.3 4.0 1.3 versicolor  
## 95 5.6 2.7 4.2 1.3 versicolor  
## 44 5.0 3.5 1.6 0.6 setosa  
## 24 5.1 3.3 1.7 0.5 setosa  
## 48 4.6 3.2 1.4 0.2 setosa  
## 68 5.8 2.7 4.1 1.0 versicolor  
## 74 6.1 2.8 4.7 1.2 versicolor  
## 47 5.1 3.8 1.6 0.2 setosa  
## 119 7.7 2.6 6.9 2.3 virginica  
## 73 6.3 2.5 4.9 1.5 versicolor  
## 91 5.5 2.6 4.4 1.2 versicolor  
## 1 5.1 3.5 1.4 0.2 setosa  
## 52 6.4 3.2 4.5 1.5 versicolor  
## 89 5.6 3.0 4.1 1.3 versicolor  
## 136 7.7 3.0 6.1 2.3 virginica  
## 145 6.7 3.3 5.7 2.5 virginica  
## 79 6.0 2.9 4.5 1.5 versicolor  
## 96 5.7 3.0 4.2 1.2 versicolor  
## 87 6.7 3.1 4.7 1.5 versicolor  
## 92 6.1 3.0 4.6 1.4 versicolor  
## 105 6.5 3.0 5.8 2.2 virginica  
## 112 6.4 2.7 5.3 1.9 virginica  
## 93 5.8 2.6 4.0 1.2 versicolor  
## 53 6.9 3.1 4.9 1.5 versicolor  
## 94 5.0 2.3 3.3 1.0 versicolor  
## 109 6.7 2.5 5.8 1.8 virginica  
## 35 4.9 3.1 1.5 0.2 setosa  
## 139 6.0 3.0 4.8 1.8 virginica  
## 14 4.3 3.0 1.1 0.1 setosa  
## 31 4.8 3.1 1.6 0.2 setosa  
## 115 5.8 2.8 5.1 2.4 virginica  
## 103 7.1 3.0 5.9 2.1 virginica  
## 36 5.0 3.2 1.2 0.2 setosa  
## 62 5.9 3.0 4.2 1.5 versicolor  
## 132 7.9 3.8 6.4 2.0 virginica  
## 81 5.5 2.4 3.8 1.1 versicolor  
## 127 6.2 2.8 4.8 1.8 virginica  
## 124 6.3 2.7 4.9 1.8 virginica  
## 16 5.7 4.4 1.5 0.4 setosa  
## 126 7.2 3.2 6.0 1.8 virginica  
## 104 6.3 2.9 5.6 1.8 virginica  
## 108 7.3 2.9 6.3 1.8 virginica  
## 32 5.4 3.4 1.5 0.4 setosa  
## 8 5.0 3.4 1.5 0.2 setosa  
## 107 4.9 2.5 4.5 1.7 virginica  
## 86 6.0 3.4 4.5 1.6 versicolor  
## 64 6.1 2.9 4.7 1.4 versicolor  
## 5 5.0 3.6 1.4 0.2 setosa  
## 85 5.4 3.0 4.5 1.5 versicolor  
## 100 5.7 2.8 4.1 1.3 versicolor  
## 12 4.8 3.4 1.6 0.2 setosa  
## 77 6.8 2.8 4.8 1.4 versicolor  
## 37 5.5 3.5 1.3 0.2 setosa  
## 10 4.9 3.1 1.5 0.1 setosa  
## 4 4.6 3.1 1.5 0.2 setosa  
## 60 5.2 2.7 3.9 1.4 versicolor  
## 70 5.6 2.5 3.9 1.1 versicolor  
## 19 5.7 3.8 1.7 0.3 setosa  
## 82 5.5 2.4 