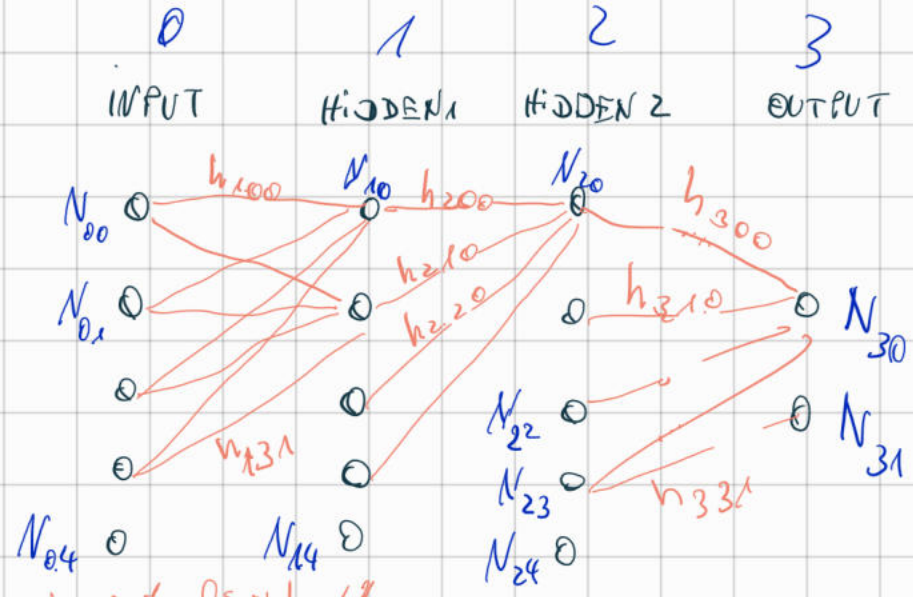
Feed forward:



h_{ijk} $i=1$: weight from input to first hidden

$i=3$: weight from last hidden to output layer

$$N_{10} = N_{00}\cdot h_{100} + N_{01}\cdot h_{110} + N_{02}\cdot h_{120} + N_{03}\cdot h_{130} + N_{04}\cdot h_{140} = \sum_{j=0}^{4} N_{0j}\cdot h_{1j0}$$

With sigmoid function $N_{10} \Rightarrow \sigma = \frac{1}{1+e^{-N_{10}}}$

$$N_{ik} = \sum_{\substack{i=1\\ j=0\\ k=0}}^{j=4} N_{i-1\,j}\cdot h_{ijk}$$

i: layer
j: from
k: to

Backpropagation:

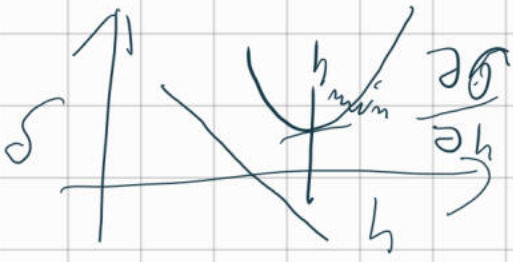
$$O_0 - N_{30} = \delta_0$$

We want $\delta_0 = 0$

$$\Rightarrow O_0 - N_{30} = 0 = O_0 - \sigma(N_{30})$$

$$O_0 = \frac{1}{1+e^{-x}} \quad /\cdot 1$$

$$\frac{1}{O_0} = 1 + e^{-x} \quad /-1$$



$$\frac{1}{O_0} - 1 = e^{-x} \quad / \ln$$

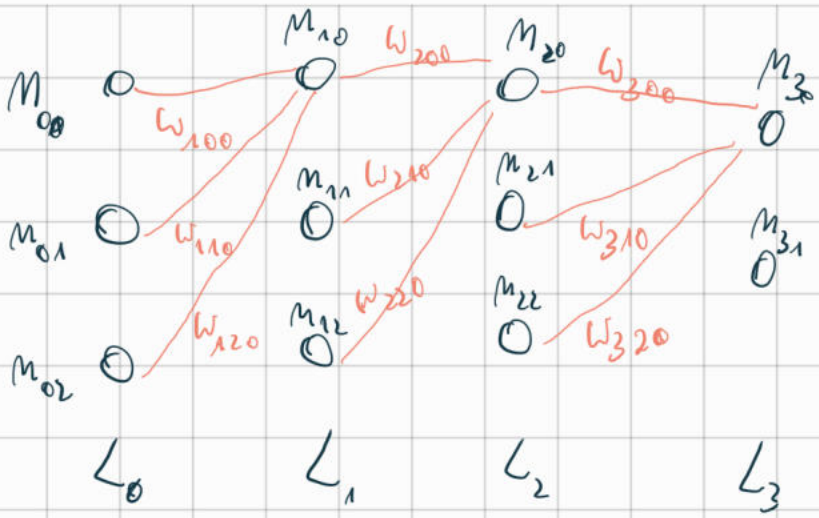
$$\ln\left(\frac{1}{O_0} - 1\right) = -x \quad / \cdot -1$$

$$-\ln\left(\frac{1}{O_0} - 1\right) = x = \sum_{j=0}^{4} N_{i-1\,j} \cdot h_{ijk} \qquad \begin{matrix} i=3 \\ k=0 \end{matrix}$$

Partial: $\frac{\partial (O_k - N_{i,k})}{\partial h_{ijk}} = 0 \qquad \begin{matrix} j = 0, \dots, 4 \\ k = 0 \\ i = 3 \end{matrix}$

$$\frac{\partial N_{ik}}{\partial h_{ijk}} = \frac{\partial}{\partial h_{ijk}} \frac{1}{1+e^{-}}$$

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} \frac{1}{1+e^{-(x+y+z)}} =$$



$\Delta Output = Expected_0 - n_{30}$

$n_{30} = f\left(\sum w_{3i0} \cdot n_{2i}\right)$

$n_{lj} = f\left(\sum w_{lij} \cdot n_{l-1\,i}\right)$

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

$\frac{\Delta Output}{\Delta w} = \frac{\partial n_{lj}}{\partial w_{lij}} = f(1-f)$

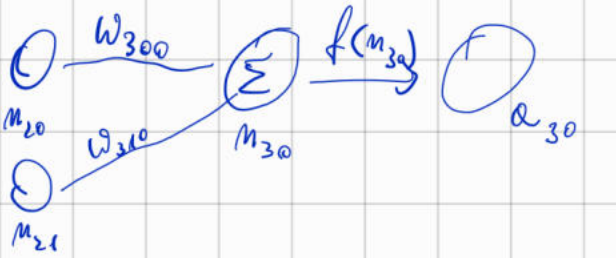
$\Delta w = \frac{\Delta Output}{f(1-f)}$

