

Introduction to deep learning 1/?

Esten Høyland Leonardsen

02.11.22

UiO:Life Science, University of Oslo



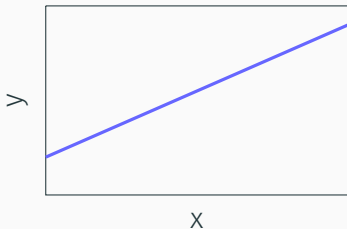
1. Building an artificial neural network (ANN)
2. Training the ANN
3. Transformation to a Convolutional Neural Network (CNN)

Building a neural network: Linear regression

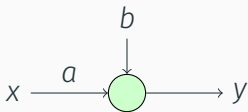
$$y = ax + b$$

Building a neural network: Linear regression

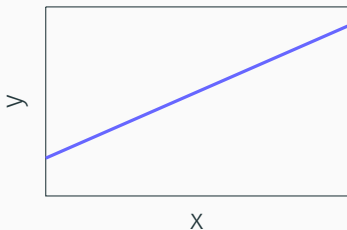
$$y = ax + b$$



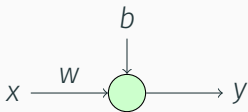
Building a neural network: Linear regression



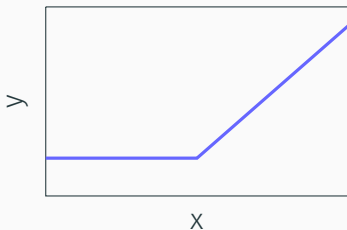
$$y = ax + b$$



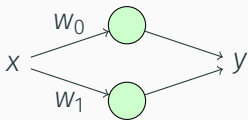
Building a neural network: Artificial neuron



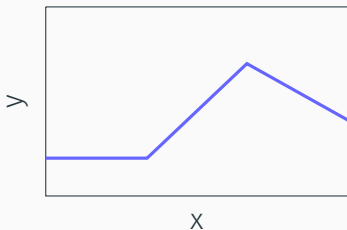
$$y = \max(0, wx + b)$$



Building a neural network: Artificial neural network (ANN)



$$y = \max(0, w_0x + b_0) + \max(0, w_1x + b_1)$$



Building a neural network: Universal approximation theorem

"Any relationship that can be described with a polynomial function can be approximated by a neural network with a single hidden layer."

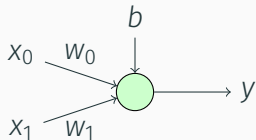
- Some guy in the 80s, probably

Building a neural network: Universal approximation theorem

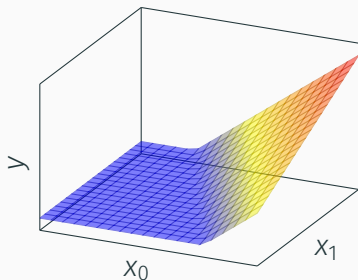
"Any relationship that can be described with a polynomial function can be approximated by a neural network with a single hidden layer."

- Some guy in the 80s, probably

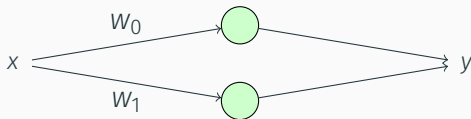
Building a neural network: Increasing dimensionality



$$y = \max(0, w_0x_0 + w_1x_1 + b)$$

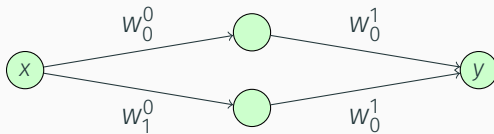


Building a neural network



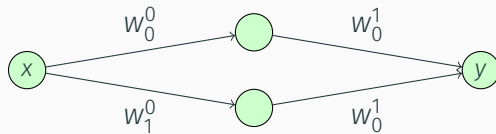
$$y = \max(0, w_0x + b_0) + \max(0, w_1x + b_1)$$

Building a neural network



$$y = \max(0, w_{0,0}^1 * \max(0, w_{0,0}^0 * x + b_{0,0}) + w_{1,0}^1 * \max(0, w_{0,1}^0 * x + b_{1,0}) + b_1)$$

