CS 1678 HW1 Report:

Given:

x = [1 1 1]

y = [0 0]

Initialize all weights to 0.05:

In other words:

$$w_{10}^{(1)} = w_{20}^{(1)} = w_{11}^{(1)} = w_{21}^{(1)} = w_{12}^{(1)} = w_{22}^{(1)} = w_{10}^{(2)} = w_{20}^{(2)} = w_{11}^{(2)} = w_{21}^{(2)} = w_{12}^{(2)} = w_{22}^{(2)} = 0.05$$

The Activations can be calculated as:

$z_0 = 1$

$z_1 = \sigma(\sum_j w^{(1)}{}_{ij}\, x) = \sigma(0.05*1 + 0.05*1 + 0.05*1) = 0.5374$

$z_2 = \sigma(\sum_j w^{(1)}{}_{2j}\, x) = \sigma(0.05*1 + 0.05*1 + 0.05*1) = 0.5374$

$y_1 = \sigma(\sum_k w^{(1)}{}_{ik}\, z) = \sigma(0.05*1 + 0.05*0.53743 + 0.05*0.53743) = 0.5259$

$y_2 = \sigma(\sum_k w^{(1)}{}_{2k}\, z) = \sigma(0.05*1 + 0.05*0.53743 + 0.05*0.53743) = 0.5259$

The Errors can be calculated as:

$\delta_{y1} = y_1 * (1 - y_1) * (y_1 - y_{true}) = 0.5259 * (1 - 0.5259) * (0.5259 - 0) = 0.1311$

$\delta_{y2} = y_2 * (1 - y_2) * (y_2 - y_{true}) = 0.5259 * (1 - 0.5259) * (0.5259 - 0) = 0.1311$

$\delta_{z1} = z_1 * (1 - z_1) * \sum_k \delta_k w_{k1} = (0.53743) * (1-0.53743) * (0.13112 * 0.05 + 0.13112 * 0.05) = 0.0033$

$\delta_{z2} = z_2 * (1 - z_2) * \sum_k \delta_k w_{k2} = (0.53743) * (1-0.53743) * (0.13112 * 0.05 + 0.13112 * 0.05) = 0.0033$

Now, update the weights:

$$w_{10}^{(2)} = w_{10}^{(2)} - 0.3 * \delta_{y1} * z_0 = 0.05 - 0.3 * 0.1311 * 1 = 0.0107$$

$$w_{20}^{(2)} = w_{20}^{(2)} - 0.3 * \delta_{y2} * z_0 = 0.05 - 0.3 * 0.1311 * 1 = 0.0107$$

$$w_{11}^{(2)} = w_{11}^{(2)} - 0.3 * \delta_{y1} * z_1 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.0289$$

$$w_{21}^{(2)} = w_{21}^{(2)} - 0.3 * \delta_{y2} * z_1 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.0289$$

$$w_{12}^{(2)} = w_{12}^{(2)} - 0.3 * \delta_{y1} * z_2 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.0289$$

$$w_{22}^{(2)} = w_{22}^{(2)} - 0.3 * \delta_{y2} * z_2 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.0289$$

$$w_{10}^{(1)} = w_{10}^{(1)} - 0.3 * \delta_{z1} * x_0 = 0.05 - 0.3 * 0.0033 * 1 = 0.0490$$

$$w_{20}^{(1)} = w_{20}^{(1)} - 0.3 * \delta_{z2} * x_0 = 0.05 - 0.3 * 0.0033 * 1 = 0.0490$$

$$w_{11}^{(1)} = w_{11}^{(1)} - 0.3 * \delta_{z1} * x_1 = 0.05 - 0.3 * 0.0033 * 1 = 0.0490$$