Error $E_0 = \frac{1}{2}\left(Expected_0 - a_{30}\right)^2$

Output

$E_{tot} = \sum E_j = \frac{1}{2}\sum \left(Expected_j - a_{3j}\right)^2$

how much effect w_{300} on E_{tot}? $\frac{\partial E_{tot}}{\partial w_{300}}$



$\frac{\partial E_{tot}}{\partial a_{30}} \frac{\partial a_{30}}{\partial n_{30}} \frac{\partial n_{30}}{\partial w_{300}} = \frac{\partial E_{tot}}{\partial w_{300}}$

$$\frac{\partial E_{tot}}{\partial a_{30}} = \frac{\partial}{\partial a_{30}}\left(\frac{1}{2}\left(Expected_0 - a_{30}\right)^2\right) = a_{30} - Expected_0$$

$$\frac{\partial a_{30}}{\partial n_{30}} = a_{30}\left(1 - a_{30}\right)$$

$$\frac{\partial n_{30}}{\partial w_{300}} = \frac{\partial}{\partial w_{300}} \sum w_{3i0}\, a_{2i} = a_{20}$$

$$w_{300}^{new} = w_{300}^{old} - \eta \frac{\partial E_{tot}}{\partial w_{300}} = w_{300}^{old} - \eta \left(a_{30} - Expected_0\right) \cdot a_{30}\left(1 - a_{30}\right) \cdot a_{20}$$

---

hidden

$$\frac{\partial E_{tot}}{\partial w_{200}} = \frac{\partial E_{tot}}{\partial a_{20}} \frac{\partial a_{20}}{\partial n_{20}} \frac{\partial n_{20}}{\partial w_{200}}$$

$$\Downarrow$$

$$\frac{\partial E_0}{\partial a_{20}} + \frac{\partial E_1}{\partial a_{20}}$$

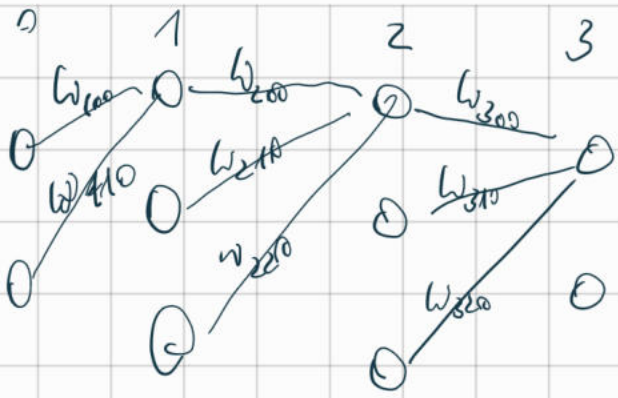
$$\frac{\partial E_j}{\partial a_{20}} = \frac{\partial E_j}{\partial n_{3j}} \frac{\partial n_{3j}}{\partial a_{20}}$$

$$\frac{\partial E_j}{\partial n_{3j}} = a_{3j} - Expected_j$$

$$\frac{\partial n_{3j}}{\partial a_{20}} = \frac{\partial}{\partial a_{20}}\left(w_{300} \cdot a_{20} + w_{310}\, a_{21} + \dots\right) = w_{300}$$

$$\frac{\partial a_{20}}{\partial n_{20}} = a_{20}\left(1 - a_{20}\right) \qquad \frac{\partial n_{20}}{\partial w_{200}} = \frac{\partial}{\partial w_{200}}\left(w_{200} \cdot a_{10} + w_{210} \cdot a_{11} + \dots\right) = a_{10}$$

$$w_{200}^{new} = w_{200}^{old} - \eta \frac{\partial E_{tot}}{\partial w_{200}}$$

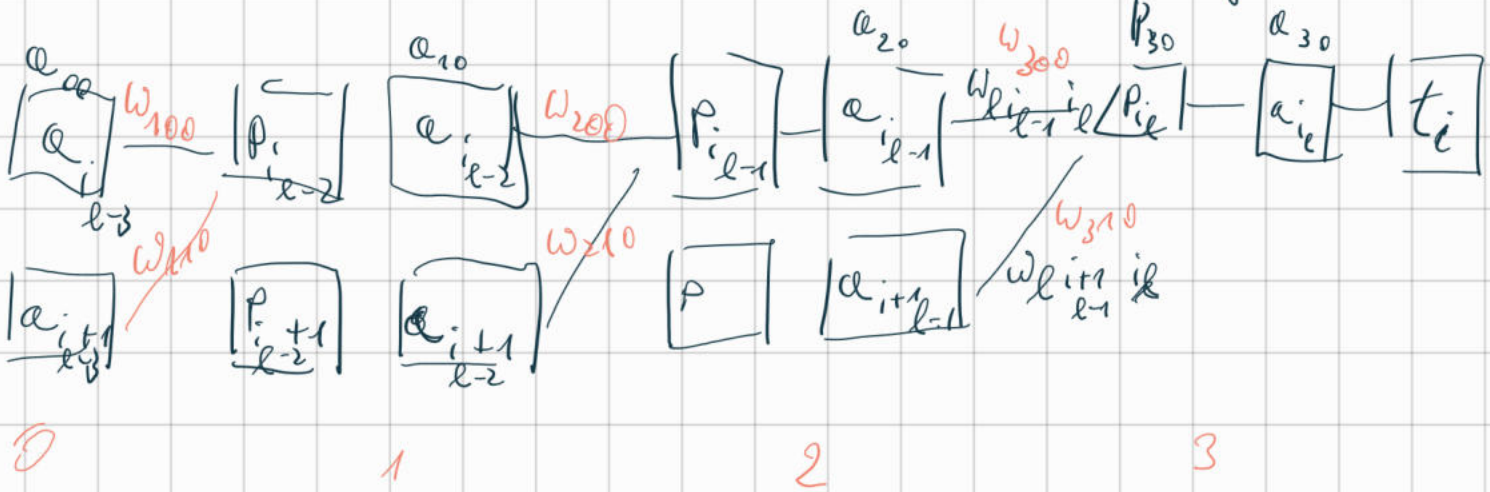


$$\frac{\partial E_{tot}}{\partial w_{100}} = \frac{\partial E_{tot}}{\partial a_{10}} \frac{\partial a_{10}}{\partial n_{10}} \frac{\partial n_{10}}{\partial w_{100}}$$

