NYCU Pattern Recognition, Homework 1

Deadline: Mar. 22, 23:59

Part. 1, Coding (70%):

In this coding assignment, you are required to implement linear regression using only NumPy. Then, train your implemented model using **Gradient Descent** by the provided dataset, and test the performance with testing data.

Please note that only <u>NumPy</u> can be used to implement your model. Therefore, you will get no points by simply calling sklearn.linear_model.LinearRegression. Moreover, please train your linear model using <u>Gradient Descent</u>, not the closed-form solution.

Allowed packages: numpy, pandas, matplotlib.

(25%) Linear Regression Model - Single Feature

Requirements:

- Use the single feature (BMI) to train your linear regression model.
- Use MSE (Mean Square Error) as your loss function.
- Please use Gradient Descent. (Not the mini-batch GD nor SGD)
- Tune the learning rate and epoch to get the same result provided on slide page 9.

Criteria:

- 1. (0%) Show the learning rate and epoch you choose.
- 2. (5%) Show the weights and intercepts of your linear model.
- 3. (5%) What's your final training loss (MSE)?
- 4. (5%) What's the MSE of your validation prediction and validation ground truth?
- 5. (5%) Plot the training curve. (x-axis=epoch, y-axis=loss)
- 6. (5%) Plot the line (using red) you find with the training data (using blue) and validation data (using orange).

(45%) Linear Regression Model - Multiple Features

Requirements:

- Use all the 6 features to train your linear regression model.
- Use MSE (<u>Mean Square Error</u>) as your loss function.
- Please use Gradient Descent. (Not the mini-batch GD nor SGD)
- Tune the learning rate and epoch to get the same result provided on slide page 10.

Criteria:

- 7. (0%) Show the learning rate and epoch you choose.
- 8. (10%) Show the weights and intercepts of your linear model.
- 9. (5%) What's your final training loss (MSE)?
- 10. (5%) What's the MSE of your validation prediction and validation ground truth?
- 11. (5%) Plot the training curve. (x-axis=epoch, y-axis=loss)

12. (20%) Train your own model and save your final testing predictions in the csv file. Try different learning rates, epochs, batch_size, and do some data analysis to choose the feature you want to use).

Points	Test MSE
20	< 30000000
15	< 40000000
10	< 50000000
5	50000000 ~ 100000000
0	> 100000000

Part. 2, Questions (30%):

- (7%) 1. What's the difference between Gradient Descent, Mini-Batch Gradient Descent, and Stochastic Gradient Descent?
- (7%) 2. How do different values of learning rate (too large, too small...) affect the convergence of optimization? Please explain in detail.
- (8%) 3. Suppose you are given a dataset with two variables, X and Y, and you want to perform linear regression to determine the relationship between these variables. You plot the data and notice that there is a strong nonlinear relationship between X and Y. Can you still use linear regression to analyze this data? Why or why not? Please explain in detail.
- (8%) 4. In the coding part of this homework, we can notice that when we use more features in the data, we can usually achieve a lower training loss. Consider two sets of features, A and B, where B is a subset of A. (1) Prove that we can achieve a non-greater training loss when we use the features of set A rather than the features of set B. (2) In what situation will the two training losses be equal?