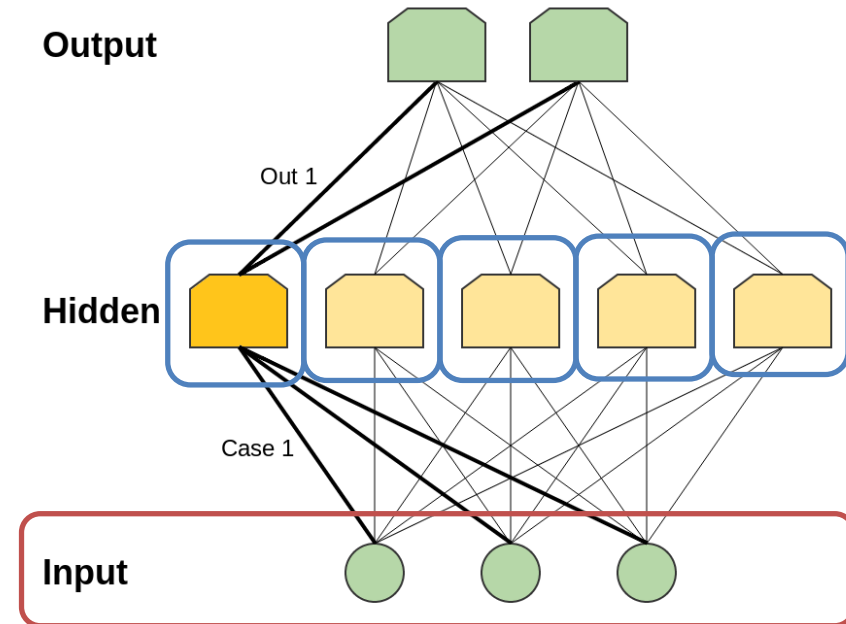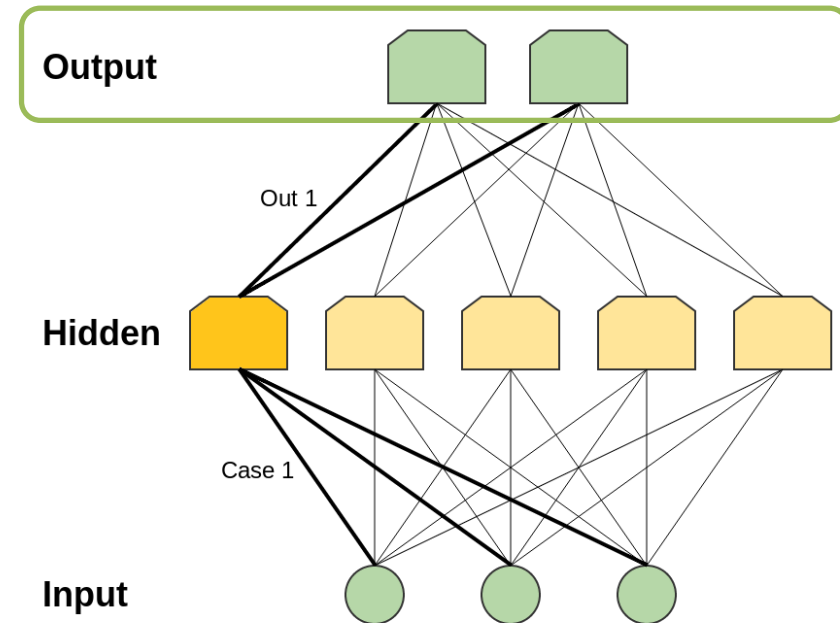# Artificial Neural Network

- Artificial Neural Networks are *biologically-inspired* models, consisting of interconnected neurons

- As a simple example, suppose each observation has three independent variables x_i

