STAT 479 HW 1 Written Portion #'s 1-3

Cory Suzuki
Stat 479
Dr. Moon
10/1/2024

① Neurons ~ Nodes that receive input, process it, or passes the result to neurons in the next layer. Neurons in the input layer take in raw data, hidden layer neurons process the input through learned transformations, and the output layer neurons provide the final prediction or output.

Weights ~ Weights represent the importance or strength of connection between neurons, which are usually learned during training and backpropagation. Weights are adjusted to minimize the error between the predicted output and true target values, allowing the network to learn patterns in the data.

Activation ~ These functions introduce nonlinearity into the network to enable it to learn complex patterns. Common functions include ReLU, tanh, and sigmoidal functions. Softmax is also widely used.

Bias ~ An additional parameter added to the dot product between weights and inputs before application to the activation function. It ensures that the neurons have some output even if all the inputs are zero, making the model more flexible.

Ex) $$\sum_{i=1}^{n} w_i x_i + b$$

weights inputs bias

$\phi(\cdot)$ is an activation function.

Ex) Sigmoidal activation looks like:

$$\phi(x) = \frac{1}{1+e^{-x}}$$

② $W_1 = \begin{pmatrix} .2 & -.3 & .4 \\ .5 & -.1 & -.4 \end{pmatrix}_{2\times 3}$ $\quad b_1 = \begin{pmatrix} .1 & .2 & -.3 \end{pmatrix}_{1\times 3}$ $\quad x = \begin{pmatrix} .7 \\ .5 \end{pmatrix}_{2\times 1}$

- $W_1^T x + b_1^T$

$$= \begin{pmatrix} .2 & .5 \\ -.3 & -.1 \\ .4 & -.4 \end{pmatrix} \begin{pmatrix} .7 \\ .5 \end{pmatrix} + \begin{pmatrix} .1 \\ .2 \\ -.3 \end{pmatrix} = \begin{pmatrix} 0.39 \\ -0.26 \\ 0.08 \end{pmatrix} + \begin{pmatrix} .1 \\ .2 \\ -.3 \end{pmatrix} = \begin{pmatrix} 0.49 \\ -0.06 \\ -0.22 \end{pmatrix}$$

Now by applying ReLU, let $a_1 = W_1^T x + b_1^T$. Then we have activation, call it ϕ,

$$\phi(a_1) = \phi(W_1^T x + b_1^T) = \begin{pmatrix} \phi(0.49) \\ \phi(-0.06) \\ \phi(-0.22) \end{pmatrix} = \boxed{\begin{pmatrix} 0.49 \\ 0 \\ 0 \end{pmatrix}}$$

③ $x = .5,\ y = 1,\ w_1 = .2,\ w_2 = .4,\ b_1 = .1,\ b_2 = .2 \quad \eta = .1$

- $z_1 = w_1 x + b_1 = .2(.5) + .1 = 0.2$

$$a_1 = \frac{1}{1+e^{-z_1}} = \frac{1}{1+e^{-0.2}} = 0.5498$$

$$z_2 = w_2 a_1 + b_2 = .4(0.5498) + .2 = 0.4199$$

$$\hat{y} = \frac{1}{1+e^{-z_2}} = \frac{1}{1+e^{-0.4199}} = 0.6035$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_2}$$

$$= \left[-\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}}\right] \cdot \hat{y}(1-\hat{y}) \cdot a_1$$

- $= \left[-\frac{1}{0.6035} + \frac{1-1}{1-0.6035}\right] \cdot 0.6035(1-0.6035) \cdot 0.5498$

$$\Rightarrow \frac{\partial L}{\partial w_2} = -0.2180$$

Since y target is assumed to be binary, let the loss function L be the binary cross entropy loss.

