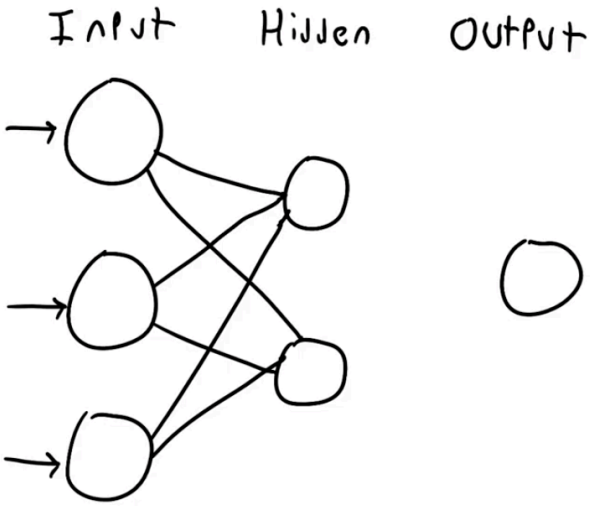
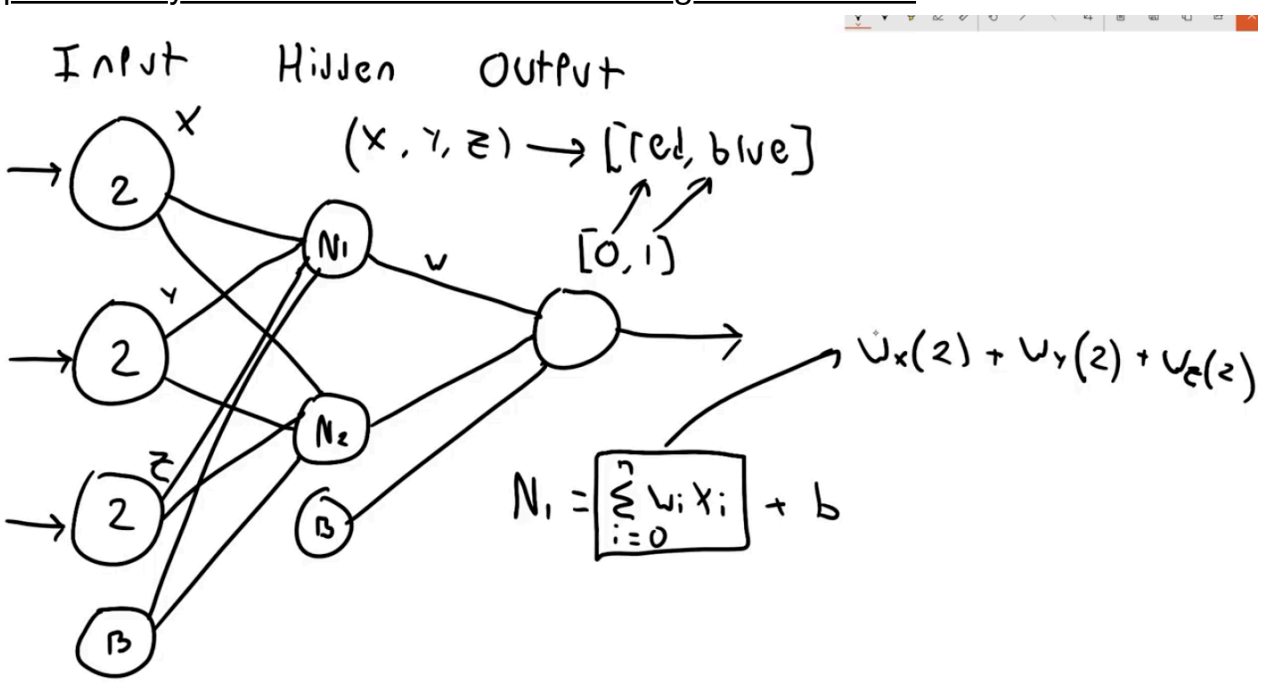
- ***Neural networks***
  - maths
  - gradient descent
  - propagation
  - using them to classify
  - explaining these individually then combining them at the end
- ***The black box model***
  - for classifications / predictions
  - we have inputs and outputs
  - you feed in information and it gives you an output (e.g a classification)
  - F(x) -> taking an input and mapping it to some output
  - the black box is taking the input and mapping it to something else -> the output
- ***Layers***
  - -> the neural network is made up of layers
  - ***-> the input layer***
    - this accepts the raw data
    - e.g the pixels in an image -> row by column -> times them together and those are the numbers of input neurons the model will have
  - ***-> the output layer***
    - -> for example if we have an image classifier
    - -> the image is either one output or another
    - -> the training and testing data
    - -> each of the different things which the image could be a picture of is given by one output neurone
      - -> each of those neurones has a probability that the image will be that
      - -> the most likely one the image is is the output of the model
      - -> linear regression is with one output neurone
  - ***-> hidden layers***
    - these are the layers in between the input and output layers
    - -> we don't observe them
  - -> each layer is connected to the next one using widths
    - -> one node can be connected to any node in the next layer
      - -> this is a dense network
        - -> the mappings (which of these nodes it is) depends on the weightings which the model gives
        - -> the trainable parameters are the parameters whose weight the model can tweak
        - -> you can look at one node and calculate the number of notes in the next layer which it can be connected to, by multiplying them out
          - -> to work out the total number of connections in the entire model



  - ***-> biases***
    - -> biases exist in the previous layers
    - -> a bias exists in the previous layer to the layer that it connects

        - -> the bias is a constant which we connect to the hidden layer
        - -> they have constant weights of 1
        - -> he biases don't connect with eachother -> biases are constants which are added to the model and then we have weights which determine how liklely the outcome that the node which the model it's connecting to is
- ***-> colour classification example***
    - -> the output of a model is red or blue
    - -> the input is a datapoint (2,2,2) and we want o know if it's red or blue
    - -> going to the next layer and finding out what the values of the nodes are at that point
    - -> calculating the weighted sum of all of the nodes which are connected to it -> in other words the expected value of a node
        - -> with the values in that equation being the different nodes which are feeding into it from the previous layer in the model
    - -> the weights are completely random
        - -> then updated and changed to make sense in the model
        - -> we are also adding in the bias -> which was connected with a weight of 1
        - -> it's for the probability equations -> adding a constant to the summation equation
        - -> N is the number of neurones in the previous layers -> we are iterating through the layers
        - -> the bias is the constant which you add onto the statistics equation
            - and the inputs of the statistic equation are the weights of the nodes from the previous layer in the model which are feeding into that node



        - -> a densely connected neural network is one in which all the neurons in the current layer are connected to every neurone in the previous layer