# Machine Learning

## Assignment 7.3

Submitted By: Ranji Raj

December 10, 2020

## Initial settings

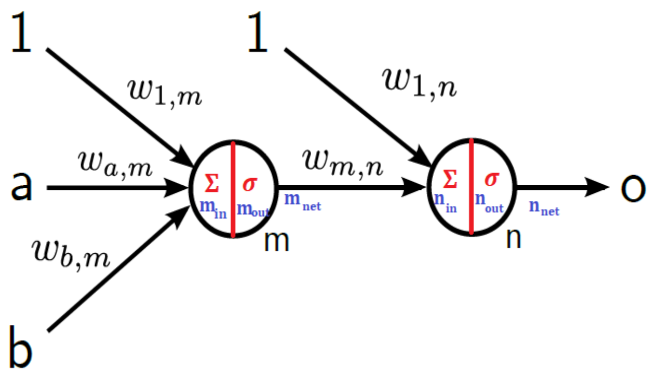


Figure 1: Sigmoid Neural Net

Given:

$$w_{a,m} = w_{b,m} = w_{1,m} = w_{m,n} = w_{1,n} = 0.1, \eta = 0.3$$

Linear computation:

$$w_0 + w_1 x_1 + w_2 x_2$$

Non-linear (sigmoid) computation:

$$\frac{1}{1 + e^{-net}}$$

## Example 1

Data: $a = 1, b = 0, t = 1$
Input at unit m,

$$m_{in} = w_{1,m} + a.w_{a,m} + b.w_{b,m}$$

$$\therefore m_{in} = 0.1 + (1)(0.1) + (0)(0.1) = 0.2$$

Output at unit m,

$$m_{out} = \frac{1}{1 + e^{-m_{in}}} = \frac{1}{1 + e^{-0.2}} = 0.5498$$

Input at unit n,

$$n_{in} = m_{out}.w_{m,n} + w_{1,n}$$

$$\therefore n_{in} = (0.5498)(0.1) + 0.1 = 0.15498$$

Output at unit n,

$$n_{out} = \frac{1}{1 + e^{-n_{in}}} = \frac{1}{1 + e^{-0.15498}} = 0.5387$$

Error,

$$(t - o) = 1 - 0.5387 = 0.4613$$

Calculating gradient $\frac{\partial E}{\partial w_{m,n}}$,

$$\frac{\partial E}{\partial w_{m,n}} = (\frac{\partial E}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{m,n}})$$

$$\therefore \frac{\partial E}{\partial w_{m,n}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{m,n}})$$

$$\therefore \frac{\partial E}{\partial w_{m,n}} = -(t - o).(n_{out}.(1 - n_{out})).m_{out}$$

$$\therefore \frac{\partial E}{\partial w_{m,n}} = -(0.4613)(0.5387)(1 - 0.5387)(0.5498) = \textbf{-0.063}$$

Updating weight $w_{m,n}$,

$$w_{m,n} = w_{m,n} - \eta \frac{\partial E}{\partial w_{m,n}}$$

$$\therefore w_{m,n} = 0.1 - 0.3(-0.063) = \mathbf{0.1189}$$

Calculating gradient $\frac{\partial E}{\partial w_{1,n}}$,

$$\frac{\partial E}{\partial w_{1,n}} = (\frac{\partial E}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{1,n}})$$

$$\therefore \frac{\partial E}{\partial w_{1,n}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{1,n}})$$

$$\therefore \frac{\partial E}{\partial w_{1,n}} = -(t-o).(n_{out}.(1-n_{out})).1$$

$$\therefore \frac{\partial E}{\partial w_{1,n}} = -(0.4613)(0.5387)(1-0.5387).1 = \textbf{-0.1146}$$

Updating weight $w_{1,n}$,

$$w_{1,n} = w_{1,n} - \eta \frac{\partial E}{\partial w_{1,n}}$$

$$\therefore w_{1,n} = 0.1 - 0.3(-0.1146) = \mathbf{0.134}$$

Calculating gradient $\frac{\partial E}{\partial w_{b,m}}$,

$$\frac{\partial E}{\partial w_{b,m}} = (\frac{\partial E}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{b,m}})$$

$$\therefore \frac{\partial E}{\partial w_{b,m}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial m_{out}})(\frac{\partial m_{out}}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{b,m}})$$

$$\therefore \frac{\partial E}{\partial w_{b,m}} = -(t-o).(n_{out}.(1-n_{out})).w_{m,n}.(m_{out}.(1-m_{out})).b$$

$$\therefore \frac{\partial E}{\partial w_{b,m}} = -(0.4613)(0.5387)(1-0.5387)(0.1)(0.5498)(1-0.5498)(0) = \mathbf{0.0}$$

