Machine Learning - 1100-ML0ENG (Ćwiczenia informatyczne Z-23/24)

Creating a basic plot

R is known to be a really powerful programming language when it comes to graphics and visualizations in addition to statistics and data science.

Graphics in R can be done in three ways

1. graphics package (the base graphics in R, loaded by default)
2. lattice package which adds more functionalities to the base package
3. ggplot2 package (which needs to be installed and loaded beforehand)

The graphics package comes with a large choice of plots (such as plot, hist, boxplot, pie, etc.). It is often the preferred way to draw plots for most R users, and in particular for beginners to intermediate users.

Data

```
house <- read.csv("http://imul.math.uni.lodz.pl/~bartkiew/ml/data/HousePrices.csv")
summary(house)
house <-
read.csv("http://imul.math.uni.lodz.pl/~bartkiew/ml/data/HousePrices.csv",stringsAsF
actors = TRUE)
summary(house)
str(house)
```

The plot() function

In R, the base graphics function to create a plot is the **plot() function**. It has many options and arguments to control many things, such as the **plot type, labels, titles and colors**.

```
plot(house$SqFt,house$Price)
```

Adding Titles and Axis Labels

You can add your own title and axis labels easily by specifying following arguments.

- main - Main plot title
- xlab - x-axis label
- ylab - y-axis label

```
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price",
 main = "Houses")
```

Change the shape, size and colors of the Points

The **pch (plotting character)** argument to specify symbols to use when plotting points.

Here's a plot of symbols you can use.

```
x=1:22
plot(x, pch=x )
```

You can change the foreground color of symbols using the argument **col**.

```
plot(x, col=x)
```

R has a number of predefined colors that you can use in graphics. Use the **colors()** function to get a complete list of available names for colors.

```
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = 15, col = 5)
```

```
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = 16, col = "tomato3")
```

For symbols **21 through 25**, you can specify border color using **col** argument and fill color using **bg** argument.

```
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = 22, col="blue", bg="lightblue")
```

To alter the size of the plotted characters, use **cex (character expansion)** argument.

```
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = 22, col="blue", bg="lightblue" , cex=1.8)
```

You can use information from a dataset to mark the characteristics of a symbol

```
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = house$Bedrooms)
```

```
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = house$Bedrooms+13, col = house$Bathrooms)
```

```
limit<-200000
house$p<-house$Price
house$p[house$Price <= limit]<-1
house$p[house$Price > limit]<-2
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = house$Bedrooms+13, col = house$Bathrooms, cex =house$p)
```

```
house$nei<-house$Neighborhood
levels(house$nei)<-c(1,2,3)
plot(house$SqFt,house$Price,
 xlab = "Area",
 ylab = "Price", main = "Houses",
 pch = house$Bedrooms+13, col = as.numeric(house$nei),cex=1.5)
```

Last modified: środa, 11 października 2023, 9:19

Accessibility settings

Przetwarzanie danych osobowych

Platformą administruje Komisja ds. Doskonalenia Dydaktyki wraz z Centrum Informatyki Uniwersytetu Łódzkiego Więcej

Informacje na temat logowania

Na platformie jest wykorzystywana metoda logowania za pośrednictwem Centralnego Systemu Logowania.

Studentów i pracowników Uniwersytetu Łódzkiego obowiązuje nazwa użytkownika i hasło wykorzystywane podczas logowania się do systemu USOSweb.

Deklaracja dostępności