Brief History of Recent Deep Learning

Formation IA Écologie

Paul Tresson, Maximilien Servajean, Benjamin Bourel

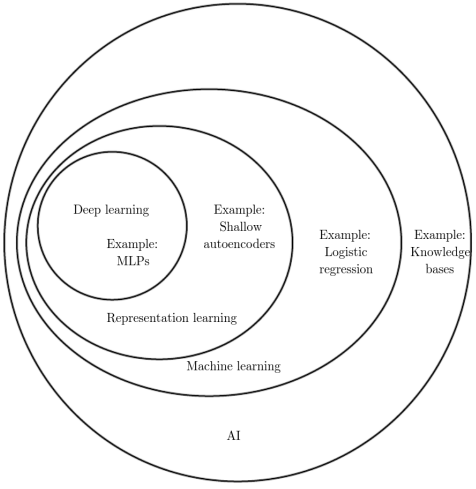
May 9, 2025

UMR AMAP

Institut de Recherche pour le Développement FRANCE

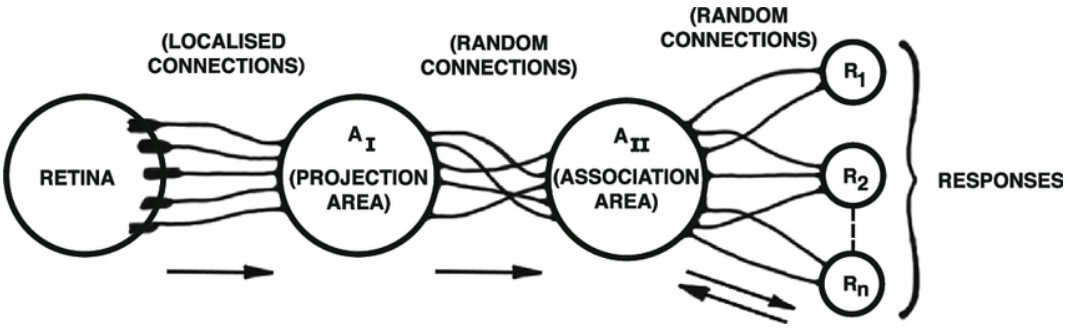
AMAPlab

Reminder



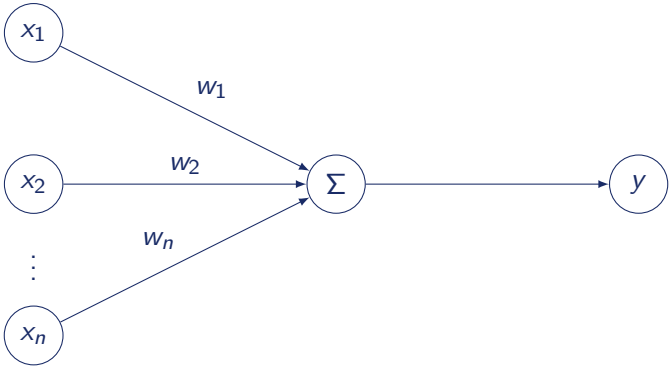
Different types of Artificial Intelligence. From Goodfellow et al., 2016

First Neural Network Model : The Perceptron



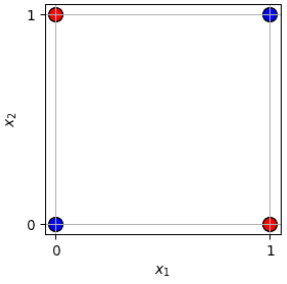
First probabilistic model of a neuron proposed by Rosenblatt, 1958

