Machine Learning
2nd-Term Semester 2016-2017
Assignment 2

SYM

October 12, 2017

General instructions: All course participants are requested to handle their exercise solutions as follows:

- Write your answers as PDF using one of the following text processing tools: MS-Word, LibreOffice, or Latex.
- Always mention your **name** and **student ID** in the PDF file.
- For programming section, write the source code using programming language that you prefer or familiar with.
- The assignment is designed to be solved in a week. However you could accomplish it less than a week when you allocate your time properly to work on it.
- The deadlines are as follows,
  - CLO1: Friday 20.10.2017 at 21.00 UTC+7,
  - CLO2: Tuesday 24.10.2017 at 21.00 UTC+7,
  - CLO3, exercise 17: Friday 27.10.2017 at 21.00 UTC+7,
  - CLO3, exercise 18: Tuesday 31.10.2017 at 21.00 UTC+7,
  - CLO3, exercise 19: Friday 3.11.2017 at 21.00 UTC+7,
- Submit your work through email to the lecturer (**sym.milo.at.gmail.com**) before the deadlines.
- For CLO3, submit your work (PDF and all your codes) into one directory (as a **zip** file) Do not include any of the data files in your solution file.
- **All forms of cheating are strictly prohibited**.

Section 1: CLO1 (Totally 70 points)

1. **(5 points)** Explain about univariate linear regression.
2. **(5 points)** Explain about univariate non-linear regression.
3. **(5 points)** Explain about multivariate linear regression.
4. **(5 points)** Explain about multivariate non-linear regression.
5. **(5 points)** Explain the advantages of Multi-Layer Perceptron (MLP).
6. **(5 points)** Explain the disadvantages of MLP.
7. **(5 points)** Explain the similarities between MLP and Support Vector Machine (SVM).
8. **(5 points)** Explain the differences between MLP and SVM.

9. **(5 points)** Describe the influences of number of neurons to the MLP model.
10. **(5 points)** Describe the influences of type of activation function to the MLP model.
11. **(5 points)** Describe the influences of learning rate to the MLP model.
12. **(5 points)** Explain how MLP obtains optimal decicion boundary.
13. **(5 points)** Explain how SVM obtains optimal hyperplane.
14. **(5 points)** Explain how SVM classifies data sets that are non-linearly separable.

Section 2: CLO2 (Totally 110 points)

Instructions: In all the exercises, do not just give your answer, but also the derivation of how you obtained it.

15. Given data set 1 (x_i, y_i) shown in Table 1.

Table 1: Data set 1.

ID	x	y
p1	1.9	2.6
p2	4.1	4.5
p3	5.1	4.1
p4	6.0	?
p5	6.9	3.2
p6	7.3	3
p7	7.9	1.8

(a) **(20 points)** Build a univariate linear regression model using data set 1 (except p4).

(b) **(5 points)** Predict the y_4 that is the y value of p6 using univariate linear regression model in 15(a).

(c) **(20 points)** Build a univariate non-linear regression model using data set 1 (except p4). (Hints: use m degree polynomial. It is up to you in selecting m value.)

(d) **(5 points)** Predict the y_4 that is the y value of p4 using univariate non-linear regression model in 15(c).

(e) **(5 points)** Between two models resulted from 15(a) and 15(b), which model that gives better prediction to p4? Why? Give your explanation.

16. Given data set 2 ($\mathbf{x}_i$, y_i), where $\mathbf{x}_i = (x_1, x_2)$, shown in Table 2.

Table 2: Data set 2.

ID	x_1	x_2	y
p1	3.2	5.8	3.2
p2	4.1	5.1	4.5
p3	5.1	4.0	4.1
p4	6.0	3.1	?
p5	6.9	2.2	3.2
p6	7.3	1.7	3
p7	7.9	0.9	1.8

(a) **(20 points)** Build a multivariate linear regression model using data set 2 (except p4).

(b) **(5 points)** Predict the y value of p4 using model in 16(a).

