

Assignment 2
CS 4783/5783
Due: 09/26/2021 11:59 pm

[Question 1]

[10 points]

Derive the update rule and show how to train a 3-layer (1 input layer, 1 hidden layer and 1 output layer) network with backpropagation for regression using the Mean Square Error loss. Assume that you are using the Sigmoid activation function for the hidden layer. Explain briefly how this is different from the update rule for the network trained for binary classification using log loss.

[Question 2]

[40 points]

For the given data in train.csv and test.csv, construct a neural network for the regression task. Your network must have 1 input layer, 1 hidden layer and 1 output layer. Use sigmoid to be your activation function for the hidden layer. Remember: since it is a regression problem, you should use a linear activation for the final (output) layer. The number of neurons in your output layer should be 1 since you are predicting only one output value.

The data is already split to have your input data for training (`X_train.csv`) and testing (`X_train.csv`) and their corresponding target values `Y_train.csv` and `Y_test.csv`, respectively.

You can load the data as follows:

```
X_train = np.loadtxt("X_train.csv")
```

Implement the backpropagation algorithm and train your network until convergence.

Answer the following questions:

1. Report the average MSE loss and the accuracy.
2. Plot the loss and accuracy as a function of the number of iterations.
3. What is the effect of the learning rate on the training process? Vary the learning rate to be between 0.001 and 1.0 and plot the resulting accuracy as a function of learning rate.
4. What is the effect of the number of neurons in the hidden layer? Vary the number of neurons from 1 to 10 and report the final loss and accuracy along with a brief description (2-3 lines) of your observation.
5. What is the effect of the activation functions in the network? Explore two different activation functions other than sigmoid such as `tanh`, `linear`, or `ReLU`.

Hint: Remember that your update rules for weights will change with change in activation function.