Machine Learning Online | Assignment

Linear Regression | Air Quality Prediction

In this problem, you are given a dataset (https://github.com/coding-blocks-archives/machine-learning-online-2018/tree/master/assignment_datasets/Regression_Data) containing information about various features on which air quality depends. Download the data set from the above link, train your Linear Regression Model without using library function for linear regression.

1. How many features you observe in the dataset.
2. Use Gradient Descent Algorithm with variouse learning rates. Use Convergence criteria as change in error. Repeat the same part with stochastic gradient descent(batch size=1).
   Plot the error for GD and SGD, report your observations.
3. Report the value of coefficients and intercept, and hypothesis function.
4. Implement a function to compute Coefficient of Determination.
   Report the 'Score' on the training dataset.
5. Make predictions on test dataset and compute score. Compare your Training and Test Scores. Which one is better?

Coding Blocks | Online