3.7 1.0 versicolor  
## 15 5.8 4.0 1.2 0.2 setosa  
## 61 5.0 2.0 3.5 1.0 versicolor  
## 116 6.4 3.2 5.3 2.3 virginica  
## 141 6.7 3.1 5.6 2.4 virginica  
## 51 7.0 3.2 4.7 1.4 versicolor  
## 102 5.8 2.7 5.1 1.9 virginica  
## 148 6.5 3.0 5.2 2.0 virginica  
## 46 4.8 3.0 1.4 0.3 setosa  
## 26 5.0 3.0 1.6 0.2 setosa  
## 69 6.2 2.2 4.5 1.5 versicolor  
## 130 7.2 3.0 5.8 1.6 virginica  
## 43 4.4 3.2 1.3 0.2 setosa  
## 17 5.4 3.9 1.3 0.4 setosa  
## 149 6.2 3.4 5.4 2.3 virginica  
## 28 5.2 3.5 1.5 0.2 setosa  
## 29 5.2 3.4 1.4 0.2 setosa  
## 49 5.3 3.7 1.5 0.2 setosa  
## 106 7.6 3.0 6.6 2.1 virginica  
## 34 5.5 4.2 1.4 0.2 setosa  
## 101 6.3 3.3 6.0 2.5 virginica  
## 88 6.3 2.3 4.4 1.3 versicolor  
## 146 6.7 3.0 5.2 2.3 virginica  
## 56 5.7 2.8 4.5 1.3 versicolor  
## 2 4.9 3.0 1.4 0.2 setosa  
## 142 6.9 3.1 5.1 2.3 virginica  
## 39 4.4 3.0 1.3 0.2 setosa  
## 72 6.1 2.8 4.0 1.3 versicolor  
## 63 6.0 2.2 4.0 1.0 versicolor  
## 84 6.0 2.7 5.1 1.6 versicolor  
## 137 6.3 3.4 5.6 2.4 virginica  
## 58 4.9 2.4 3.3 1.0 versicolor  
## 133 6.4 2.8 5.6 2.2 virginica  
## 150 5.9 3.0 5.1 1.8 virginica  
## 6 5.4 3.9 1.7 0.4 setosa  
## 25 4.8 3.4 1.9 0.2 setosa  
## 21 5.4 3.4 1.7 0.2 setosa  
## 71 5.9 3.2 4.8 1.8 versicolor  
## 129 6.4 2.8 5.6 2.1 virginica  
## 97 5.7 2.9 4.2 1.3 versicolor  
## 125 6.7 3.3 5.7 2.1 virginica  
## 42 4.5 2.3 1.3 0.3 setosa  
## 7 4.6 3.4 1.4 0.3 setosa  
## 38 4.9 3.6 1.4 0.1 setosa  
## 78 6.7 3.0 5.0 1.7 versicolor  
## 40 5.1 3.4 1.5 0.2 setosa  
## 114 5.7 2.5 5.0 2.0 virginica  
## 67 5.6 3.0 4.5 1.5 versicolor  
## 140 6.9 3.1 5.4 2.1 virginica  
## 22 5.1 3.7 1.5 0.4 setosa  
## 83 5.8 2.7 3.9 1.2 versicolor  
## 135 6.1 2.6 5.6 1.4 virginica  
## 110 7.2 3.6 6.1 2.5 virginica  
## 18 5.1 3.5 1.4 0.3 setosa  
## 33 5.2 4.1 1.5 0.1 setosa  
## 11 5.4 3.7 1.5 0.2 setosa  
## 45 5.1 3.8 1.9 0.4 setosa  
## 117 6.5 3.0 5.5 1.8 virginica  
## 76 6.6 3.0 4.4 1.4 versicolor

