11020EE655000　Machine Learning

Homework 2

Announced: April 14 at 01:00 pm

Deadline:　　May 5 at 11:59 pm (eeclass)

TA: 陳冠宇 {seed0123456@gmail.com}

　　　林辰安 {strike860930@gmail.com}

　　　賴清元 {y8574317@gmail.com}

Grading Policy:

1. In the handwriting assignment, you need to provide detailed derivations. Partial points will be credited when a wrong answer is accompanied by correct reasoning. Please hand in the handwriting assignment in class on 5/5.

2. In the programming assignment, there are two files you need to prepare: the code (hw2.py) and report (hw2report_studentID.pdf) and they should be compressed into a ZIP file (hw2_studentID.zip) then uploaded to eeclass website.

3. The only programming language that can be used is Python. Built-in machine learning libraries or functions are NOT allowed to use.

4. Discussions are encouraged, but plagiarism is strictly prohibited (changing variable names, etc.). You can use any open source with clearly mentioned in your report. If there is any plagiarism, you will get 0 in this homework.

Part 1. Handwriting

You can find the corresponding problems from the textbook.

1. (6 points) Exercise 4.7

2. (8 points) Exercise 4.13

3. (6 points) Exercise 4.14

Part 2. Implementation

In this problem, you need to apply the Maximum Likelihood (ML) and Bayesian linear regression methods to train a linear model in order to predict the chance of being admitted to graduate admissions.

Data

Data, contained in the two csv files Training_set.csv and Validation_set.csv. In this problem is the prediction of being admitted to graduate admissions. More detailed descriptions are given below:

In the Training_set.csv total have 300 pieces of data

In the Validation_set.csv total have 100 pieces

  Column1: GRE Scores

  Column2: TOEFL Scores

  Column3: Research Experience (either 0 or 1)

  Column4: Chance of Admit (ranging from 0 to 1)

Feature Vector

In this problem, we utilize the Gaussian basis function and the Research Experience to form the feature vector, denote d as

$$\phi(\mathbf{x}) = [\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \ldots, \phi_P(\mathbf{x}), \phi_{P+1}(\mathbf{x}), \phi_{P+2}(\mathbf{x})]^T$$

where we place P Gaussian basis functions uniformly over the spatial domain with P = $O_1 \times O_2$, x = (x_1, x_2, x_3) is the input data (the scores together with the Research Experience), and O_1 and O_2 denote the number of locations along the horizontal and vertical directions, respectively, that you choose for your model in the prediction.

More specifically, for $1 \le k \le P$, the Gaussian basis function is defined as

$$\phi_k(\mathbf{x}) = \exp\left\{ -\frac{(x_1 - \mu_i)^2}{2s_1^2} - \frac{(x_2 - \mu_j)^2}{2s_2^2} \right\}, \quad \text{for} \quad 1 \le i \le O_1, 1 \le j \le O_2,$$

where

$$k = O_2 \times (i-1) + j$$

$$\mu_i = s_1 \times (i-1) + x_{1\_min}, \qquad \mu_j = s_2 \times (j-1) + x_{2\_min}$$

$$s_1 = \frac{x_{1\_max} - x_{1\_min}}{O_1 - 1}, \qquad s_2 = \frac{x_{2\_max} - x_{2\_min}}{O_2 - 1}$$

Finally, the last two components of the feature vector are $\varphi_{P+1}(x) = x_3$(Research Experience) and $\varphi_{P+2}(x) = 1$ (bias).

Problem

Please employ the linear model to predict the chance of being admitted to graduate admissions given in the testing_set.csv.

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=1}^{P+2} w_j \phi_j(\mathbf{x})$$

Take hw2_template.py as reference, there should be at least two functions BLR() and MLR() in your hw2.py.

In data evaluation, main() will call BLR() and MLR() to calculate the predicted chance of admit.

1) (20 points for implementation + 15 points for MSE Screenshot) Please use Maximum Likelihood and Least Squares to train the model. Then, use your trained linear model to predict the chance of admit and compute the mean squared error for each data in Validation_set to get MSE < 0.01.

2) (20 points for implementation + 15 points for MSE Screenshot) Please use Bayesian Linear Regression to estimate w. Then, use your estimated parameter to predict the chance of admit and compute the mean squared error for each data in Validation_set to get MSE < 0.01.

   Mean squared error:

   $$\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2$$

3) (10 points) Please discuss the difference between Maximum Likelihood and Bayesian Linear Regression, and the impact of different choices of O1 and O2 and results in your report.