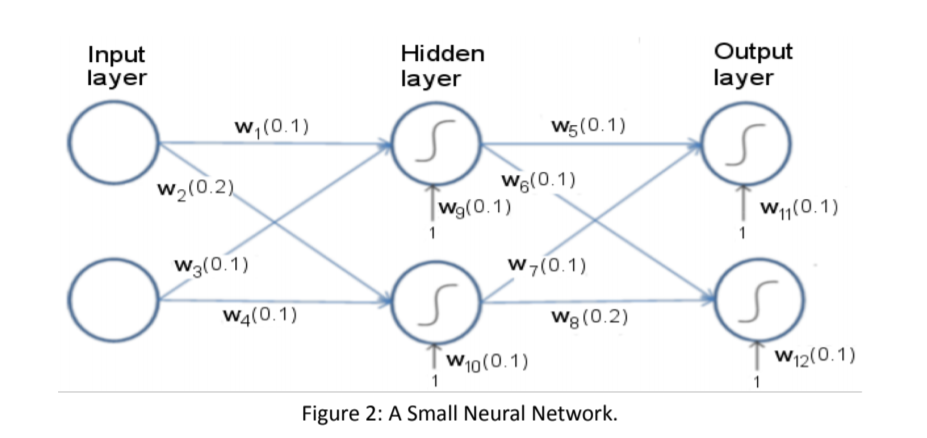
Intelligent systems assignment 2

Sushant Karki, s5170503

convener : liat

**Part A**



First let’s name the nodes as: i1, i2 for nodes at input layer; h1, h2 for hidden layer and o1, o2 for output layer. We would use the same labels for rest of this assignment. We would use **X1 = (0.1,0.1) labelled 0 and X2 = (0.1,0.2) labelled 1**. The corresponding output for them is **Y1 = (1,0), Y2 = (0,1).**

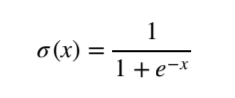
## To find the input for hidden layer:

Net Input for hidden layer h1 = (w1\*0.1 + w3\*0.1 +w9) = 0.12

Net Input for hidden layer h2 = (w2\*0.1 + w4\*0.1 + w10) = 0.13

Now, to find the output for the hidden layer:

Using the basic sigmoid function:



Output for h1 = 0.5299640518

Output for h2 = 0.5324543064

## To find the input for the output layer:

Net input for output layer o1 = (w5\*outh1 + w7\* outh2+ w11) = 0.2062418358

Net input for output layer o2 = (w6\*outh1 + w8\* outh2+ w12) = 0.2594872665

Now, to find the output for the output layer using the sigmoid function:

Output for o1 = 0.5513784697

Output for o2 = 0.5645102464

## Calculating error:

The output was supposed to be (1,0) but we got (0.5519368127, 0.564513173)

The error is:

.5\*(1-0.5513784697)2 = 0.10063063873 and

.5(-0.5645102464)2 = 0.1593359091

Total error is 0.2599665479

# Backtracking:

## For w5:

To find the derivative of total error in terms of w5:

= -(desired\_output – output of o1) = **-0.4486215303**

Because total error (E-total) = 0.5\*(desired\_output – output of o1)^2 + 0.5\*(desired\_output – output of o2)^2

= o1 \* (1 – output of o1) = **0.2473602529**

Because Output = 1/(1+e-x) where x is the net input.

= outh1 = **0.5299640518**

Because Net input = (w5\* Output of h1 + w7\*. Output of h2 + w11)

Putting it all together = **-0.0588107124**

*Will only be updated after second batch*

## For W7:

To find the derivative of total error in terms of w7:

The only difference from W5 is = Output of h2 = **0.5324543064**

Because Net input = (w5\* Output of h1 + w7\*. Output of h2 + w11)

Putting it all together = **-0.0590870589**

*Will only be updated after second batch*

## For W6:

To find the derivative of total error in terms of w6:

= -(desired\_output – output of o2) = **0.5645102464**

Because total error (E-total) = 0.5\*(desired\_output – output of o1)^2 + 0.5\*(desired\_output – output of o2)^2

= output of o2 \* (1 – output of o2) = **0.2458384281**

Because Output = 1/(1+e-x) where x is the net input.