(c) **(20 points)** Build a multivariate non-linear regression model using data set 2 (except p4). (Hints: use some interaction terms as given in the slide of Regression page 17.)

(d) **(5 points)** Predict the y value of p4 using univariate non-linear regression model in 15(c).

(e) **(5 points)** Between two models resulted from 16(a) and 16(b), which model that gives better prediction to p4? Why? Give your explanation.

Section 3: CLO3 (Totally 105 points)

Instructions:

- Write a report (as PDF) of this section.
- We use the report as the main basis for grading: All your results should be in the report. We also look at the code, but we won't however go fishing for results in your code.
- The code needs to be submitted as a runnable file or set of files (command to run it given in the report).
- In your report, the results will be mostly either in the form of a figure or program output. In both cases, add some sentences which explain what you are showing and why the results are the answer to the question.
- If we ask you to test whether an algorithm works, always give a few examples showing that the algorithm indeed works

We have 10 datasets for each of exercises (17, 18, and 19). Each data set has 3 columns where 1st and 2nd columns are attributes while 3rd column is a class label.

For exercise 17, the datasets are linear_0, linear_1, ..., linear_9. For exercise 18, the datasets are non-linear_0, non-linear_1, ..., non-linear_9. For exercise 19, the datasets are (0) Aggregation, (1) Compound, (2) D31, (3) Flame, (4) Heart-1, (5) Heart-2, (6) Jain, (7) Pathbased, (8) R15, and (9)Spiral.

Use your student ID to select the data sets as follows: let t is **the last digit of your student ID**, then use the data set linear_t, non-linear_t, and (t) for exercise 17, 18, and 19, respectively.

17. In this exercise we will implement SVM for linearly separable data.

    (a) **(5 points)** Load the selected data set. Visualize all data points using scatter plot. Use different color or symbol for each class. Use attribute 1 as x -axis, attribute 2 as y -axis.

    (b) **(20 points)** Using quadratic programming library, find $\mathbf{w}$ and b that construct the hyperplane for classifying the selected data set.

    (c) **(5 points)** Now visualize the hyperplane on the scatter plot that is created by 17(a).

18. In this exercise we will implement SVM for non-linearly separable data.

    (a) **(5 points)** Load the selected data set. Visualize all data points using scatter plot. Use different color or symbol for each class. Use attribute 1 as x -axis, attribute 2 as y -axis.

    (b) **(10 points)** Create a function that implements polynomial kernel. Inputs for the function are attributes (x_1 and x_2). The function transforms data from original feature space into new feature space. Here we transform data from 2 dimensional space into 3 dimensional space, $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$,that is transforming $\mathbf{x} = (x_1, x_2)$ into $\mathbf{z} = (z_1, z_2, z_3) = (x_1, x_2, x_1^2 + x_2^2)$.

    (c) **(5 points)** In a new feature space, $\mathbb{R}^3$, visualize all data points using scatter plot in one color, and use different symbols for each class, for example class '-1' uses symbol 'o' while class '1' uses symbol '+').

    (d) **(20 points)** Using quadratic programming library, find the vector $\mathbf{w}$ and b that construct the hyperplane for classifying the data set in $\mathbb{R}^3$.

    (e) **(5 points)** Now visualize the hyperplane on the scatter plot that is created by 18(d).

19. In this exercise we will implement Probabilistic Neural Network (PNN) to classify data.

    (a) **(5 points)** Load the selected data set. Visualize all data points using scatter plot. Use different color or symbol for each class. Use attribute 1 as x -axis, attribute 2 as y -axis.

    (b) **(5 points)** Select randomly three data for test set ($\mathbf{x}'$, y'), while remaining data as training set ($\mathbf{x}$, y).

    (c) **(15 points)** Create a function that implements PNN. Inputs for the function are training set ($\mathbf{x}$, y) and the attributes $\mathbf{x}'$ of test set. The function outputs the predicted class $\hat{y}$ for test set.

    (d) **(5 points)** How good is the classification result on test set? Give your opinion.