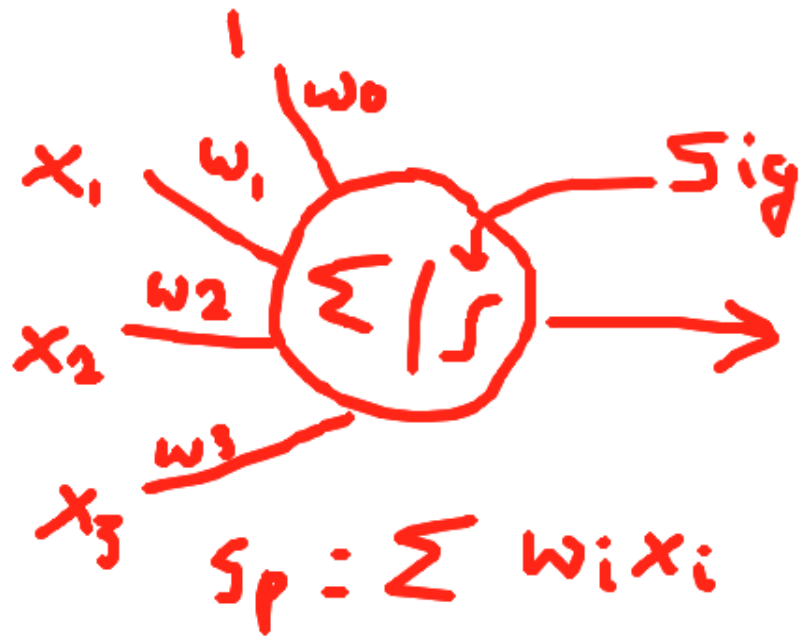


Training an Artificial Neuron

Notes



$$g_p = \text{Sig}(\sum w_i x_i)$$

$$\hat{y} = g(\sum w_i x_i)$$

$$\text{SSE} = \sum (y - \hat{y})^2$$

$$E = \sum (y - \hat{y})^2$$

$$= \sum (y - g(\sum w_i x_i))^2$$

$$\begin{matrix} x_1 & w_1 \\ x_2 & w_2 \\ x_3 & w_3 \end{matrix} \rightarrow \textcircled{\sum / s} \rightarrow \hat{y} = g(\sum w_i x_i)$$

$$E = \sum (y - g(\sum w_i x_i))^2$$

$$E \rightarrow 0, \quad v \rightarrow 0 \quad \text{brake} = -\frac{dv}{dt} \leftarrow \text{mult}$$

$$-\frac{dE}{dw_i}$$

$$\Delta w_i = -\eta \frac{dE}{dw_i}$$

Multiple

SF(LR)

$$\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} w_1 \\ w_2 \\ w_3 \end{array} \begin{array}{c} \diagup \\ \diagdown \\ \diagup \end{array} \textcircled{\Sigma} \rightarrow \hat{y} = g(\sum w_i x_i) \quad E = \sum (y - g(\sum w_i x_i))^2$$

$$\frac{\partial E}{\partial w_i} = \frac{\partial}{\partial w_i} \sum (y - g(\sum w_i x_i))^2 \quad \frac{\partial u^2}{\partial u} = 2u \cdot \frac{\partial u}{\partial w_i}$$

$$= 2(y - g(\sum w_i x_i)) \cdot \frac{\partial}{\partial w_i} (y - g(\sum w_i x_i))$$

$$= 2 \sum (y - g(\sum w_i x_i)) \cdot (0 - g'(\sum w_i x_i) (1 - g(\sum w_i x_i))) \cdot \frac{\partial}{\partial w_i} (\sum w_i x_i)$$

$$= -2 \sum (y - g(w_i x_i)) \cdot g'(\sum w_i x_i) \cdot (1 - g(\sum w_i x_i)) \cdot x_i$$

$$\begin{array}{ccc} \uparrow & & \uparrow \\ \text{Dev from Truth} * & \text{Der(Act)} * & x_i \\ \hline & & \text{Input} \end{array}$$

$$\text{Sig}(x) = \frac{1}{1+e^{-x}}, \frac{d}{dx}(\text{Sig}(x)) =$$

$$\frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right) = \frac{d}{dx}\left(\frac{1+e^{-x}}{u}\right)' = -1 \cdot (1+e^{-x})^{-1-1} \cdot \frac{d}{dx}(1+e^{-x})$$

$$= -1 \cdot (1+e^{-x})^{-2} \cdot (0 + e^{-x} \cdot -1) = \frac{e^{-x}}{(1+e^{-x})^2}$$

$$= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} = \text{Sig}(x) \cdot \left(\frac{e^{-x} + 1 - 1}{1+e^{-x}}\right)$$

$$= \text{Sig}(x) \cdot \left(1 - \frac{1}{1+e^{-x}}\right) = \underline{\text{Sig}(x)} \cdot \underline{(1 - \text{Sig}(x))}$$

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i} \text{ multiple times}$$

$$\frac{\partial E}{\partial w_i} = -2 \sum (\underbrace{y - \hat{y}}_{\text{Dev from Truth}}) \cdot \underbrace{\text{Der}(\text{Act})}_{\text{Gradient}} \cdot \underbrace{x_i}_{\text{Decent}}$$

$$\text{Dev from Truth} \rightarrow 0 \Rightarrow \frac{\partial E}{\partial w_i} \rightarrow 0$$

