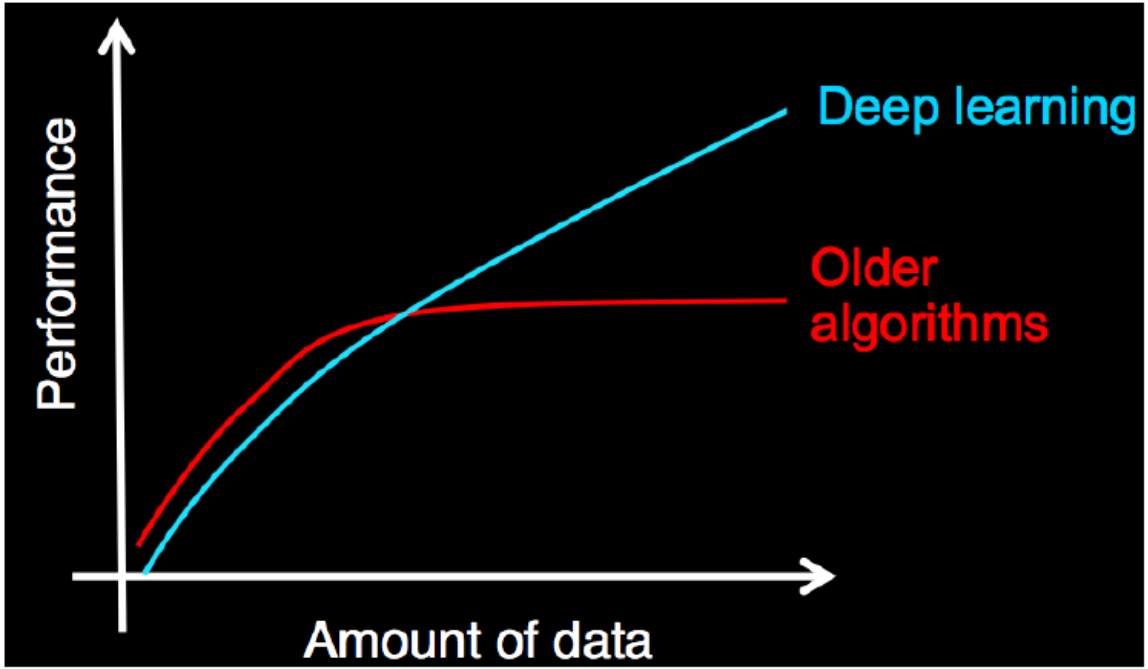
Deep Learning

Part I

History

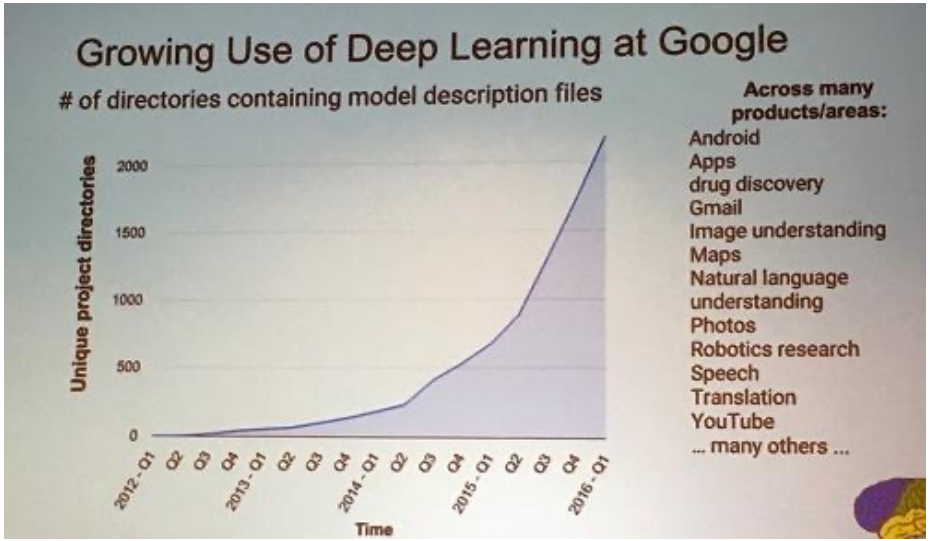
- Nearly everything we talk about today existed during 1990
- What changed?
  - – More data
  - – Faster computers (GPUs)
  - – Some improvements:
    - relu
    - Drop-out
    - adam
    - batch-normalization
    - residual networks