i_ℓ = neuron in layer ℓ

ℓ = layer

t_i = target output i



$$E_i = \frac{1}{2}(t_i - a_{i\,\ell_{max}})^2 \quad \Rightarrow E_0 = \frac{1}{2}(t_0 - a_{30})^2$$

$$a_{30} = f(p_{30}) \qquad p_{30} = a_{20} \cdot w_{300} + a_{21} \cdot w_{310}$$

$$a_{31} = f(p_{31}) \qquad p_{31} = a_{20} \cdot w_{301} + a_{21} \cdot w_{311}$$

$$a_{20} = f(p_{20}) \qquad p_{20} = a_{10} \cdot w_{200} + a_{11} w_{210}$$

$$a_{21} = f(p_{21}) \qquad p_{21} = a_{10} \cdot w_{201} + a_{11} \cdot w_{211}$$

$$a_{10} = f(p_{10}) \qquad p_{10} = a_{00} \cdot w_{100} + a_{01} \cdot w_{110}$$

$$a_{11} = f(p_{11}) \qquad p_{11} = a_{00} \cdot w_{101} + a_{01} \cdot w_{111}$$

(output)

$$\frac{\partial E_0}{\partial w_{300}} = \frac{\partial E_0}{\partial a_{30}} \frac{\partial a_{30}}{\partial p_{30}} \frac{\partial p_{30}}{\partial w_{300}}$$

$$\frac{\partial E_0}{\partial a_{30}} = \frac{2}{2}\left(t_0 - a_{30}\right)\cdot -1 = a_{30} - t_0$$

$$\frac{\partial a_{30}}{\partial p_{30}} = \frac{d f(p_{30})}{d p_{30}} = a_{30}\left(1 - a_{30}\right)$$

$$\frac{\partial p_{30}}{\partial w_{300}} = \frac{\partial}{\partial w_{300}}\left(a_{20} w_{300} + a_{21} \cdot w_{310}\right) = a_{20}$$

$$\frac{\partial E_K}{\partial w_{ijk}} = \frac{\partial E_k}{\partial a_{ik}} \frac{\partial a_{ik}}{\partial p_{ik}} \frac{\partial p_{ik}}{\partial w_{ijk}}$$

$$\frac{\partial E_K}{\partial a_{ik}} = a_{ik} - t_k$$

$$\frac{\partial a_{ik}}{\partial p_{ik}} = a_{ik}\left(1 - a_{ik}\right)$$

$$\frac{\partial p_{ik}}{\partial w_{ijk}} = a_{i-1\,j}$$

i = layer (output)

j = from $(i-1)$

k = to (i)

$$w_{ijk}^{new} = w_{ijk}^{old} - \eta \frac{\partial E_K}{\partial w_{ijk}}$$

Hidden

$$\frac{\partial E_{tot}}{\partial w_{200}} = \frac{\partial E_{tot}}{\partial a_{20}} \frac{\partial a_{20}}{\partial \rho_{20}} \frac{\partial \rho_{20}}{\partial w_{200}}$$

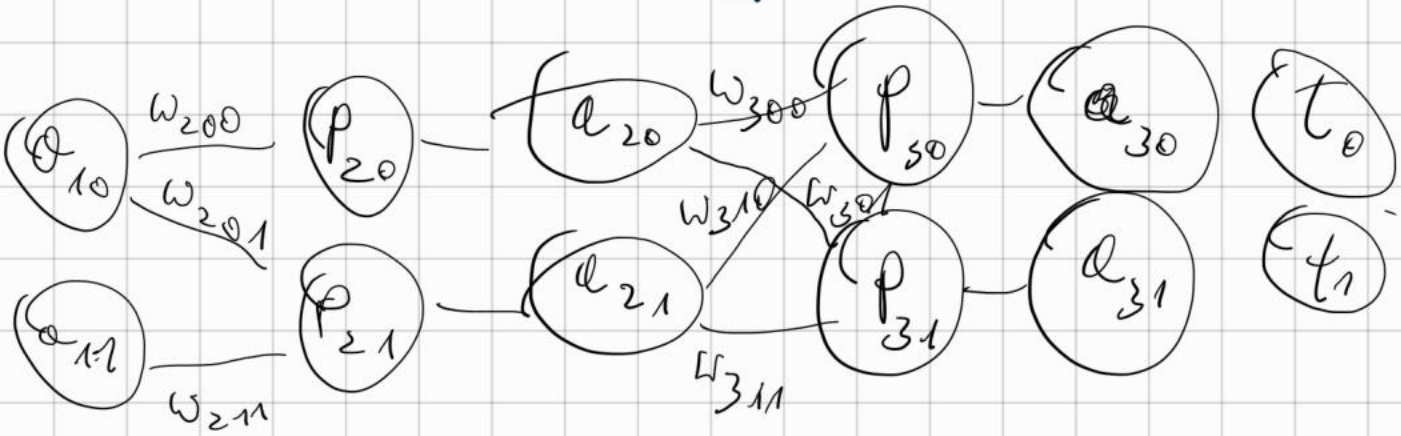
$$\frac{\partial E_{tot}}{\partial a_{20}} = \frac{\partial E_0}{\partial a_{20}} + \frac{\partial E_1}{\partial a_{20}} = \frac{\partial E_0}{\partial \rho_{30}} \frac{\partial \rho_{30}}{\partial a_{20}} + \frac{\partial E_1}{\partial \rho_{31}} \frac{\partial \rho_{31}}{\partial a_{20}}$$

$$\frac{\partial E_{tot}}{\partial a_{i-1\,j}} = \sum_k \frac{\partial E_k}{\partial \rho_{ik}} \frac{\partial \rho_{ik}}{\partial a_{i-1\,j}}$$

$$\frac{\partial E_{tot}}{\partial w_{i-1\,jk}} = \sum_k \frac{\partial E_k}{\partial \rho_{ik}} \frac{\partial \rho_{ik}}{\partial a_{i-1\,j}} \cdot \frac{\partial a_{i-1\,j}}{\partial \rho_{i-1\,j}} \frac{\partial \rho_{i-1\,j}}{\partial w_{i-1\,jk}}$$

$$\frac{\partial E_k}{\partial \rho_{ik}} = \frac{\partial}{\partial \rho_{ik}} \frac{1}{2}\left(t_k - a_{ik}\right)^2 \Rightarrow a_{ik} = \frac{1}{1+e^{-\rho_{ik}}} = f(\rho_{ik})$$

$$= \frac{2}{2}\left(t_k - f(\rho_{ik})\right)(-1)\frac{df(\rho_{ik})}{d\rho_{ik}} = -(t_k - f)$$