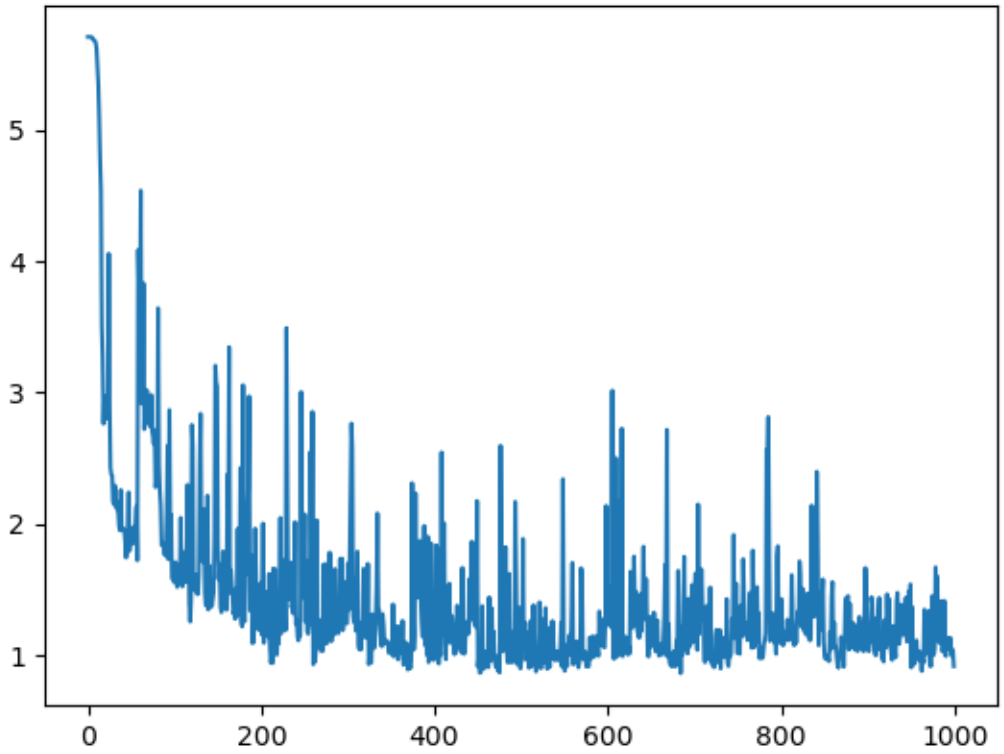
$$w_{21}^{(1)} = w_{21}^{(1)} - 0.3 * \delta_{z2} * x_1 = 0.05 - 0.3 * 0.0033 * 1 = 0.0490$$

$$w_{12}^{(1)} = w_{12}^{(1)} - 0.3 * \delta_{z1} * x_2 = 0.05 - 0.3 * 0.0033 * 1 = 0.0490$$

$$w_{22}^{(1)} = w_{22}^{(1)} - 0.3 * \delta_{z2} * x_2 = 0.05 - 0.3 * 0.0033 * 1 = 0.0490$$

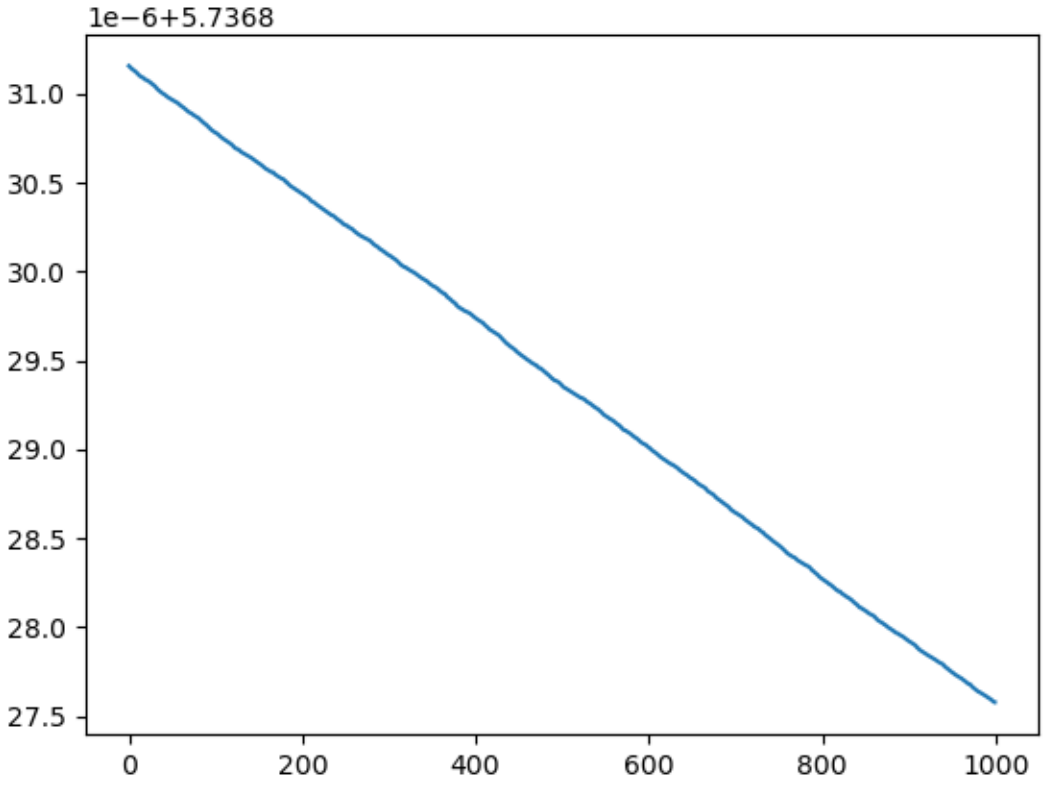
Learning Rate: 0.07

RMSE: 0.9753058006939729



Learning Rate: 0.00001

RMSE: 5.650004498903635



Learning Rate: 0.0064

RMSE: 0.9256773890094454