Deep learning basics

Grégory Châtel

Disaitek
Intel Software Innovator

@rodgzilla
github.com/rodgzilla

June 4th, 2018

Machine learning is a subfield of artificial intelligence.

Intuitively We want to *learn from* and *make predictions on* data.

Technically We want to build a model that approximate well (*e.g.* minimize a loss function) an unknown function.

Application examples

Supervised learning

- Regression

Polynomial	$(x, y, z) \to f(x, y, z)$
House price	(surface, nb rooms, city) $\to$ price

- Classification

Image classification	pixel values $\to$ cat or dog
Text classification	list of words $\to$ spam or valid email

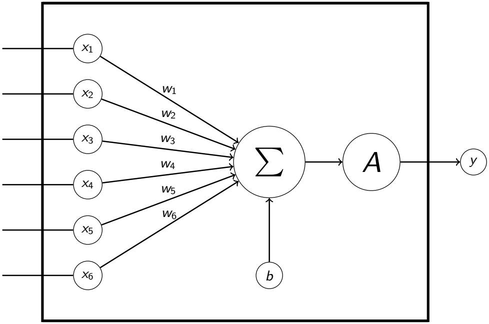
Deep learning

Deep learning is a subfield of machine learning in which we use artificial neural networks to make predictions.

An artificial neural networks is a computation model *loosely* based on the human brain. It aims to mimic electric signals travelling through neurons in order to make computations.

Artificial neural network

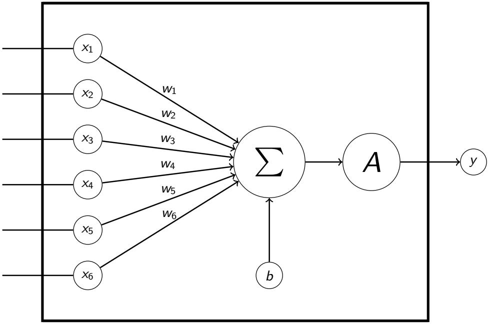
Neuron



$$A(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{otherwise} \end{cases}$$

Artificial neural network

Neuron

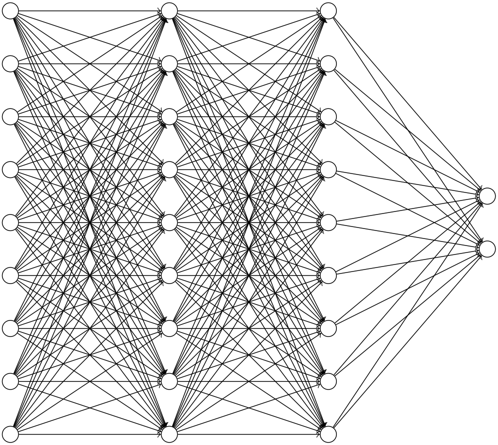


$$A(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{otherwise} \end{cases}$$