- The values of these three variables are fed to a number of **hidden neurons**, which combine them linearly and transform them with an **activation function** $F(\cdot)$

$$F\left(b_j + \sum w_{ji} x_i\right)$$

- The activation function is either logistic, tanh or most recently, *rectified linear unit*:

$$F(z) = \max(0, z)$$

- b_j and w_{ji} need to be fitted

Source: colah's blog
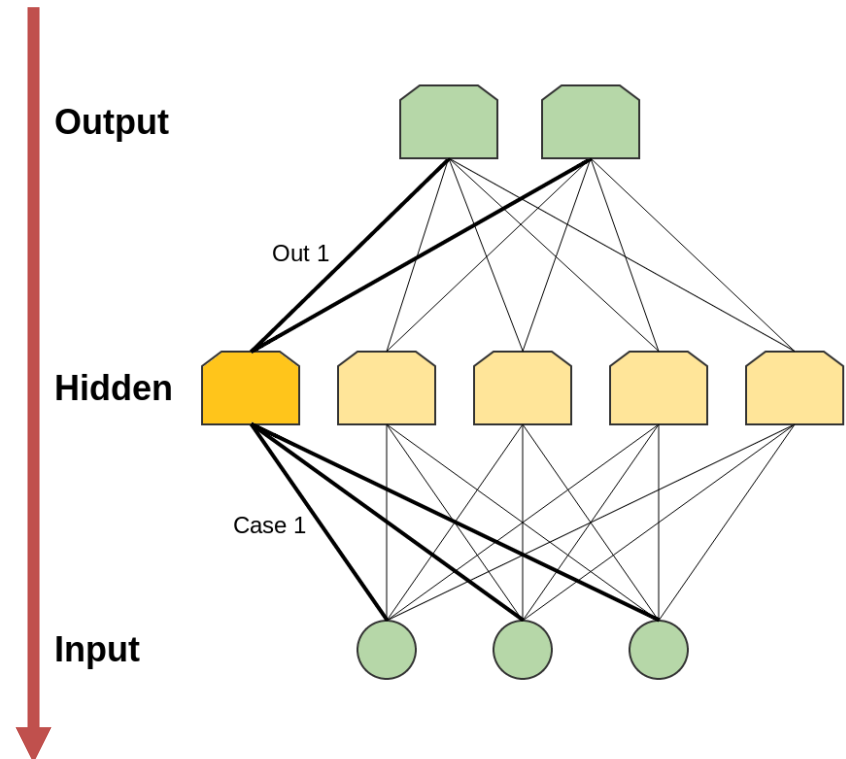
Artificial Neural Network

- The outputs from the hidden neurons are fed into the output neurons, which combine them linearly and transform them again

- The number of output neurons depends on the nature of the dependent variable

  - Single output neuron for linear or binary dependent variable

  - Multiple output neurons for categorical variable, each representing a score for a category. The outputs of all output neurons would be combined through a softmax function— i.e. multinomial logit

Source: colah's blog

Artificial Neural Network

- Parameter estimate is conducted through **back propagation**

    - The residual $(\hat{y} - y)$ is used to correct the parameters in each layer through repeated use of chain rule

    - This process could become unstable as the number of layers increase

    - Techniques developed to overcome this problem: carefully chosen initial values, variable learning rates and normalize output values after every layer

Output

Out 1

Hidden

Case 1

Input

Source: colah's blog

Concept

- An econometrician's view of artificial neural network: a bunch of regression stacked together
  - Halbert White made significant contribution to the theoretical foundation of ANN in the 1980s
  - Application was rare because the lack of computational power

