4105931 機器學習

Assignment #2 – Regression
Deadline : 2023/4/11 11:59 pm

1. Use the linear model $y = 2x + \varepsilon$ with zero-mean Gaussian noise $\varepsilon \sim N(0, 1)$ to generate 15 data points with (equal spacing) $x \in [-3, 3]$.

2. Perform Linear Regression. Show the fitting plot, the training error, and the five-fold cross-validation errors.

3. Perform Polynomial Regression with degree 5, 10 and 14, respectively. For each case, show the fitting plot, the training error, and the five-fold cross-validation errors. (Hint: Arrange the polynomial regression equation as follows and solve the model parameter vector w.)

$$y = \begin{bmatrix} x_1^5 & x_1^4 & x_1^3 & x_1^2 & x_1^1 & 1 \\ x_2^5 & x_2^4 & x_2^3 & x_2^2 & x_2^1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{15}^5 & x_{15}^4 & x_{15}^3 & x_{15}^2 & x_{15}^1 & 1 \end{bmatrix} \bullet \begin{bmatrix} w_5 \\ w_4 \\ w_3 \\ w_2 \\ w_1 \\ w_0 \end{bmatrix}$$

4. Change the model to $y = \sin(2\pi x) + \varepsilon$ with the noise $\varepsilon \sim N(0, 0.04)$ and (equal spacing) $x \in [0, 1]$. Then repeat those stated in 2) and 3). Compare the results with linear/polynomial regression on different datasets.

5. Following 4), perform polynomial regression with degree 14 by varying the number of training data points $m = 10, 80, 320$. Show the five-fold cross-validation errors and the fitting plots. Compare the results to those in 4).

6. Following 4), perform polynomial regression of degree 14 via **regularization**. Compare the results by setting $\lambda = 0, 0.001/m, 1/m, 1000/m$, where $m = 15$ is the number of data points (with $x = 0, 1/(m-1), 2/(m-1), \ldots, 1$). Show the five-fold cross-validation errors and the fitting plots.

Note:

• The assignment should be implemented by Python.
• You need to hand in the python code and the report.
• In your report, it should contain: (請以中文撰寫)
• **Execution description**: steps how to execute your codes.
• **Experimental results**: As specified in the assignment.
• **Conclusion**: The observation from your results.
• **Discussion**: The questions or the difficulties you met during the implementation.
• Assignment format
    • Zip all your files into a single one and upload it to the E-Course2 website.
• Please format the file name as: Student ID_proj1_verNo, ex：602410143_proj1_v1 • No copy! Late policy applies.