This file will describe how I derived the equations to change the weights in the neural network and how the statistical method will work.

Neural Network Backpropagation:

Due to the fact that I do not have much experience in creating neural networks, the equation for calculating the value for each node will only be a weighted sum of the previous layer.

$$a_j^k = \sum_{i=1}^{12} a_i^{k-1} w_{ij}^k . k - layer\ of\ the, destination\ of\ the\ weight. i, j - indexes\ of\ the\ nodes$$

The error will be calculated as the average of the error of each of the output nodes.

$$E = \frac{1}{3}\sum_{d=1}^{3} E_d$$

And the error for each of the output node will be calculated as such.

$$E_d = (a_d^4 - \hat{a}_d)^2, \hat{a}\ being\ the\ desired\ output\ at\ node\ d$$

The change to the weights will be calculated as such

$$\Delta w_{ij}^k = -\alpha \frac{\partial E}{\partial w_{ij}^k}$$

Alpha being the learning step so that we don't overshoot, I and j being the indexes where the node is going from and to, k being the layer where the node is going to. The equation can be further manipulated

$$\frac{\partial E}{\partial w_{ij}^k} = \frac{1}{3}\sum_{d=1}^{3} \frac{\partial E_d}{\partial w_{ij}^k}$$

$$\frac{\partial E_d}{\partial w_{ij}^k} = \frac{\partial E_d}{\partial a_j^4}\frac{\partial a_j^4}{\partial w_{ij}^k}$$

This will have to be done on a layer per layer basis, starting at the last layer. So for the case w_{id}^4.

$$\frac{\partial E_d}{\partial w_{ij}^4} = \frac{\partial E_d}{\partial a_j^4}\frac{\partial a_j^4}{\partial w_{ij}^4}$$

Because the case when d is not equal to j, a_j^4 does not affect the error in E_d, hence those other terms will be 0.

$$\frac{\partial E_d}{\partial a_j^4} = \frac{\partial}{\partial a_j^4}\left(a_d^4 - \hat{a}_d\right)^2 = 2(a_j^4 - \hat{a}_j)$$

$$\frac{\partial a_j^4}{\partial w_{ij}^4} = \frac{\partial}{\partial w_{ij}^4} \sum_{l=1}^{12} a_l^3 w_{lj}^4$$

Only one of those terms will have w_{ij}^4, when l=i, hence all but that term will end up as 0.

$$\frac{\partial a_j^4}{\partial w_{ij}^4} = \frac{\partial}{\partial w_{ij}^4} a_i^3 w_{ij}^4 = a_i^3$$

$$\frac{\partial E_d}{\partial w_{ij}^4} = 2(a_j^4 - \hat{a}_j)\, a_i^3$$

$$\frac{\partial E}{\partial w_{ij}^4} = \frac{2}{3}(a_d^4 - \hat{a}_d)\, a_i^3$$

$$\Delta w_{ij}^4 = -\frac{2\alpha}{3}(a_j^4 - \hat{a}_j)\, a_i^3$$

Now with k=4 done, we can move on to k=3 – the weights going from layer 2 to layer 3.

$$\frac{\partial E}{\partial w_{ij}^3} = \frac{1}{3} \sum_{d=1}^{3} \frac{\partial E_d}{\partial w_{ij}^3}$$

$$\frac{\partial E_d}{\partial w_{ij}^3} = \frac{\partial E_d}{\partial a_d^4} \frac{\partial a_d^4}{\partial w_{ij}^3} = \frac{\partial E_d}{\partial a_d^4} \frac{\partial a_d^4}{\partial a_j^3} \frac{\partial a_j^3}{\partial w_{ij}^3}$$

$$\frac{\partial a_d^4}{\partial a_j^3} = \frac{\partial}{\partial a_j^3} \sum_{l=1}^{12} a_l^3 w_{ld}^4 = \frac{\partial}{\partial a_j^3} a_j^3 w_{jd}^4 = w_{jd}^4$$

$$\frac{\partial a_j^3}{\partial w_{ij}^3} = \frac{\partial}{\partial w_{ij}^3} \sum_{l=1}^{12} a_l^2 w_{ij}^3 = \frac{\partial}{\partial w_{ij}^3} a_i^2 w_{ij}^3 = a_i^2$$

$$\frac{\partial E_d}{\partial w_{ij}^3} = 2(a_d^4 - \hat{a}_d)\, w_{jd}^4 a_i^2$$

$$\frac{\partial E}{\partial w_{ij}^3} = a_i^2 \frac{2}{3} \sum_{d=1}^{3} (a_d^4 - \hat{a}_d)\, w_{jd}^4$$

$$\Delta w_{ij}^3 = -\alpha \frac{2a_i^2}{3} \sum_{d=1}^{3} (a_d^4 - \hat{a}_d)\, w_{jd}^4$$

