# Programming Assignment: Logistic 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 logistic regression by using maximum likelihood estimation and gradient ascent. 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 whether a person has diabetes or not by using the dataset downloaded from <http://networkrepository.com/pima-indians-diabetes.php>

You are required to use all eight input features (first eight columns) to build the logistic regression model.

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. Normalize input features so that the mean value of each feature is 0 and the standard deviation is 1.
3. Run gradient ascent to learn the logistic regression model using the training data
4. Evaluate the performance of the model on the testing data

To get started, first open the main script assignmentLogisticR.m. You are required to modify this script as well as the following seven scripts:

* loadData.m – Function to load and split the dataset into training and testing sets
* featureNormalize.m – Function to normalize features
* gradientAscent.m – Function to run gradient ascent
* likelihoodFunction.m – Function to compute the log likelihood and its gradient w.r.t. to the parameters
* evaluateAccuracy.m – Function to evaluate the performance of the logistic regression model
* predict.m – Function to predict the output
* sigmoid.m – Function to compute sigmoid

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.**