First Neural Network Model : The Perceptron



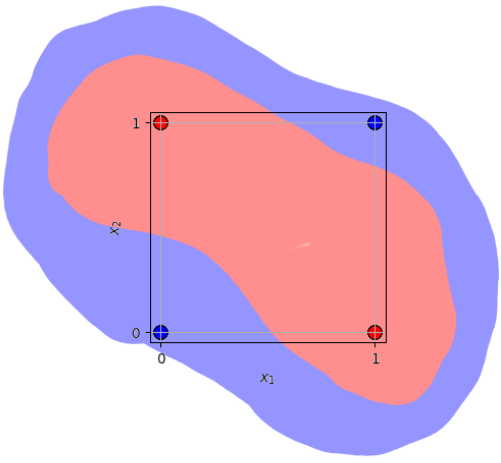
First Roadblock : XOR problem

x_1	x_2	y
0	0	0
1	0	1
0	1	1
1	1	0



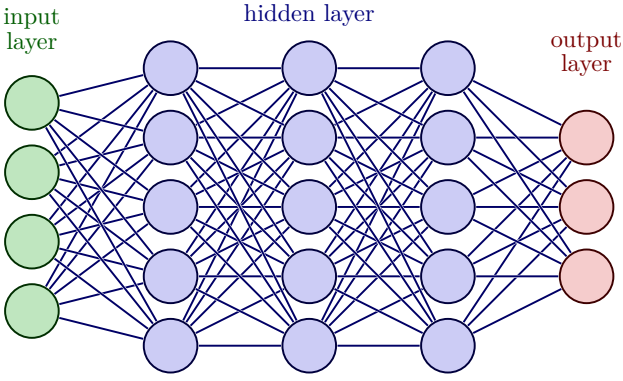
First Roadblock : XOR problem

x_1	x_2	y
0	0	0
1	0	1
0	1	1
1	1	0



1980's : First Deep Neural Networks

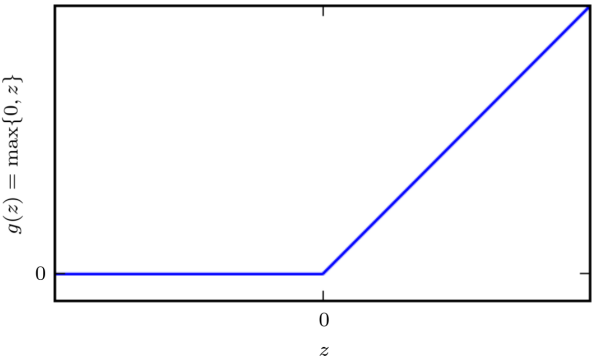
- Multi Layer Perceptron



Multi Layer Perceptron

1980's : First Deep Neural Networks

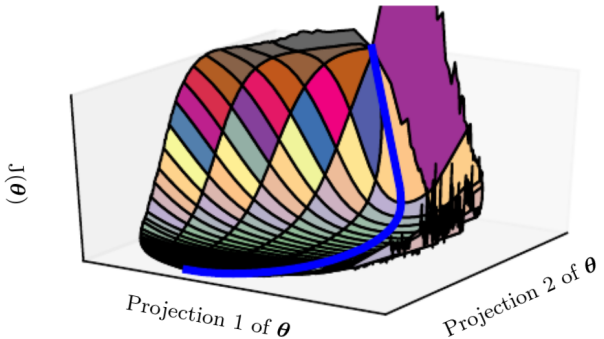
- Multi Layer Perceptron
- Non-linearity (*e.g.* ReLu)



Rectified Linear Unit. Figure from Goodfellow et al., 2016

1980's : First Deep Neural Networks

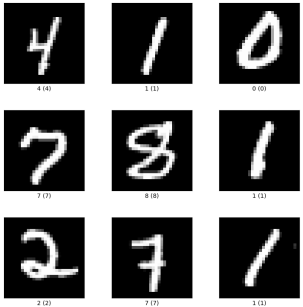
- Multi Layer Perceptron
- Non-linearity (*e.g.* ReLu)
- Gradient Descent and Backpropagation
  Rumelhart, 1986