#kilka pierwszych przypadków  
head(iris\_dataset)

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 118 7.7 3.8 6.7 2.2 virginica  
## 131 7.4 2.8 6.1 1.9 virginica  
## 80 5.7 2.6 3.5 1.0 versicolor  
## 113 6.8 3.0 5.5 2.1 virginica  
## 55 6.5 2.8 4.6 1.5 versicolor  
## 143 5.8 2.7 5.1 1.9 virginica

#wykres cech irysów  
ggplot(iris\_dataset, aes(x = Sepal.Width, y = Petal.Length)) + geom\_point(aes(color = Species))



#przekodowaie danych  
iris\_dataset$Species <- factor(iris\_dataset$Species, labels=c("A","B","C"))  
  
#normalizacja min-max  
normalize <- function(x) {return( (x-min(x)) / (max(x)-min(x)) )}  
normalized\_iris\_dataset <- as.data.frame(lapply(iris\_dataset[1:4], normalize))  
normalized\_iris\_dataset

## Sepal.Length Sepal.Width Petal.Length Petal.Width  
## 1 0.94444444 0.75000000 0.96610169 0.87500000  
## 2 0.86111111 0.33333333 0.86440678 0.75000000  
## 3 0.38888889 0.25000000 0.42372881 0.37500000  
## 4 0.69444444 0.41666667 0.76271186 0.83333333  
## 5 0.61111111 0.33333333 0.61016949 0.58333333  
## 6 0.41666667 0.29166667 0.69491525 0.75000000  
## 7 0.94444444 0.33333333 0.96610169 0.79166667  
## 8 0.22222222 0.75000000 0.08474576 0.08333333  
## 9 0.66666667 0.45833333 0.57627119 0.54166667  
## 10 0.55555556 0.54166667 0.62711864 0.62500000  
## 11 0.61111111 0.50000000 0.69491525 0.79166667  
## 12 0.19444444 0.54166667 0.06779661 0.04166667  
## 13 0.58333333 0.45833333 0.76271186 0.70833333  
## 14 0.02777778 0.37500000 0.06779661 0.04166667  
## 15 0.33333333 0.20833333 0.50847458 0.50000000  
## 16 0.36111111 0.33333333 0.66101695 0.79166667  
## 17 0.11111111 0.50000000 0.10169492 0.04166667  
## 18 0.50000000 0.41666667 0.66101695 0.70833333  
## 19 0.19444444 0.62500000 0.05084746 0.08333333  
## 20 0.22222222 0.20833333 0.33898305 0.41666667  
## 21 0.52777778 0.37500000 0.55932203 0.50000000  
## 22 0.36111111 0.37500000 0.44067797 0.50000000  
## 23 0.58333333 0.37500000 0.55932203 0.50000000  
## 24 0.08333333 0.66666667 0.00000000 0.04166667  
## 25 0.55555556 0.20833333 0.67796610 0.75000000  
## 26 0.19444444 0.58333333 0.10169492 0.12500000  
## 27 0.13888889 0.41666667 0.06779661 0.00000000  
## 28 0.11111111 0.50000000 0.05084746 0.04166667  
## 29 0.55555556 0.33333333 0.69491525 0.58333333  
## 30 0.63888889 0.37500000 0.61016949 0.50000000  
## 31 0.72222222 0.50000000 0.79661017 0.91666667  
## 32 0.69444444 0.50000000 0.83050847 0.91666667  
## 33 0.47222222 0.08333333 0.67796610 0.58333333  
## 34 0.33333333 0.12500000 0.50847458 0.50000000  
## 35 0.36111111 0.29166667 0.54237288 0.50000000  
## 36 0.19444444 0.62500000 0.10169492 0.20833333  
## 37 0.22222222 0.54166667 0.11864407 0.16666667  
## 38 0.08333333 0.50000000 0.06779661 0.04166667  
## 39 0.41666667 0.29166667 0.52542373 0.37500000  
## 40 0.50000000 0.33333333 0.62711864 0.45833333  
## 41 0.22222222 0.75000000 0.10169492 0.04166667  
## 42 0.94444444 0.25000000 1.00000000 0.91666667  
## 43 0.55555556 0.20833333 0.66101695 0.58333333  
## 44 0.33333333 0.25000000 