DECSAI
Departamento de Ciencias de la Computación e I.A.
Universidad de Granada

Universidad de Jaén

CEATIC

V Jornadas Doctorales del programa TIC
Universidad Internacional de Andalucía
Campus Antonio Machado - Baeza

Deep Learning

Fernando Berzal, berzal@acm.org

Deep Learning

- Breve historia de las redes neuronales artificiales.
- Técnicas de deep learning:
  Entrenamiento de redes neuronales artificiales.
- En la práctica: Implementación de sistemas basados en redes neuronales artificiales.
- Aplicaciones de las técnicas de deep learning.
- Limitaciones de las técnicas de deep learning.

Redes neuronales artificiales

Redes neuronales artificiales

El cerebro humano

Inspiración de las redes neuronales artificiales

Las RNA intentan modelar la estructura y funcionamiento de algunas partes del sistema nervioso animal.

Redes neuronales artificiales

¿Por qué estudiar redes neuronales?

- Para comprender cómo funciona realmente el cerebro.
- Para diseñar un modelo de cómputo paralelo inspirado en las neuronas y sus sinapsis [conexiones] adaptativas.
- **Para resolver problemas prácticos utilizando algoritmos de aprendizaje inspirados en el cerebro.**

NOTA: Incluso aunque no sepamos realmente cómo funciona el cerebro, los algoritmos de aprendizaje nos serán muy útiles.

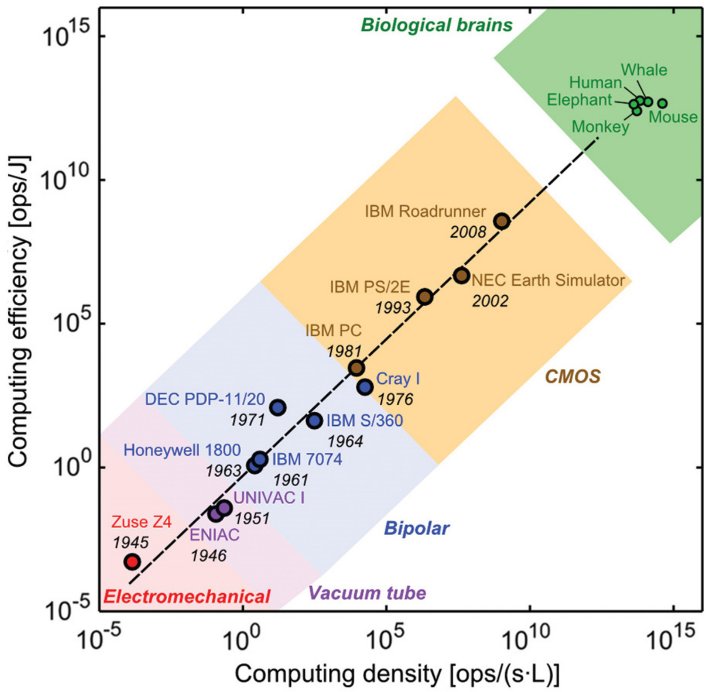
Redes neuronales artificiales

El cerebro humano

Diferencias entre un ordenador y el cerebro humano

Ordenador	Cerebro humano
Computación en serie	Computación en paralelo
Poco robusto	Tolerancia a fallos
Programable	Aprendizaje autónomo
Digital	Analógico
10^9 transistores	**10^{11} neuronas** **$10^{14} \sim 10^{15}$ sinapsis**
Nanosegundos (3.6GHz)	Milisegundos (4~90Hz)
51.2 GB/s	10 spikes/s
210,000,000 m/s	1 ~ 100 m/s
2.3×10^{13} TEPS	**6.4×10^{14} TEPS**