$$\frac{\partial E_{tot}}{\partial a_{20}} = \frac{\partial}{\partial a_{20}} \left(\frac{1}{2} \left[t_0 - a_{30} \right]^2 + \frac{1}{2} \left[t_1 - a_{31} \right]^2 \right)$$

$$a_{30} = f(p_{30}) \qquad p_{30} = a_{20} w_{300} + a_{21} \cdot w_{310}$$

$$a_{31} = f(p_{31}) \qquad p_{31} = a_{20} \cdot w_{301} + a_{21} \cdot w_{311}$$

$$\frac{\partial}{\partial a_{20}} \frac{1}{2} \left(t_0 - a_{30} \right)^2 = \frac{\partial}{\partial a_{20}} \frac{1}{2} \left(t_0 - f(p_{30}) \right)^2$$

$$= \frac{2}{2} \left(t_0 - f(p_{30}) \right) (-1) \frac{\partial f}{\partial a_{20}} = \left(f(p_{30}) - t_0 \right) f'(p_{30})$$

$$= \left(a_{30} - t_0 \right) \frac{\partial}{\partial a_{20}} \left(a_{20} w_{300} \right) = w_{300} \left(a_{30} - t_0 \right) = w_{ijk} \left(a_{ik} - t_k \right)$$

$$\frac{\partial E_1}{\partial a_{20}} = \frac{\partial}{\partial a_{20}} \frac{1}{2} \left(t_1 - a_{31} \right)^2 = \frac{\partial}{\partial a_{20}} \frac{1}{2} \left(t_1 - f(p_{31}) \right)^2$$

$$= w_{301} \left(a_{31} - t_1 \right)$$

$$\frac{\partial E_{tot}}{\partial a_{i-1\,j}} = \sum_k w_{ijk} \left(a_{ik} - t_k \right)$$

$$\frac{\partial a_{i-1j}}{\partial \rho_{i-1j}} = \frac{\partial}{\partial \rho} f = f' = a_{i-1j}\left(1 - a_{i-1j}\right)$$

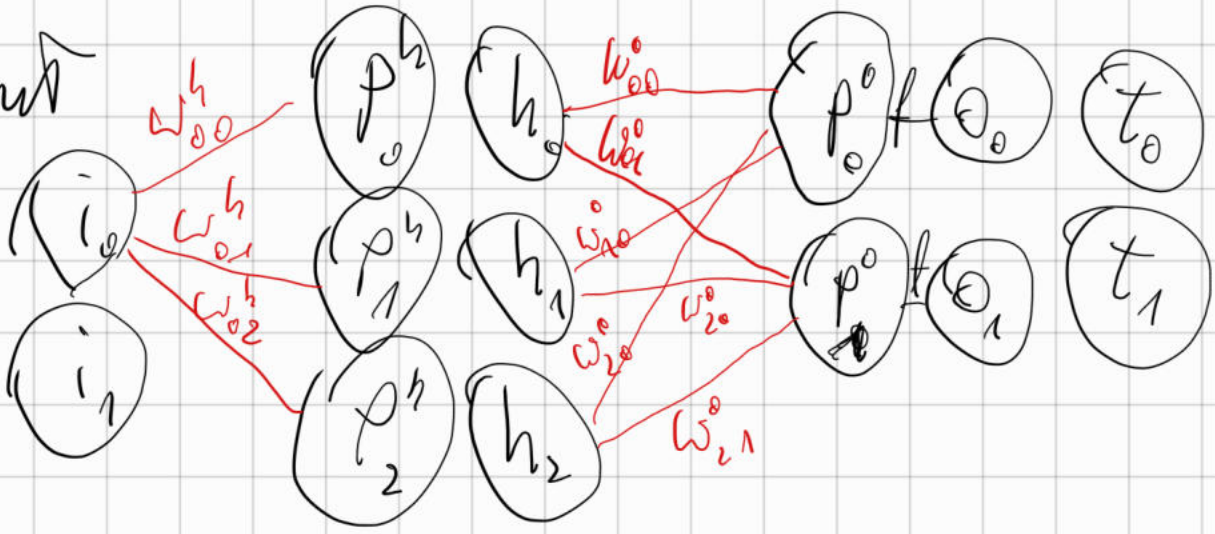
$$\frac{\partial \rho_{20}}{\partial \omega_{200}} = \frac{\partial}{\partial \omega_{200}}\left(a_{10}\cdot \omega_{200} + a_{11}\omega_{210}\right) \equiv a_{10}$$

$$\frac{\partial \rho_{i-1k}}{\partial \omega_{i-1jk}} = a_{i-2j}$$

$$\frac{\partial E_{tot}}{\partial \omega_{ijk}} = \frac{\partial E_{tot}}{\partial a_{ik}} \frac{\partial a_{ik}}{\partial \rho_{ik}} \frac{\partial \rho_{ik}}{\partial \omega_{ijk}}$$

$$= \left[\sum_l \omega_{imaxl}\left(a_{imaxl} - t_l\right)\right] \cdot a_{ik}\left(1 - a_{ik}\right) \cdot a_{i-1j}$$

Output



$$\frac{\partial E_{tot}}{\partial \omega^o_{21}} = \frac{\partial}{\partial \omega^o_{21}} \left(\frac{1}{2}\left[-O_0+t_0\right]^2 \right) + \frac{\partial}{\partial \omega^o_{21}} \frac{1}{2}\left(O_1+t_1\right)^2$$

$$O_0 = f(p^o_0) = f_0 \qquad p^o_0 = h_0\omega^o_{00} + h_1\omega^o_{10} + h_2\omega^o_{20}$$

$$O_1 = f(p^o_1) = f_1 \qquad p^o_1 = h_0\omega^o_{01} + h_1\omega^o_{11} + h_2\omega^o_{21}$$

$$\frac{\partial E_0}{\partial \omega^o_{21}} = \frac{\partial}{\partial \omega^o_{21}} \frac{1}{2}\left(t_0-O_0\right)^2 = \frac{1}{2}\left(t_0-O_0\right)\cdot 0 = 0$$

$$\frac{\partial E_1}{\partial \omega^o_{21}} = \frac{\partial}{\partial \omega^o_{21}} \frac{1}{2}\left(t_1-O_1\right)^2 = -\left(t_1-O_1\right)\cdot \frac{dO_1}{d\omega^o_{21}}$$

$$\frac{dO_1}{d\omega^o_{21}} = \frac{1}{1+e^{-p^o_1}} = O_1(1-O_1)\cdot h_2$$

$$\frac{\partial E_i}{\partial \omega_{ijk}} = \left(-t_k + o_{ik}\right)\cdot o_{ik}\cdot\left(1-o_{ik}\right)\cdot o_{i-1\,j}$$

i = output idx
j = from (hidden)
k = to (output)

$$\frac{\partial E_k}{\partial w_{ijk}} = \left(-t_k + t_k a_{ik} + a_{ik} - a_{ik}^2\right) \cdot a_{ik} \cdot a_{i-1\,j}$$

$$= \left(-a_{ik}^3 + (t_k+1)a_{ik}^2 - t_k a_{ik}\right) a_{i-1\,j}$$

$$\frac{\partial E_{tot}}{\partial w_{ijk}} = \sum_k \frac{\partial E_k}{\partial w_{ijk}}$$

$$w_{ijk}^{new} = w_{ijk}^{old} - \eta \frac{\partial E_{tot}}{\partial w_{ijk}}$$

