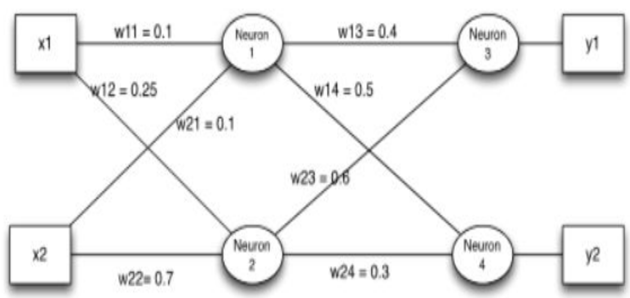
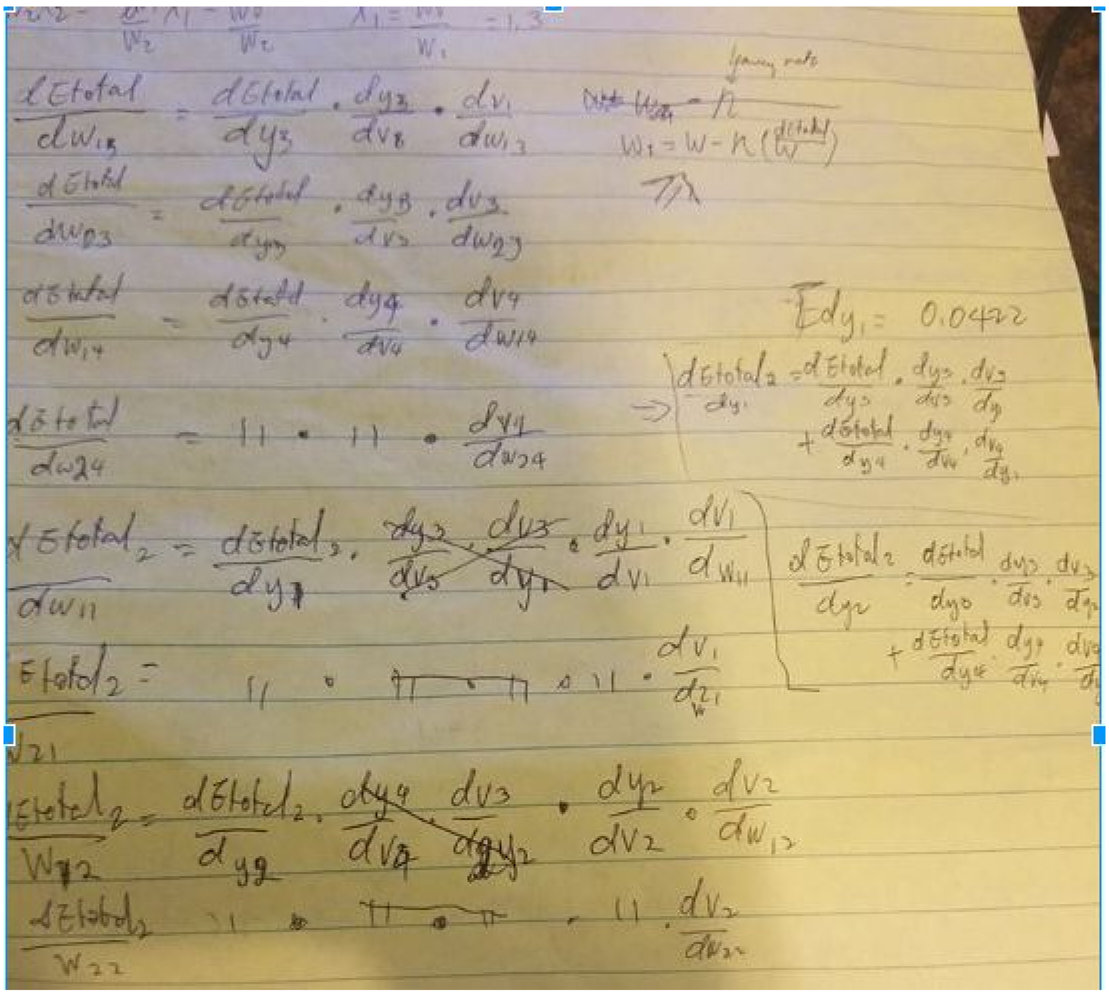
1)
Result of the New Weights

nw11	0.0990
nw12	0.2500
nw13	0.4042
nw14	0.4920
nw21	0.0990
nw22	0.7000
nw23	0.6056
nw24	0.2895

