AI1072: Machine learning, exercise sheet 2

MNIST can be obtained via

http:// filr .hs−fulda . de / ssf / s / readFile / share /206/ −8986560519763980414/ publicL

numpy& matplotlib with MNIST

Load the MNIST data as follows:
import gzip, pickle ;
with gzip.open("MNIST.pkl.gz") as f:
 data = pickle.load(f) ;
 traind = data["data_train"]
 trainl = data["labels_train"]

'traind' contains the samples, 'trainl' the target values (labels) in one-hot format.
a) Inspect both arrays and especially print out their shape. How many samples and targets are there? How many numbers constitute a sample? Why are there 10 dimensions to a single target value?
b) Slice out the 1000th sample and disaplay it. Print its class as well!
c) Display a histogram over the classes of all samples!
d) Print the number of samples having class 5
e) Display the classes of the first 100 samples as a bar plot!
f) Generate the following variations of the 10th sample and display them together:
- flip along x axis
- flip along y axis
g) Extract from the training data just the samples having class 4 and display the last 4 of them!
h) Copy out 100 randomly chosen samples and display the first 3!

Matplotlib

a) plot the sine function between -5 and 5 using 100 support points! Use numpy.sin or math.sin to generate the values!
b) generate a scatter plot of the same data as in a)!
c) generate a bar plot of the same data as in a)!
d) plot sine and cosine together in a single plot