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

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

$$\Rightarrow \frac{df}{dx} + f = \frac{1}{1-e^{-x}} \Rightarrow f' = f(1-f)$$

Widoden $\frac{\partial E_{tot}}{\partial w^h_{01}} = \frac{\partial E_0}{\partial w^h_{01}} + \frac{\partial E_1}{\partial w^h_{01}} =$

$\frac{\partial}{\partial w^h_{01}} \frac{1}{2}(t_0 - o_0)^2 + \frac{\partial}{\partial w^h_{01}} \frac{1}{2}(t_1 - o_1)^2 =$

$o_0 = f(p^o_0) \qquad p^o_0 = h_0 w^o_{00} + h_1 w^o_{10} + h_2 w^o_{20}$

$o_1 = f(p^o_1) \qquad p^o_1 = h_0 w^o_{01} + h_1 w^o_{11} + h_2 w^o_{21}$

$p^h_0 = i_0 w^h_{00} + i_1 w^h_{10} \qquad p^h_1 = i_0 w^h_{01} + i_1 w^h_{11} \qquad p^h_2 = i_0 w^h_{02} + i_1 w^h_{12}$

$p^o_0 = f(p^h_0) w^o_{00} + f(p^h_1) w^o_{10} + f(p^h_2) w^o_{20}$

$\frac{\partial E_0}{\partial w^h_{01}} = (t_0 - o_0) \cdot (-1) \frac{\partial o_0}{\partial w^h_{01}} = (o_0 - t_0) \frac{\partial}{\partial w^h_{01}} f(p^o_0) =$

$(o_0 - t_0) f(p^o_0)(1 - f(p^o_0)) \frac{\partial p^o_0}{\partial w^h_{01}} = (o_0 - t_0) o^o_0 (1 - o^o_0) \frac{\partial p^o_0}{\partial w^h_{01}}$

$\frac{\partial p^o_0}{\partial w^h_{01}} = \frac{\partial}{\partial w^h_{01}} \left[f(p^h_0) \cdot w^o_{00} + f(p^h_1) w^o_{10} + f(p^h_2) w^o_{20} \right]$

$\frac{\partial}{\partial w^h_{01}} \left[f(p^h_1) w^o_{10} \right] = w^o_{10} \left[h_1 (1 - h_1) \cdot \frac{\partial p^h_1}{\partial w^h_{01}} \right]$

$$\frac{\partial p_1^h}{\partial w_{01}^h} = \frac{\partial}{\partial w_{01}^h}\left(i_0 w_{01}^h + i_1 w_{11}^h\right) = i_0$$

i = layer m = idx of output layer

j = from n = hidden idx

k = to

l = output idx

$$\frac{\partial E_{tot}}{\partial w_{ijk}} = \sum_l \frac{\partial E_l}{\partial w_{ijk}} = \sum_l \left(a_{ml} - t_l\right) a_{ml}\left(1 - a_{ml}\right) \frac{\partial p_{ml}}{\partial w_{ijk}}$$

$$\frac{\partial p_{ml}}{\partial w_{ijk}} = \sum_n \frac{\partial}{\partial w_{ijk}} f\left(p_{m-1\,n}\right) w_{mnl} = \sum_n w_{mnl} \cdot a_{m-1\,n}\left(1 - a_{m-1\,n}\right) \frac{\partial p_{m-1\,n}}{\partial w_{ijk}}$$

E in Hidden layer:

$$\frac{\partial E_{tot}}{\partial w_{1jk}} = \sum_i \left(a_{2i} - t_i\right) a_{2i}\left(1 - a_{2i}\right) = \left(a_{2i}^2 - t_i a_{2i}\right)\left(1 - a_{2i}\right) \frac{\partial p_{2i}}{\partial w_{1jk}}$$

$$\left(a_{2i}^2 - a_{2i}^3 - t_i a_{2i} + t_i a_{2i}^2\right) \frac{\partial p_{2i}}{\partial w_{1jk}} = \left(-a_{2i}^3 + \left(1 + t_i\right) a_{2i}^2 - t_i a_{2i}\right) \frac{\partial p_{2i}}{\partial w_{1jk}}$$

$$p_{2i} = \sum_k a_{1k} w_{2ki} \qquad \frac{\partial p_{2i}}{\partial w_{1jk}} = \sum_k w_{2ki} \frac{\partial a_{1k}}{\partial w_{1jk}}$$

$$a_{1k} = \sum_j a_{0j} \cdot w_{1jk} \qquad \frac{\partial a_{1k}}{\partial w_{1jk}} = a_{0j}$$

---

3 layer

$$\underset{j}{\text{input}} \rightarrow \underset{k}{\text{hidden}} \rightarrow \underset{i}{\text{output}}$$

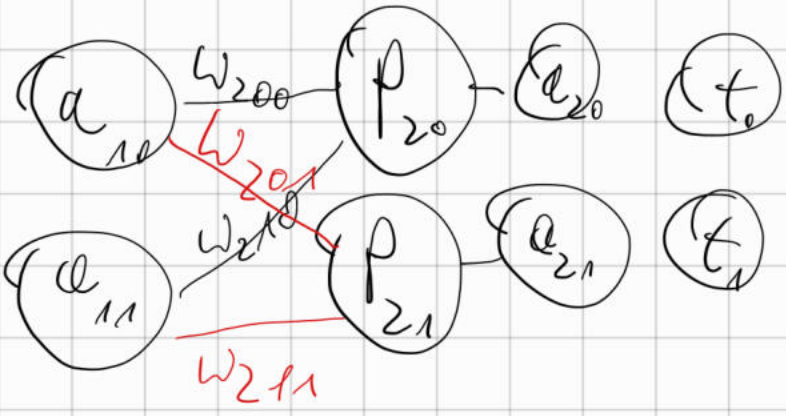
w_{1xx} w_{2xx}

$$w_{1jk}^{new} = w_{1jk}^{old} - \eta \sum_i \left(-e_{2i}^3 + (1+t_i) a_{2i}^2 - t_i a_{2i}\right) \frac{\partial p_{2i}}{\partial w_{1jk}}$$