ARTIFICIAL NEURAL NETWORKS
Approximation & Learning Theory
HALBERT WHITE
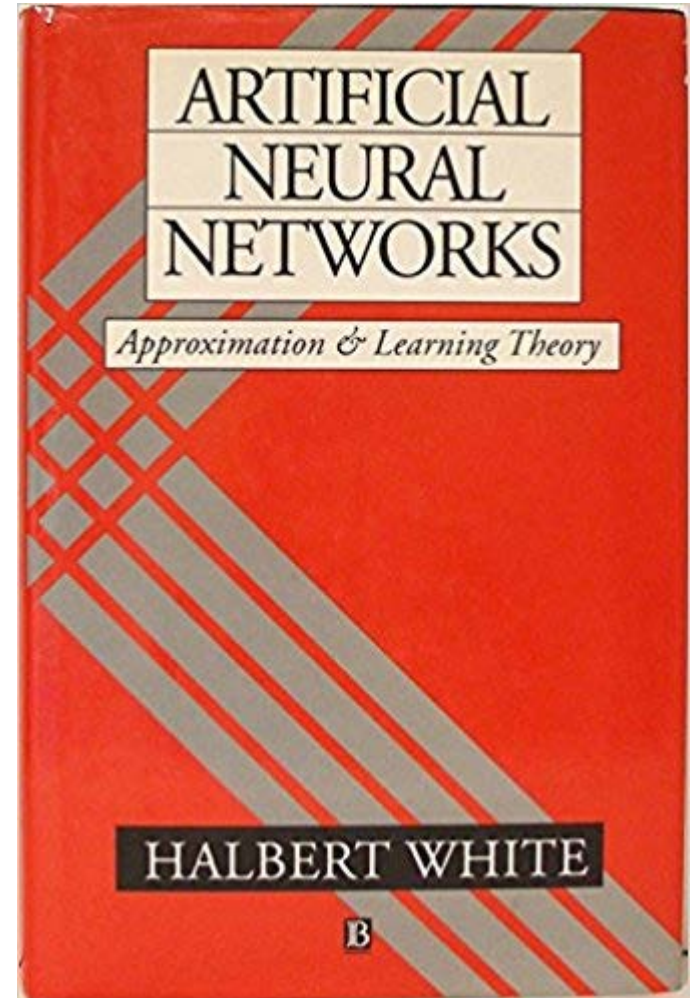
Deep Learning

