**CPC251: Machine Learning and Computational Intelligence**
**Academic Session: Semester 2, 2022/2023**
**School of Computer Sciences, USM, Penang**

**Assignment 1**

## Description and Requirements

This is a group assignment. In the assignment, you are given a dataset. The dataset is available in eLearn@USM. You need to implement the gradient descent algorithm to estimate (train) the weights of the linear regression model. You must use Python programming language with Jupyter Notebook.

Split the given dataset into two training and test with a ratio of 8:2. Use the training set to estimate the weights and the test set to evaluate the linear regression model with the estimated weights.

Define the (3) functions as follows

```
def train_model(X, y, alpha, max_epoch):
    """ Pass four arguments
    Arguments:
        X: input features
        y: responses
        alpha: learning rate
        max_epoch: maximum epochs
    Returns:
        w: estimated weights
        hist_loss: training loss history
    """


def predict(w, X):
    """ Pass two arguments
    Arguments:
        w: weights
        X: input features
    Returns:
        yhat: predicted values
    """


def loss_fn(y, yhat):
    """ Pass two arguments
    Arguments:
        y: responses
        yhat: predicted value
    Returns:
        loss: loss value
    """
```

Display the training loss value for each epoch of the training loop.

Display the estimated weights (after model training).

Display the training loss against epoch graph (after model training).

Evaluate the linear regression model with the estimated weights on the test set and display at least R-squared, mean squared error and mean absolute error measures.

Document the codes using **comments**.

Submission requirements

- Due date: **19 May 2022 (Friday)**, 11:59 p.m. (Week 8).
- Use the given jupyter notebook template. Do **not** change the function name and its parameters.
- The jupyter notebook must be executed to show the outputs.
- Submission must be made in **ipynb** format (submitted online).
- The filename **must** follow these naming conventions.
  - <CPC251_Assignment1_GroupNo.ipynb>
- Plagiarism (using other people's ideas and text without proper acknowledgment and using them as your own) is a serious academic offence. The consequences for plagiarism are severe.

## Rubric for Part 1

| Component | 10-9 (Excellent) | 8-6 (Good) | 5-3 (Average) | 2-1 (Poor) | Weightage |
|---|---|---|---|---|---|
| Requirements and Delivery | The three functions are implemented completely, including the specified parameters.<br><br>All the outputs (weights, training loss value and graph, and model evaluation) are displayed. | The three functions are partially implemented (with the specified parameters)<br><br>One or more of the outputs (weights, loss value and graph, and model evaluation) are not displayed. | One or two functions are implemented.<br><br>One or more of the outputs (weights, loss value and graph, and model evaluation) are not displayed. | The three functions are not implemented.<br><br>The weights, training loss value and graph, and model evaluation are not displayed. | 4 |
| Runtime and Algorithm | Executes without errors<br><br>The algorithm and outputs are correct. | Executes without errors<br><br>The algorithm and/or outputs have minor errors. | Executes without errors<br><br>The algorithm and/or outputs are partially correct. | Does not execute due to error<br><br>Algorithm is incorrect and no output. | 2 |
| Efficiency | Solution is efficient, easy to understand, and maintain. | The code is fairly efficient without sacrificing readability and understanding. | A logical solution that is easy to follow but it is not the most efficient. | A difficult to understand and inefficient solution.<br><br>Code is huge and appears to be patched together. | 2 |
| Documentation | The source codes are well documented and commented. | The source codes are partially documented and commented. | The source codes are minimally documented and commented | The source codes are not documented and commented | 2 |