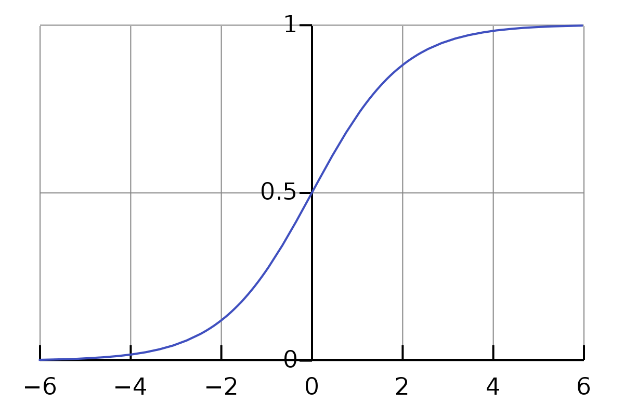
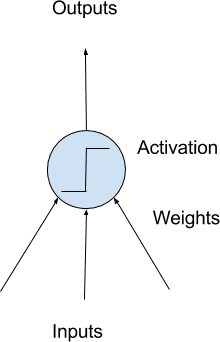
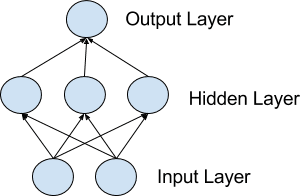
<https://www.sciencedirect.com/science/article/abs/pii/S0168900205022679>

* inspired by simple model of how the brain works in nature:
  + a neuron”fires” if the stimuli received from other neurons exceed a certain threshold
  + described by:
    - 
    - g(t) is a sigmoid
      * 
      * ^sigmoid
    - determines the threshold
    - node j in layer n is given by the weighted sum of all nodes in layer n-1
    - network training is minimising a loss function by adjusint the weights such that the deviation of the actual network output from the desired output is minimized
    - popular loss functions: sum of quadratic deviations or a measure of the entropy
  + do not require that all input variables are filled for eachevent
  + able to learn correlations between the input variables
  + can incorporate information from quality variables

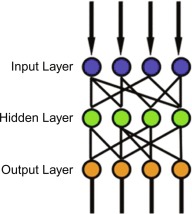
<https://machinelearningmastery.com/neural-networks-crash-course/>

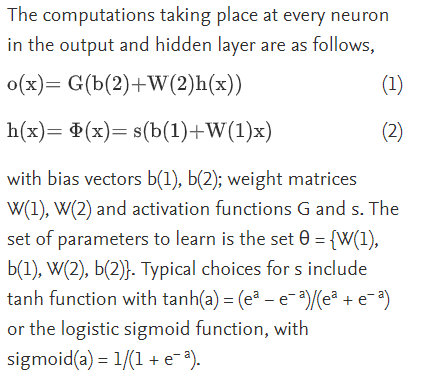
* Multi-Layer Perceptrons
  + perceptron=single neuron model that was a precursor to larger neural networks
  + simple models of biological brains can be used to solve difficult computalional tasks ie machine learning
  + learn the representation in training data and how to best relate it to the output variable you want to predict
    - neural networks learn mapping
* neurons
  + 
  + have weighted input signals and produce an output signal using an activation function
* neuron weights
  + each neuron has a bias, thought of as an input that always has a value of 1.0 and must be weighted
  + 2 input=3 weights
    - one for each input and one for the bias
  + weights often intialized to small random values (such as values from range 0 to 0.3)
  + larger weights indicate increased complexity and fragility
* activation
  + weighted inputs summed and passed through an activation function (sometimes called a transfer function)
  + activation function
    - mapping of summed weighted input to the output of the neuron
    - governs the threshold at which the neuron is activated and strength of the output signal
  + non-linear functions called the sigmoid function used that output a value between 0 and 1 with a s-shaped distribution
* network of neurons
  + neurons arranged in network
  + row of neurons=layer
  + network can have multiple layers
  + architecture of the neurons called the network topology
  + 
* input or visible layers
  + bottom layer that takes the input from the data set
  + exposed part of the netowork
  + pass the input value through to the next layer
* hidden layers
  + not directly exposed to the input
* output layer
  + ouput a value or vector of values
  + choice of activation function in the output layer is constrained by the type of problem being modeled
    - “A binary classification problem may have a single output neuron and use a sigmoid activation function to output a value between 0 and 1 to represent the probability of predicting a value for the class 1. This can be turned into a crisp class value by using a threshold of 0.5 and snap values less than the threshold to 0 otherwise to 1.”

