703642, Advanced Machine Learning course 29th October, 2018

**Assignment 1**

**Programming Project: Ridge Regression and Model Selection**

Kamal Zakieldin, Yusuf Ipek

**Overview:**

In this report we are describing our implementation for the ridge regression model and our experiments we have made, first we introduce the used technology and the libraries, then we discuss our tasks illustrating what we have made then showing the results and the values we have got.

**Technologies:**

**Python** as the programming language.

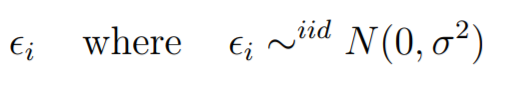
**Libraries:**

This section will give a short overview of the libraries we used.

* **numpy** provide help working with matrices.
* **matplotlib** for all ploting and drawing the figures to visualize our results.

**Discussion:**

We start our implementation by generating the data set of N points, choosing a quadratic function f(x) using the following equation:



We have used epsilon to add a normal distributed uniform noise calculated using the mean and the variance of points.

in the beginning we have created dataset with high X values that create a very high Y values, and depending on high Y values were really difficult and increased the error, that’s why we used a smaller X values that subsequently produced better Y values, generating the data set was one of the most underestimated step that affect the performance of our model.

Notices:

1. The way of choosing the initial values of the weights vector always affect the final weights vector, choosing a constant initial values like 1 makes the model adjust them slightly without huge difference in the final weights.

On the other hand, choosing the initial weights randomly makes it produces different weights every trial, specially when we have few number of input points (x values), increasing the number of points in the training set affect the weights values and reduce the differences arises from the random selection of the initial values