Deep Learning and Data



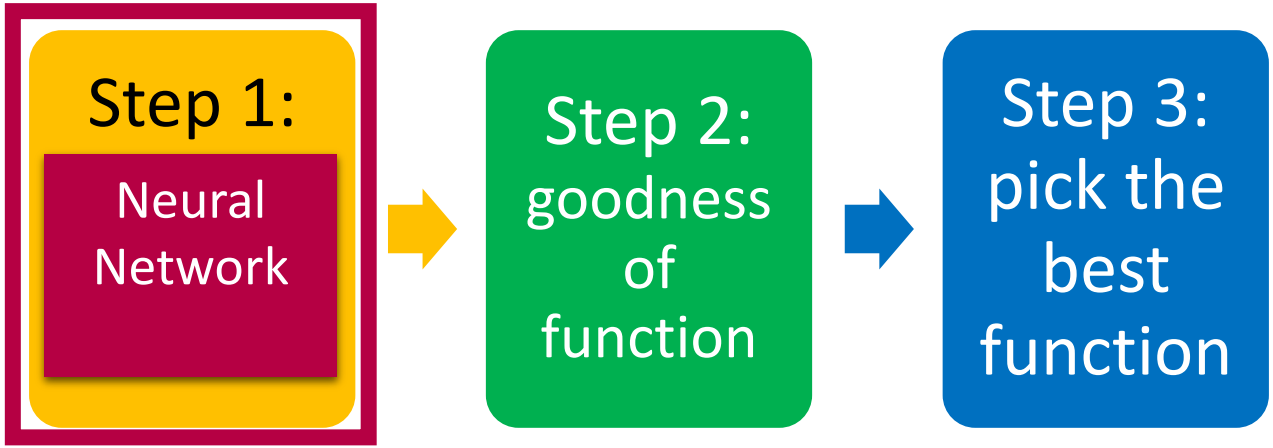
Deep Learning Attracts a lot of Attention

You may have seen lots of exciting results before.

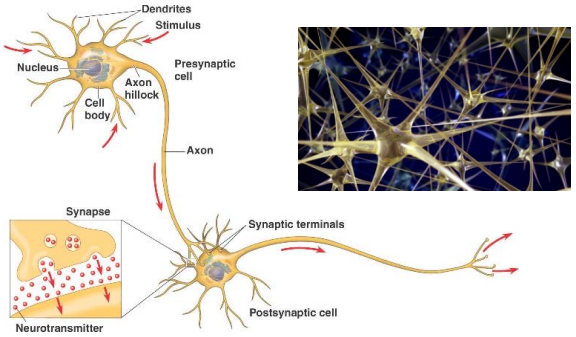


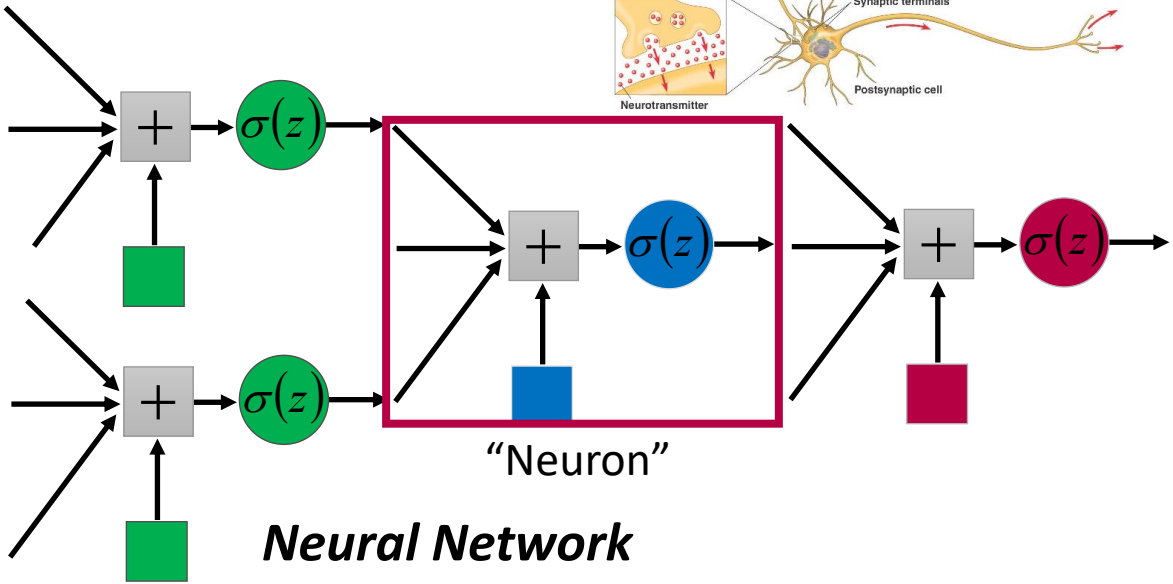
Deep learning trends at Google. Source: SIGMOD 2016/Jeff Dean

