Exercise 8.1

Scikit-Learn includes some real world datasets. One of those is the California Housing dataset. It has 20640 samples and 8 predictive attributes and the target, which is the median house value for California districts, expressed in hundreds of thousands of dollars ($100,000). This data can be fetched with the function `sklearn.datasets.fetch_california_housing`.

Standardize the predictive data, i.e. the X, with `sklearn.preprocessing.StandardScaler`.

Split the data to train and test data with ratio 80/20. Use the seed value 827 in the `random_state` parameter of the function `sklearn.model_selection.train_test_split`.

Create a custom dataset and a dataloader for mini-batch prosessing. Use a batch size of 64 and don't use shuffling for the dataloader.

Create the following articifial neural network with PyTorch:

- input layer contains 8 values.

- hidden layer with 16 neurons with activation function ReLU. Initialize the weight values using the He initialization.

- hidden layer with 8 neurons with activation function ReLU. Initialize the weight values using the He initialization.

- output layer with 1 neuron and no activation.

Set the seed for neural network weight initialization to value 41 just before creating the model (neural network).

What is the maximum value for the weights in the second hidden layer? Give the answer rounded to two decimals.

[   ]

Use MSE as a loss function and Momentum as an optimizer. Note that PyTorch includes the SGD optimizer for implementing the Momentum optimization. Use a learning rate of 0.015 and a momentum of 0.9 for the SGD-optimizer.

Train the model with 10 epochs. Calculate and print the average loss for a epoch, i.e. total loss sum of every sample divided by the number of samples.

Note that for a batch the loss (MSELoss) object return the average loss with function `loss.item()`. So to get the loss sum of all samples in a batch you can multiply this average loss with batch size. And to get to total loss sum of every sample you can add the loss sums of batches.

What is the Loss value after the 10th epoch? Give the answer rounded to three decimals.

[   ]

Evaluate the model on the test dataset. What is the value of the test loss? Give the answer rounded to three decimals.

[   ]