$$w_{2jk}^{new} = w_{2jk}^{old} - \eta \sum_i \left(-e_{2i}^3 + (1+t_i) a_{2i}^2 - t_i a_{2i}\right) \frac{\partial p_{2i}}{\partial w_{2jk}}$$

$$\frac{\partial p_{2i}}{\partial w_{2jk}} = a_{2i}(1-a_{2i}) a_{1j}$$

Output



$$\frac{\partial E_{total}}{\partial w_{210}} = \frac{\partial E_0}{\partial w_{210}} + \frac{\partial E_1}{\partial w_{210}} = \sum_m \frac{\partial E_m}{\partial w_{210}}$$

$$E_m = \frac{1}{2}(t_m - a_{2m})^2 \qquad \frac{\partial E_m}{\partial w_{210}} = (t_m - a_{2m})(-1)\frac{\partial a_{2m}}{\partial w_{210}}$$

$$a_{20} = f(a_{10} w_{200} + a_{11} w_{210})$$

$$a_{2m} = f\left(\sum_n a_{1n} w_{2mn}\right)$$

$$\frac{\partial a_{2m}}{\partial w_{2xy}} = a_{2m}(1 - a_{2m})\, a_{1m} \quad (\text{für } m = x)$$

Loop 5 times (z)
  Loop to (m)

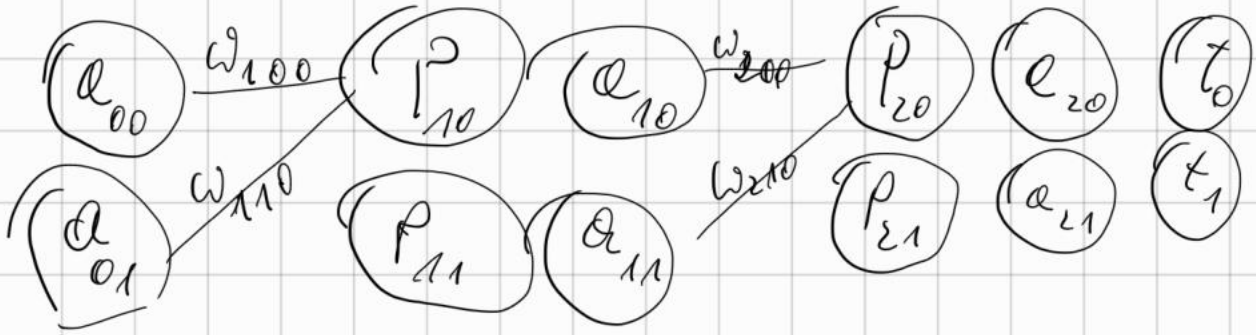
m : Neuron Output

n : Neuron hidden

    dedw = $a_{2m}(1 - a_{2m})\, a_{1n}$ (für $m = x$)

  → dEdw += $(a_{2m} - t_m) \cdot$ dedw

→ $w_{2xy}^{new} = w_{2xy}^{old} - \eta \cdot$ dEdw



$$\frac{\partial E_{tot}}{\partial w_{110}} = \sum_m \frac{\partial E_m}{\partial w_{110}} = \sum_n (a_{2n} - t_n) \frac{\partial a_{2n}}{\partial w_{110}}$$

$$a_{2n} = f(p_{2n}) \qquad \frac{\partial a_{2n}}{\partial w_{1xy}} = a_{2n}(1 - a_{2n}) \frac{\partial p_{2n}}{\partial w_{1xy}}$$

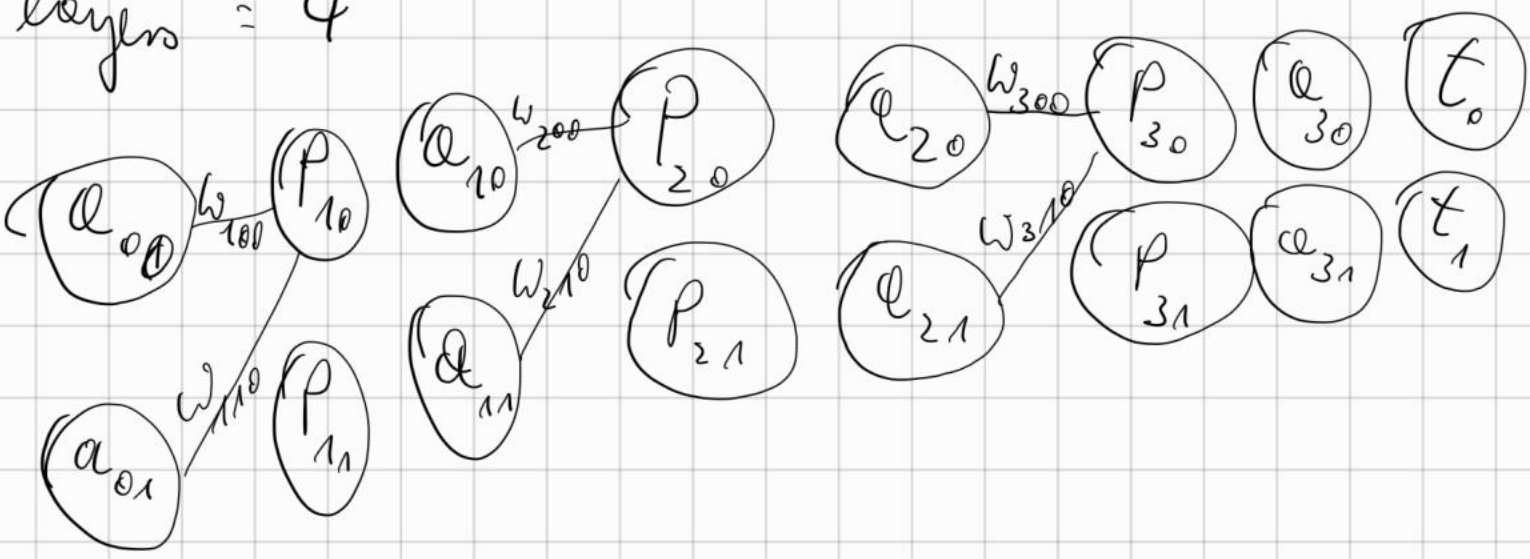
$$p_{2n} = \sum_m a_{1m} \cdot w_{2mn} \qquad \frac{\partial p_{2n}}{\partial w_{1xy}} = \sum_m w_{2mn} \frac{\partial a_{1m}}{\partial w_{1xy}}$$

$$a_{1m} = f(p_{1m}) \qquad \frac{\partial a_{1m}}{\partial w_{1xy}} = a_{1m}(1 - a_{1m}) \cdot \frac{\partial p_{1m}}{\partial w_{1xy}}$$

$$p_{1m} = \sum_l a_{0l} w_{1lm} \qquad \frac{\partial p_{1m}}{\partial w_{1xy}} = a_{0l} \quad (\text{für } l = x)$$

