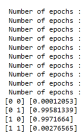
Homework 2

Tipu Sultan

October 2022

1 Problem 2:

For Xor output



For OR output

```
 Number of epochs :  0
 Number of epochs :  10001
 Number of epochs :  20002
 Number of epochs :  30003
 Number of epochs :  40004
 Number of epochs :  50005
 Number of epochs :  60006
 Number of epochs :  70007
 Number of epochs :  80008
 Number of epochs :  90009
[0 0] [0.00869339]
[0 1] [0.99704534]
[1 0] [0.99694604]
[1 1] [0.99876308]
```

For AND output

```
 Number of epochs :  0
 Number of epochs :  10001
 Number of epochs :  20002
 Number of epochs :  30003
 Number of epochs :  40004
 Number of epochs :  50005
 Number of epochs :  60006
 Number of epochs :  70007
 Number of epochs :  80008
 Number of epochs :  90009
[0 0] [2.52581305e-06]
[0 1] [6.48702943e-05]
[1 0] [3.07493354e-06]
[1 1] [0.9958047]
```

2 Problem 3:

We know,

$$Errore = \frac{1}{2}\sum_{d}^{D}(t_d - y_d)^2$$

Apply differantion

$$\frac{de}{dx} = \frac{1}{2} * 2(t_d - y_d) * \frac{d}{dx}(t_d - y_d)$$

$$= \sum(t_d - y_d) - 1 - - - - - - - - - - - - - - - eq1$$

$$\Delta w = \frac{de}{dw}$$

$$\frac{de}{dw} = \frac{de}{dx} * \frac{dx}{dw} - - - - - - - - - eq2$$

Sigmoid Funcation

$$f(x) = \frac{1}{1 + e^{-x}}$$

we can assume:

$$1 + e^{-x} = z$$

The equation will look like:

$$\frac{df(x)}{dx} = \frac{d}{dx} z^{-1}$$

$$= -z^{-1-1}$$

$$= -1(\frac{1}{z})^2 * \frac{d}{dx} z$$

$$= -1(\frac{1}{1 + e^{-x}})^2 * \frac{d}{dx} z$$

Now put the value of z

$$= -1(\frac{1}{1 + e^{-x}})^2 * \frac{d}{dx}(1 + e^{-x})$$

$$= -1(\frac{1}{1 + e^{-x}})^2 * [0 + e^{-1} * -1]$$

$$= \frac{1}{1 + e^{-x}} * \frac{1}{1 + e^{-x}} * e^{-x}$$

$$= \frac{1}{1 + e^{-x}} * 1 - \frac{1}{1 + e^{-x}}$$

$$= f(x) * (1 - f(x)) - - - - - eq3$$

put eq 1 and 3 in eq2

$$\Delta w = \sum(t_d - y_d) * y_d(1 - y_d)$$

How many iterations did it take to learn the weights for XOR, AND, OR operators?

= In my program, It takes 100000 epochs to reach the considerable output.

Can it be improved by increasing the learning rate, ETA?

=Yes, It is possible to improve. But The learning rate should be between 0 to 1. It is a hyperparameter.Choosing the appropriate learning rate is very important. If the value is too small. It may result in a long training process that could get stuck. On the other hand, If our value is too large may result in learning an unstable training process.