0.57627119 0.45833333  
## 45 0.22222222 0.62500000 0.06779661 0.04166667  
## 46 0.58333333 0.50000000 0.59322034 0.58333333  
## 47 0.36111111 0.41666667 0.52542373 0.50000000  
## 48 0.94444444 0.41666667 0.86440678 0.91666667  
## 49 0.66666667 0.54166667 0.79661017 1.00000000  
## 50 0.47222222 0.37500000 0.59322034 0.58333333  
## 51 0.38888889 0.41666667 0.54237288 0.45833333  
## 52 0.66666667 0.45833333 0.62711864 0.58333333  
## 53 0.50000000 0.41666667 0.61016949 0.54166667  
## 54 0.61111111 0.41666667 0.81355932 0.87500000  
## 55 0.58333333 0.29166667 0.72881356 0.75000000  
## 56 0.41666667 0.25000000 0.50847458 0.45833333  
## 57 0.72222222 0.45833333 0.66101695 0.58333333  
## 58 0.19444444 0.12500000 0.38983051 0.37500000  
## 59 0.66666667 0.20833333 0.81355932 0.70833333  
## 60 0.16666667 0.45833333 0.08474576 0.04166667  
## 61 0.47222222 0.41666667 0.64406780 0.70833333  
## 62 0.00000000 0.41666667 0.01694915 0.00000000  
## 63 0.13888889 0.45833333 0.10169492 0.04166667  
## 64 0.41666667 0.33333333 0.69491525 0.95833333  
## 65 0.77777778 0.41666667 0.83050847 0.83333333  
## 66 0.19444444 0.50000000 0.03389831 0.04166667  
## 67 0.44444444 0.41666667 0.54237288 0.58333333  
## 68 1.00000000 0.75000000 0.91525424 0.79166667  
## 69 0.33333333 0.16666667 0.47457627 0.41666667  
## 70 0.52777778 0.33333333 0.64406780 0.70833333  
## 71 0.55555556 0.29166667 0.66101695 0.70833333  
## 72 0.38888889 1.00000000 0.08474576 0.12500000  
## 73 0.80555556 0.50000000 0.84745763 0.70833333  
## 74 0.55555556 0.37500000 0.77966102 0.70833333  
## 75 0.83333333 0.37500000 0.89830508 0.70833333  
## 76 0.30555556 0.58333333 0.08474576 0.12500000  
## 77 0.19444444 0.58333333 0.08474576 0.04166667  
## 78 0.16666667 0.20833333 0.59322034 0.66666667  
## 79 0.47222222 0.58333333 0.59322034 0.62500000  
## 80 0.50000000 0.37500000 0.62711864 0.54166667  
## 81 0.19444444 0.66666667 0.06779661 0.04166667  
## 82 0.30555556 0.41666667 0.59322034 0.58333333  
## 83 0.38888889 0.33333333 0.52542373 0.50000000  
## 84 0.13888889 0.58333333 0.10169492 0.04166667  
## 85 0.69444444 0.33333333 0.64406780 0.54166667  
## 86 0.33333333 0.62500000 0.05084746 0.04166667  
## 87 0.16666667 0.45833333 0.08474576 0.00000000  
## 88 0.08333333 0.45833333 0.08474576 0.04166667  
## 89 0.25000000 0.29166667 0.49152542 0.54166667  
## 90 0.36111111 0.20833333 0.49152542 0.41666667  
## 91 0.38888889 0.75000000 0.11864407 0.08333333  
## 92 0.33333333 0.16666667 0.45762712 0.37500000  
## 93 0.41666667 0.83333333 0.03389831 0.04166667  
## 94 0.19444444 0.00000000 0.42372881 0.37500000  
## 95 0.58333333 0.50000000 0.72881356 0.91666667  
## 96 0.66666667 0.45833333 0.77966102 0.95833333  
## 97 0.75000000 0.50000000 0.62711864 0.54166667  
## 98 0.41666667 0.29166667 0.69491525 0.75000000  
## 99 0.61111111 0.41666667 0.71186441 0.79166667  
## 100 0.13888889 0.41666667 0.06779661 0.08333333  
## 101 0.19444444 0.41666667 0.10169492 0.04166667  
## 102 0.52777778 0.08333333 0.59322034 0.58333333  
## 103 0.80555556 