Building a neural network



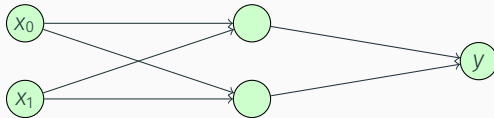
Input

Hidden

Output

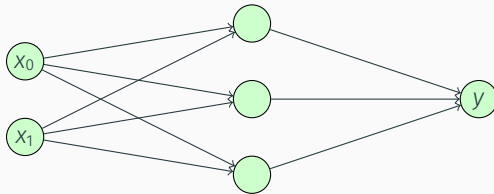
$$y = \max(0, w_{0,0}^1 * \max(0, w_{0,0}^0 * x + b_{0,0}) + w_{1,0}^1 * \max(0, w_{0,1}^0 * x + b_{1,0}) + b_1)$$

Building a neural network



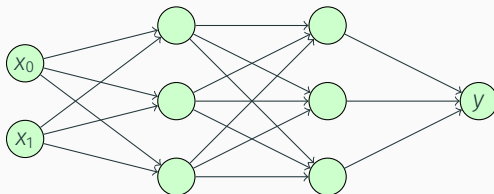
$$y = \max(0, w_{0,0}^1 * \max(0, w_{0,0}^0 * x_0 + w_{1,0}^0 * x_1 + b_{0,0}) + w_{1,0}^1 * \max(0, w_{0,1}^0 * x_0 + w_{1,1}^0 * x_1 + b_{0,1}) + b_1)$$

Building a neural network



$$y = \max(0, w_{0,0}^1 * \max(0, w_{0,0}^0 * x_0 + w_{1,0}^0 * x_1 + b_{0,0}) + w_{1,0}^1 * \max(0, w_{0,1}^0 * x_0 + w_{1,1}^0 * x_1 + b_{0,1}) + w_{2,0}^1 * \max(0, w_{0,2}^0 * x_0 + w_{1,2}^0 * x_1 + b_{0,2}) + b_1)$$

Building a neural network



$$\begin{aligned}
 y = & \max(0, w_{0,0}^2 * \max(0, w_{0,0}^1 * \max(0, w_{0,0}^0 * x_0 + w_{1,0}^0 * x_1 + b_{0,0}) + \\
 & w_{1,0}^1 * \max(0, w_{0,1}^0 * x_0 + w_{1,1}^+ * w_1 + b_{0,1}) + \\
 & w_{2,0}^1 * \max(0, w_{0,2}^0 * x_0 + w_{1,2}^+ * w_1 + b_{0,2}) + \\
 & b_{1,0}) + \\
 & w_{1,0}^2 * \max(0, w_{0,1}^1 * \max(0, w_{0,0}^0 * x_0 + w_{1,0}^0 * x_1 + b_{0,0}) + \\
 & w_{1,1}^1 * \max(0, w_{0,1}^0 * x_0 + w_{1,1}^+ * w_1 + b_{0,1}) + \\
 & w_{2,1}^1 * \max(0, w_{0,2}^0 * x_0 + w_{1,2}^+ * w_1 + b_{0,2}) + \\
 & b_{1,1}) + \\
 & w_{2,0}^2 * \max(0, w_{0,2}^1 * \max(0, w_{0,0}^0 * x_0 + w_{1,0}^0 * x_1 + b_{0,0}) + \\
 & w_{1,2}^1 * \max(0, w_{0,1}^0 * x_0 + w_{1,1}^+ * w_1 + b_{0,1}) + \\
 & w_{2,2}^1 * \max(0, w_{0,2}^0 * x_0 + w_{1,2}^+ * w_1 + b_{0,2}) + \\
 & b_{1,2}) + \\
 & b_2)
 \end{aligned}$$

Building a neural network: Summary

- Artificial neurons are (linear) weighted sums wrapped in non-linear activation functions
- Multiple artificial neurons stacked together in a layerwise fashion comprise an artificial neural network
- Artificial neural networks allow us to model arbitrarily complex relationships between inputs and outputs