CS4990 Summer 2019 Project Assignment 1

Total points: 100

Due date: Monday, June 10, 2019

Purpose:

1. Warm up your Python programming skills.
2. Understand linear regression and gradient descent method.

Task Description:

Please use linear regression to predict house price in Boston suburbs. An iPython notebook is provided ("Assignment 1.ipynb"), where some of the early steps that prepare the training data and testing data have been implemented for you.

Please follow the instructions to implement and run each step. The specific tasks are

- Use LinearRegression() from the scikit-learn library
- Implement analytical solution (based on normal equation) to perform linear regression
- Implement numerical solution (based on gradient descent) to perform linear regression

In particular, in your gradient descent implantation, please tune the parameters to get close to the accuracy of the linear regression model from scikit-learn library.

In each task, you also need to

- Show all weights (intercept and coefficients)
- Show the resulting error $\ell(w) = \frac{1}{2N}\sum_{i=1}^{N}[t^{(i)} - y(x^{(i)})]^2$ on training data and test data, respectively.
- Use matplotlib to plot your predictions and real prices on the testing data.

What to Submit?

1. A completed iPython notebooks (Note: properly comment each code line)
2. Please zip them into a file (yourname_assignment1.zip) and submit the zipped file in blackboard