Updating weight $w_{b,m}$,

$$w_{b,m} = w_{b,m} - \eta \frac{\partial E}{\partial w_{b,m}}$$

$$\therefore w_{b,m} = 0.1 - 0.3(0.0) = \mathbf{0.1}$$

Calculating gradient $\frac{\partial E}{\partial w_{a,m}}$,

$$\frac{\partial E}{\partial w_{a,m}} = (\frac{\partial E}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{a,m}})$$

$$\therefore \frac{\partial E}{\partial w_{a,m}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial m_{out}})(\frac{\partial m_{out}}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{a,m}})$$

$$\therefore \frac{\partial E}{\partial w_{a,m}} = -(t-o).(n_{out}.(1-n_{out})).w_{m,n}.(m_{out}.(1-m_{out})).a$$

$$\therefore \frac{\partial E}{\partial w_{a,m}} = -(0.4613)(0.5387)(1-0.5387)(0.1)(0.5498)(1-0.5498)(1) = \mathbf{-2.83 \times 10^{-3}}$$

Updating weight $w_{a,m}$,

$$w_{a,m} = w_{a,m} - \eta \frac{\partial E}{\partial w_{a,m}}$$

$$\therefore w_{a,m} = 0.1 - 0.3(-0.00283) = \mathbf{0.100849}$$

Calculating gradient $\frac{\partial E}{\partial w_{1,m}}$,

$$\frac{\partial E}{\partial w_{1,m}} = \left(\frac{\partial E}{\partial m_{net}}\right)\left(\frac{\partial m_{net}}{\partial w_{1,m}}\right)$$

$$\therefore \frac{\partial E}{\partial w_{1,m}} = \left(\frac{\partial E}{\partial n_{out}}\right)\left(\frac{\partial n_{out}}{\partial n_{net}}\right)\left(\frac{\partial n_{net}}{\partial m_{out}}\right)\left(\frac{\partial m_{out}}{\partial m_{net}}\right)\left(\frac{\partial m_{net}}{\partial w_{1,m}}\right)$$

$$\therefore \frac{\partial E}{\partial w_{1,m}} = -(t-o).(n_{out}.(1-n_{out})).w_{m,n}.(m_{out}.(1-m_{out})).1$$

$$\therefore \frac{\partial E}{\partial w_{1,m}} = -(0.4613)(0.5387)(1-0.5387)(0.1)(0.5498)(1-0.5498)(1) = \mathbf{-2.83 \times 10^{-3}}$$

Updating weight $w_{1,m}$,

$$w_{1,m} = w_{1,m} - \eta \frac{\partial E}{\partial w_{1,m}}$$

$$\therefore w_{1,m} = 0.1 - 0.3(-0.00283) = \mathbf{0.100849}$$

| **weights** | **values** |
|---|---|
| $w_{m,n}$ | 0.1189 |
| $w_{1,n}$ | 0.134 |
| $w_{b,m}$ | 0.1 |
| $w_{a,m}$ | 0.100849 |
| $w_{1,m}$ | 0.100849 |

Table 1: Weights after one presentation for Example 1

| $w_{m,n}$ | $w_{1,n}$ | $w_{b,m}$ | $w_{a,m}$ | $w_{1,m}$ |
|---|---|---|---|---|
| 0.1189 | 0.134 | 0.1 | 0.100849 | 0.100849 |

Table 2: Weights for Example 2

## Example 2

Data: $a = 0, b = 1, t = 0$
Input at unit m,

$$m_{in} = w_{1,m} + a.w_{a,m} + b.w_{b,m}$$

$$\therefore m_{in} = 0.100849 + (0)(0.100849) + (1)(0.1) = 0.200849$$

Output at unit m,

$$m_{out} = \frac{1}{1 + e^{-m_{in}}} = \frac{1}{1 + e^{-0.200849}} = 0.55$$

Input at unit n,

$$n_{in} = m_{out}.w_{m,n} + w_{1,n}$$

$$\therefore n_{in} = (0.55)(0.1189) + 0.134 = 0.199$$

Output at unit n,

$$n_{out} = \frac{1}{1 + e^{-n_{in}}} = \frac{1}{1 + e^{-0.199}} = 0.549$$

Error,

$$(t - o) = 0 - 0.549 = -0.549$$

Calculating gradient $\frac{\partial E}{\partial w_{m,n}}$,

$$\frac{\partial E}{\partial w_{m,n}} = (\frac{\partial E}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{m,n}})$$

$$\therefore \frac{\partial E}{\partial w_{m,n}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{m,n}})$$

$$\therefore \frac{\partial E}{\partial w_{m,n}} = -(t - o).(n_{out}.(1 - n_{out})).m_{out}$$

$$\therefore \frac{\partial E}{\partial w_{m,n}} = -(-0.549)(0.549)(1 - 0.549)(0.55) = \mathbf{0.074}$$

