Example 1

w1=0.1

w2=0.2

w3=0.2

w4= (-0.1)

w5=0.2

w6=0.1

w7= (-0.3)

w8=0.2

w9= (-0.1)

x1 = 0, x2 = 1, out = 1

$\eta = 0.2$

1. **Forward Pass:**

   $o = \sigma(s) = 1 / (1 + exp(-s))$

   $o_1 = \sigma(w_1x1) = 1/ (1 + \exp(-(w_1x1))) = 1/ (1 + \exp(-(0.1*0))) = 0.5000$

   $o_2 = \sigma(w_3x1+w_4x2) = 1/ (1+\exp(-(0.2*0 +(-0.1*1)))) = 0.4750$

   $o_3 = \sigma(w_6x2) = 1/ (1 + \exp(-(0.1*1))) = 0.5250$

   $out = w_9o1 + w_8o2 + w_7o3 + w_2x1 + w_5x2 = (-0.1) *0.5 + (0.2) *0.4750 + 0.5250*(-0.3) + 0.2*0 + 0.2*1 = 0.0875$

2. **Backward Pass – Error Derivatives β:**

   $\beta_{out} = (output\ error) = (1-0.0875) = 0.9125$

   $\beta_1 = o_1*(1- o_1) *\beta_{out}*w_9 = 0.5*(1-0.5) *0.9125*(-0.1) = -0.0228$

   $\beta_2 = o_2*(1- o_2) *\beta_{out}*w_8 = 0.475*(1-0.475) *0.9125*(0.2) = 0.0455$

   $\beta_3 = o_3*(1- o_3) *\beta_{out}*w_7 = 0.525*(1-0.525) *0.9125*(-0.3) = -0.0683$

3. **Backward Pass – Weight Updates:**

   $w_1 = w_1 + \eta*\beta_1*x1 = 0.1 + 0.2*(-0.0228) *0 = 0.1000$

   $w_2 = w_2 + \eta*\beta_{out}*x1 = 0.2 + 0.2*(0.9125) *0 = 0.2000$

   $w_3 = w_3 + \eta*\beta_2*x1 = 0.2 + 0.2*(0.0455) *0 = 0.2000$

   $w_4 = w_4 + \eta*\beta_2*x2 = -0.1 + 0.2*(0.0455) *1 = -0.0909$

   $w_5 = w_5 + \eta*\beta_{out}*x2 = 0.2 + 0.2*(0.9125) *1 = 0.3825$

   $w_6 = w_6 + \eta*\beta_3*x2 = 0.1 + 0.2*(-0.0683) *1 = 0.0863$

   $w_7 = w_7 + \eta*\beta_{out}* o_3 = -0.3 + 0.2*(0.9125) * 0.525 = -0.2042$

   $w_8 = w_8 + \eta*\beta_{out}* o_2 = 0.2 + 0.2*(0.9125) * 0.475 = 0.2867$

   $w_9 = w_9 + \eta*\beta_{out}* o_1 = -0.1 + 0.2*(0.9125) * 0.5 = -0.0088$

Example 2

Online learning: weights updated from Example 1

w1=0.1

w2=0.2

w3=0.2

w4= (-0.0909)

w5=0.3825

w6=0.0863

w7= (-0.2042)

w8=0.2867

w9= (-0.0088)

x1 = 1, x2 = 0, out = 1

η = o.2

1. Forward Pass:

$o = \sigma(s) = 1 / (1 + \exp(-s))$

$o_1 = \sigma(w_1x1) = 1/ (1 + \exp(-(w_1x1))) = 1/ (1 + \exp(-(0.1*1))) = 0.5250$

$o_2 = \sigma(w_3x1+w_4x2) = 1/ (1+\exp(-(0.2*1 + (-0.0909) *0))) = 0.5498$

$o_3 = \sigma(w_6x2) = 1/ (1 + \exp(-(0.0863*0))) = 0.5000$

$out = w_9o1 + w_8o2 + w_7o3 + w_2x1 + w_5x2 =$

(-0.0088) *0.5250 + 0.2867*0.5498 + (-0.2042) *0.5 + 0.2*1 + 0.3825*0 = 0.2509

2. Backward Pass – Error Derivatives β:

$\beta_{out} =$ *(output error)* = (1-0.2509) = 0.7491

$\beta_1 = o_1*(1- o_1) *\beta_{out}*w_9 = 0.5250*(1-0.5250) *0.7491*(-0.0088) = -0.0016$

$\beta_2 = o_2*(1- o_2) *\beta_{out}*w_8 = 0.5498*(1-0.5498) *0.7491*0.2867 = 0.0532$

$\beta_3 = o_3*(1- o_3) *\beta_{out}*w_7 = 0.5*(1-0.5) *0.7491*(-0.2042) = -0.0382$

3. Backward Pass – Weight Updates:

$w_1 = w_1 + \eta*\beta_1*x1 = 0.1 + 0.2*(-0.0016) *1 = 0.0997$

$w_2 = w_2 + \eta*\beta_{out}*x1 = 0.2 + 0.2*(0.7491) *1 = 0.3498$

$w_3 = w_3 + \eta*\beta_2*x1 = 0.2 + 0.2*(0.0532) *1 = 0.2106$

$w_4 = w_4 + \eta*\beta_2*x2 = (-0.0909) + 0.2*(0.0532) *0 = -0.0909$

$w_5 = w_5 + \eta*\beta_{out}*x2 = 0.3825 + 0.2*(0.7491) *0 = 0.3825$

$w_6 = w_6 + \eta*\beta_3*x2 = 0.0863 + 0.2*(-0.0382) *0 = 0.0863$

$w_7 = w_7 + \eta*\beta_{out}* o_3 = (-0.2042) + 0.2*(0.7491) * 0.5000 = -0.1293$

$w_8 = w_8 + \eta*\beta_{out}* o_2 = 0.2867 + 0.2*(0.7491) * 0.5498 = 0.3691$

$w_9 = w_9 + \eta*\beta_{out}* o_1 = (-0.0088) + 0.2*(0.7491) * 0.5250 = 0.0699$