- **Deep Learning** refers to the stacking of multiple hidden layers
  - Typically in the single digit, but can go as high as a hundred layers
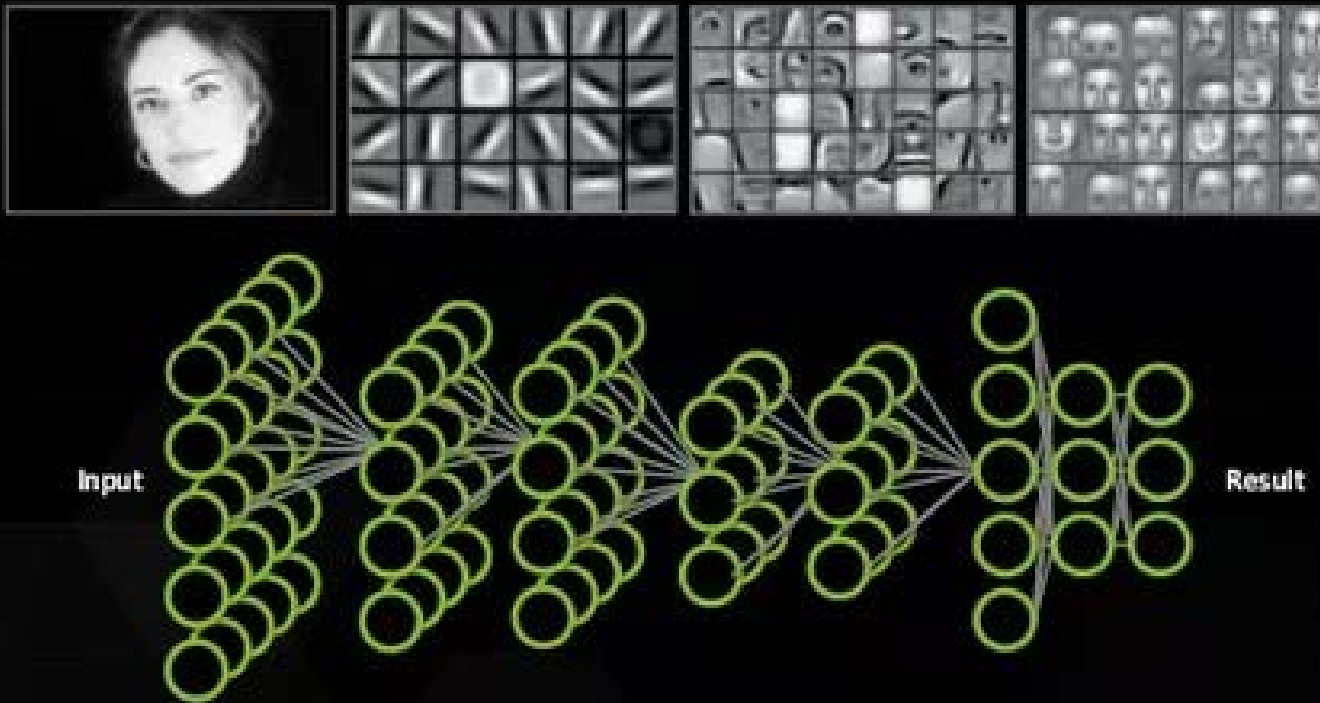
Different Types of ANN
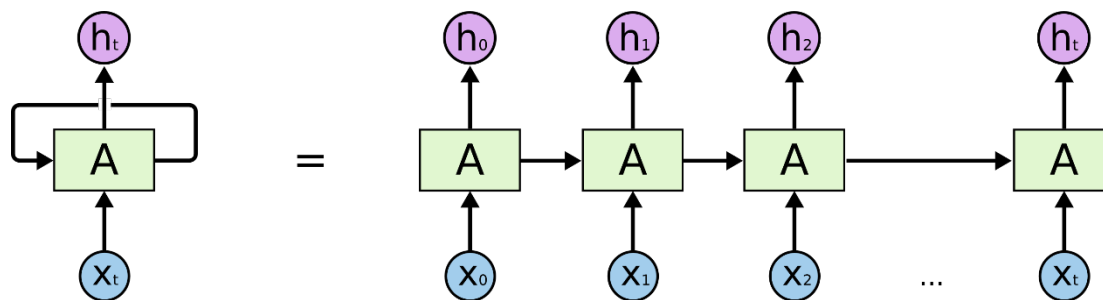
- Convolutional Neural Networks
  - Each neuron is only connected to neighboring neurons
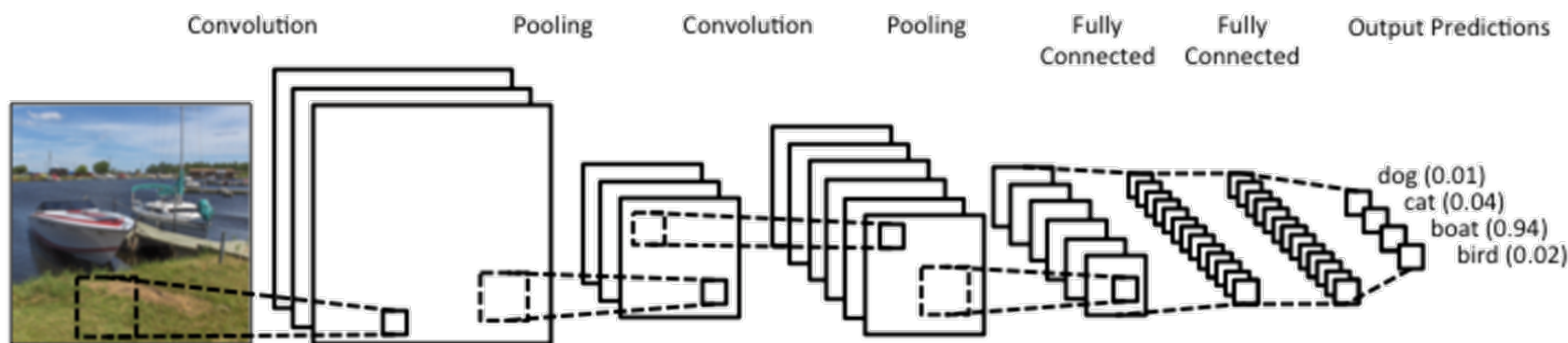
Source: WILDML