Three Steps for Deep Learning: Part I



Neural Network: Part I





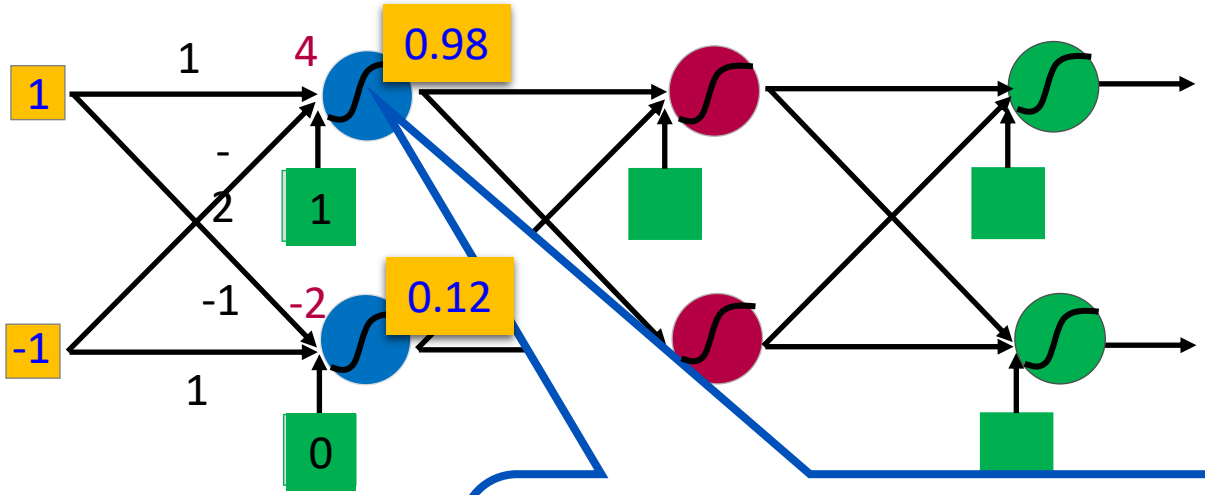
"Neuron"

Neural Network

Different connection leads to different network structures

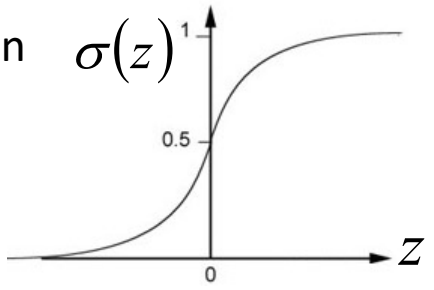
Network parameter θ: all the weights and biases in the "neurons"