<https://www.sciencedirect.com/topics/computer-science/multilayer-perceptron>

multilayer perceptron

* supplement of feed forward neural network
* 3 types of layers:
  + input layer, output layer, hidden layer
* input layer recieves the input signal
* prediction and classification performed by the output layer
* arbitrary number of hidden layers are placed between the input and output layer
* neurons trained with the back propagation learning algorithm

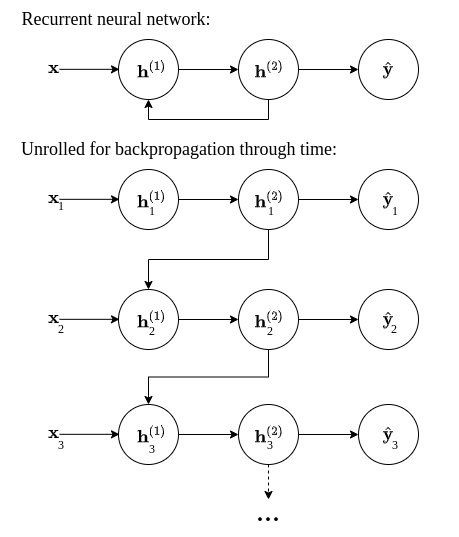


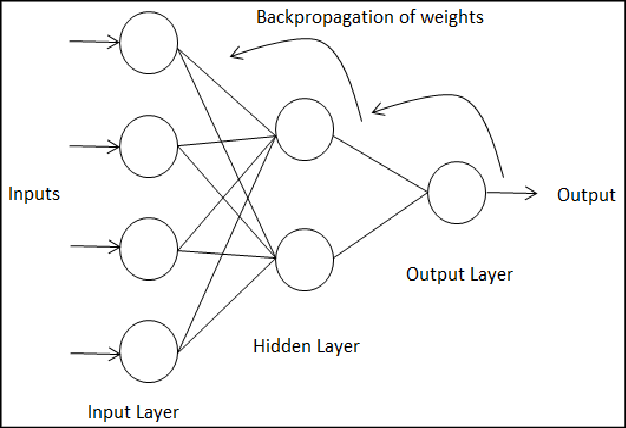


<https://deepai.org/machine-learning-glossary-and-terms/backpropagation>

backpropagation

* backward propagation of errors
* calculating derivatives inside deep feedforward neural networks
* calculate the gradient of the loss function with respect to each of the weights of the network
* enables every weight to be updated individually to gradually reduce the loss function
* calculation of the gradient proceeding backwards through the feedforward network from the last layer through to the first
* to calculate the gradient of a layer, the gradients of all following layers are combined via the chain rule (of calculus)





<https://deepai.org/machine-learning-glossary-and-terms/hidden-layer-machine-learning#:~:text=Hidden%20layers%2C%20simply%20put%2C%20are,specific%20to%20an%20intended%20result.&text=Hidden%20layers%20allow%20for%20the,to%20produce%20a%20defined%20output.>

hidden layer

* provides nonlinear transofmrations of the inputs entered into the network
* vary depending on the function of the neural network
* vary depending on their associated weights
* layers of mathematical functions each designed to produce an output specific to an intended result
* allow for the function of a neural network to be broken down into specifc transformation of the data

videos

* ANNS capable of learning
  + use experience to improve performance
* neurons connected by weighted links passing signals from one neuron to another
* each neuron recieves a number of input signals through its connections, only produces a single output signal