Home Assignment – 6

Aarav Aryaman, 220012

Problem:

Develop an artificial neural network (ANN), with one hidden layer, that will approximate the function given below in the interval [-1, 1]

$$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$$

Write a computer program (preferably in python), from scratch, to compute the necessary weights and biases of the ANN. The computer program must NOT use scikitlearn / scipy / statistics / Tensroflow / Keras / Pytorch or similar packages/libraries. You can only use packages for vector / matrix / array operations and plotting (numpy, matplotlib etc.).

1. Generate 200 training data by randomly selecting x_1, x_2 within the interval [-1, 1].

2. Train the network and find the weights and biases.

3. Generate 100 test data by randomly selecting x_1, x_2 within the interval [-1, 1].

4. Calculate the training and test errors (cost functions).

Solution Algorithm:

Step 1: Data Generation

- The training data is generated using a uniform distribution in the interval [−1,1]. The output y is calculated using the function f.

Step 2: Neural Network Architecture

The ANN consists of:

- Input Layer: 2 neurons (for x_1 and x_2).
- Hidden Layer: 20 neurons (taken arbitrarily) with ReLU activation.
- Output Layer: 1 neuron (for predicting $f(x_1, x_2)$)

Step 3: Forward and Backward Propagation

- The forward pass calculates the output of the network, while the backward pass updates the weights and biases based on the error.

Step 4: Training the Network

- The network is trained using batch (10) gradient descent over some number of epochs (5000).

Step 5: Evaluation

- The mean squared error is calculated for both training and test sets to evaluate the performance of the ANN.

Step 6: Implementation and Results

- After implementing the ANN, we train it and evaluate its performance.
  - Weights and Biases after Training:

    Weights Input-Hidden:

    [[...]]

    Bias-Hidden:

    [[...]]

    Weights Hidden-Output:

    [[...]]

    Bias-Output:

    [[...]]

  - Training Error (MSE): 6.561068

    Test Error (MSE): 12.360928

    (for the submitted set of randomly generated data)

- Plot