NYCU Introduction to Machine Learning, Homework 1

[Student ID]**, [**Name]

The screenshot and the figures we provided below are just examples. **The results below are not guaranteed to be correct.** Please make sure your answers are clear and readable, or no points will be given.

Please also remember to convert it to a pdf file before submission.

You should use English to answer the questions (-5pt if not report in English).

**After reading this paragraph, you should delete this paragraph.**

**Part. 1, Coding (60%)**:

**(10%) Linear Regression Model - Closed-form Solution**

1. (10%) Show the weights and intercepts of your linear model.



**(40%) Linear Regression Model - Gradient Descent Solution**

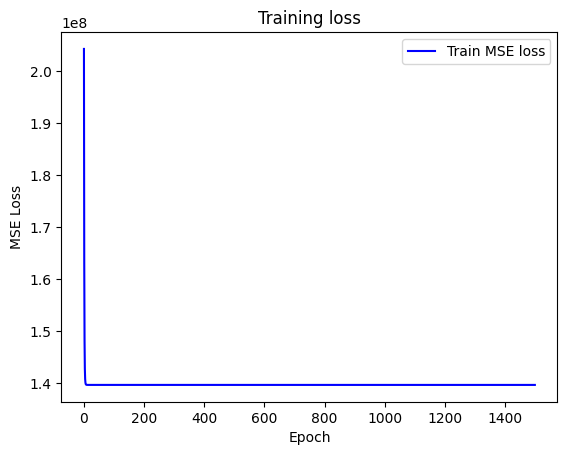
1. (10%)

* Show the hyperparameters of your setting (e.g., learning rate, number of epochs, batch size, etc.).
* Show the weights and intercepts of your linear model.





1. (10%) Plot the learning curve. (x-axis=epoch, y-axis=training loss)



1. (20%) Show your MSE.cf, MSE.gd, and error rate between your closed-form solution and the gradient descent solution.



**(10%) Code Check and Verification**

5. (10%) Lint the code and show the PyTest results.





**Part. 2, Questions (40%):**

1. (10%) Linear models 𝑦 = 𝑤⊤ 𝑥 + 𝑏 have limited fitting power.
2. In one sentence, explain why a single linear model is limited.
3. Give one concrete task that a single linear model cannot solve, and state why no single hyperplane/affine function solves it.

[Your answer here.]

1. (15%) Why do we add a regularization term in linear regression? What are the differences between L2 regularization (Ridge) and L1 regularization (Lasso)? Please explain in detail.

[Your answer here.]

1. (15%)

* What is overfitting? Under what conditions can a model overfit? (List two) How can overfitting be alleviated? (List two)

[Your answer here.]