ECE 364: Assignment #5

1. (10 pts) Consider a neuron with three binary inputs (which can take values 0 or 1): x_1, x_2, and x_3. It uses a logistic activation function. Hence, $z = w_0 + w_1x_1 + w_2x_2 + w_3x_3$ and $a = logistic(z)$. Assign values to weights w_0, w_1, w_2, and w_3 so that the output of the neuron is greater than 0.5 if an only if the following is true: (x_1 AND x_2) OR x_3.

2. (10 pts) You would like to train a fully-connected neural network with 5 hidden layers, each with 10 hidden neurons. The input is 20-dimensional and the output is a scalar. What is the total number of trainable weights (also called parameters) in your network?

3. (10 pts) Consider a two-layer feed-forward neural network with an input layer with two inputs, 1 and 2, a hidden layer with one neuron, 3, and an output layer with one neuron, 4. This neural network has five weights. Suppose the activation function is logistic for the hidden layer and linear for the output layer. Initialize all weights to 0.1, then determine their values after one training iteration of the backpropagation algorithm. Assume a learning rate of 0.3 and the following training data instance: (1,0;1), where the first two values denote the inputs and the third value the desired output.

4. (10 pts) Suppose a network architecture has four neurons in a softmax output layer. If the one-hot encoding of the target for the current training example is $\mathbf{t}$ = [0,0,1,0] and the logits for the four neurons in the softmax output layer for this example are [0, 0.5, 0.25, 0.75], then what is the δ (i.e., error gradient) value for each of the four neurons?

5. (20 pts) **Coding project**

   In this project, you will train multi-layer neural networks to determine whether transactions made by credit cards are fraudulent or not. The dataset consists of 30 different descriptive features. The target feature is a binary variable, where 1 indicates that the transaction is fraudulent and 0 that it is not. This dataset is highly imbalanced, the positive class (fraud) accounts for just 0.172% of all transactions.

   When you train multi-layer neural networks, you will investigate the impact of different hyperparameters, including layer width, network depth, and the choice of the hidden activation function. You will also experiment with early stopping as a regularization strategy.

   Please see the Jupyter notebook for more details.

   GitHub repository for ECE364 coding projects: https://github.com/JHA-Lab/ece364_2024