

## Binary Classification

- Single logistic output
- Cross entropy error

### ① Forward Pass

$$z_j^1 = \sum_i (x_i \cdot \omega_{j,i}^1) + b_j^1$$

$$a_j^1 = \frac{1}{1 + e^{-z_j^1}}$$

$$z^2 = \sum_i (a_i^1 \cdot \omega_{i,i}^2) + b^2$$

$$y = \frac{1}{1 + e^{-z^2}}$$

### ② Error

$$\begin{aligned} E &= -(t \cdot \ln(y) + (1-t) \cdot \ln(1-y)) \\ &= -t \cdot \ln(y) - (1-t) \cdot \ln(1-y) \end{aligned}$$

### ③ Back Prop.

$$\frac{\partial E}{\partial \omega_i^2} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial z^2} \cdot \frac{\partial z^2}{\partial \omega_i^2}$$

$$\frac{\partial E}{\partial y} = -\frac{t}{y} - \frac{(1-t)(-1)}{1-y}$$

$$= \frac{-t(1-y) + y(1-t)}{y(1-y)}$$

$$= \frac{-t + y \cdot t + y - y \cdot t}{y(1-y)}$$

$$= \frac{y - t}{y(1-y)}$$

$$\frac{\partial y}{\partial z^2} = (-t) \cdot \frac{(-e^{-z^2})}{(1 + e^{-z^2})^2}$$

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

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

$$= \frac{1}{1 + e^{-z^2}} - \left( \frac{1}{1 + e^{-z^2}} \right)^2$$

$$= y - y^2$$

$$= y(1-y)$$

$$\left| \begin{array}{l} \frac{d}{dx} (\ln[f(x)]) \\ = \frac{f'(x)}{f(x)} \end{array} \right.$$

$$\left| \begin{array}{l} \left(\frac{f}{g}\right)' = \frac{f'g - f \cdot g'}{g^2} \end{array} \right.$$

$$\left| \begin{array}{l} \frac{d}{dx} (e^{f(x)}) \\ = f'(x) e^{f(x)} \end{array} \right.$$

$$\frac{\partial \bar{z}^2}{\partial w^2_i} = a^1_i$$

$$\begin{aligned}\frac{\partial \bar{E}}{\partial w^2_i} &= \frac{y-t}{\cancel{y(1-y)}} \cdot \cancel{y(1-y)} \cdot a^1_i \\ &= (y-t)a^1_i\end{aligned}$$

$$\begin{aligned}\frac{\partial \bar{E}}{\partial b^2} &= \frac{\partial \bar{E}}{\partial y} \cdot \frac{\partial y}{\partial z^2} \cdot \frac{\partial z^2}{\partial b^2} \\ &= (y-t) \cdot (1)\end{aligned}$$

$$\begin{aligned}\frac{\partial \bar{E}}{\partial w^1_{ij}} &= \frac{\partial \bar{E}}{\partial y} \cdot \frac{\partial y}{\partial z^2} \cdot \frac{\partial z^2}{\partial a^1_i} \cdot \frac{\partial a^1_i}{\partial z^1_j} \cdot \frac{\partial z^1_j}{\partial w^1_{ij}} \\ &\quad \begin{matrix} \downarrow & \leftarrow \\ (y-t) & ? \end{matrix} \quad \begin{matrix} \downarrow \\ a^1_i(1-a^1_i) \end{matrix} \quad \begin{matrix} \downarrow \\ \rightarrow x_i \end{matrix} \\ &\quad \downarrow \\ &\quad w^2_i \\ &= (y-t) \cdot w^2_i \cdot a^1_i(1-a^1_i) \cdot x_i\end{aligned}$$

$$\begin{aligned}\frac{\partial \bar{E}}{\partial b^1_j} &= \frac{\partial \bar{E}}{\partial y} \cdot \frac{\partial y}{\partial z^2} \cdot \frac{\partial z^2}{\partial a^1_i} \cdot \frac{\partial a^1_i}{\partial z^1_j} \cdot \frac{\partial z^1_j}{\partial b^1_j} \\ &= (y-t) \cdot w^2_i \cdot a^1_i(1-a^1_i) \cdot (1)\end{aligned}$$

#### ④ Variables update

$$\omega_{ij}^t = \omega_{ij}^t - \alpha \cdot \frac{\partial E}{\partial \omega_{ij}^t}$$

learning  
rate