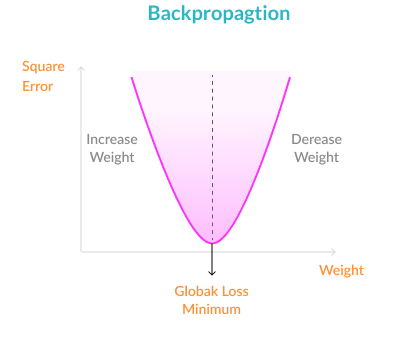
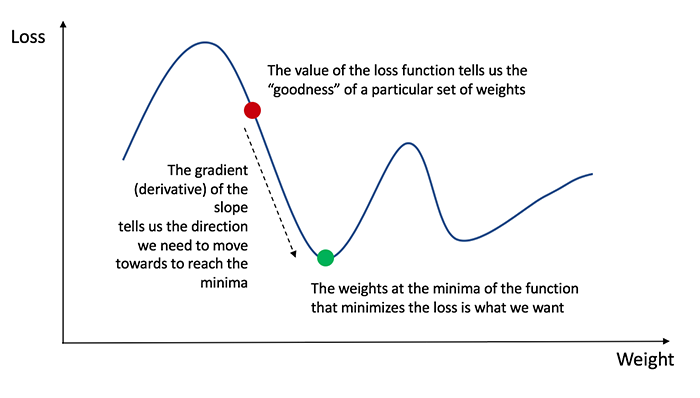
**Ví dụ về giải thuật Backpropagation ( BG)**

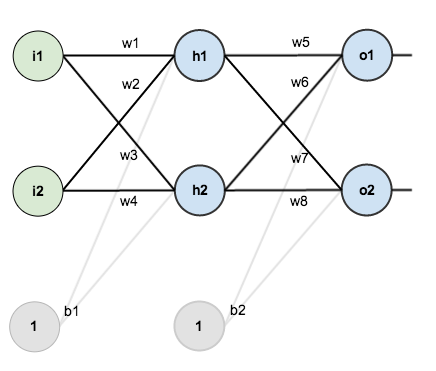
Backpropagation là một giải thuật phổ biến để huấn luyện 1 neural network(NN).

Mục đích của việc training là tìm ra tập weights va biases để giảm error sao cho kết quả đầu ra có giá trị gần giống với kết quả mong muốn tương ứng với giá trị đầu vào.





Vi dụ : 1 mạng (NN) gồm : 02 inputs , 02 hidden perceptrons , 02 biases, 02 outputs như hình vẽ :



Các giá trị khởi động ( weights , biases va inputs , outputs ) để mạng có thể vận hành :

i1 = 0.05 ; w1 = 0.15 ; w2 = 0.2

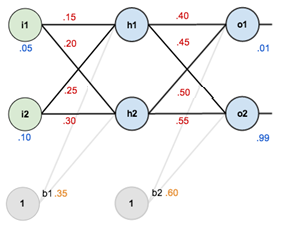
i2 = 0.1 ; w3 = 0.2 ; w4 = 0.3

b1 = 0.35

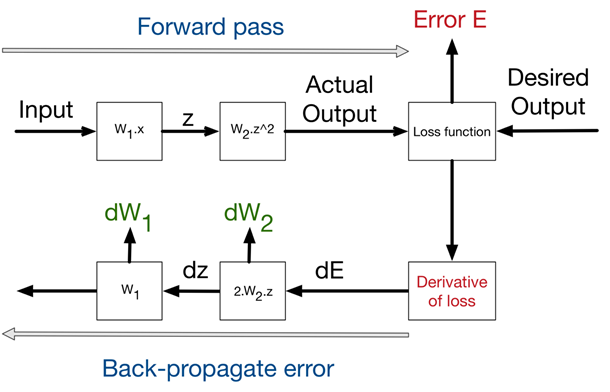
b2 = 0.6

o1 = 0.1

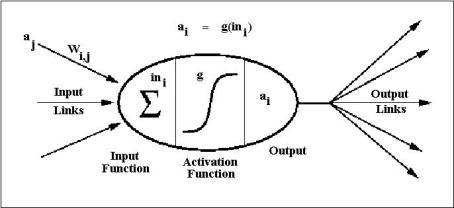
o2 = 0.99

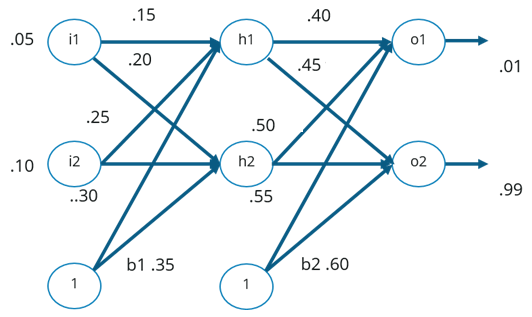


Trong ví dụ này : input : i1 = 0.05 , i2 = 0.10 mong muốn output : o1=0.01 va o2 = 0.99