$$y = A(w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6 + b)$$

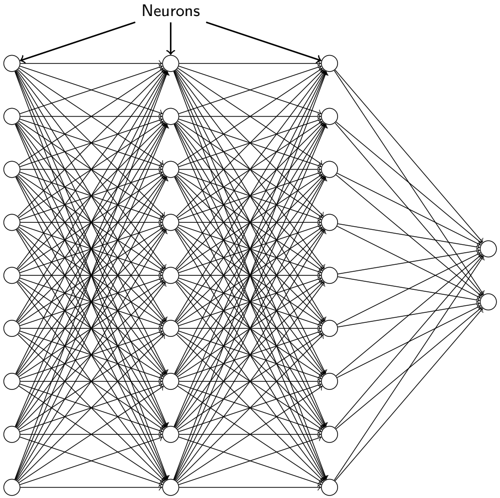
Artificial neural network

Network



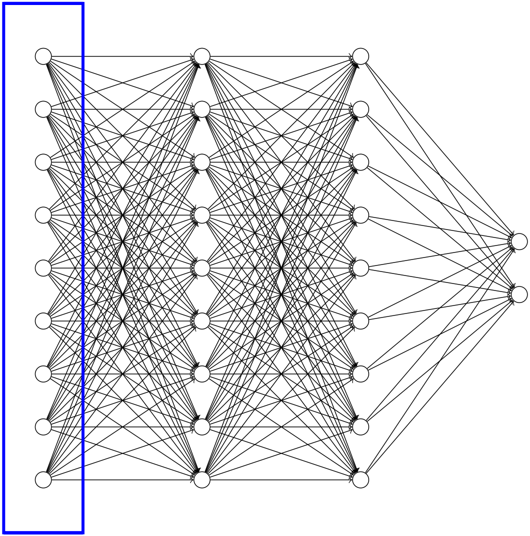
Artificial neural network

Network



Artificial neural network

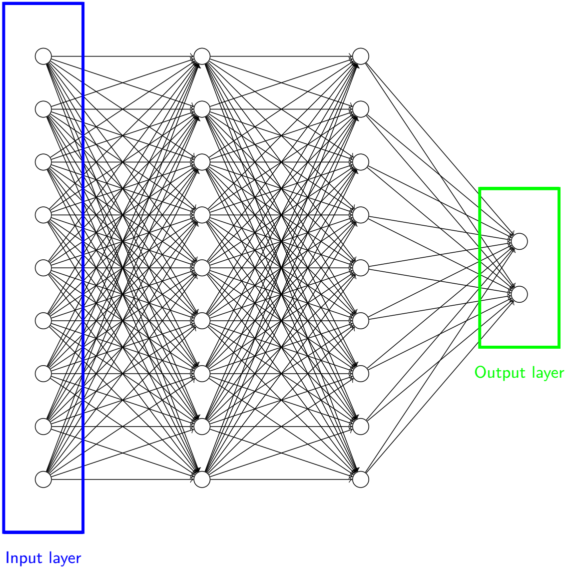
Network



Input layer

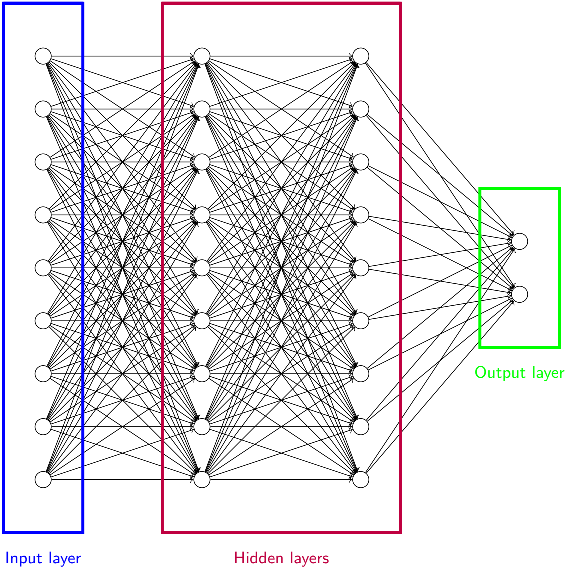
Artificial neural network

Network



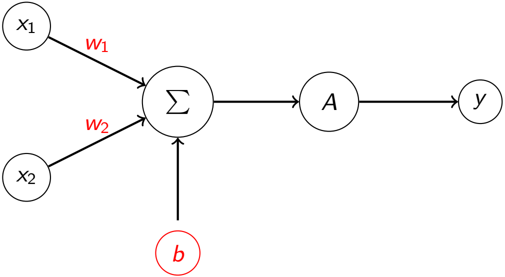
Artificial neural network

Network



Computation example

Binary AND gate

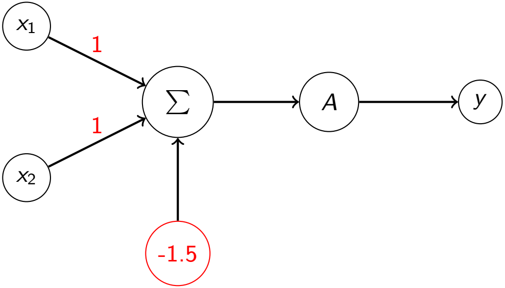


We want to set w_1, w_2 and b such that:

$$A(w_1x_1 + w_2x_2 + b) = x_1 \text{ AND } x_2$$

Computation example

Binary AND gate

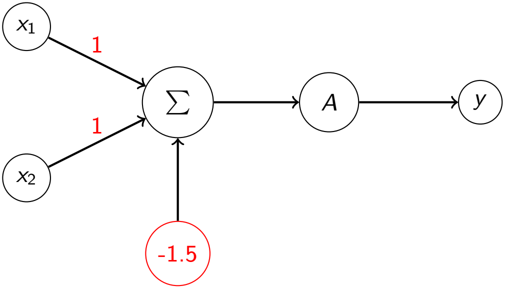


We want to set w_1, w_2 and b such that:

$$A(w_1 x_1 + w_2 x_2 + b) = x_1 \text{ AND } x_2$$

Computation example

Binary AND gate



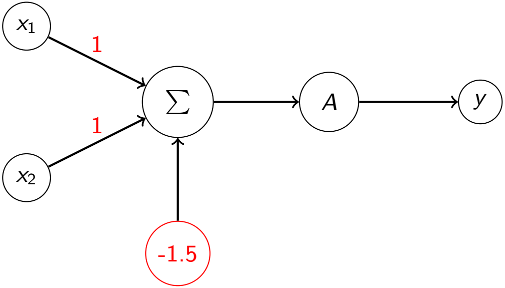
We want to set w_1, w_2 and b such that:

$$A(w_1 x_1 + w_2 x_2 + b) = x_1 \text{ AND } x_2$$

$$x_0 = 0, x_1 = 1. \quad y = A(0 + 1 - 1.5) = A(-0.5) = 0$$

Computation example

Binary AND gate



We want to set w_1, w_2 and b such that:

$$A(w_1 x_1 + w_2 x_2 + b) = x_1 \text{ AND } x_2$$

$$x_0 = 0, x_1 = 1. \quad y = A(0 + 1 - 1.5) = A(-0.5) = 0$$

$$x_0 = 1, x_1 = 1. \quad y = A(1 + 1 - 1.5) = A(0.5) = 1$$

Model complexity

One way to measure the complexity of a neural network is its number of parameters.

Model complexity

One way to measure the complexity of a neural network is its number of parameters.

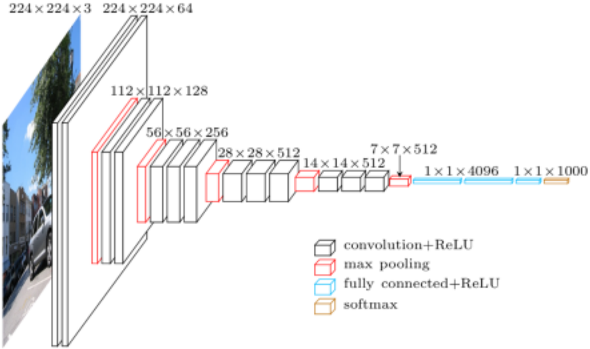
- AND network: 3 parameters

Model complexity

One way to measure the complexity of a neural network is its number of parameters.

- AND network: 3 parameters
- dogs vs cats pictures (VGG16 network): 138,357,544 parameters

Convolutional neural network
VGG network (2014)

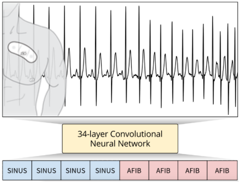


Classify natural images into 1000 categories. (92.7% top-5 accuracy).

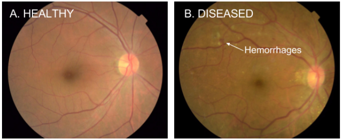
Image from `https://www.cs.toronto.edu/~frossard/post/vgg16/`

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Deep learning and medicine
Arrhythmia and diabetic retinopathy

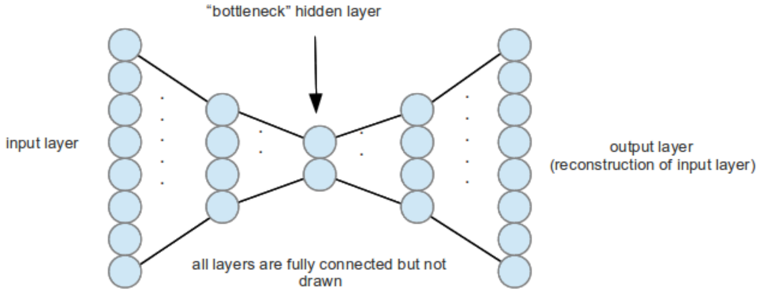


Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C., & Ng, A. Y. (2017). Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv preprint arXiv:1707.01836.



Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Kim, R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22), 2402-2410.

Architecture examples
Deep autoencoder



If we cut this autoencoder at the bottleneck, we get two parts: an encoder and a decoder. The encoder is an encoder highly specific to the content the network has been trained with.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. science, 313(5786), 504-507.

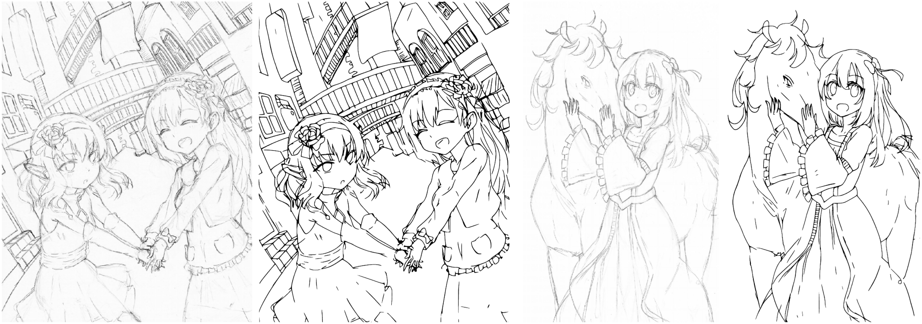
Architecture examples

Deep autoencoder: Compression

If we force the output image to be realistic, we lose *semantic information* rather than resolution.

Santurkar, S., Budden, D., & Shavit, N. (2017). Generative compression. arXiv preprint arXiv:1703.01467.

Architecture examples
Drawing cleanup



Simo-Serra, E., Iizuka, S., Sasaki, K., & Ishikawa, H. (2016). Learning to simplify: fully convolutional networks for rough sketch cleanup. ACM Transactions on Graphics (TOG), 35(4), 121.

Architecture examples
Neural style transfer

Let's take a random image from the internet.

Architecture examples

Neural style transfer

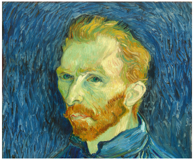
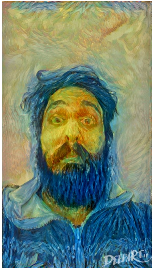
α content() $+ \beta$ style() $=$ 

Image created using `https://deepart.io`

Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.

Architecture examples

Neural style transfer

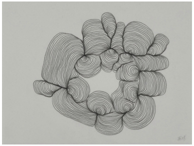
α content() $+ \beta$ style() $=$ 

Image created using `https://deepart.io`

Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.

Architecture examples

Neural style transfer

α content() $+ \beta$ style() $=$ 

Image created using `https://deepart.io`

Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.

Architecture examples

Neural style transfer

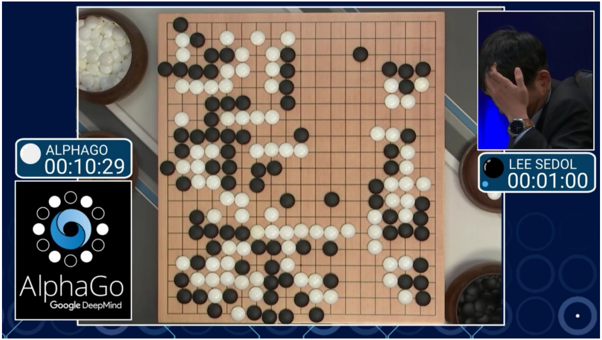
α content() $+ \beta$ style() = 

Image created using `https://deepart.io`

Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.

Reinforcement learning

DeepMind's AlphaGo



In 2016, DeepMind showcased an artificial intelligence for the game of Go that managed to beat Lee Sedol, one of the strongest player in the world.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. nature, 529(7587), 484-489.