Repeating the process for k=2, for weights going from layer 1 to layer 2

$$\frac{\partial E}{\partial w_{ij}^2} = \frac{1}{3} \sum_{d=1}^{3} \frac{\partial E_d}{\partial w_{ij}^2}$$

$$\frac{\partial E_d}{\partial w_{ij}^2} = \frac{\partial E_d}{\partial a_d^4}\frac{\partial a_d^4}{\partial w_{ij}^2} = \frac{\partial E_d}{\partial a_d^4}\frac{\partial a_d^4}{\partial a_j^2}\frac{\partial a_j^2}{\partial w_{ij}^2} = \frac{\partial E_d}{\partial a_d^4}\frac{\partial a_d^4}{\partial a_j^2}\frac{\partial a_j^2}{\partial w_{ij}^2}$$

$$\frac{\partial a_d^4}{\partial a_j^2} = \frac{\partial}{\partial a_j^2}\sum_{p=1}^{12} a_p^3 w_{pd}^4 = \sum_{p=1}^{12} w_{pd}^4 * (\sum_{o=1}^{12}\frac{\partial}{\partial a_j^2} a_o^2 w_{op}^3) = \sum_{p=1}^{12} w_{pd}^4 * (\frac{\partial}{\partial a_j^2} a_j^2 w_{jp}^3) = \sum_{p=1}^{12} w_{pd}^4 w_{jp}^3$$

$$\frac{\partial a_j^2}{\partial w_{ij}^2} = \frac{\partial}{\partial w_{ij}^2}\sum_{p=1}^{12} a_p^1 w_{pj}^2 = \frac{\partial}{\partial w_{ij}^2} a_i^1 w_{ij}^2 = a_i^1$$

$$\frac{\partial E_d}{\partial w_{ij}^2} = 2(a_d^4 - \hat{a}_d) * (\sum_{p=1}^{12} w_{pd}^4 w_{jp}^3) * \ a_i^1$$

$$\frac{\partial E}{\partial w_{ij}^k} = \frac{1}{3}\sum_{d=1}^{3}\frac{\partial E_d}{\partial w_{ij}^k} = \frac{2a_i^1}{3}\sum_{d=1}^{3}(a_d^4 - \hat{a}_d) * (\sum_{p=1}^{12} w_{pd}^4 w_{jp}^3)$$

$$\Delta w_{ij}^2 = -\alpha\frac{\partial E}{\partial w_{ij}^2} = -\alpha\frac{2a_i^1}{3}\sum_{d=1}^{3}(a_d^4 - \hat{a}_d) * (\sum_{p=1}^{12} w_{pd}^4 w_{jp}^3)$$

Hence we get the equations for all of the 3 layers of weights:

$$\Delta w_{ij}^2 = -\alpha\frac{2a_i^1}{3}\sum_{d=1}^{3}(a_d^4 - \hat{a}_d) * (\sum_{p=1}^{12} w_{pd}^4 w_{jp}^3)$$

$$\Delta w_{ij}^3 = -\alpha\frac{2a_i^2}{3}\sum_{d=1}^{3}(a_d^4 - \hat{a}_d)\, w_{jd}^4$$

$$\Delta w_{ij}^4 = -\frac{2\alpha}{3}(a_j^4 - \hat{a}_j)\, a_i^3$$

Statistical method algorithm:

1. Take the data of which we want to produce a forecast, an array, a, of size 12 and type double.
2. Create an array, d, of type double of size 11
3. $d[i] = \frac{a[i+1]-a[i]}{a[i]} = \frac{a[i+1]}{a[i]} - 1$
4. For each product which has more than 14 entries
    a. Create an array, e, of type double and size one smaller than the total number of entries in that product.
    b. $e[i] = \frac{a[i+1]-a[i]}{a[i]} = \frac{a[i+1]}{a[i]} - 1$, a being the entries within the product
5. Get the number of possible 15 sized consecutive chunks, c, by formula $c = n - 14$, n being the size of array e.
6. For each c create a new array f of size 15 and populate it by consecutive values in e, starting from c
7. Get a value E via equation $E = \sum_{i=1}^{12}(d[i] - f[i])^2$
8. Find the sequence with the smallest value of E.
9. Get the values f[13], f[14] and f[15].

10. Return the values a[12]*(f[13]+1), a[12]*(f[13]+1) *(f[14]+1) and a[12]*(f[13]+1) *(f[14]+1) *(f[15]+1) as the forecast for that product.

Bibliography

[1]"Backpropagation | Brilliant Math & Science Wiki", *Brilliant.org*, 2018. [Online]. Available: https://brilliant.org/wiki/backpropagation/. [Accessed: 19- Mar- 2018].