Step 1 : **Forward pass**

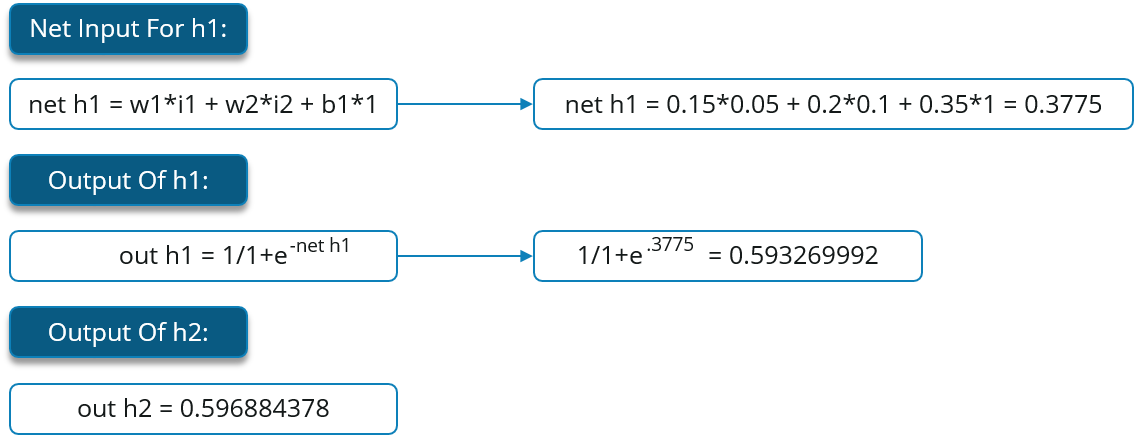




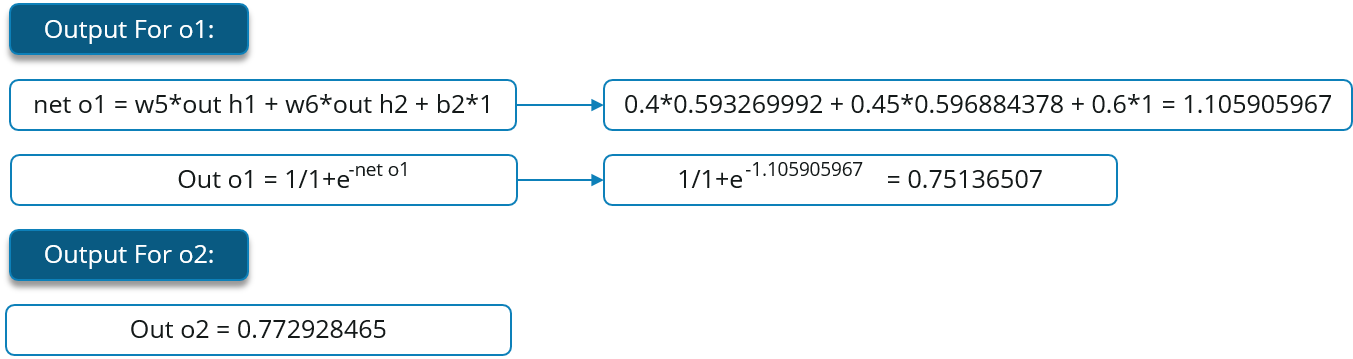
Tổng giá trị đầu vào\ra tai 1 perceptron được tinh theo công thức :

net = http://staff.itee.uq.edu.au/janetw/cmc/chapters/BackProp/sum.gifiwiIi. + bi\*1

 ( hàm activation sigmoid )