1. Consider the Neural Network below.





The code : given x1=1 and x2=1 . Done in MATLAB
w11=0.1;
w12=0.25;
w21=0.1;

```
w22=0.7;
w13=0.4;
w14=0.5;
w23=0.6;
w24=0.3;
v1=w11*x1+w21*x2;
v2=w12*x1+w22*x2;
y1=1/(1+exp(-1*v1));
y2=1/(1+exp(-1*v2));
v3=w13*y1+w23*y2;
v4=w14*y1+w24*y2;
y3=1/(1+exp(-1*v3));
y4=1/(1+exp(-1*v4));
d=[1;0];
pred=[y3;y4];
E=sum((pred-d).^2/2);
Edy3=y3-1;
Edy4=y4-0;
y3dv3=y3^2*exp(-v3);
y4dv4=y4^2*exp(-v4);
v3dw13=y1;
v3dw23=y2;
v4dw14=y1;
v4dw24=y2;
Edw13=Edy3*y3dv3*v3dw13;
Edw14=Edy4*y4dv4*v4dw14;
Edw23=Edy3*y3dv3*v3dw23;
Edw24=Edy4*y4dv4*v4dw24;
nw13=w13-0.1*Edw13;
nw14=w14-0.1*Edw14;
nw23=w23-0.1*Edw23;
nw24=w24-0.1*Edw24;
v3dy1=w13;
v3dy2=w23;
v4dy1=w14;
v4dy2=w24;
Ey3dv3=Edy3*y3dv3;
Ey4dv4=Edy4*y4dv4;
Ey3dy1=Ey3dv3*v3dy1;
Ey4dy1=Ey4dv4*v4dy1;
Ey3dy2=Ey3dv3*v3dy2;
Ey4dy2=Ey4dv4*v4dy2;
Edy1=Ey3dy1+Ey4dy1;
```

```
Edy2=Ey3dy2+Ey4dy2;
y1dv1=y1^2*exp(-v1);
y2dv2=y2^2*exp(-v2);
v1dw11=x1;
v2dw12=x1;
v1dw21=x2;
v2dw22=x2;
Edw11=Edy1*y1dv1*v1dw11;
Edw12=Edy2*y2dv2*v2dw12;
Edw21=Edy1*y1dv1*v1dw21;
Edw22=Edy2*y2dv2*v2dw22;

nw11=w11-0.1*Edw11;
nw12=w12-0.1*Edw12;
nw21=w21-0.1*Edw21;
nw22=w22-0.1*Edw22;
```

(next page)

2-3)

2) $y = \tanh x = \dfrac{\sinh x}{\cosh x}$

$\dfrac{(\frac{d}{dx}\sinh x)\cosh x - \sinh x \frac{d}{dx}\cosh x}{\cosh^2 x}$

$\dfrac{\cosh^2 x + \sinh^2 x}{\cosh^2 x}$

$= 1 + (\tanh x)^2$

$\boxed{= (1+y^2)}$

#2 is straightforward, the derivative of tanh x results to $1+\tanh(x)^2$ or $\operatorname{sech}^2(x)$ so we can also express the answer as y' as $(1+y^2)$

3) $y = \dfrac{e^{x_2}}{e^{x_1}+e^{x_2}+e^{x_3}}$

Assume

$\dfrac{dy}{dx_1} = \dfrac{(\frac{d}{dx_1}e^{x_2})(e^{x_1}+e^{x_2}+e^{x_3}) - (\frac{d}{dx_1}[e^{x_1}+e^{x_2}+e^{x_3}])(e^{x_2})}{(e^{x_1}+e^{x_2}+e^{x_3})^2}$

$= \dfrac{-e^{x_2} \cdot e^{x_1}}{(e^{x_1}+e^{x_2}+e^{x_3})^2} = \boxed{\dfrac{-y \cdot e^{x_1}}{e^{x_1}+e^{x_2}+e^{x_3}}}$

$\dfrac{dy}{dx_2} = \dfrac{\frac{d}{dx_2}e^{x_2}(e^{x_1}+e^{x_2}+e^{x_3}) - \frac{d}{dx_2}(e^{x_1}+e^{x_2}+e^{x_3})(e^{x_2})}{(e^{x_1}+e^{x_2}+e^{x_3})^2}$

$= \dfrac{e^{x_2}(e^{x_1}+e^{x_2}+e^{x_3}) - e^{x_2}e^{x_2}}{(e^{x_1}+e^{x_2}+e^{x_3})^2}$

$\boxed{= \dfrac{e^{x_2}e^{x_1}+e^{x_3}(e^{x_2})}{(e^{x_1}+e^{x_2}+e^{x_3})^2}} = \dfrac{y(e^{x_1}+e^{x_3})}{e^{x_1}+e^{x_2}+e^{x_3}}$

$\dfrac{d}{dx_3} = \dfrac{-e^{x_2}e^{x_3}}{(e^{x_1}+e^{x_2}+e^{x_3})^2}$

$= \boxed{\dfrac{-y \cdot e^{x_3}}{e^{x_1}+e^{x_2}+e^{x_3}}}$

Explanation

1) We can think of this as having 3 different derivatives assuming there are 3 denominators collected from the total system, each will yield a result. however we know that the numerator can be treated as a constant or a variable when differentiating the equation

please replace all x with x_s to satisfy homework conditions.

typed solutions: dy/dx= y=tanh(x)=sinhx/coshx
dy/dx=(d/dx(sinhx)*coshx-d/dx(coshx)*sinhx)/(coshx)^2=1+tanhx^2
=(1+y^2)
The update equation is straightforward. Just differentiate it and relating it to the output, the results look like the final expression above..