0.41666667 0.81355932 0.62500000  
## 104 0.02777778 0.50000000 0.05084746 0.04166667  
## 105 0.30555556 0.79166667 0.05084746 0.12500000  
## 106 0.52777778 0.58333333 0.74576271 0.91666667  
## 107 0.25000000 0.62500000 0.08474576 0.04166667  
## 108 0.25000000 0.58333333 0.06779661 0.04166667  
## 109 0.27777778 0.70833333 0.08474576 0.04166667  
## 110 0.91666667 0.41666667 0.94915254 0.83333333  
## 111 0.33333333 0.91666667 0.06779661 0.04166667  
## 112 0.55555556 0.54166667 0.84745763 1.00000000  
## 113 0.55555556 0.12500000 0.57627119 0.50000000  
## 114 0.66666667 0.41666667 0.71186441 0.91666667  
## 115 0.38888889 0.33333333 0.59322034 0.50000000  
## 116 0.16666667 0.41666667 0.06779661 0.04166667  
## 117 0.72222222 0.45833333 0.69491525 0.91666667  
## 118 0.02777778 0.41666667 0.05084746 0.04166667  
## 119 0.50000000 0.33333333 0.50847458 0.50000000  
## 120 0.47222222 0.08333333 0.50847458 0.37500000  
## 121 0.47222222 0.29166667 0.69491525 0.62500000  
## 122 0.55555556 0.58333333 0.77966102 0.95833333  
## 123 0.16666667 0.16666667 0.38983051 0.37500000  
## 124 0.58333333 0.33333333 0.77966102 0.87500000  
## 125 0.44444444 0.41666667 0.69491525 0.70833333  
## 126 0.30555556 0.79166667 0.11864407 0.12500000  
## 127 0.13888889 0.58333333 0.15254237 0.04166667  
## 128 0.30555556 0.58333333 0.11864407 0.04166667  
## 129 0.44444444 0.50000000 0.64406780 0.70833333  
## 130 0.58333333 0.33333333 0.77966102 0.83333333  
## 131 0.38888889 0.37500000 0.54237288 0.50000000  
## 132 0.66666667 0.54166667 0.79661017 0.83333333  
## 133 0.05555556 0.12500000 0.05084746 0.08333333  
## 134 0.08333333 0.58333333 0.06779661 0.08333333  
## 135 0.16666667 0.66666667 0.06779661 0.00000000  
## 136 0.66666667 0.41666667 0.67796610 0.66666667  
## 137 0.22222222 0.58333333 0.08474576 0.04166667  
## 138 0.38888889 0.20833333 0.67796610 0.79166667  
## 139 0.36111111 0.41666667 0.59322034 0.58333333  
## 140 0.72222222 0.45833333 0.74576271 0.83333333  
## 141 0.22222222 0.70833333 0.08474576 0.12500000  
## 142 0.41666667 0.29166667 0.49152542 0.45833333  
## 143 0.50000000 0.25000000 0.77966102 0.54166667  
## 144 0.80555556 0.66666667 0.86440678 1.00000000  
## 145 0.22222222 0.62500000 0.06779661 0.08333333  
## 146 0.25000000 0.87500000 0.08474576 0.00000000  
## 147 0.30555556 0.70833333 0.08474576 0.04166667  
## 148 0.22222222 0.75000000 0.15254237 0.12500000  
## 149 0.61111111 0.41666667 0.76271186 0.70833333  
## 150 0.63888889 0.41666667 0.57627119 0.54166667

normalized\_iris\_dataset$Species <- iris\_dataset$Species  
  
training\_dataset <- normalized\_iris\_dataset[1:100,]  
test\_dataset <- normalized\_iris\_dataset[101:150,]  
  
model <- naiveBayes(Species ~ . , data=training\_dataset)  
  
predictions <- predict(model, test\_dataset)  
actual\_labels <- test\_dataset[,5]  
#macierz błedów  
table(predictions,actual\_labels)

## actual\_labels  
## predictions A B C  
## A 21 0 0  
## B 0 10 0  
## C 0 3 16

#dokładność modelu  
mean(actual\_labels==predictions)

## [1] 0.94