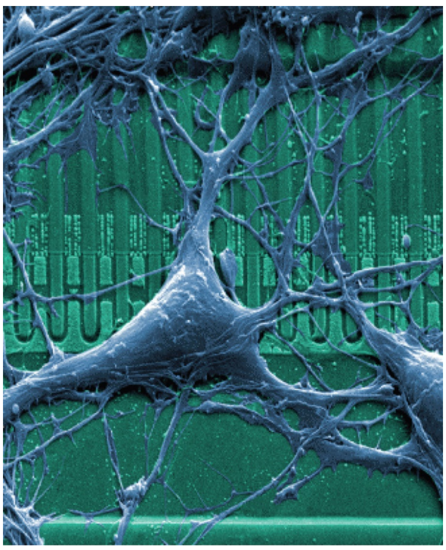
Redes neuronales artificiales



IBM Journal of Research and Development, 2011

Redes neuronales artificiales

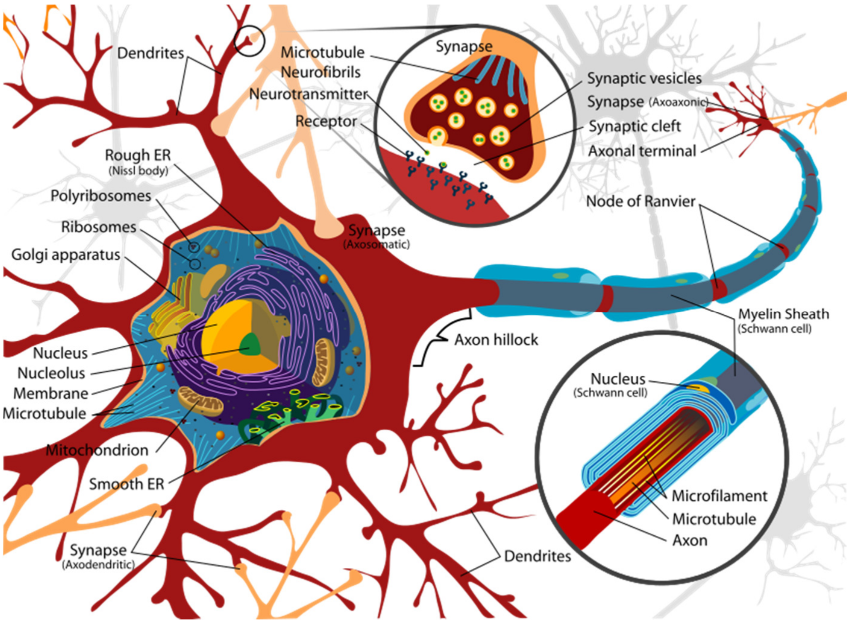
Neuronas



Microfotografía de una neurona “cultivada” sobre una oblea de silicio.
[Peter Fromherz, Max Planck Institute]

Introducción

Neuronas

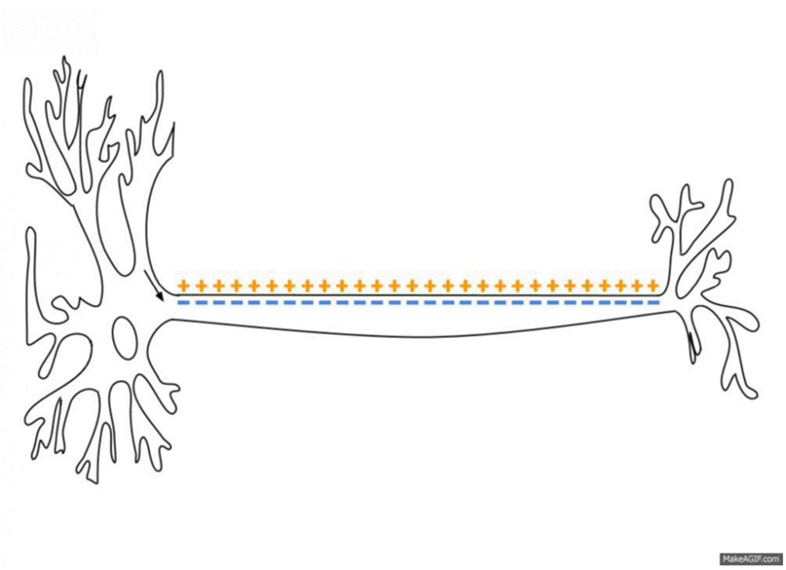


[Wikipedia]

