Introduction to Machine Learning Laboratory Exercise

Dr. Shuvendu Rana

Simple Linear Regression

1. Implement Linear Regression and calculate sum of residual error on the following Datasets.

```
x = [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]

y = [1, 3, 2, 5, 7, 8, 8, 9, 10, 12]
```

- Compute the regression coefficients using analytic formulation and calculate Sum Squared Error (SSE) and R² value.
- Implement gradient descent (both Full-batch and Stochastic with stopping criteria) on Least Mean Square loss formulation to compute the coefficients of regression matrix and compare the results using performance measures such as R² SSE etc.
- 2. Download Boston Housing Rate Dataset. Analyse the input attributes and find out the attribute that best follow the linear relationship with the output price. Implement both the analytic formulation and gradient descent (Full-batch, stochastic) on LMS loss formulation to compute the coefficients of regression matrix and compare the results.

#Downloading Boston Housing Rate Dataset

```
#from sklearn.datasets import load boston #import sklearn
from sklearn import datasets
boston data = datasets.load boston()
#boston data = Load boston()
df = pd.DataFrame(data=boston_data['data'])
#Spliting dataset into train and test set
from sklearn import datasets, linear_model
# load the boston dataset
boston = datasets.load_boston(return_X_y=False)
# defining feature matrix(X) and response vector(y)
X = boston.data
y = boston.target
# splitting X and y into training and testing sets
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.4,
                                                      random state=1)
```