Gradient descent example. Figure from Goodfellow et al., 2016

1990's : First Successes

- MNIST
  LeCun, 1998

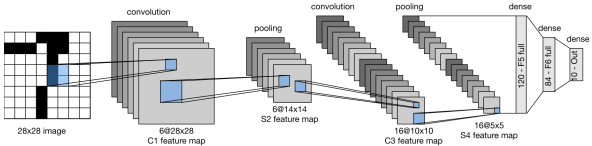


Images from MNIST

Modified National Institute of Standards and Technology

1990's : First Successes

- MNIST
  LeCun, 1998
- LeNet
  LeCun et al., 1989b

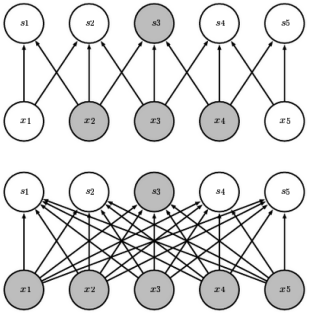


LeNet5.
Figure from Zhang et al. - `https://github.com/d2l-ai/d2l-en`

Modified National Institute of Standards and Technology

1990's : First Successes

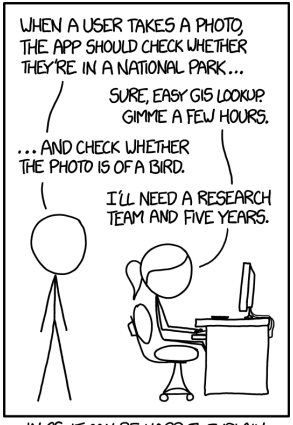
- MNIST
  LeCun, 1998
- LeNet
  LeCun et al., 1989b
- Convolutional Neural Network (CNN)
  LeCun et al., 1989a



sparse connectivity with CNNs.
Figure from Goodfellow et al., 2016

Modified National Institute of Standards and Technology

Good proof of concept but too costly in computing power and datasets



XKCD comic from 2014

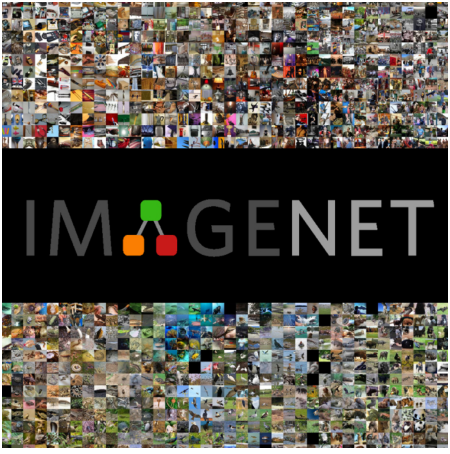
Good proof of concept but too costly in computing power and datasets ...until

Good proof of concept but too costly in computing power and datasets ...until

Good proof of concept but too costly in computing power and datasets ...until

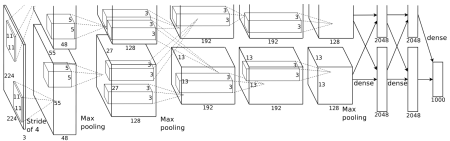
2010's : CNN revolution

- ImageNet
  Deng et al., 2009



2010's : CNN revolution

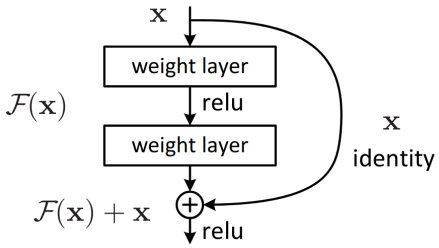
- ImageNet
  Deng et al., 2009
- AlexNet
  Krizhevsky et al., 2012



AlexNet. Figure from Krizhevsky et al., 2012

2010's : CNN revolution

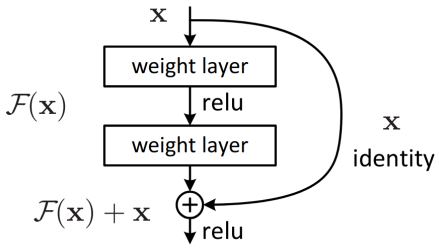
- ImageNet
  Deng et al., 2009
- AlexNet
  Krizhevsky et al., 2012
- ResNet
  He et al., 2016



