Homework 1

Principles of Artificial Intelligence (5083)

12th October, 2023

Question 1.

(**Curious professor**) At the end of the semester, Jerry has finished marking everyone's score on the final exam. He is very curious about how to predict the final exam scores from the homework scores for future students. Moreover, he happens to have inside information about whether these students are admitted into NYCU master program. He is also curious about how to use the class performance for predicting the application results. Please help Jerry train machines to do the tasks.

Part 1: Linear Regression + Stochastic Gradient Descent (SGD) (35pt)

- Use the averaged homework scores and the scores of the final exam for 500 students in the attached dataset.
- Choose 400 data pairs as training dataset, the remaining 100 data pair as testing dataset.
- Employ linear regression with square loss for the problem.
- Update your weights and bias via SGD with $T = 1000$ iterations and learning rate $\eta = 0.01$.
- Plot a figure including the *testing* dataset and the result of linear regression obtained by the training data. An example is shown in Fig. 1.
- Compute the mean square error (MSE) of the testing dataset.

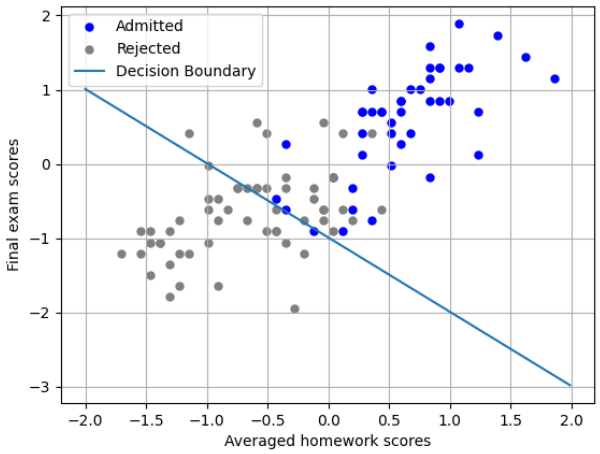

Figure 1: Example for problem 1. The horizontal axis represents the averaged homework scores. The vertical axis represents the scores for the final exams.

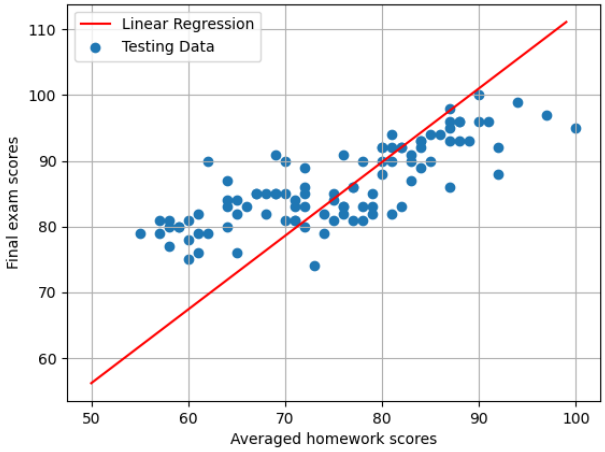

Figure 2: Example for problem 2. The horizontal axis represents the averaged homework scores. The vertical axis represents the scores for the final exams. Note that in this example, all scores are normalized to zero mean and unit variance.

Part 2: Logistic Regression + Stochastic Gradient Descent (SGD) (35pt)

- Use the homework scores, final exam scores and the labels of whether the students are admitted into NYCU master program.
- Choose 400 data pair as training dataset, the remaining 100 data pairs as testing dataset.
- Employ logistic regression with logistic loss for the problem.
- Update your weights and bias via SGD with $T = 1000$ iterations and learning rate $\eta = 0.75$.
- Apply the predictor you train to the *testing* dataset and compute the logistic loss.
- Make hard decisions based on the predicted probability and plot a figure including these hard decisions and the decision boundary obtained by the training data. An example is shown in Fig. 2.

Part 3: Hyperparameter selection (30pt)

- For the previous two parts, play around with various T, η, and mini-batch size and discuss your results.

Rules:

- Do not use high level toolbox/module functions. (e.g. sklearn, scipy)

Hints:

- For logistic regression, normalize your dataset to zero mean and unit variance before learning.
- TAs are your friends. Please **DO NOT** hesitate to contact the TAs if you have questions.

Grading: Please submit your program and a short report (less than **5 pages**). In your report, please provide the required figures as mentioned, followed by some discussions about these regressions. Your grades will be decided based on both the demo and report.