Updating weight $w_{m,n}$,

$$w_{m,n} = w_{m,n} - \eta \frac{\partial E}{\partial w_{m,n}}$$

$$\therefore w_{m,n} = 0.1189 - 0.3(0.074) = \mathbf{0.096}$$

Calculating gradient $\frac{\partial E}{\partial w_{1,n}}$,

$$\frac{\partial E}{\partial w_{1,n}} = (\frac{\partial E}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{1,n}})$$

$$\therefore \frac{\partial E}{\partial w_{1,n}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial w_{1,n}})$$

$$\therefore \frac{\partial E}{\partial w_{1,n}} = -(t-o).(n_{out}.(1-n_{out})).1$$

$$\therefore \frac{\partial E}{\partial w_{1,n}} = -(-0.549)(0.549)(1-0.549).1 = \mathbf{0.135}$$

Updating weight $w_{1,n}$,

$$w_{1,n} = w_{1,n} - \eta \frac{\partial E}{\partial w_{1,n}}$$

$$\therefore w_{1,n} = 0.134 - 0.3(0.135) = \mathbf{0.093}$$

Calculating gradient $\frac{\partial E}{\partial w_{b,m}}$,

$$\frac{\partial E}{\partial w_{b,m}} = (\frac{\partial E}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{b,m}})$$

$$\therefore \frac{\partial E}{\partial w_{b,m}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial m_{out}})(\frac{\partial m_{out}}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{b,m}})$$

$$\therefore \frac{\partial E}{\partial w_{b,m}} = -(t-o).(n_{out}.(1-n_{out})).w_{m,n}.(m_{out}.(1-m_{out})).b$$

$$\therefore \frac{\partial E}{\partial w_{b,m}} = -(-0.549)(0.549)(1-0.549)(0.1189)(0.55)(1-0.55)(1) = \mathbf{4 \times 10^{-3}}$$

Updating weight $w_{b,m}$,

$$w_{b,m} = w_{b,m} - \eta \frac{\partial E}{\partial w_{b,m}}$$

$$\therefore w_{b,m} = 0.1 - 0.3(0.004) = \mathbf{0.0988}$$

Calculating gradient $\frac{\partial E}{\partial w_{a,m}}$,

$$\frac{\partial E}{\partial w_{a,m}} = (\frac{\partial E}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{a,m}})$$

$$\therefore \frac{\partial E}{\partial w_{a,m}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial m_{out}})(\frac{\partial m_{out}}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{a,m}})$$

$$\therefore \frac{\partial E}{\partial w_{a,m}} = -(t-o).(n_{out}.(1-n_{out})).w_{m,n}.(m_{out}.(1-m_{out})).a$$

$$\therefore \frac{\partial E}{\partial w_{a,m}} = -(-0.549)(0.549)(1-0.549)(0.1189)(0.55)(1-0.55)(0) = \mathbf{0.0}$$

Updating weight $w_{a,m}$,

$$w_{a,m} = w_{a,m} - \eta\frac{\partial E}{\partial w_{a,m}}$$

$$\therefore w_{a,m} = 0.100849 - 0.3(0.0) = \mathbf{0.100849}$$

Calculating gradient $\frac{\partial E}{\partial w_{1,m}}$,

$$\frac{\partial E}{\partial w_{1,m}} = (\frac{\partial E}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{1,m}})$$

$$\therefore \frac{\partial E}{\partial w_{1,m}} = (\frac{\partial E}{\partial n_{out}})(\frac{\partial n_{out}}{\partial n_{net}})(\frac{\partial n_{net}}{\partial m_{out}})(\frac{\partial m_{out}}{\partial m_{net}})(\frac{\partial m_{net}}{\partial w_{1,m}})$$

$$\therefore \frac{\partial E}{\partial w_{1,m}} = -(t-o).(n_{out}.(1-n_{out})).w_{m,n}.(m_{out}.(1-m_{out})).1$$

$$\therefore \frac{\partial E}{\partial w_{1,m}} = -(0.549)(0.549)(1-0.549)(0.1189)(0.55)(1-0.55)(1) = \mathbf{4 \times 10^{-3}}$$

Updating weight $w_{1,m}$,

$$w_{1,m} = w_{1,m} - \eta\frac{\partial E}{\partial w_{1,m}}$$

$$\therefore w_{1,m} = 0.100849 - 0.3(0.004) = \mathbf{0.099}$$

| **weights** | **values** |
|---|---|
| $w_{m,n}$ | 0.096 |
| $w_{1,n}$ | 0.093 |
| $w_{b,m}$ | 0.0988 |
| $w_{a,m}$ | 0.100849 |
| $w_{1,m}$ | 0.099 |

Table 3: Weights after one presentation for Example 2