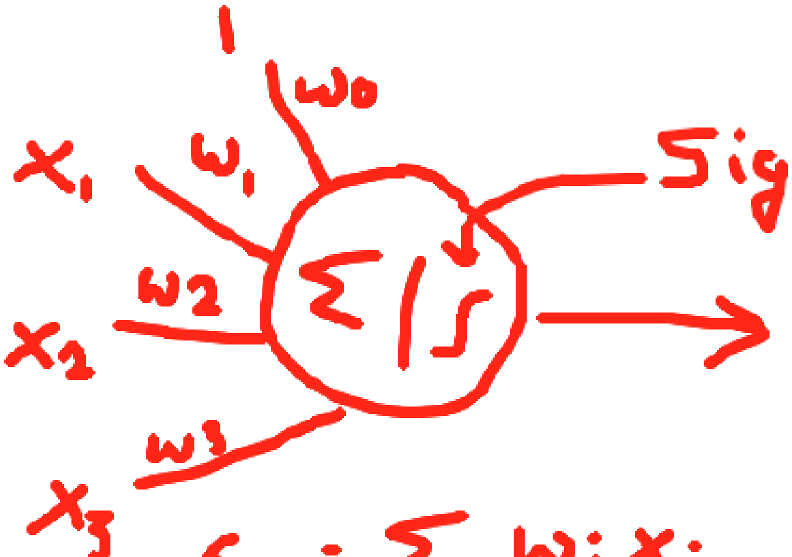
Training an Artificial Neuron

Notes



$$o_p = Sig(\Sigma w_i x_i)$$

$$S_p = \Sigma w_i x_i$$

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

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

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

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

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

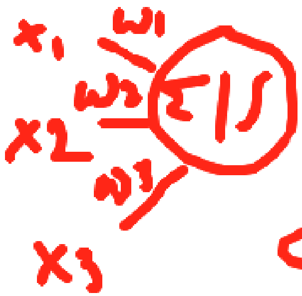
$E = \sum (y - g(\sum w_i x_i))^2$

$E \to 0$, $v \to 0$ brake $= -\frac{dv}{dt}$ ← mult

$-\frac{dE}{dw_i}$

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

SF(LR)



$$\hat{y} = g\left(\sum w_i x_i\right) \qquad E = \sum \left(y - g\left(\sum w_i x_i\right)\right)^2$$

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

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

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

$$= -2\sum \underline{\left(y - g(w_i x_i)\right)} \cdot g\underline{\left(\sum w_i x_i\right) \cdot \left(1 - g(w_i x_i)\right)} \cdot x_i$$

<u>Dev from Truth</u> * <u>Der(Act)</u> * X_i

Input

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

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

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

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

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

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

$$\boxed{\frac{\partial E}{\partial w_i} = -2 \sum \underline{(y - \hat{y})} \cdot \underline{\text{Der}(\text{Act})} \cdot \underline{x_i}}$$

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