**Walchand College of Engineering, Sangli**

## **Machine Learning Lab (6CS372)**

**TY BTech | AY 2023-2024 | Even Sem**

**Assignment 3**

1. **Simulation based assignment on python built-in functions**

Visit link <https://python-iitk.vlabs.ac.in/exp/built-in-functions/>

Click on tabs on left hand side from Aim to References and read / complete the activity.

1. **Linear regression**

**Part 1: Linear regression *without* scikit-learn (Refer linReg demo)**

* 1. Load diabetes dataset from sklearn.
  2. Preprocessing: Null value handling, standardization
  3. Data splitting: Split data as 70% train and 30% test.
  4. Select a single input feature. Plot input feature against target variable.
  5. Write functions for computing cost, gradients and gradient descent algorithm. (Save cost values of each iteration).
  6. Plot regression line on scatter plot of feature vs target.
  7. Plot cost vs #iterations.
  8. Report parameter values, training error, test error and model accuracy.

**Part 2: Linear regression *with* scikit-learn**

1. Download dataset as per your batch.
2. Preprocessing: Null value handling, standardization, replace categorical values with numeric values (e.g. 0, 1, 2 etc.)
3. Data splitting: Split data as 70% train and 30% test using train\_test\_split function.
4. Feature selection: Dimensionality reduction / select manually
5. Fit model using fit function.
6. Report parameter values, training error and test error and model accuracy.

**Batch 1 -** <https://www.kaggle.com/datasets/spittman1248/cdc-data-nutrition-physical-activity-obesity>

**Batch 2-** <https://archive.ics.uci.edu/ml/datasets/Air+Quality>

**Batch 3-** <https://archive.ics.uci.edu/ml/datasets/Appliances+energy+prediction>

**Batch 4-** <https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset>

**Batch 5-** <https://archive.ics.uci.edu/dataset/186/wine+quality>

**Batch 6-** <https://archive.ics.uci.edu/dataset/29/computer+hardware>

**Batch 7-** <https://archive.ics.uci.edu/dataset/477/real+estate+valuation+data+set>

**Batch 8-** <https://archive.ics.uci.edu/dataset/162/forest+fires>