Loop Synapses layer 1
  Loop n
    Loop m
      $da_1 dw_1 = a_{1m}(1 - a_{1m}) \cdot a_{0l} \quad (l = x)$
      → $dp_2 dw_1 += w_{2mn} \cdot da_1 dw_1$
    $da_2 dw_1 = a_{2n}(1 - a_{2n}) \cdot dp_2 dw_1$
    → $dE dw_1 += (a_{2n} - t_n) \cdot da_2 dw_1$
  → $w_{1xy}^{neu} = w_{1xy}^{old} - \eta \, dE dw_1$

layers = 4



$$\frac{\partial E_{tot}}{\partial w_{210}} = \sum_k \frac{\partial E_k}{\partial w_{210}} \qquad E_k = \frac{1}{2}\left(t_k - a_{3k}\right)^2$$

$$\frac{\partial E_k}{\partial w_{210}} = \left(a_{3k} - t_k\right) \frac{\partial a_{3k}}{\partial w_{210}} \qquad a_{3k} = f(p_{3k})$$

$$\frac{\partial a_{3k}}{\partial w_{210}} = a_{3k}(1 - a_{3k}) \frac{\partial p_{3k}}{\partial w_{210}}$$

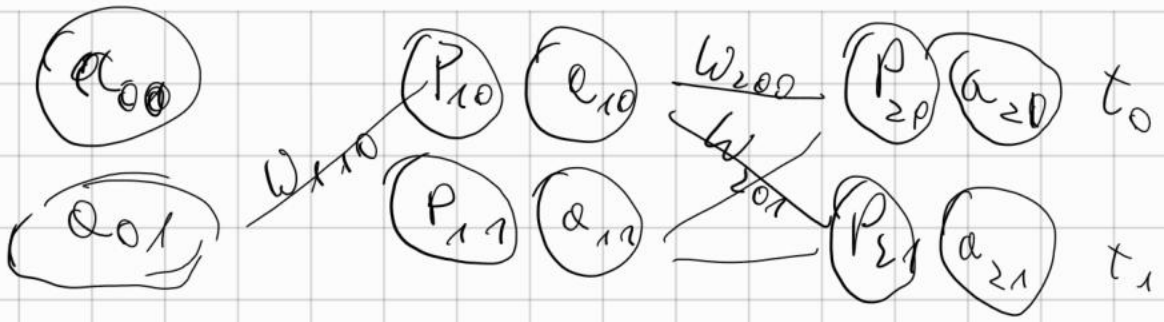
$$p_{3k} = \sum_m a_{2m} \cdot w_{3mk} \qquad \frac{\partial p_{3k}}{\partial w_{210}} = \sum_m w_{3mk} \frac{\partial a_{2m}}{\partial w_{210}}$$

$$a_{2m} = f(p_{2m}) \qquad \frac{\partial a_{2m}}{\partial w_{210}} = a_{2m}(1 - a_{2m}) \frac{\partial p_{2m}}{\partial w_{210}}$$

$$p_{2m} = \sum_i a_{1i} w_{2im} \qquad \frac{\partial p_{2m}}{\partial w_{210}} = \sum_i w_{2im} \frac{\partial a_{1i}}{\partial w_{210}} = w_{2im} \cdot a_{1i}$$

$i=1 \quad m=0$

3 layers:



$$\frac{\partial E_0}{\partial w_{110}} = \frac{\partial a_{20}}{\partial w_{110}} (a_{20} - t_0) = a_{20}(1-a_{20})(a_{20}-t_0)\frac{\partial p_{20}}{\partial w_{110}}$$

$$\frac{\partial p_{20}}{\partial w_{110}} = w_{200}\frac{\partial a_{10}}{\partial w_{110}} + w_{210}\frac{\partial a_{11}}{\partial w_{110}}$$

$$\frac{\partial E_1}{\partial w_{110}} = \frac{\partial a_{21}}{\partial w_{110}} (a_{21} - t_1) = a_{21}(1-a_{21})(a_{21}-t_1)\frac{\partial p_{21}}{\partial w_{110}}$$

$$\frac{\partial p_{21}}{\partial w_{110}} = w_{201}\frac{\partial a_{10}}{\partial w_{110}} + w_{211}\frac{\partial a_{11}}{\partial w_{110}}$$

$$\frac{\partial a_{10}}{\partial w_{110}} = a_{10}(1-a_{10})\cdot a_{01}$$

$$\partial E_{tot} = a_{20}(1-a_{20})(a_{20}-t_0)w_{200}\cdot a_{10}(1-a_{10})\cdot a_{01} + a_{21}(1-a_{21})(a_{21}-t_1)w_{201}\cdot a_{10}(1-a_{10})a_{01}$$

$$\frac{\partial E_{tot}}{\partial w_{1xy}} = \sum_m a_{2m}(1-a_{2m})(a_{2m}-t_m)w_{2ym}\cdot a_{1y}(1-a_{1y})\cdot a_{0x}$$

4 layers

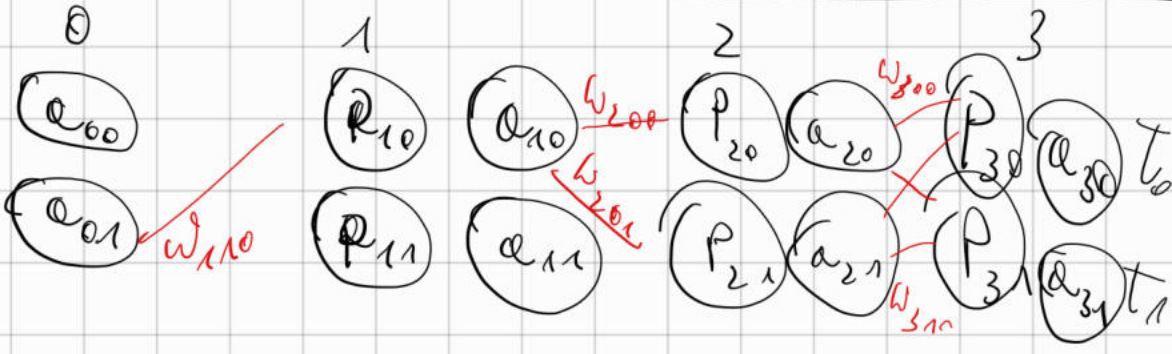
layeridx = 2

1	2	3
a_{10}	$p_{20}\ a_{20}$	$p_{30}\ a_{30}\ t_0$
a_{11} ω_{210}	$p_{21}\ a_{21}$	$p_{31}\ a_{31}\ t_1$

$$\frac{\partial E_{tot}}{\partial \omega_{2xy}} = \sum_m a_{3m}(1-a_{3m})(a_{3m}-t_m)\,\omega_{3ym}\,a_{2y}(1-a_{2y})\,a_{1x}$$

