CS4210 Summer 2023 Project Assignment 2

Total points: 100

Due date: Friday, 07/28/2023

Purposes:

1. Develop neural networks using deep learning libraries
2. Apply neural networks to solve real-world problems

Task Description:

The California Housing dataset has contains 20,640 samples and 10 features (longitude, latitude, housing_median_age, total_rooms, total_bedrooms, population, households, median_income, ocean_proximity, and median house price). More details can be found at https://www.kaggle.com/datasets/camnugent/california-housing-prices

In this assignment, you will use PyTorch to build a neural network to predict "median house price" based on the other 9 features.

An iPython notebook is provided ("Assignment_2.ipynb"), where some of the early steps that prepare the training data and validation data have been implemented for you. Also, for comparison, the notebook contains the code of using scikit-learn's Linear Regression model to fit the data. (Note: due to the simplicity of this linear regression example, we just simply treat testing data as validation data.)

Please to implement the following tasks 1-2:

- **Task 1:** Using PyTorch's nn.Sequential() to build a neural network to fit the data. Note that it is ok to use any network structure (such as the number of layers, the number of neurons per layer, and activation functions) that you want.
- **Task 2:** Subclassing nn.Module to build a neural network (the same network structure as that of Task 1) to fit the data

What to Submit?

1. A completed iPython notebook for tasks 1-2 (Note: properly comment your programs)
2. Please zip them into a file (yourname_assignment2.zip) and submit the zipped file in Canvas