So updating the weight equation the results would end up looking like:
DError/dw=-(d-y)*(1+y^2)*xni
w := w-(learning rate)*DError/dw

2) $1+y^2$

$\cancel{dE}$

$$\boxed{\frac{dE}{dw} = -(d-y)(1+y^2)\, x_{ni} \qquad W := W - \eta \frac{dE}{dw}}$$

3) i made up my own example just to show the expected result.

Given the numerator is not similar to the differentiated value.
We consider the numerator to be some constant
the results will end up looking like:
dy/dv1= e^v2/(e^v1+e^v2+e^v3)
-e^v2*e^v1/(e^v1+e^v2+e^v3)^2 = -y*ev1/(e^v1+e^v2+e^v3)

dy/dv3=-e^v2*e^v3/(e^v1+e^v2+e^v3)^2 =- y*ev3/(e^v1+e^v2+e^v3)

lastly if the numerator is similar to the differentiated value, it should be differentiated accordingly as a variable of interest. The result will look like this.
dy/dv2=-e^v2*e^v3/(e^v1+e^v2+e^v3)^2 = y*(e^v1+e^v3)/(e^v1+e^v2+e^v3

Generally To express myresult:

y=e^(vj)/(sum[e^vi])

we will need to get dy/dva

There will be 2 cases:
case 1) va=vj

dy/dva=dy/dvj=y=e^(vj)/(sum[e^vi])
=(de^(vj)/dvj*(sum[e^vi]) - e^(vj)*d(sum[e^vi])/dvj))/(sum[e^vi]^2)
=(e^vj*sum[e^vi]-(e^vj)^2)/sum[e^vi]^2)
**=y(1-y)
DError/dw=-(d-y)*(y*(1-y))*xni
w := w-(learning rate)*DError/dw

case 2) va!=vj
dy/dva=dy/d(v-somethingelse)=y=e^(vj)/(sum[e^vi])
=(de^(vj)/dvk*(sum[e^vi]) - e^(vj)*d(sum[e^vi])/dvk))/(sum[e^vi]^2)
=(-(e^vj)*e^vk)/sum[e^vi]^2)
**=-yi*yk
DError/dw=-(d-y)*(yi*yk)*xni
w := w-(learning rate)*DError/dw

just replace the equations here for the weight update equation

DError/dw=-(d-y)*(1+y^2)*xni
w := w-(learning rate)*DError/dw

case 2 va != vj case 1 va= vj

$$\frac{d}{dv_k} \frac{e^{v_j}}{\Sigma e^{v_k}}$$

$$\frac{\left(\frac{d}{dv_k} e^{v_j}\right) \Sigma e^{v_k} - e^{v_j} \cdot \left(\frac{d}{dv_k} \cdot \Sigma e^{v_k}\right)}{(\Sigma e^{v_k})^2}$$

$$\boxed{\frac{0 - e^{v_j} \cdot e^{v_k}}{(\Sigma e^{v_k})^2}}$$

$$= -(y_j \cdot y_k)$$

$$\frac{d}{dv_j} \frac{e^{v_j}}{\Sigma e^{v_k}}$$

$$\frac{\left(\frac{d}{dv_j} e^{v_j}\right) \Sigma e^{v_k} - \left(\frac{d}{dv_j} (\Sigma e^{v_k})\right) \cdot e^{v_j}}{(\Sigma e^{v_k})^2}$$

$$\frac{e^{v_j} \Sigma e^{v_k} - (e^{v_j})^2}{(\Sigma e^{v_k})^2}$$

$$\frac{e^{v_j} (\Sigma e^{v_k} - e^{v_j})}{e^{v_k^2}}$$

$$y \cdot \frac{\Sigma e^{v_k}}{e^{v_k}} - \frac{e^{v_j}}{e^{v_k}}$$

$$= y_k(1 - y_k)$$

correction: yj(1-yj)

DE/dw=DE/Dy*Dy/Dv*Dv/Dw

nfo

$$W := \cancel{\frac{dG}{dy}} \quad W - n\left(\frac{dy}{dv}\frac{dG}{dW}\right)$$

#2 $\frac{dy}{dv} = (1+y^2)$

$$\frac{dG}{dw} = \boxed{-(d-y) \cdot (1+y^2) \cdot X_{n_i}}$$

$$W := W - n \frac{dG}{dw}$$

#3

$$\frac{dG}{dw} = \overbrace{-(d-y)}^{\frac{dG}{dy}} \cdot \frac{dy}{dv} \cdot \frac{dv}{dw}$$

case 2

$$\cancel{\frac{dy}{dv_{k \neq j}}} \quad \frac{dy}{dv_{a \neq j}} = -(y_j \cdot y_k)$$

case 1

$$\cancel{\frac{dy}{dv_{k=j}}} \quad \frac{dy}{dv_{a=j}} = y_j(1-y_j)$$

case 2

$$\frac{dG}{dw} = -(d-y) \cdot -(y_j \cdot y_k) \cdot X_n.$$

case 1

$$\frac{dG}{dw} = -(d-y) \cdot y_j(1-y_j) \cdot X_{n_i}$$

$$W := W - n \frac{dG}{dw}$$