# Programming Assignment: Linear Regression

## Machine Learning

## **Total points: 100**

**Note: This assignment is for each individual student to complete on his or her own.**

In this assignment, you will implement linear regression by using both normal equations and the gradient descent. To get started, you will need to download the starter code and unzip its contents to the directory where you wish to complete the assignment.

The problem considered in this assignment is to predict the price of a house using the real estate valuation data set downloaded from <https://archive.ics.uci.edu/ml/datasets/Real+estate+valuation+data+set>

You are required to use three input features to build the linear regression model:

1) X2=the house age

2) X3=the distance to the nearest MRT station (unit:degree)

3) X4=the number of convenience stores in the living circle on foot (integer).

When implementing normal equations, you are required the complete the following steps:

1. Split the dataset into two, one for training (70%) and another one for testing (30%).
2. Use normal equations and the training data to estimate parameters in the linear regression model
3. Evaluate the performance of the model on the testing data
4. Predict the price of a house that is 2 years old, 500 meters to the nearest MRT station, and has 8 convenience stores in the living circle

When implementing gradient descent, you are required to complete the following steps:

1. Normalize input features so that the mean value of each feature is 0 and the standard deviation is 1.
2. Run gradient descent to learn the linear regression model using the training data
3. Evaluate the performance of the model on the testing data
4. Predict the price of a house that is 2 years old, 500 meter to the nearest MRT station, and has 8 convenience stores in the living circle

To get started, first open the main script assignmentLinearR.m. You are required to modify this script as well as all the other six scripts, including

* loadData.m – Function to load and split the dataset into training and testing sets
* normalEqn.m – Function to compute the normal equations
* evaluateAccuracy.m – Function to evaluate the performance of the linear regression model
* featureNormalize.m – Function to normalize features
* computCost.m – Function to compute the cost
* gradientDescent.m – Function to run gradient descent

You can download Matlab by following the instructions provided in this link: <https://library.sdsu.edu/computers-technology/software/matlab>

**Your submission should be a zip file that includes all codes.**