A residual connection. Figure from He et al., 2016

2010's : CNN revolution

- ImageNet
  Deng et al., 2009
- AlexNet
  Krizhevsky et al., 2012
- ResNet
  He et al., 2016

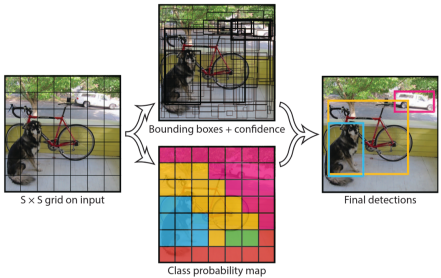


A residual connection. Figure from He et al., 2016

His 2016 paper *Deep Residual Learning for Image Recognition* is the most cited research paper in 5 years according to Google Scholar's reports in 2020 and 2021.[7][8]

2010's : CNN revolution

- ImageNet
  Deng et al., 2009
- AlexNet
  Krizhevsky et al., 2012
- ResNet
  He et al., 2016
- YOLO, mask-RCNN...
  He et al., 2017; Redmon et al., 2016



YOLO. Figure from Redmon et al., 2016

2010's : NLP

- IMBD

  Maas et al., 2011

If you like adult comedy cartoons, like South Park, then this is nearly a similar format about the small adventures of three teenage girls at Bromwell High. Keisha, Natella and Latrina have given exploding sweets and behaved like bitches, I think Keisha is a good leader. There are also small stories going on with the teachers of the school. There's the idiotic principal, Mr. Bip, the nervous Maths teacher and many others. The cast is also fantastic, Lenny Henry's Gina Yashere, EastEnders Chrissie Watts, Tracy-Ann Oberman, Smack The Pony's Doon Mackichan, Dead Ringers' Mark Perry and Blunder's Nina Conti. I didn't know this came from Canada, but it is very good. Very good!

Extract from IMDB database

2010's : NLP

- IMBD
  Maas et al., 2011
- LSTM
  Hochreiter and Schmidhuber, 1997

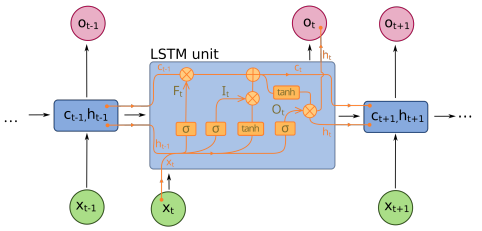


Figure by fdeloche
https://commons.wikimedia.org/w/index.php?curid=60149410

2010's : NLP

- IMBD
  Maas et al., 2011
- LSTM
  Hochreiter and Schmidhuber, 1997
- RNN's
  e.g. Sutskever et al., 2014

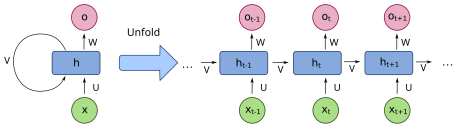
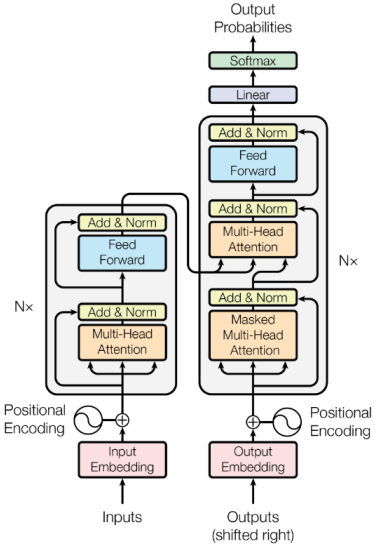


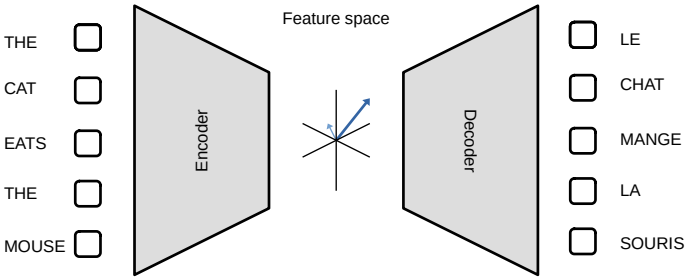
Figure by fdeloche
https://commons.wikimedia.org/w/index.php?curid=60109157

2020's : Transformers ...

