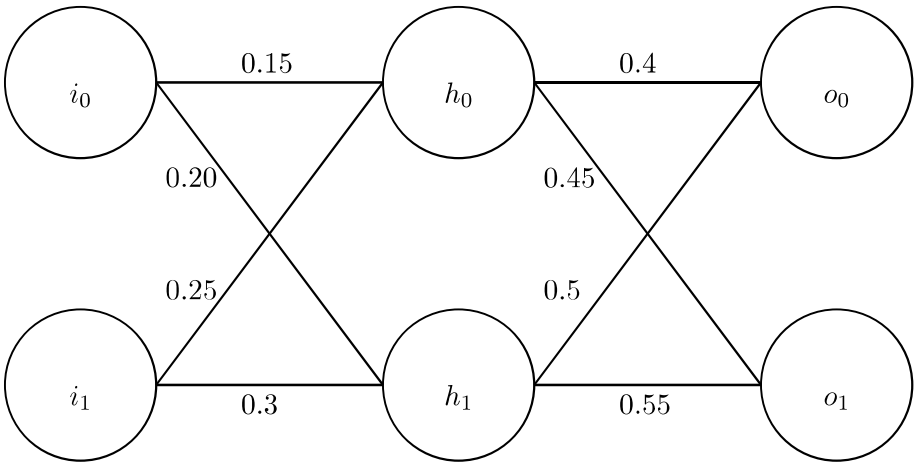
Neural Network

The neural network below contains an input layer, one hidden layer, and an output layer.



The bias for h_0 and h_1 is 0.35 and the bias for o_0 and o_1 is 0.6.

We will give the inputs 0.05 and 0.10 to i_0 and i_1 respectively, and expect outputs 0.01 and 0.99 from outputs o_0 and o_1 respectively.

1. Show that the activation value at h_0 is 0.3825.
2. Find the activation values for the output nodes.
3. Calculate the sum of the squares of the errors.
4. Suppose we use scary multivariable calculus to find the the rate of change of the total error with respect to the weight of i_0^0 (which has a value of 0.15) is -0.2, and the rate of change of the total error with respect to the weight of i_1^0 is 0.9. Which should we change, and in which direction? (i.e. add or subtract). Briefly explain your answer.
5. In the absence of multivariable calculus, which heuristic(s) that you have studied previously could be used to optimise your neural network.

```
h0 = (i0 * 0.15 + i1 + 0.25) + 0.35
   = 0.0325 + 0.35
   = 0.3825, as required
h1 = (i0 * 0.20 + i1 * 0.3) + 0.35
   = 0.04 + 0.35
   = 0.39
o0 = (h0 * 0.4 + h1 * 0.5) + 0.6
   = 0.348 + 0.6
   = 0.948
o1 = (h0 * 0.45 + h1 * 0.55) + 0.6
   = 0.386625 + 0.6
   = 0.986625
```

```
Sum of Squares of Errors:
SSE = (0.948 - 0.01)^2 + (0.986625 - 0.99)^2
    = (0.938)^2 + (-0.003375)^2
    = 0.879844 + 0.000011
    = 0.879855

i0,0 should be increased slightly because it has a
negative gradient, meaning increasing will result in
a lower error
i0,1 should be decreased slightly as it has a
positive gradient, and decreasing will result in a
lower error.
```

```
We could use the Hill Climbing Heuristic or Simulated Annealing to optimise the NN, since
they try to find local minima and global minima respectively.
```