$$L = -\left[y\ln(\hat{y}) + (1-y)\ln(1-\hat{y})\right]$$
$$= -y\ln(\hat{y}) - (1-y)\ln(1-\hat{y})$$

$$\frac{\partial L}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}}\left[-y\ln(\hat{y}) - (1-y)\ln(1-\hat{y})\right]$$
$$= -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}}$$

$$\frac{\partial \hat{y}}{\partial z_2} = \frac{\partial}{\partial z_2}\left[\frac{1}{1+e^{-z_2}}\right] = \frac{0(1+e^{-z_2}) + e^{-z_2}}{(1+e^{-z_2})^2}$$

$$= \frac{e^{-z_2}}{(1+e^{-z_2})^2} = \frac{1}{(1+e^{-z_2})} \cdot \frac{e^{-z_2}}{(1+e^{-z_2})} \quad \hat{y}(1-\hat{y})$$

$$= \frac{1}{1+e^{-z_2}}\left[\frac{(1+e^{-z_2} - 1)}{1+e^{-z_2}}\right] = \frac{1}{1+e^{-z_2}}\left(\frac{1+e^{-z_2}}{1+e^{-z_2}} - \frac{1}{1+e^{-z_2}}\right.$$

$$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial b_2} = \left[-\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} \right] \cdot \hat{y}(1-\hat{y}) \cdot 1$$

$$= \left[-\frac{1}{0.6035} + \frac{1-1}{1-0.6035} \right] \cdot 0.6035(1-0.6035) = -0.3965$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial a_1} \cdot \frac{\partial a_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_1}$$

$$= \left[-\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} \right] \cdot \hat{y}(1-\hat{y}) \cdot w_2 \cdot a_1(1-a_1) \cdot x$$

$$= \left[-\frac{1}{0.6035} + \frac{1-1}{1-0.6035} \right] \cdot 0.6035(1-0.6035) \cdot .4 \cdot 0.5498(1-0.5498) \cdot .5$$

$$= -0.0196$$

$$\frac{\partial a_1}{\partial z_1} = \frac{\partial}{\partial z_1}\left[\frac{1}{1+e^{-z_1}} \right]$$

$$= \frac{e^{-z_1}}{(1+e^{-z_1})^2}$$

$$= \frac{1}{(1+e^{-z_1})} \cdot \frac{e^{-z_1}}{(1+e^{-z_1})}$$

$$= \frac{1}{1+e^{-z_1}} \cdot \frac{1+e^{-z_1}-1}{1+e^{-z_1}}$$

$$= \frac{1}{1+e^{-z_1}}\left[1 - \frac{1}{1+e^{-z_1}} \right]$$

$$= a_1(1-a_1)$$

$$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial a_1} \cdot \frac{\partial a_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial b_1}$$

$$= \left[-\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} \right] \cdot \hat{y}(1-\hat{y}) \cdot w_2 \cdot a_1(1-a_1) \cdot 1$$

$$= \left[-\frac{1}{0.6035} + \frac{1-1}{1-0.6035} \right] \cdot 0.6035(1-0.6035) \cdot .4 \cdot 0.5498(1-0.5498) \cdot 1$$

$$= -0.0393$$

Updating w_2, b_2, W_1, b_1.

$$w_{2(new)} = w_{2(old)} - \eta \frac{\partial L}{\partial w_2} \qquad b_{2(new)} = b_{2(old)} - \eta \frac{\partial L}{\partial b_2}$$

$$W_{1(new)} = W_{1(old)} - \eta \frac{\partial L}{\partial w_1} \qquad b_{1(new)} = b_{1(old)} - \eta \frac{\partial L}{\partial b_1}$$

$$w_{2(new)} = .4 - .1(-0.2180) = \boxed{0.4218}$$

$$W_{1(new)} = .2 - .1(-0.0196) = \boxed{0.202}$$

$$b_{2(new)} = .2 - .1(-0.3965) = \boxed{0.2397}$$

$$b_{1(new)} = .1 - .1(-0.0393) = \boxed{0.1039}$$