CSE 6005- Machine Learning
Lab Practice Sheet 1 (Regression)

Date of Completion : 31/01/2020

- You are supposed to answer the following questions after doing the appropriate experiments.
- Every answer should be supported by the experiment(s) with the details: Objective of the experiment, Design of the experiment and the inference from the experiment.
- Choice of the data-set for any experiment is your choice, but the data set should be a multi-variate data set.
- Answer for every questions should be in the form of a report with the details of the experiments performed with justification.

1. Fit a regression based supervised learning model for a data set and do the required optimization for learning the parameters of your model in the following two different ways.
   (a) Optimization through Normal equations Method.

   (b) Optimization through Gradient-descent Method.

   Compute the Performance Measure of the models with the validation data. Based on your dataset, decide the optimization method which converges fast.

2. Without using any library function, Implement a linear regression model which optimizes the parameters through the method of normal equations.

3. 'Size of the data used to train a linear regression model is directly proportional to the performance of the model'. Take a data set with different sizes, and do a minimum of 10 experiments and produce a scatter-plot (size of the data Vs Performance of the model) and comment on the above statement, based on your experiment.

4. Do a minimum of 50 experiments with the same data set for the linear regression model, by varying the learning rate (α) and compute the minimum number of iterations (n_i) required for the convergence of the parameters involved in the learning model. With the data generated by through your experiments (learning rate, Minimum number of iterations for the convergence), Fit a learning model for that data to bring out the relationship between α and 'n_i', in two different ways, first by taking α as the independent variable and then by taking n_i as the independent variable an conclude whether the two models remain the same or not.

5. Take a data set. Fit a linear regression model with three different hypothesis that involves different degrees of polynomials of the components of the input vector. Conclude which fits the given data set best.

6. Take a data set, fit a non-linear model and a linear model for the same data. Based on your observations, conclude which model suits best for the chosen data set.