= Output of h1 = **0.5299640518**

Because Net input = (w6\* Output of h1 + w8\*. Output of h2 + w12)

Putting it all together = **0.0735475163**

*Will only be updated after second batch*

## For W8:

To find the derivative of total error in terms of w8:

The only difference between W6 and W8 is = Output of h2 = **0.5324543064**

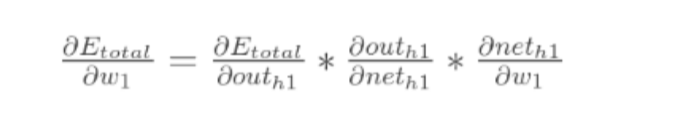
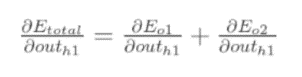
Because Net input = (w6\* Output of h1 + w8\*. Output of h2 + w12)

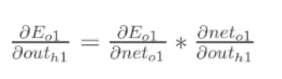
Putting it all together = **0.0738931097**

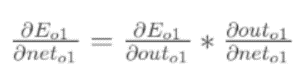
*Will only be updated after second batch*

# BACKTRACKING FOR HIDDEN LAYER:

For w1 and w3:







Putting it all together :

∂neto1/∂outh1 = w5 = 0.1 & ∂neto2/∂outh1 = w6 = 0.1

∂Eo1/∂outo1 = -0.4486215303 ∂Eo2/∂outo2 =0.5645102464

∂outo1/∂neto1 = 0.2473602529 ∂outo2/∂neto2 = 0.2458384281

∂outh1/∂neth1 = outh1(1-outh1) = 0.2491021556

∂neth1/∂w1 = 0.1

For w1:

0.0027807176 \*0.2491021556\*0.1= **0.0000692683**

For w3: the formula is the same except for ∂neth1/∂w1 which is now ∂neth1/∂w3:

0.0027807176 \*0.2491021556\*0.1= **0.0000692683**

For w2 and w4:

For w2:

∂neto1/∂outh2 = w7 = 0.1 & ∂neto2/∂outh1 = w8 = 0.2

∂Eo1/∂outo1 = -0.4486215303 ∂Eo2/∂outo2 = 0.5645102464

∂outo1/∂neto1 = 0.2473602529 ∂outo2/∂neto2 = 0.2458384281

∂outh1/∂neth1 = outh2(1-outh2) = 0.2458384281

∂neth1/∂w1 = 0.1

0.0166585488 \*0.2458384281\*0.1= **0.0004095311**

For w3: same but for ∂neth2/∂w2 its ∂neth2/∂w3

0.0166585488 \*0.2458384281\*0.1 = **0.0004095311**

Batch 2

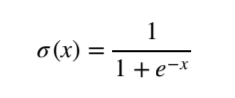
To find the input for hidden layer:

Net Input for hidden layer h1 = (w1\*0.1 + w3\*0.2 +w9) = 0.13

Net Input for hidden layer h2 = (w2\*0.1 + w4\*0.2 + w10) = 0.14

Now, to find the output for the hidden layer:

Using the basic sigmoid function:



Output for h1 = .5324543064

Output for h2 = .5349429452

To find the input for the output layer:

Net input for output layer o1 = (w5\*.5324543064+ w7\*.5349429452 + w11) = 0.2067397252

Net input for output layer o2 = (w6\*.5324543064 + w8\*.5349429452+ w12) = 0.2602340197

Now, to find the output for the output layer using the sigmoid function:

Output for o1 = 0.5515016246

Output for o2 = 0.5646938182

Calculating error:

The output was supposed to be (0,1) but we got (0.5515016246, 0.5646938182)

The error is:

.5\*(1-0.5515016246)2 = 0.1005753964 and

.5(-0.5646938182)2 = 0.1594395542

Total error is 0.2600149506

Backtracking:

For w5:

To find the derivative of total error in terms of w5:

= -(desired\_output – output of o1) = **0.5515016246**

Because total error (E-total) = 0.5\*(desired\_output – output of o1)^2 + 0.5\*(desired\_output – output of o2)^2

= o1 \* (1 – output of o1) = **0.2473475827**

Because Output = 1/(1+e-x) where x is the net input.

= outh1 = **.5324543064**

Because Net input = (w5\* Output of h1 + w7\*. Output of h2 + w11)

Putting it all together = **0.072633473**

*Will only be updated after second batch*

For W7:

To find the derivative of total error in terms of w7:

The only difference from W5 is = Output of h2 = .**5349429452**

Because Net input = (w5\* Output of h1 + w7\*. Output of h2 + w11)

Putting it all together = **0.0572413894**

*Will only be updated after second batch*

For W6:

To find the derivative of total error in terms of w6:

= -(desired\_output – output of o2) = **-0.4353061818**

Because total error (E-total) = 0.5\*(desired\_output – output of o1)^2 + 0.5\*(desired\_output – output of o2)^2

= output of o2 \* (1 – output of o2) =  **0.2458147099**

Because Output = 1/(1+e-x) where x is the net input.

