**Neural network**

A diagram of a diagram

Description automatically generated

Figure 1 ตัวอย่าง Neural Network

Overview of all steps of Neural network.

Our aim is to minimize an error of the model by following these steps:

Node calculation to get linear equation & Take activation function to convert it into non-linear equation.

Training data by taking loss function.

Using Optimizer, The Backwards Pass, to differential loss function.

Take gradient to Optimizer to minimize loss function.

Backpropagation เริ่มจากการเอาค่า error จาก output ของ Neural Network มาเทียบกับ Expected target เมื่อได้ค่า error ก็เอาค่า error กลับไปยังผู้ที่เกี่ยวข้องก็คือค่า weight ต่างๆ weight ที่มีค่ามากก็ปรับมาก weight ที่มีค่าน้อยก็ปรับน้อย เพื่อหาผู้รับผิดชอบที่ทำให้ค่า output มันไม่เท่ากับ target.

First step:

ในตัวอย่างนี้มี input 2 ค่า คือ i1, i2 มีค่า weight คือ w1-w8 และตั้ง output target ไว้ที่ output1(o1) output2(o2) โดย w = weight, b = bias

Using the Activation Function which is Sigmoid Function. ซึ่งเป็นฟังก์ชันที่เป็น Curve รูปตัว S เห็นแล้วเข้าใจได้ง่าย และเนื่องจาก Output ของ Sigmoid Function มีค่าระหว่าง 0 – 1 จึงเหมาะที่จะถูกใช้ในงานที่ต้องการ Output เป็นความน่าจะเป็น (Probability) หรือใช้เป็น Output ว่า 1=Yes, 0=No

สมการของ Sigmoid Function :

สูตรการคำนวณแต่ละ Node :

When: a = activation function, w = weight vector, = input, b = bias

First equation is the linear equation which is then we put it in the activation function which is a() to convert it into non-linear equation to make it can be explain every single data in the world.

คำนวณค่า output จริงๆที่ได้จาก Neural Network เริ่มจาก h1:

คำนวน h2

คำนวณแบบเดิมเพื่อหา o1 กับ o2 แล้ว จะกลายเป็น input ของชั้นถัดไป

netout\_h1= 0.59326  
netout\_h2= 0.59688

Second step:

หลังจากที่เราได้ค่า output จาก Node ที่คำนวณได้แล้ว ให้เอาไปเข้า loss function เพื่อหาว่า model ของเรามี error มากน้อยแค่ไหน โดยในกรณีนี้เป็นการทำนายข้อมูลหรือ regression model เราจึงเลือกใช้ loss function ที่เหมาะสมกับมัน มีชื่อว่า Mean squared error (MSE) โดย

When

เมื่อได้ output ของแต่ละกิ่งแล้วคือ o1 และ o2 จากนั้น ค่า Error หาได้จาก

จากในรูป Expected target ของ o1 = 0.01, ค่า output ที่คำนวณได้

Expected target ของ o2 = 0.99, ค่า output ที่คำนวณได้

ถือว่า error น้อยมาก ถ้าเป็น 0 คือ ไม่มี error เลย

Third step:

Using Optimizer, The Backwards Pass, to differential loss function. เป็นการหาผู้รับผิดชอบที่ทำให้เกิด error ถ้าทำ error มาก (weight มาก) ก็รับกรรมมาก ถ้าทำ error น้อย(weight น้อย) ก็รับกรรมน้อย

ทั้งหมดที่เราต้องปรับคือ 8 weight และ 2 bias ที่จะมารับผิดชอบผลของ error แต่ขอยกตัวอย่างการคำนวณแค่ w8 ดังนั้น เราจึงหาค่า gradient ที่จะต้องปรับดังนี้

Consider . We want to know how much a change in affects the total error, aka is read as “the partial derivative of with respect to ”. You can also say “the gradient with respect to ”.

By applying the chain rule, we know that:

We need to figure out each piece in this equation. First, how much does the total error change with respect to the output?

Calculate first term:

Calculate second term:

Calculate last term:

Putting it all together:

in other words, Optimizer

Fourth step:

Take SGD (stochastic gradient descent) to Optimizer to minimize loss function.

When

The reason why is that we must subtract the weights with the optimizer to make them counter-gradient to minimize error.

Given Learning rate =

ดังนั้น ใหม่ของเราจะมีค่า ซึ่งมากกว่าเดิม มานิดนึงเพราะ ค่า netout\_O2 มันน้อย ไม่ถึง 0.9 ก็เลยต้อง upweight ขึ้นมาทีละหน่อย

จากนั้นก็ทำการหารค่า W7 ,W6, W5 เมื่อได้ค่าครบ ก็ ลงไปทำชั้นถัดไปคือหา W4 W3 W2 และ W1 ในชั้นที่ลึกๆ กว่า