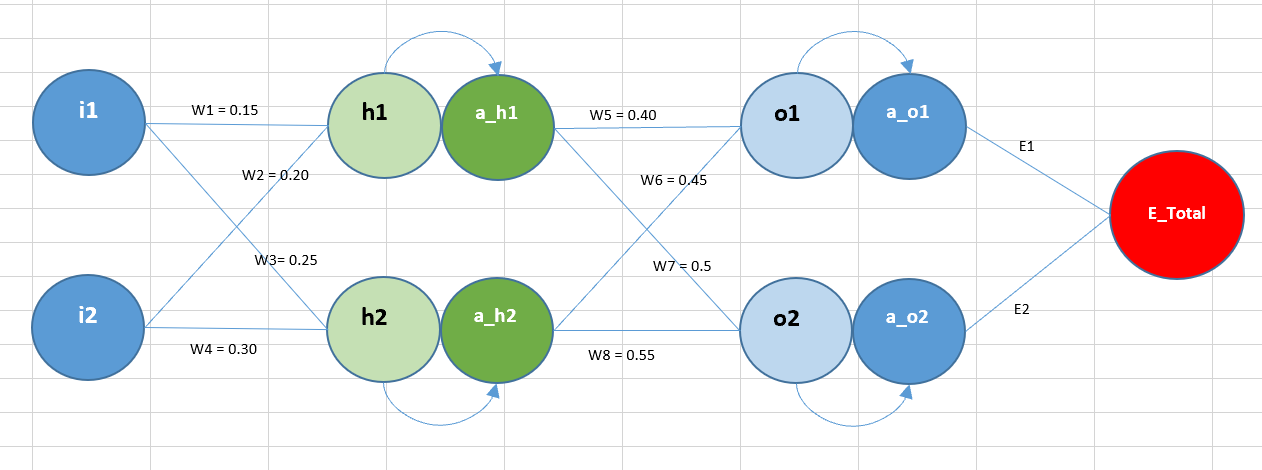
**Neural Network Back Propagation**



**NN parameters**

|  |
| --- |
| **h1 = w1\*i1+w2\*i2** |
| **h2 = w3\*i1+w4\*i2** |
| **a\_h1 = σ(h1) = 1/(1+exp(-h1))** |
| **a\_h2 = σ(h2)** |
| **o1 = w5\*a\_h1+w6\*a\_h2** |
| **o2 = w7\*a\_h1+w8\*a\_h2** |
| **a\_o1 = σ(o1)** |
| **a\_o2 = σ(o2)** |
| **E1 = ½\*(t1-a\_o1)²** |
| **E2 = ½\*(t2-a\_o2)²** |

**Partial Derivation on total loss wrt Weights of network**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

1. **∂E\_T/∂w5 =** ∂(E1+E2)/∂w5 = ∂E1/∂w5 = (∂E1/∂a\_o1) \* (∂a\_o1/∂o1) \* (∂o1/∂w5) ---- (1)

Now, ∂E1/∂a\_o1 = (t1 - a\_o1) \* (-1) = ***a\_o1 – t1*** *---- x*

∂a\_o1/∂o1 = ∂(σ(o1))/ ∂o1 = σ(o1) \* (1 - σ(o1)) = ***a\_o1 \* (1 – a\_o1)*** *---- y*

∂o1/∂w5 = ∂(w5\*a\_h1+w6\*a\_h2)/ ∂w5 = **a\_h1** ---- z

Substituting x, y and z in (1)

**∂E\_T/∂w5 = (*a\_o1 – t1) \* (a\_o1 \* (1 – a\_o1)) \* (*a\_h1)**

Similarly -

1. **∂E\_T/∂w6 = (*a\_o1 – t1) \* (a\_o1 \* (1 – a\_o1)) \* (*a\_h2)**
2. **∂E\_T/∂w7 = (*a\_o2 – t2) \* (a\_o2 \* (1 – a\_o2)) \* (*a\_h2)**
3. **∂E\_T/∂w8 = (*a\_o2 – t2) \* (a\_o2 \* (1 – a\_o2)) \* (*a\_h1)**
4. **∂E\_T/∂a\_h1 =** ∂(E1+E2)/∂a\_h1 = ∂E1/∂a\_h1 + ∂E2/∂a\_h1 ---- (1)

∂E1/∂a\_h1 = ∂E1/∂a\_o1 \* ∂a\_o1/∂o1 \* ∂o1/∂a\_h1 ----l

∂E1/∂a\_o1 = ***a\_o1 – t1*** *---- m*

∂a\_o1/∂o1 = ***a\_o1 \* (1 – a\_o1)*** *---- n*

∂o1/∂a\_h1 = ∂(w5\*a\_h1+w6\*a\_h2) /∂a\_h1 = **w5** ----o

Substituting m, n, o in l

∂E1/∂a\_h1 = **(*a\_o1 – t1) \* (a\_o1 \* (1 – a\_o1)) \* w5*** *--- x*

Similarly,

∂E2/∂a\_h1 = **(*a\_o2 – t2) \* (a\_o2 \* (1 – a\_o2)) \* w7*** *---- y*

Substituting x, y in (1)

**∂E\_T/∂a\_h1 = (*a\_o1 – t1) \* (a\_o1 \* (1 – a\_o1)) \* w5 +*** (***a\_o2 – t2) \* (a\_o2 \* (1 – a\_o2)) \* w7***

1. **∂E\_T/∂a\_h2 = (*a\_o2 – t2) \* (a\_o2 \* (1 – a\_o2)) \* w8 +*** (***a\_o1 – t1) \* (a\_o1 \* (1 – a\_o1)) \* w6***
2. **∂E\_T/∂w1 =** ∂E\_T/∂a\_h1 \* ∂a\_h1/∂h1 \* ∂h1/∂w1 ---- (1)

∂a\_h1/∂h1 = a\_h1 \* ( 1 – a\_h1) -----x

∂h1/∂w1 = i1 ---- y

Substituting x, y in (1)

**∂E\_T/∂w1 = ((*a\_o1 – t1) \* (a\_o1 \* (1 – a\_o1)) \* w5 +*** (***a\_o2 – t2) \* (a\_o2 \* (1 – a\_o2)) \* w7) \* (a\_h1 \* ( 1 – a\_h1) ) \* i1***

Or

**∂E\_T/∂w1 = *∂E\_T/∂a\_h1 \* (a\_h1 \* ( 1 – a\_h1) ) \* i1***

1. **∂E\_T/∂w2 = *∂E\_T/∂a\_h1 \* (a\_h1 \* ( 1 – a\_h1) ) \* i2***
2. **∂E\_T/∂w3 = *∂E\_T/∂a\_h2 \* (a\_h2 \* ( 1 – a\_h2) ) \* i1***
3. **∂E\_T/∂w4 = *∂E\_T/∂a\_h2 \* (a\_h2 \* ( 1 – a\_h2) ) \* i2***