Homework 1

MTH 496 – Machine Learning
Due date: Friday, Sept 30, 2022

(5 problems/2 pages)

1 Handwritten Homework

Note Submit your homework on D2L → Assessments → Assignment. You may submit a pdf file, a markdown file, or a jupyter notebook for this section.

Problem 1 (10pts). Assume the training data is given as follows: $(x_1^{(1)}, x_2^{(1)}, y^{(1)})$, $(x_1^{(2)}, x_2^{(2)}, y^{(2)})$, $\ldots$, $(x_1^{(M)}, x_2^{(M)}, y^{(M)})$. The model of the linear regression takes the form

$$p_{\mathbf{c}}(x) = c_0 + c_1 x_1 + c_2 x_2$$

a) What is the loss function associated with $p_{\mathbf{c}}(x)$.

b) What is the gradient of the loss function.

c) Find the optimal values c_0, c_1, and c_2 (**Note: show all of your steps to receive a full credit.**).

Problem 2 (10pts). Assume the training data for the classification task is given as follows: $(x^{(1)}, y^{(1)})$, $(x^{(2)}, y^{(2)})$, $\ldots$, $(x^{(M)}, y^{(M)})$, with $y^{(i)} \in \{0, 1\}$, $i = 1, 2, \ldots, M$. The logistic regression is employed to learn this dataset.

a) What is the prediction for a given input x?

b) What is the loss function of the logistic regression method?

Problem 3 (10pts). a) What is the purpose of regularization?

b) State the loss functions of linear regression and logistic regression under regularization (choose any regularization method you like).

Problem 4 (5pts). a) What is the advantage of Stochastic Gradient Descent (SGD)?

2 Programming Homework

Note Submit your homework on D2L → Assessments → Assignment. Please only submit a jupyter notebook, and name it as `HW1_FirstName_LastName`. For example: `HW1_Xiaoqi_Wei.ipynb`

Problem 1 (65pts). Given the Iris dataset. It has been split into training data: `Iris_X_train.csv` (features), `Iris_y_train.csv` and test data `Iris_X_test.csv` (features), `Iris_y_test.csv`. File `Iris_description.txt` describes the meaning of each column in the data set.

a) Program a regularized logistic regression model to predict the test data. (sklearn.linear_model is not allowed to use.)

b) Comment on the choice of regularized parameters.