- Recurrent Neural Networks
 - Auto-regressive neurons with the ability to forget

Source: colah's blog

Computation

- The idea of artificial neural network can be traced back to the 1940s

- Due to the large number of parameters and large data size involved, effective use of ANN is prohibitive until recently

- ANN took off in recent years due to massive increase in computational capabilities, particularly in the use of graphic processing unit (GPU) for computation

(3 variables + intercept) × 5 hidden neurons
= 20 parameters to fit

(30 variables + intercept) × 1000 hidden neurons
× 5 layers
= 155,000 parameters to fit

NVIDIA® DGX STATION™
YOUR PERSONAL
AI SUPERCOMPUTER
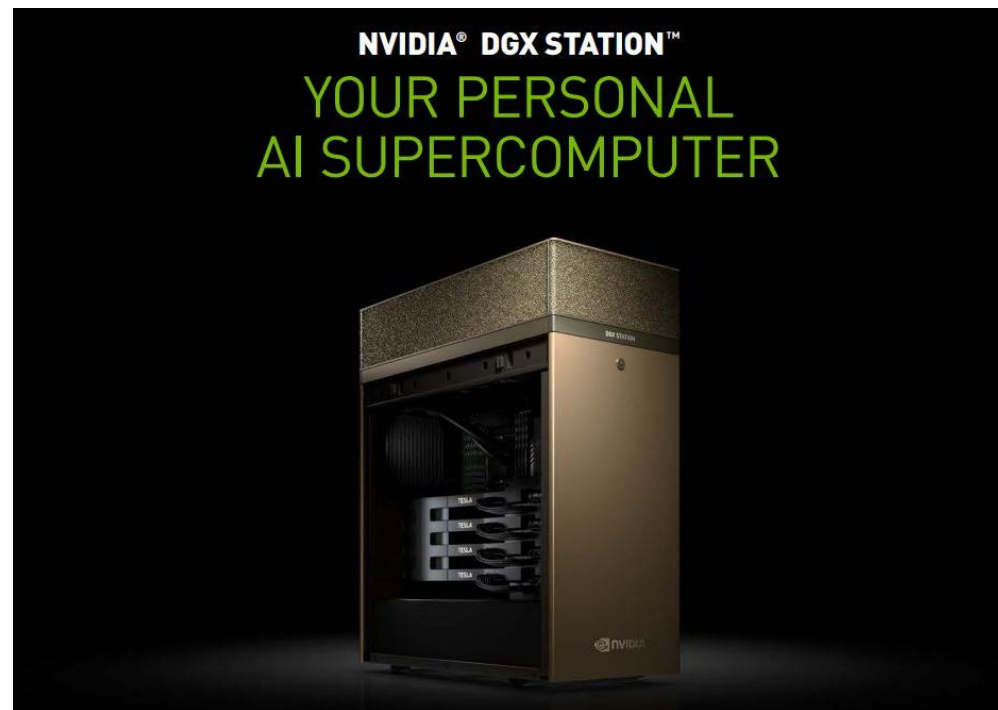
Source: Nvidia

Hyperparameters

- The number of neurons per layer, the number and types of layers to use as well as the rate of learning has to be hand picked. These are called **hyperparameters**

- Hyperparameters are chosen through **cross validation**
  1. Separate data into 3 sets: train, validation and test
  2. The train set is used to train the model. This is repeated for every combination of hyperparameters
  3. The combination of hyperparameters that best predicts the validation set is chosen
  4. The test set is only used for reporting the goodness-of-fit of the chosen hyperparameters