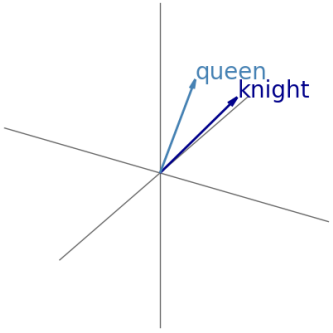
- Attention is all you need
  Vaswani et al., 2017



Sidestep : Transformers

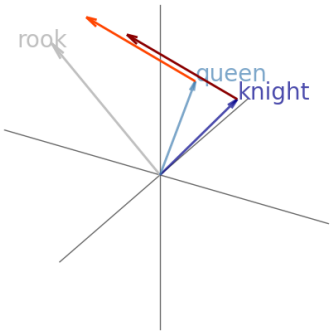


Sidestep : Transformers



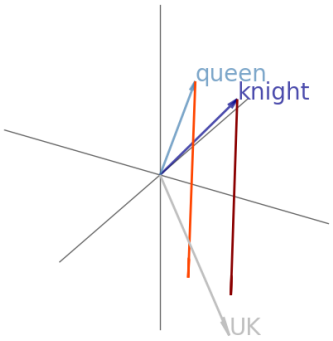
The knight saves the queen

Sidestep : Transformers



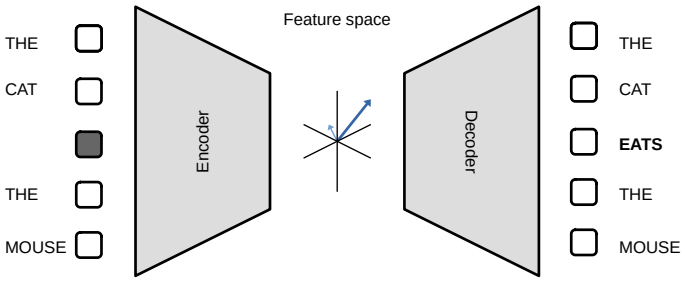
Queen takes knight , checkmate !

Sidestep : Transformers



He was knighted by Queen Elizabeth II

Sidestep : Transformers



2020's : Transformers ... and Self Supervised Learning

- Attention is all you need
  Vaswani et al., 2017
- LLMs (GPT, BERT...)
  Devlin et al., 2019

2020's : Transformers ... and Self Supervised Learning

- Vision Transformer
  Dosovitskiy et al., 2020

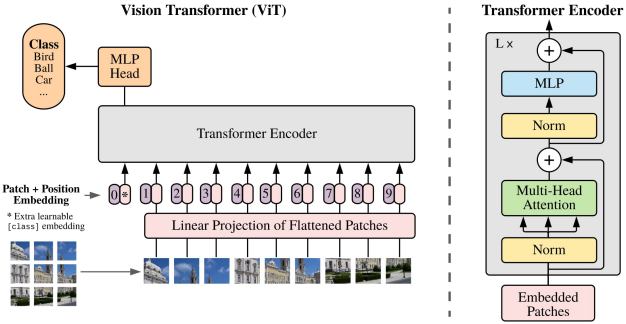


Figure from Dosovitskiy et al., 2020

2020's : Transformers ... and Self Supervised Learning

- Vision Transformer
  Dosovitskiy et al., 2020
- Masked Auto Encoder
  He et al., 2022

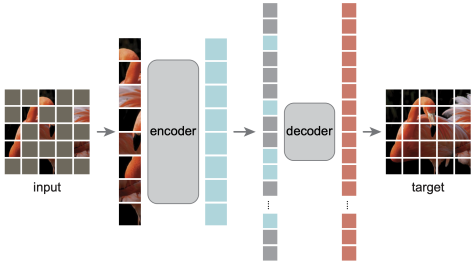


Figure from He et al., 2022

2020's : Transformers ... and Self Supervised Learning

- Vision Transformer
  Dosovitskiy et al., 2020
- Masked Auto Encoder
  He et al., 2022
- DINO
  Caron et al., 2021