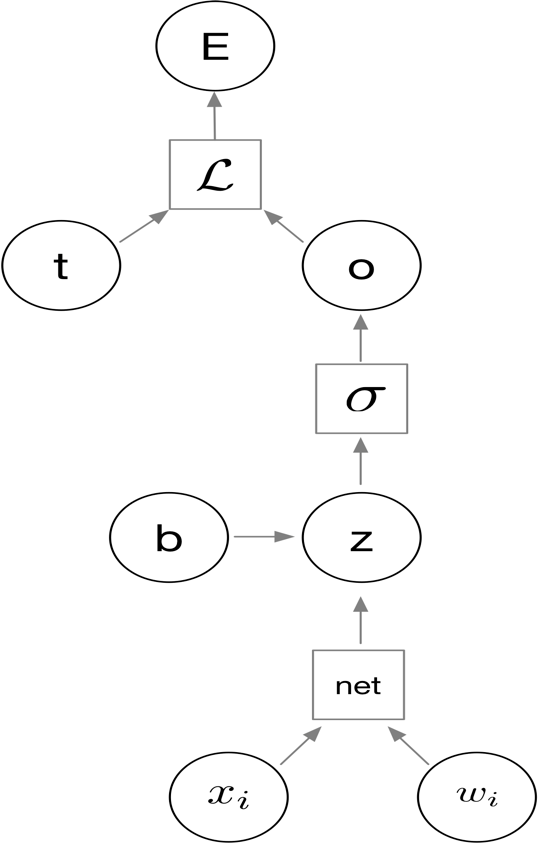


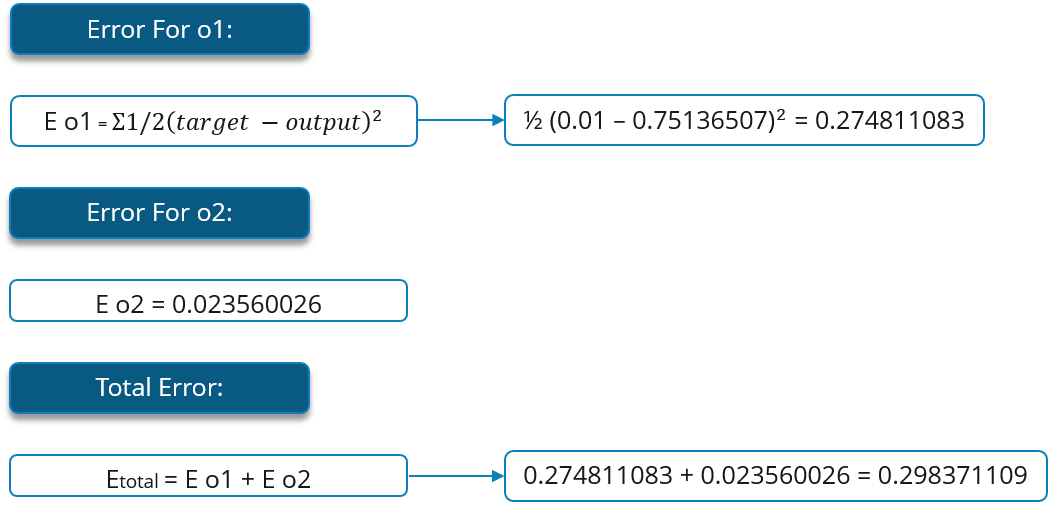
Tương tự cho lớp hidden :



Ta tính được **Total error** theo công thức **squared error** function :







Step 2 : Cập nhật biến trọng số ( **Backwards pass** )

1. Tại lớp đầu ra ( Output Layer ) : xem xét tại w5 : thay đổi w5 bao nhiêu để ảnh hưởng đến tổng error 

  gọi la đạo hàm 1 phần của of  với biến thiên theo  . (Hay còn gọi là độ dốc theo w5)

Ứng dụng chain rule :



Trực quan bằng hình ảnh :

A close up of text on a white background

Description automatically generated

Tính giá trị từng thành phần của phương trình :



1. Đạo hàm Total error đối với biến o1:





1. Đạo hàm ouput o1 đối với biến net 1:





1. Đạo hàm input net 1 đối với biến w5 :



Kết quả :



Vậy giá trị cập nhất của trọng số weight w5 là :

Với n = learning rate = 0.5



Tương tự cho w6,7,8 : ta tính được :

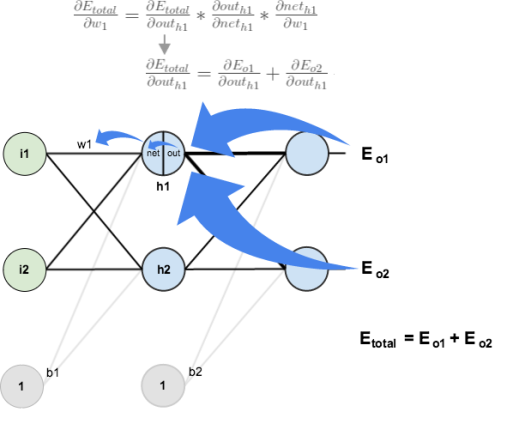


1. Lớp ẩn ( Hidden Layer ) : ta tiếp tụcc tính new w1,2,3,4

* New w1



Trực quan bằng hình ảnh :



Quá trình tính toán tương tự nhưng có chút xíu khác do output của lớp hidden phân phối đến output của các perceptrons phía sau nên ta phải tách ra như sau :



Với :





Tương tự , ta có : 



Bây giớ , ta tính phần còn lại :



Kết quả cuối cùng :



Bây giờ, new w1 : voi n= learnig rate = 0.5



Tương tự , ta cũng tính được :



Tổng kết , sau khi ta cập nhật tất cả các trọng số weights vào fordward pass , error từ 0,298371109 giảm còn 0.298371109 . Lặp lại 10.000 lần . Error giảm mạnh xuống còn 0.0000351085 và output la 0.015912196 (so vớii 0.01 target) và 0.984065734 (so với 0.99 target).