Wargen Guittap 5004493060
Austin Merrick 2001003855

Assignment #2 Report

- The dataset you used, its source and characteristics.

Data source:
https://archive.ics.uci.edu/ml/datasets/Real+estate+valuation+data+set

We use the follow as features:
the house age (x2)
the distance to the nearest MRT station (x3)
the number of convenience stores within walking distance (x4)

To predict:
house price of unit area (Y)

- The solution "$\hat{w}$" for both algorithms.

Using Ordinary Least Squares:

$$\begin{bmatrix} 42.68254401355236 \\ -0.23779864 \\ -0.00543852 \\ 1.27951194 \end{bmatrix}$$

Intercept:
w0
42.68254401355236

Coefficents:

w1	w2	w3
[-0.23779864	-0.00543852	1.27951194]

Using Gradient Descent:

$$\begin{bmatrix} 37.87614143 \\ -2.70194556 \\ -6.75175513 \\ 3.8461217 \end{bmatrix}$$

Wargen Guittap 5004493060
Austin Merrick 2001003855

Intercept:
w0
37.87614143

Coefficents:

w1	w2	w3
[-2.70194556	-6.75175513	3.8461217]

- The learning rate(s) you used for gradient descent and how many iterations it took for gradient descent to converge.

Using an inverse scaling learning rate (.01/pow(t,power_t)), it converged after 25 iterations

- Relevant evaluation metrics for the training dataset for both algorithms.

Using Ordinary Least Squares:

The coefficient of determination R^2 of the prediction for the training data: 0.5157049069833045

Using Gradient Descent:

The coefficient of determination R^2 of the prediction for the training data: 0.5156990992106267

- Relevant evaluation metrics for the test dataset for both algorithms.

Using Ordinary Least Squares:

The coefficient of determination R^2 of the prediction for the test data: 0.6483532977702768

Using Gradient Descent:

The coefficient of determination R^2 of the prediction for the test data: 0.6484441468648924