Architecture examples

Sequence to sequence network: Neural Machine Translation

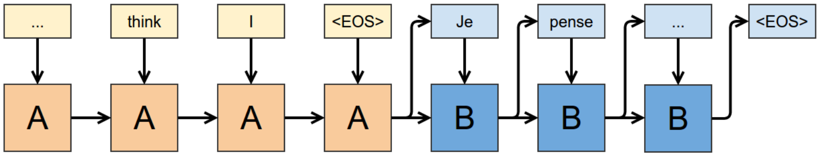


Image from `https://colah.github.io`

Google, September 2016: "The Google Translate mobile and web apps are now using GNMT (Google NMT) for 100% of machine translations from Chinese to English—about 18 million translations per day."

Architecture examples

Generative adversarial network

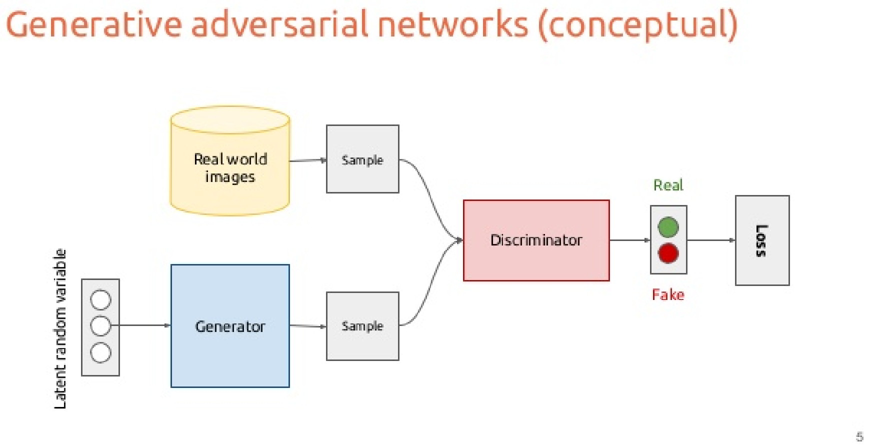


Image from `http://wiki.tum.de/`

Architecture examples

Generative adversarial network: Celebrity faces generation

Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196.

Architecture examples

Image to sequence: automatic captioning

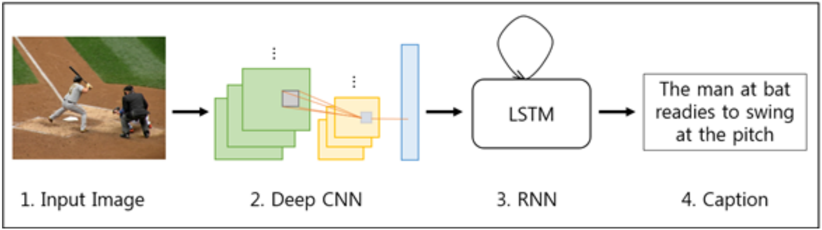
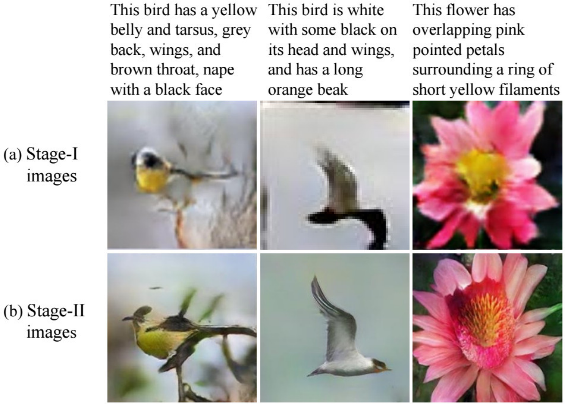


Image from `http://brain.kaist.ac.kr/`

Architecture examples

Generative adversarial network: text to image



Zhang, H., Xu, T., Li, H., Zhang, S., Huang, X., Wang, X., & Metaxas, D. (2017, October). Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. In IEEE Int. Conf. Comput. Vision (ICCV) (pp. 5907-5915).