

Binary Classification

- Single logistic output
- Cross entropy error

① Forward Pass

$$z^1_j = \sum_i (x_i \cdot w^1_{ji}) + b^1_j$$

$$a^1_i = \frac{1}{1 + e^{-z^1_j}}$$

$$z^2 = \sum_i (a^1_i \cdot w^2_i) + b^2$$

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

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.

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

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

$$\begin{aligned} \frac{\partial y}{\partial z^2} &= \frac{(-1) \cdot (-e^{-z^2})}{(1+e^{-z^2})^2} \\ &= \frac{1+e^{-z^2}-1}{(1+e^{-z^2})^2} \\ &= \frac{\cancel{1+e^{-z^2}}}{(1+e^{-z^2})^2} - \frac{1}{(1+e^{-z^2})^2} \\ &= \frac{1}{1+e^{-z^2}} - \left(\frac{1}{1+e^{-z^2}} \right)^2 \\ &= y - y^2 \\ &= y(1-y) \end{aligned}$$

$$\left| \frac{d}{dx} (\ln[f(x)]) \right| = \frac{f'(x)}{f(x)}$$

$$\left| \left(\frac{f}{g} \right)' \right| = \frac{f' \cdot g - f \cdot g'}{g^2}$$

$$\left| \frac{d}{dx} (e^{f(x)}) \right| = f'(x) e^{f(x)}$$

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

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

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

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

$$= (y-t) \cdot w^2_i \cdot a^1_i(1-a^1_i) \cdot x_i$$

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

④ Variables update

$$w'_{ij} = w_{ij} - \underset{\substack{\uparrow \\ \text{learning} \\ \text{rate}}}{\alpha} \cdot \frac{\partial E}{\partial w'_{ij}}$$