= Output of h1 = **.5324543064**

Because Net input = (w6\* Output of h1 + w8\*. Output of h2 + w12)

Putting it all together = **-0.0569750935**

*Will only be updated after second batch*

For W8:

To find the derivative of total error in terms of w8:

The only difference between W6 and W8 is = Output of h2 = **.5349429452**

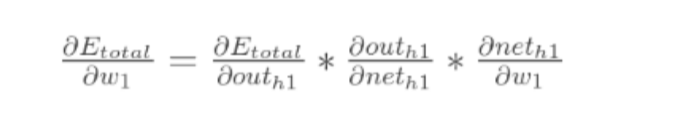
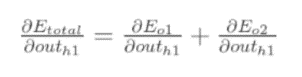
Because Net input = (w6\* Output of h1 + w8\*. Output of h2 + w12)

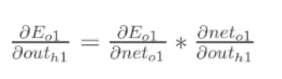
Putting it all together = **-0.0572413895**

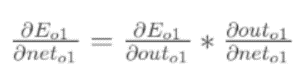
*Will only be updated after second batch*

BACKTRACKING FOR HIDDEN LAYER:

For w1 and w3:







Putting it all together :

∂neto1/∂outh1 = w5 = 0.092736653 & ∂neto2/∂outh1 = w6 = 0.1056975094

∂Eo1/∂outo1 = 0.5515016246 ∂Eo2/∂outo2 = -0.4353061818

∂outo1/∂neto1 = 0.2473025675 ∂outo2/∂neto2 = 0.2458147099

∂outh1/∂neth1 = outh1(1-outh1) = 0.2473475827

∂neth1/∂w1 = 0.1

For w1:

(0.0013380187) \* 0.2473475827\*0.1 = **0.0000330956**

For w3: the formula is the same except for ∂neth1/∂w1 which is now ∂neth1/∂w3:

(0.0013380187) \* 0.2473475827\*0.2 = **0.0000661911**

For w2 and w4:

For w2:

∂neto1/∂outh2 = w7 = 0.1 & ∂neto2/∂outh2 = w8 = 0.2

∂Eo1/∂outo1 = 0.5515016246 ∂Eo2/∂outo2 = -0.4353061818

∂outo1/∂neto1 = 0.2473025675 ∂outo2/∂neto2 = 0.2458147099

∂outh2/∂neth2 = outh2(1-outh2) = 0.2487789906

∂neth1/∂w2 = 0.1

(0.0927027045\*0.5515016246\*0.2473025675 + 0.105724139\*-0.4353061818\*0.2458147099) \* 0.2487789906\* 0.1 = -**0.0001931061**

For w4: same but for ∂neth2/∂w4 its ∂neth2/∂w4

(0.0927027045\*0.5515016246\*0.2473025675 + 0.105724139\*-0.4353061818\*0.2458147099) \* 0.2487789906\* 0.2 = **-0.0003862123**

Updating :

# Batch 1 + Batch 2

W1 - learning\_rate\*(0.0000692683 + 0.0000330956)/2= 0.099974409

W2 - learning\_rate \*(0.0004095311 +-0.0001931061)/2 =0.1998493407

W3 - learning\_rate \*(0.0000692683 + 0.0000661911)/2 =0.09997

W4 - learning\_rate \*(0.0004095311 + -0.0003862123)/2= 0.09999

W5 - learning\_rate \*(-0.0588107124 + 0.5515016246)/2=-0.0231727281

W6 - learning\_rate \*(0.0735475163 + -0.0569750935)/2=0.0958568943

W7 - learning\_rate \*(-0.0590870589 + 0.0572413894)/2=0.1004614174

W8 - learning\_rate \*(0.0738931097 + -0.0572413895)/2=0.19583707

Format :

Weight1 = [[W1],[W2]

[W3],[W4]]

Weight2 = [[W5],[W6]

[W7],[W8]]

Epoch 1 :

Weight 1 = [[0.09998575216453286, 0.1999778335261234],

[0.09997843115666014, 0.09999713796276742]]

Weight 2 = [[0.09861772394777771, 0.09834275771909186],

[0.09861141041906764, 0.19833482798113933]]

Epoch 2 :

Weight1 = [[0.09997181711646458, 0.19995603214286786],

[0.09995725241977808, 0.09999475107462791]]

Weight2 = [[0.09724483500546822, 0.09669650936039943],

[0.09723224489113419, 0.19668070552123562]]

Epoch 3 :

Weight1 = [[0.09995818993760441, 0.19993459062003616],

[0.0999364567540813, 0.09999283174548222]]

Weight2 = [[0.09588127503383505, 0.0950611918850871],

[0.09586244498383525, 0.1950375693216742]]