

HW7 Part 1: Computing weight updates by hand

$x = [1 \ 1]$ $y = [0 \ 0]$ $y_{\text{pred}} = [1 \ 1]$ $w = 0.05$ $\text{learning_rate} = 0.3$
 $z_0 = 1$ $z_1 = 0.5374$ $z_2 = 0.5374$ $y_1 = 0.5259$ $y_2 = 0.5259$

Compute errors

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

$$\delta_{y2} = y_2 * (1 - y_2) * (y_2 - y_{2\text{-true}}) = 0.5259 * (1 - 0.5259) * (0.5259 - 0) = 0.1311$$

$$\delta_{z1} = z_1 * (1 - z_1) * (\sum w \delta) = 0.5374 * (1 - 0.5374) * ([0.05 * 0.1311] + [0.05 * 0.1311]) = 0.0032597$$

$$\delta_{z2} = z_2 * (1 - z_2) * (\sum w \delta) = 0.5374 * (1 - 0.5374) * ([0.05 * 0.1311] + [0.05 * 0.1311]) = 0.0032597$$

Update Weights

$$w_{10}^{(2)} = w_{10}^{(2)} - 0.3 * \delta_{y1} * z_0 = 0.05 - 0.3 * 0.1311 * 1 = 0.01067$$

$$w_{20}^{(2)} = w_{20}^{(2)} - 0.3 * \delta_{y2} * z_0 = 0.05 - 0.3 * 0.1311 * 1 = 0.01067$$

$$w_{11}^{(2)} = w_{11}^{(2)} - 0.3 * \delta_{y1} * z_1 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.02886$$

$$w_{21}^{(2)} = w_{21}^{(2)} - 0.3 * \delta_{y2} * z_1 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.02886$$

$$w_{12}^{(2)} = w_{12}^{(2)} - 0.3 * \delta_{y1} * z_2 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.02886$$

$$w_{22}^{(2)} = w_{22}^{(2)} - 0.3 * \delta_{y2} * z_2 = 0.05 - 0.3 * 0.1311 * 0.5374 = 0.02886$$

$$w_{10}^{(1)} = w_{10}^{(1)} - 0.3 * \delta_{z1} * x_0 = 0.05 - 0.3 * 0.0032597 * 1 = 0.04902$$

$$w_{20}^{(1)} = w_{20}^{(1)} - 0.3 * \delta_{z2} * x_0 = 0.05 - 0.3 * 0.0032597 * 1 = 0.04902$$

$$w_{11}^{(1)} = w_{11}^{(1)} - 0.3 * \delta_{z1} * x_1 = 0.05 - 0.3 * 0.0032597 * 1 = 0.04902$$

$$w_{21}^{(1)} = w_{21}^{(1)} - 0.3 * \delta_{z2} * x_1 = 0.05 - 0.3 * 0.0032597 * 1 = 0.04902$$

$$w_{12}^{(1)} = w_{12}^{(1)} - 0.3 * \delta_{z1} * x_2 = 0.05 - 0.3 * 0.0032597 * 1 = 0.04902$$

$$w_{22}^{(1)} = w_{22}^{(1)} - 0.3 * \delta_{z2} * x_2 = 0.05 - 0.3 * 0.0032597 * 1 = 0.04902$$

HW7 Part 3

I did not add a bias term to any of my data while doing this. My plots looks a little off from Adriana's example.