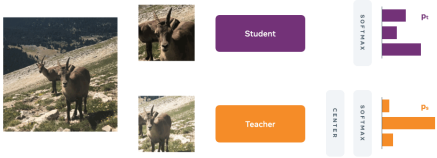


Figure from Caron et al., 2021

A lot left out !

Finetuning

A lot left out !

Finetuning

Games

A lot left out !

Reinforcement learning

Finetuning

Games

A lot left out !

Reinforcement learning

Finetuning

Games

Time series

A lot left out !

Reinforcement learning

Finetuning

Games

Time series

Speech recognition

A lot left out !

Reinforcement learning

Finetuning

Games

Image generation

Time series

Speech recognition

A lot left out !

Reinforcement learning

Finetuning

Games

Image generation

Diffusion models

Time series

Speech recognition

A lot left out !

Reinforcement learning

Finetuning

Games

Image generation

Diffusion models

Time series

Speech recognition

...

Usefull ressources

State of the art

- Huggingface
- PapersWithCode

Getting started

- Pytorch
- Keras

Understanding papers

- YannicKilcher
- AI coffe break

Understanding visually

- 3blue1brown
- deepia

Thanks for you attention !

Let's practice !

References i

Caron, Mathilde et al. (2021). **"Emerging properties in self-supervised vision transformers".** In: *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 9650–9660.

Deng, Jia et al. (2009). **"Imagenet: A large-scale hierarchical image database".** In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee, pp. 248–255.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2019). **"Bert: Pre-training of deep bidirectional transformers for language understanding".** In: *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186.

References ii

Dosovitskiy, Alexey et al. (2020). **"An image is worth 16x16 words: Transformers for image recognition at scale".** In: *arXiv preprint arXiv:2010.11929*.

Goodfellow, Ian, Yoshua Bengio, Aaron Courville, and Yoshua Bengio (2016). ***Deep learning*.** Vol. 1. 2. MIT press Cambridge.

He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick (2017). **"Mask r-cnn".** In: *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2016). **"Deep residual learning for image recognition".** In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.

He, Kaiming et al. (2022). **"Masked autoencoders are scalable vision learners".** In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 16000–16009.

References iii

Hochreiter, Sepp and Jürgen Schmidhuber (1997). **"Long short-term memory".** In: *Neural computation* 9.8, pp. 1735–1780.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton (2012). **"Imagenet classification with deep convolutional neural networks".** In: *Advances in neural information processing systems* 25.

LeCun, Yann et al. (1989a). **"Generalization and network design strategies".** In: *Connectionism in perspective* 19.143-155, p. 18.

LeCun, Yann (1998). **"The MNIST database of handwritten digits".** In: *http://yann. lecun. com/exdb/mnist/*.

LeCun, Yann et al. (1989b). **"Backpropagation applied to handwritten zip code recognition".** In: *Neural computation* 1.4, pp. 541–551.

References iv

Maas, Andrew et al. (2011). **"Learning word vectors for sentiment analysis".** In: *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*, pp. 142–150.

Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi (2016). **"You only look once: Unified, real-time object detection".** In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788.

Rosenblatt, Frank (1958). **"The perceptron: a probabilistic model for information storage and organization in the brain.".** In: *Psychological review* 65.6, p. 386.

Rumelhart, DE (1986). **"Learning representations by error propagation".** In: *Parallel distributed processing* 1, pp. 318–362.

References v

Sutskever, Ilya, Oriol Vinyals, and Quoc V Le (2014). **"Sequence to sequence learning with neural networks".** In: *Advances in neural information processing systems* 27.

Vaswani, Ashish et al. (2017). **"Attention is all you need".** In: *Advances in neural information processing systems* 30.