Problem 3

In this assignment, you are asked to implement and train a two-class classification model given by the equation

$$f(\mathbf{x}; \mathbf{w}, b) = \sigma(\mathbf{w}^T \mathbf{x} + b)$$

where **x** is a D dimensional input vector, **w** is a D dimensional weight vector and b is a scalar, and σ(·) is a custom activation function:

$$\sigma(t) = \begin{cases} -\alpha t, & \text{if } t < 0 \\ 3t^3 - 4t^2 + 2t, & \text{if } 0 \leq t \leq 1 \\ t & \text{if } t > 1 \end{cases}$$

The classification rule for the classifier, i.e. the estimated label for a given input vector **x** is

$$\hat{y} = \begin{cases} 1, & \text{if } f(\mathbf{x}) \geq T \\ 0, & \text{otherwise} \end{cases}$$

where 1 is considered to be the positive label and 0 is the negative label. T is a threshold value.

While training the model use the following loss function:

$$L(\mathbf{w}, b) = \frac{1}{N} \sum_{1}^{N} (y_i - f(\mathbf{x}_i; \mathbf{w}, b))^2$$

You are provided with two Python files: **main.py** and **solution.py**. You are asked to implement your solution in **solution.py**. **You should not change main.py.**

Detailed Requirements
1. Follow the instructions in the provided Python files for your implementation.
2. Running details of the **main.py**
   python main.py path_to_train_data path_to_test_data hyper_parameters_for_optimizer
   e.g. python main.py data/linear_data_train.pickle.gz data/linear_data_test.pickle.gz 100 0.2 3
3. α is a hyper parameter that you are asked to set by hardcoding in the solution.
4. Minimum required test accuracy for linear_data is 100%.
5. Minimum required test accuracy for mnist_01 is 99%.
6. Minimum required test accuracy for moon_data is 84%.
7. Explain your optimization strategy, comment on the datasets and improvement suggestions in a short report.
8. **Deliver only the solution.py and the report.**