---

layeridx = 1



$$\frac{\partial E_a}{\partial \omega_{110}} = \frac{\partial a_{30}}{\partial a_{110}}\left(a_{30}-t_a\right) = a_{30}(1-a_{30})(a_{30}-t_a)\frac{\partial p_{30}}{\partial \omega_{110}}$$

$$\frac{\partial p_{30}}{\partial \omega_{110}} = \frac{\partial}{\partial \omega_{110}}\left(a_{20}\cdot\omega_{300} + a_{21}\cdot\omega_{310}\right) = \omega_{300}\frac{\partial a_{20}}{\partial \omega_{110}} + \omega_{310}\frac{\partial a_{21}}{\partial \omega_{110}}$$

$$\frac{\partial a_{20}}{\partial \omega_{110}} = a_{20}(1-a_{20})\frac{\partial p_{20}}{\partial \omega_{110}} \qquad \frac{\partial a_{21}}{\partial \omega_{110}} = a_{21}(1-a_{21})\frac{\partial p_{21}}{\partial \omega_{110}}$$

$$\frac{\partial p_{20}}{\partial \omega_{110}} = \frac{\partial}{\partial \omega_{110}}\left(a_{10}\cdot\omega_{200} + a_{11}\cdot\omega_{210}\right) = \omega_{200}\frac{\partial a_{10}}{\partial \omega_{110}} + \omega_{210}\frac{\partial a_{11}}{\partial \omega_{110}}$$

$$\frac{\partial p_{21}}{\partial \omega_{110}} = \omega_{201}\frac{\partial a_{10}}{\partial \omega_{110}} + \omega_{211}\frac{\partial a_{11}}{\partial \omega_{110}}$$

$$\frac{\partial a_{10}}{\partial w_{110}} = a_{10}(1-a_{10}) \frac{\partial p_{10}}{\partial w_{110}} \qquad \frac{\partial a_{11}}{\partial w_{110}} = a_{11}(1-a_{11}) \frac{\partial p_{11}}{\partial w_{110}}$$

$$\frac{\partial p_{10}}{\partial w_{110}} = a_{01} \qquad \frac{\partial p_{11}}{\partial w_{110}} = 0$$

$$\Rightarrow \frac{\partial p_{21}}{\partial w_{110}} = w_{201} \cdot a_{10}(1-a_{10}) a_{01}$$

$$\frac{\partial p_{20}}{\partial w_{110}} = w_{200} \cdot a_{10}(1-a_{10}) a_{01}$$

$$\frac{\partial a_{20}}{\partial w_{110}} = a_{20}(1-a_{20}) w_{200} \cdot a_{10}(1-a_{10}) a_{01}$$

$$\frac{\partial a_{21}}{\partial w_{110}} = a_{21}(1-a_{21}) w_{201} \cdot a_{10}(1-a_{10}) a_{01}$$

$$\frac{\partial p_{30}}{\partial w_{110}} = w_{300} \cdot a_{20}(1-a_{20}) w_{200} \cdot a_{10}(1-a_{10}) a_{01} +$$
$$w_{310} \cdot a_{21}(1-a_{21}) w_{201} \cdot a_{10}(1-a_{10}) \cdot a_{01}$$

$$\frac{\partial p_{3n}}{\partial w_{1xy}} = \sum_m w_{3mn} \cdot a_{2m}(1-a_{2m}) w_{2qm} \cdot a_{1y}(1-a_{1y}) a_{0x}$$

$$\frac{\partial E_{tot}}{w_{1xy}} = \sum_n a_{3n}(1-a_{3n})(a_{3n}-t_n) \frac{\partial p_{3n}}{\partial w_{1xy}}$$

layer idx = 2 0 1 2 3

a_{00} p_{10} a_{10} w_{200} p_{20} a_{20} w_{300} p_{30} a_{30} t_0

a_{01} w_{210} w_{310}

p_{11} a_{11} p_{21} a_{21} p_{31} a_{31} t_1

$$\frac{\partial E_{TOT}}{\partial w_{210}} = \frac{\partial E_0}{\partial w_{210}} + \frac{\partial E_1}{\partial w_{210}}$$

$$\frac{\partial E_0}{\partial w_{210}} = \frac{\partial}{\partial w_{210}} \frac{1}{2}(t_0 - a_{30})^2 = (t_0 - a_{30})(-1)\frac{\partial a_{30}}{\partial w_{210}} = (a_{30} - t_0)\frac{\partial a_{30}}{\partial w_{210}}$$

$$\frac{\partial a_{30}}{\partial w_{210}} = \frac{\partial}{\partial w_{210}} f(p_{30}) = a_{30}(1 - a_{30})\frac{\partial p_{30}}{\partial w_{210}}$$

$$p_{30} = a_{20} w_{300} + a_{21} w_{310} \qquad \frac{\partial p_{30}}{\partial w_{210}} = w_{300}\frac{\partial a_{20}}{\partial w_{210}}$$

$$\frac{\partial a_{20}}{\partial w_{210}} = a_{20}(1 - a_{20})\frac{\partial p_{20}}{\partial w_{210}} \qquad p_{20} = a_{10} w_{200} + a_{11} w_{210}$$

$$\frac{\partial p_{20}}{\partial w_{210}} = a_{11}$$

$$\frac{\partial E_0}{\partial w_{210}} = (a_{30} - t_0)\, a_{30}(1 - a_{30})\, w_{300}\, a_{20}(1 - a_{20})\, a_{11}$$

$$\frac{\partial E_m}{\partial w_{xy}} = (a_{3m} - t_m)\, a_{3m}(1 - a_{3m})\, w_{3ym}\, a_{2y}(1 - a_{2y})\, a_{1x}$$