

# Introduction to Machine Learning, Fall 2014 - Exercise session III

Rodion “rodde” Efremov, student ID 013593012

November 10, 2015

## **Problem 1 (3 points)**

## **Problem 2 (3 points)**

## **Problem 3 (3 points)**

## **Problem 4 (9 points)**

**(a)**

Download the MNIST data from the course web page. In addition to the actual data, the package contains some functions for easily loading the data into Python/Matlab/Octave/R and for displaying digits. See the README files for details. Load the first  $N = 5,000$  images using the provided function.

## **Solution to (a)**

Check!

**(b)**

Use the provided functions to plot a random sample of 100 handwritten digits, and show the associated labels. Verify that the labels match the digit images. (This is a sanity check that you have the data [is] in the right format.)

Solution to (b)

0	2	3	9	2	0	6	6	8	6
9	4	7	5	4	0	4	6	3	8
7	2	4	5	8	1	1	8	7	6
9	5	4	6	3	1	3	4	9	3
0	5	4	2	2	7	8	7	5	4
5	6	3	7	2	8	9	0	5	9
0	9	8	1	6	9	3	6	5	8
9	9	6	8	4	7	2	4	4	4
7	8	0	0	8	7	8	2	0	8
6	3	0	1	8	7	3	8	2	1

The labels are:

```

0 2 3 9 2 0 6 6 8 6
9 4 7 5 4 0 4 6 3 8
7 2 4 5 8 1 1 8 7 6
9 5 4 6 3 1 3 4 9 3
0 5 4 2 2 7 8 7 8 4
5 6 3 7 2 8 9 0 5 9
0 9 8 1 6 9 3 6 5 8
9 9 6 8 4 7 2 4 4 4
7 8 0 0 8 7 8 2 0 8
6 3 0 1 8 7 3 8 2 1

```

(c)

Divide the data into two parts: A ‘training set’ consisting of the first 2,500 images (and associated labels), and a ‘test set’ containing the remaining 2,500 images (and their associated labels).

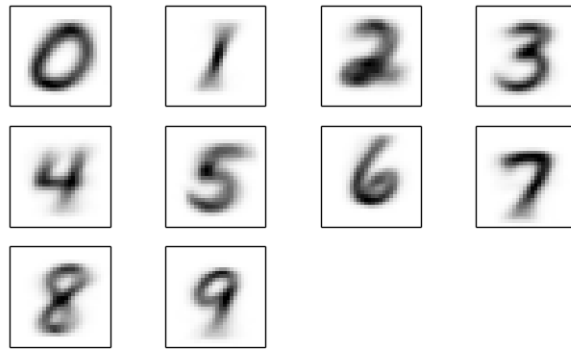
Solution to (c)

Check!

(d)

For each of the ten classes (digits 0-9), compute a class *prototype* given by the mean of all the images in the training set that belong to this class. That is, select from the training set all images of class '0' and compute the mean image of these; this should look sort of like a zero. Do this for all ten classes, and plot the resulting images. Do they look like what you would expect?

**Solution to (d)**



Obviously, the prototypes look good.