%% EED-NITJ Machine Learning Class - Exercise 1: UNIVARIANT Linear Regression

%% Initialization

clear ; close all; clc

%% ======================= Plotting =======================

fprintf('Plotting Data ...\n')

data = load('ex1data1.txt');

X = data(:, 1); y = data(:, 2);

m = length(y); % number of training examples

% Plot Data

% Note: You have to complete the code in plotData.m

plot(X, y,'rx');

xlabel('Population in ten 10,000s')

ylabel('Revenue Data in 10,000 dollars')

fprintf('Program paused. Press enter to continue.\n');

pause;

%% =================== Cost and Gradient descent ===================

X = [ones(m, 1), data(:,1)]; % Add a column of ones to x

theta = zeros(2, 1); % initialize fitting parameters

% Some gradient descent settings

iterations = 2500;

alpha = 0.01;

fprintf('\nTesting the cost function ...\n')

% compute and display initial cost

J = computeCost(X, y, theta);

fprintf('With theta = [0 ; 0]\nCost computed = %f\n', J);

fprintf('Expected cost value (approx) 32.07\n');

fprintf('Program paused. Press enter to continue.\n');

pause;

% further testing of the cost function

J = computeCost(X, y, [-1 ; 2]);

fprintf('\nWith theta = [-1 ; 2]\nCost computed = %f\n', J);

fprintf('Expected cost value (approx) 54.24\n');

fprintf('Program paused. Press enter to continue.\n');

pause;

fprintf('\nRunning Gradient Descent ...\n')

% run gradient descent

theta = gradientDescent(X, y, theta, alpha, iterations);

% print theta to screen

fprintf('Theta found by gradient descent:\n');

fprintf('%f\n', theta);

fprintf('Expected theta values (approx)\n');

fprintf(' -3.6303\n 1.1664\n\n');

% Plot the linear fit

hold on; % keep previous plot visible

plot(X(:,2), X\*theta, '-')

legend('Training data', 'Linear regression')

hold off % don't overlay any more plots on this figure

% Predict values for population sizes of 35,000 and 70,000

predict1 = [1, 3.5] \*theta;

fprintf('For population = 35,000, we predict a profit of %f\n',...

predict1\*10000);

predict2 = [1, 7] \* theta;

fprintf('For population = 70,000, we predict a profit of %f\n',...

predict2\*10000);