HOMEWORK 2 WEEK 2

- Anish Panicker
- Ap2938

Gradient descent code :

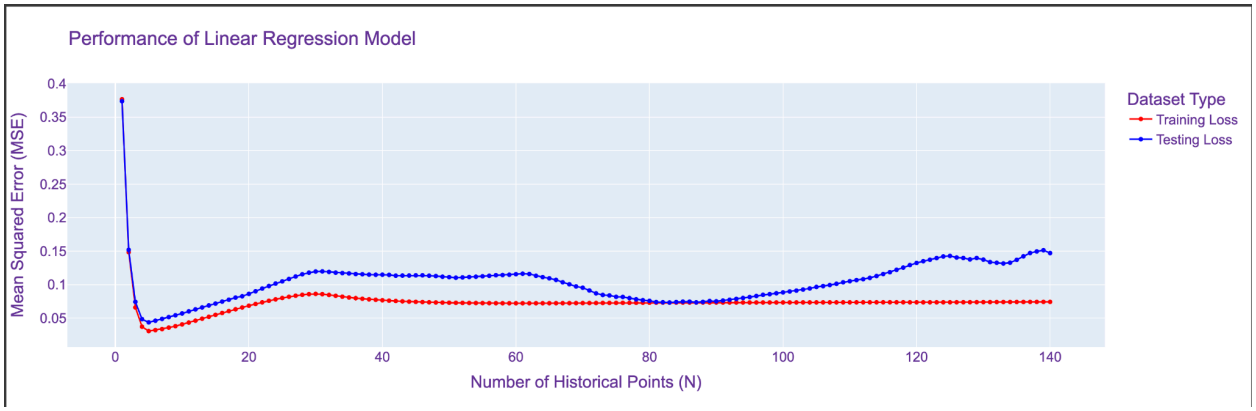
```
def gradient_descent(X, y, learning_rate=1e-4, regularization=1e-3, max_iters=5000):

    B = np.zeros(X.shape[1])
    mse_prev = float('inf')
    threshold = 1e-5
    for iteration in range(max_iters):

        y_pred = X.dot(B)
        error = y_pred - y
        gradient = X.T.dot(error) / len(y) + regularization * B
        B -= learning_rate * gradient
        mse_current = MSE(y_pred, y)
        mse_prev = mse_current


    return B, mse_current
```

Output :

```
Processing Sequence Length=1  Done. Train Loss: 0.376586, Test Loss: 0.373643
Processing Sequence Length=2  Done. Train Loss: 0.148476, Test Loss: 0.151961
Processing Sequence Length=3  Done. Train Loss: 0.066095, Test Loss: 0.074436
Processing Sequence Length=4  Done. Train Loss: 0.037360, Test Loss: 0.048597
Processing Sequence Length=5 Converged after 4690 iterations.
 Done. Train Loss: 0.030873, Test Loss: 0.043847
Processing Sequence Length=6 Converged after 4082 iterations.
 Done. Train Loss: 0.032314, Test Loss: 0.046303
Processing Sequence Length=7 Converged after 3629 iterations.
 Done. Train Loss: 0.033820, Test Loss: 0.048989
Processing Sequence Length=8 Converged after 3278 iterations.
 Done. Train Loss: 0.036020, Test Loss: 0.051726
Processing Sequence Length=9 Converged after 2999 iterations.
 Done. Train Loss: 0.038180, Test Loss: 0.054389
Processing Sequence Length=10 Converged after 2771 iterations.
 Done. Train Loss: 0.040775, Test Loss: 0.057126
Processing Sequence Length=11 Converged after 2582 iterations.
 Done. Train Loss: 0.043517, Test Loss: 0.060254
Processing Sequence Length=12 Converged after 2423 iterations.
 Done. Train Loss: 0.046355, Test Loss: 0.063204
Processing Sequence Length=13 Converged after 2289 iterations.
 Done. Train Loss: 0.049306, Test Loss: 0.066114
Processing Sequence Length=14 Converged after 2173 iterations.
 Done. Train Loss: 0.052179, Test Loss: 0.069084
Processing Sequence Length=15 Converged after 2074 iterations.
 Done. Train Loss: 0.055078, Test Loss: 0.071844
```

```
Processing Sequence Length=119 Converged after 3672 iterations.
 Done. Train Loss: 0.073754, Test Loss: 0.129322
Processing Sequence Length=120 Converged after 3672 iterations.
 Done. Train Loss: 0.073738, Test Loss: 0.132438
Processing Sequence Length=121 Converged after 3671 iterations.
 Done. Train Loss: 0.073741, Test Loss: 0.135004
Processing Sequence Length=122 Converged after 3670 iterations.
 Done. Train Loss: 0.073743, Test Loss: 0.137234
Processing Sequence Length=123 Converged after 3669 iterations.
 Done. Train Loss: 0.073758, Test Loss: 0.139698
Processing Sequence Length=124 Converged after 3668 iterations.
 Done. Train Loss: 0.073771, Test Loss: 0.142143
Processing Sequence Length=125 Converged after 3667 iterations.
 Done. Train Loss: 0.073791, Test Loss: 0.142915
Processing Sequence Length=126 Converged after 3666 iterations.
 Done. Train Loss: 0.073816, Test Loss: 0.140581
Processing Sequence Length=127 Converged after 3665 iterations.
 Done. Train Loss: 0.073851, Test Loss: 0.139821
Processing Sequence Length=128 Converged after 3664 iterations.
 Done. Train Loss: 0.073893, Test Loss: 0.137733
Processing Sequence Length=129 Converged after 3663 iterations.
 Done. Train Loss: 0.073947, Test Loss: 0.139797
Processing Sequence Length=130 Converged after 3663 iterations.
 Done. Train Loss: 0.073999, Test Loss: 0.137424
Processing Sequence Length=131 Converged after 3662 iterations.
 Done. Train Loss: 0.074062, Test Loss: 0.133729
Processing Sequence Length=132 Converged after 3662 iterations.
 Done. Train Loss: 0.074113, Test Loss: 0.132759
Processing Sequence Length=133 Converged after 3663 iterations.
 Done. Train Loss: 0.074148, Test Loss: 0.131661
Processing Sequence Length=134 Converged after 3663 iterations.
 Done. Train Loss: 0.074191, Test Loss: 0.132844
Processing Sequence Length=135 Converged after 3664 iterations.
 Done. Train Loss: 0.074226, Test Loss: 0.137092
Processing Sequence Length=136 Converged after 3665 iterations.
 Done. Train Loss: 0.074269, Test Loss: 0.142345
Processing Sequence Length=137 Converged after 3666 iterations.
 Done. Train Loss: 0.074307, Test Loss: 0.147235
Processing Sequence Length=138 Converged after 3668 iterations.
 Done. Train Loss: 0.074339, Test Loss: 0.149636
Processing Sequence Length=139 Converged after 3671 iterations.
 Done. Train Loss: 0.074365, Test Loss: 0.151400
Processing Sequence Length=140 Converged after 3673 iterations.
 Done. Train Loss: 0.074400, Test Loss: 0.147154
```

Graph result of model prediction



Sequence Length and y pred equation:

```
The best sequence length is 5 with minimal test loss.
The best model is
y_pred = 0.004 + (0.175*x1) + (0.176*x2) + (0.179*x3) + (0.182*x4) + (0.185*x5)
```