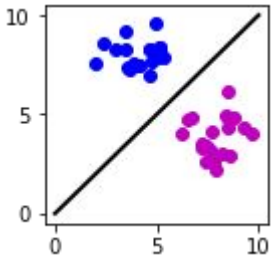
Data Science Survival Skills

Homework 7

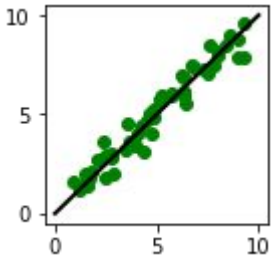
Description of the Homework

In this homework assignment, you will work with several data sets that we provide via StudOn. Your task is to fit regression lines and a decision boundary. In the exercise we will give you some methods, but you are free to use other methods as well.

CLASSIFICATION



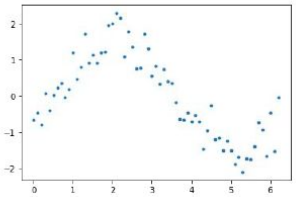
REGRESSION



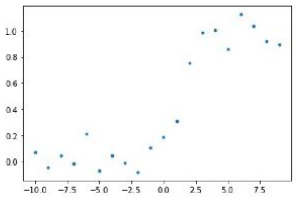
Homework 7: Tasks 1/2

- Download the data from StudOn, load it in your python script or Jupyter Notebook, and analyze it.
- Ponder the question: What kind of data do I have? Maybe a noisy signal?

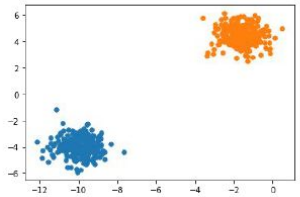
➜ **Slide: Your answers for each of the datasets (regression_1, regression_2 and classification)**



regression_1



regression_2



classification

Homework 7: Task 2/2

- For each of the three given datasets implement one machine learning method that performs regression or classification to fit a line/decision boundary for the data.
- Plot each dataset together with your predicted line/decision boundary in one plot (in total there should be **three** plots then ➜ see the example).
- Explain why your fitted line/decision boundary is appropriate for the given data. Also explain a little about what you did and what your thoughts were.

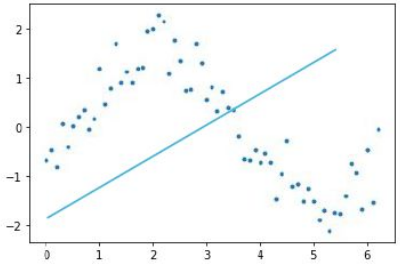
➜ **Slide: Your code where we can see your implementation of your machine learning method**

➜ **Slide: Your plots**

➜ **Slide: Your explanations**

Homework 7: Example

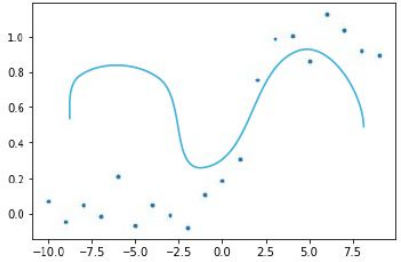
a) The data is a non-noisy linear function.



```
# Here comes YOUR code doing whatever you
think is adequate for this task.
# Plot
plt.plot(,   'ko', label = 'original signal or
data')
plt.plot(some data  label = 'results obtained
using X method' )
plt.legend()
# Use some labels: remember Homework 3
```

a) I fitted a linear function with values, slope, etc. Therefore, I maximized the distance between the given datapoints and my linear function.

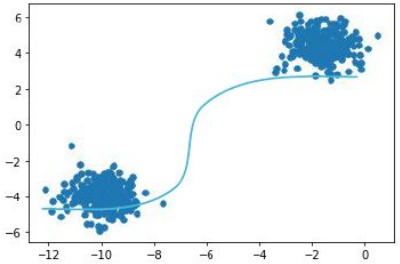
b) The data is a noisy ECG signal.



```
# Here comes YOUR code doing whatever you
think is adequate for this task.
# Plot
plt.plot(,   'ko', label = 'original signal or
data')
plt.plot(some data  label = 'results obtained
using X method' )
plt.legend()
# Use some labels: remember Homework 3
```

b) Values, slope, type of function.

c) The data is…

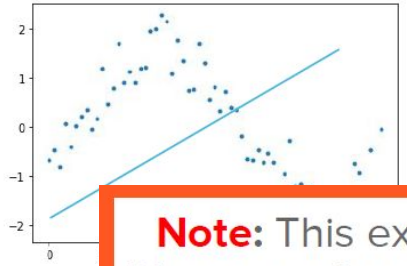


```
# Here comes YOUR code doing whatever you
think is adequate for this task.
# Plot
# Do you want to use plt.plot()?
plt.legend()
# Use some labels: remember Homework 3
```

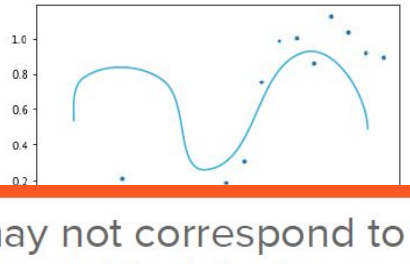
c) Values, slope, type of function.

Homework 7: Example

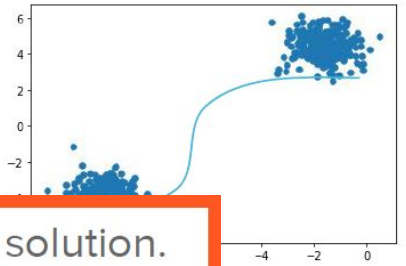
a) The data is a non-noisy linear function.



b) The data is a noisy ECG signal.



c) The data is…



Note: This example may not correspond to the actual solution. Please evaluate the data and find the best possible answer. If you lack space, do not paste the code for plotting, the most significant part is solving the task, i.e., the machine learning algorithms.

```
# Here comes Y
think is adequ
# Plot
plt.plot(,  'k
data')
plt.plot(some
using X method
plt.legend()
# Use some labels: remember Homework 3
```

a) I fitted a linear function with values, slope, etc. Therefore, I maximized the distance between the given datapoints and my linear function.

```
plt.legend()
# Use some labels: remember Homework 3
```

b) Values, slope, type of function.

```
g whatever you
task.
lot()?
Homework 3
```

c) Values, slope, type of function.

Homework: Requirements

You must complete **all** homework assignments (**unless otherwise specified**) following these guidelines:

- **One** slide/page.
- **PDF** file format only.
- It has to contain your **name, student** (**matriculation**) **number** and **IdM** in the down-left corner.
- Font: **Arial,** Font-size: **> 10 Pt.**
- Answer **all** the questions and solve all the tasks requested.
- Be careful with **plagiarism**. Repeated solutions will not be accepted!