

Submission Deadline: 22.05.2020 at 10:00 CEST**1. Before you get started, set up ILIAS**

This first assignment mostly sets up the infrastructure we will use in the rest of the course. Please execute the following steps:

1. Set up our virtual machine as described here.
2. If you haven't already done so, please also set up ILIAS:
 - (a) Log in to ILIAS (using your RZ account + password for authentication).
 - (b) Download the input data for this assignment from ILIAS by opening the drop-down menu of assignment 1 in the assignment folder and unzip it to a new folder. This already contains all other notebooks referenced in this assignment.
3. We recommend that you work on the exercises in groups of two or three. Once you have found your team partners, please inform us about your team by adding the RZ account names to the corresponding list in the ILIAS wiki. (Even if you want to work alone, please put your single name into the list.)

2. Make a forum post about a machine learning application

To get into the habit of posting frequently, let's make a forum post in ILIAS right away:

- Find an interesting application of machine learning, post a link to it and describe it in 2 sentences.
- Please also feel free to comment on each others' posts.

3. Go through our tutorial notebooks

We provide a little tutorial for you to grow more familiar with the basics necessary to successfully apply machine learning:

- Go through our introductory tutorial on Jupyter notebooks: `IntroductionToJupyterNotebooks.ipynb` (download from ILIAS)
- Go through our introductory tutorial on data loading and plotting: `IntroductionToReadingAnd-PlottingData.ipynb` (download from ILIAS)

4. Manual data analysis

In this exercise, we'll use the Iris data set from our intro tutorial on data loading and plotting. We have prepared a fully functional Jupyter notebook called `question_04.ipynb` (download from ILIAS), to let you focus on the machine learning part. (But you can also do this exercise in a python script if you wish.)

The notebook contains three sections:

1. loading the data
2. classifying the data
3. testing the data on an unseen test set

Our current implementation classifies each data point at random; it is your task to define a set of rules which improves upon this. You can use insights gathered by going through our tutorial notebooks (question 3) in order to come up with a useful set of rules.

5. Setup GitLab

Each team will have a private repository on our GitLab server with the name of the corresponding team members (or individual name if you work alone). There you will be able to submit your solution for each assignment as instructed below.

Register with our submission system GitLab (using the account that we email to you after you fill in the information in the Wiki in question 1.3). Check out a working copy of your folder in the git repository of the course, and add your code to a new subdirectory `assignment-01`, and commit it.

6. Refresh your linear algebra

You all have different backgrounds, and in this first exercise we're trying to level the playing field a bit. If, for example, you can't explain what eigenvalues and eigenvectors are, and how to compute them, we recommend that you refresh your linear algebra using, e.g., <https://www.khanacademy.org/>).

7. Your first code submission

Submit your answer.

Compute the matrix product

$$\begin{bmatrix} 5 & 8 \\ 2 & 1 \\ 6 & 7 \end{bmatrix} \cdot \begin{bmatrix} 5 & 8 & 3 \\ 2 & 1 & 9 \end{bmatrix}$$

and find the eigenvalues of the matrix

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

in a Jupyter notebook or python script. Add, commit and push it to the same subdirectory as above in your repository.

8. Submit feedback

Also commit, in that subdirectory, a text file called feedback.txt, in which you briefly describe your experiences with this first assignment and the corresponding lecture(s). As a minimum, say how much time you invested and if you had major problems, and if yes, where. If any question was particularly useful or not useful, please also let us know.

9. General comments on the coding exercises

These will be valid throughout the course:

1. Make sure that you do not upload any "by-products" to our git (e.g., class files or executables or any stuff from your local environment). Also never ever upload large data files to our git; this is considered sin and will result in excommunication.
2. When you encounter implementation problems proceed as follows. First do the obvious Google search: very often, this leads to a page describing exactly the problem you are having and the solution for it. Then search/read our forum (link on the Wiki) to see if someone has already asked a similar question. Then you are very welcome to ask a question on the forum yourself. In general: please ask before you spent a lot of time on minor implementation issues.