**How many nodes/layers?**

Theoretically only need 1 hidden layer for any non-linear problem, but in practice you need more

Look at existing work and start from that

More for more complex problems

Output neurons same as - number of outputs

Number of inputs – number of features? Or is that the parameters?

Each node in network is performing logistic regression, modifying weights of parameters it has as inputs (which are outputs from previous nodes, or features) which form an equation, which provides an output from the parameter values given