Fully Connect Feedforward Network



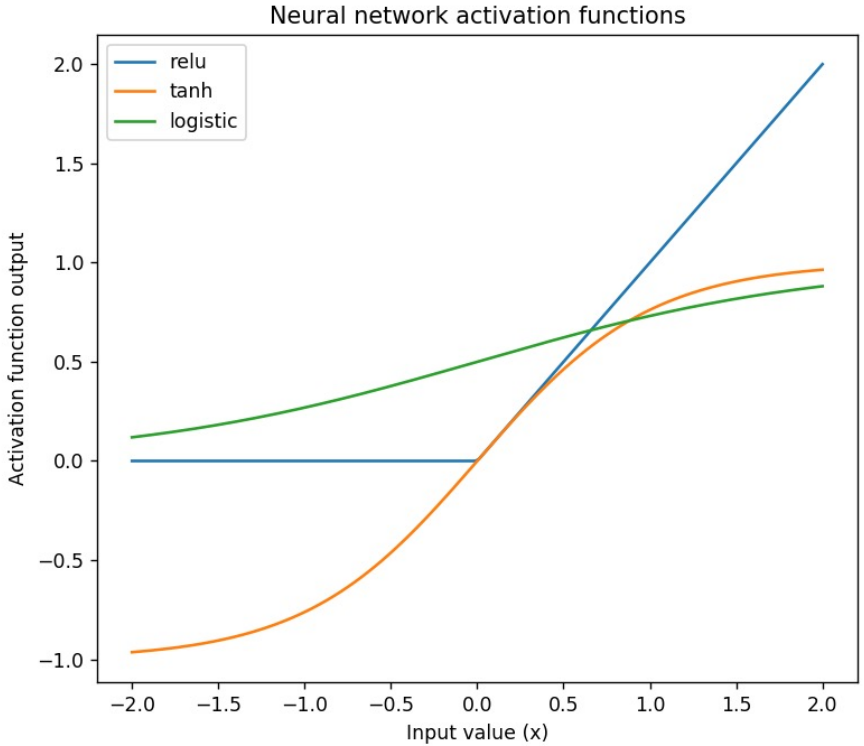
Sigmoid Function

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

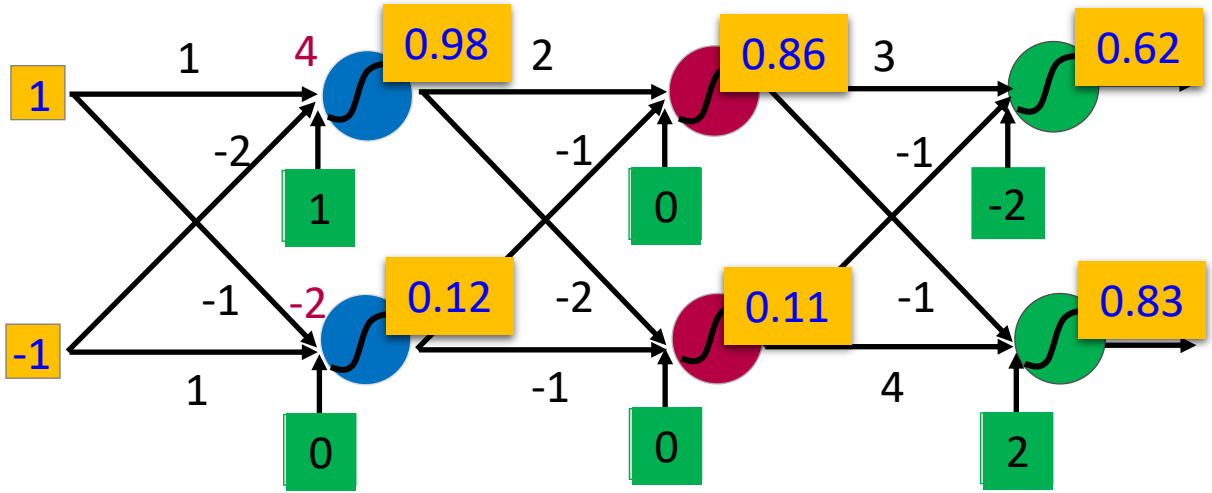


Activation

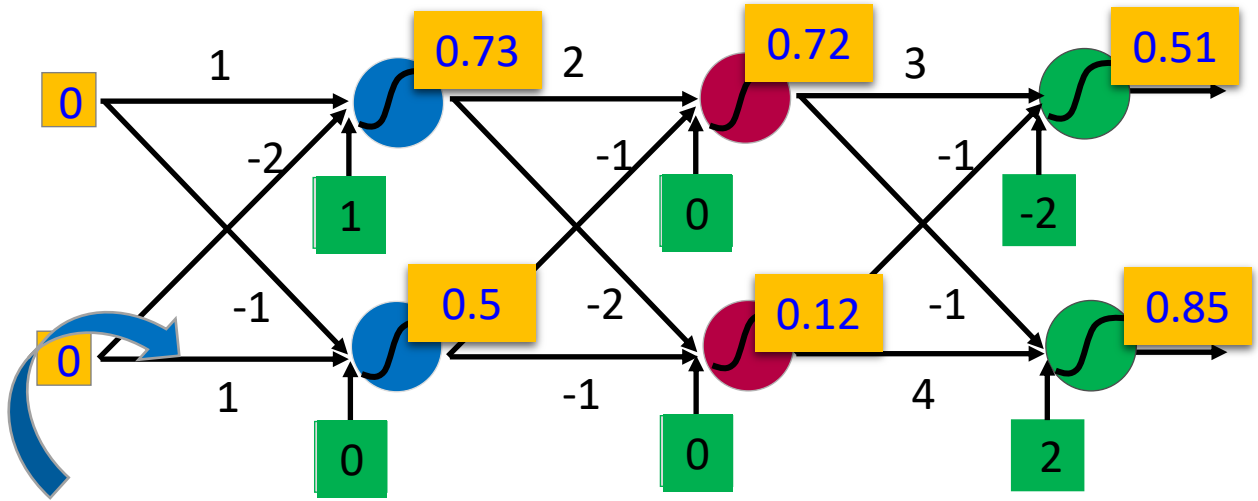
Source:
https://en.wikipedia.org/wiki/Activation_function



Fully Connect Feedforward Network: Part I



Fully Connect Feedforward Network: Part II



This is a function.

Input vector, output vector

$$f\left(\begin{bmatrix}1\\-1\end{bmatrix}\right)=\begin{bmatrix}0.62\\0.83\end{bmatrix} \quad f\left(\begin{bmatrix}0\\0\end{bmatrix}\right)=\begin{bmatrix}0.51\\0.85\end{bmatrix}$$

Given network structure, define ***a function set***

Fully Connect Feedforward Network: Part III