1 Data loading and preparation

Download the data using the link on the e-learning site of the course. On LinuxLab computers, pass the provided link to the wget command to download the file. Unzip the file, it contains a directory called DATA.

2 Data to numpy

The DATA directory contains a lot of files, all with filenames of the form "X_Y.png", where X is the class and Y is a global file number. From these files, create numpy arrays X and T. Hints (google these functions!):

- use the method os.listdir() in the *os* module to get a list of file names in a directory!
- use the string method ""split" to process the filenames
- use the function "imread" from the *imageio* module to read a png file into a numpy array (you may need to install that first: *python3 -m pip install imageio*).

3 Exploring the data 1

Using numpy: compute the min and max values of each image pixel over all images/samples.

4 Exploring the data 2

Using numpy: compute the occurrences of each class in the data

5 Exploring the data 3

Using numpy: check whether the data are block-sorted and shuffle them if necessary!