School of Computer Science and Engineering, VIT Chennai.

BCSE209P Machine Learning Lab

Assignment-2 Linear Regression

Faculty : Dr. R. Jothi

**Due Date : 24/12/2024**

Submit your python code (Colab notebook): with output for all the questions.

Q1. Suppose you are asked to build a machine learning model for predicting the output of manufacturing machine based on its operational time as given in Machine.csv dataset file.

- Print all the predictor variables and dependent variables in the given dataset (Machine.csv)
- Understand the relationship between each predictor variable and the dependent variable; draw the plot.
- Implement (write your code) linear regression algorithm with gradient descent optimization. Print the regression parameters after 2 epochs of training. Predict the output of the manufacturing if the machine runs for 13 hours. Also print the error of your regression model. (you may verify your answers with ML theory class work )
- Change the epochs and print the change in the model performance.

Q2. Use Sklearn model `LinearRegression` to solve the above problem (Machine.csv) and compare results of your implementation with Sklearn model.

Q3. Consider house price prediction problem based on linear regression. Use appropriate Scikit/SKlearn Library function to apply linear regression on the a small variant of original house price prediction dataset (housePrice_small).

- Print all the predictor variables and dependent variables in the given dataset
- Understand the relationship between each predictor variables and the dependent variable; Use heatmap to understand correlation between the dependent and independent variables.
- Preprocess the dataset (like encoding of feature values, removing nulls, scaling, etc).
- Keep 80% of samples for training and rest for testing
- Print the regression parameters after training.
- Show the accuracy on the test set.

Bonus Mark Question:

Apply linear regression (using Scikit library) on the original house price prediction dataset (housePrice.csv). You need to show all the pre-processing steps.