Redes neuronales artificiales

Neuronas

Spike, a.k.a. action potential [potencial de acción]

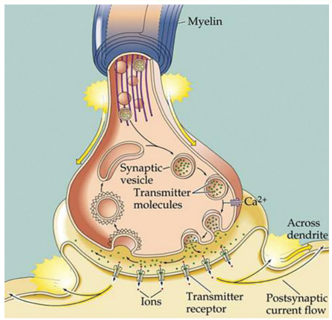


https://en.wikipedia.org/wiki/Action_potential

Redes neuronales artificiales

Neuronas

Sinapsis



Las sinapsis son lentas (en comparación con los transistores de un ordenador), pero…

- Son muy pequeñas y consumen muy poca energía.
- Se adaptan utilizando señales locales.

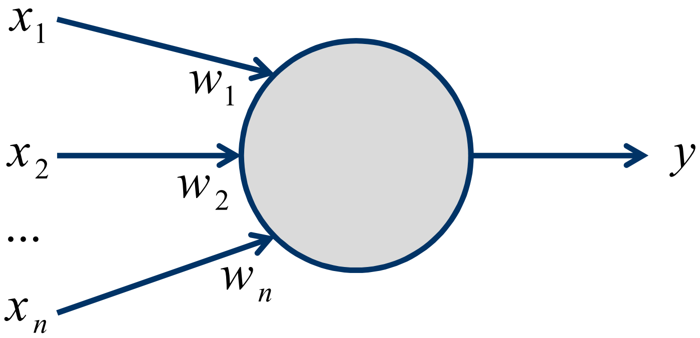
Como tenemos cerca de 10^{11} neuronas y de 10^{14} a 10^{15} sinapsis, muchas sinapsis pueden influir en un "cálculo" en un período de tiempo muy breve:

Ancho de banda muy superior al de un ordenador.

Redes neuronales artificiales

Neuronas

El modelo computacional más simple de una neurona



$$y = \sum_i x_i w_i = x_1 w_1 + x_2 w_2 + \ldots + x_n w_n$$

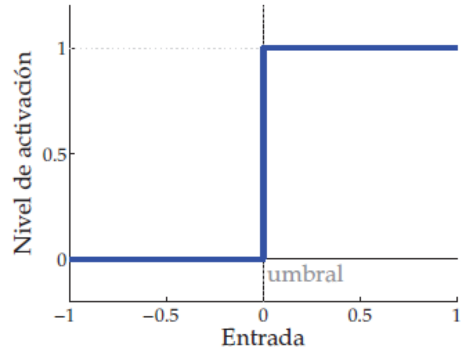
Redes neuronales artificiales

Modelo de neurona de McCulloch & Pitts

Neuronas binarias con umbral

$$z = \sum_i x_i w_i$$

$$y = \begin{cases} 1 & si\ z \geq 0 \\ 0 & en\ otro\ caso \end{cases}$$



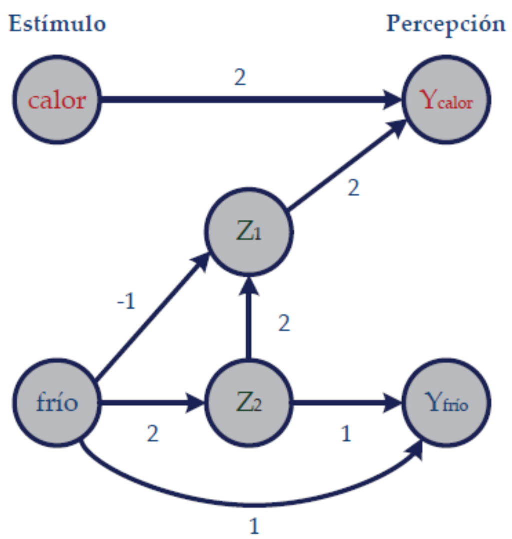
1943

Warren McCulloch & Walter Pitts:
"A logical calculus of the ideas immanent in nervous activity."
Bulletin of Mathematical Biophysics, 5:115-133.

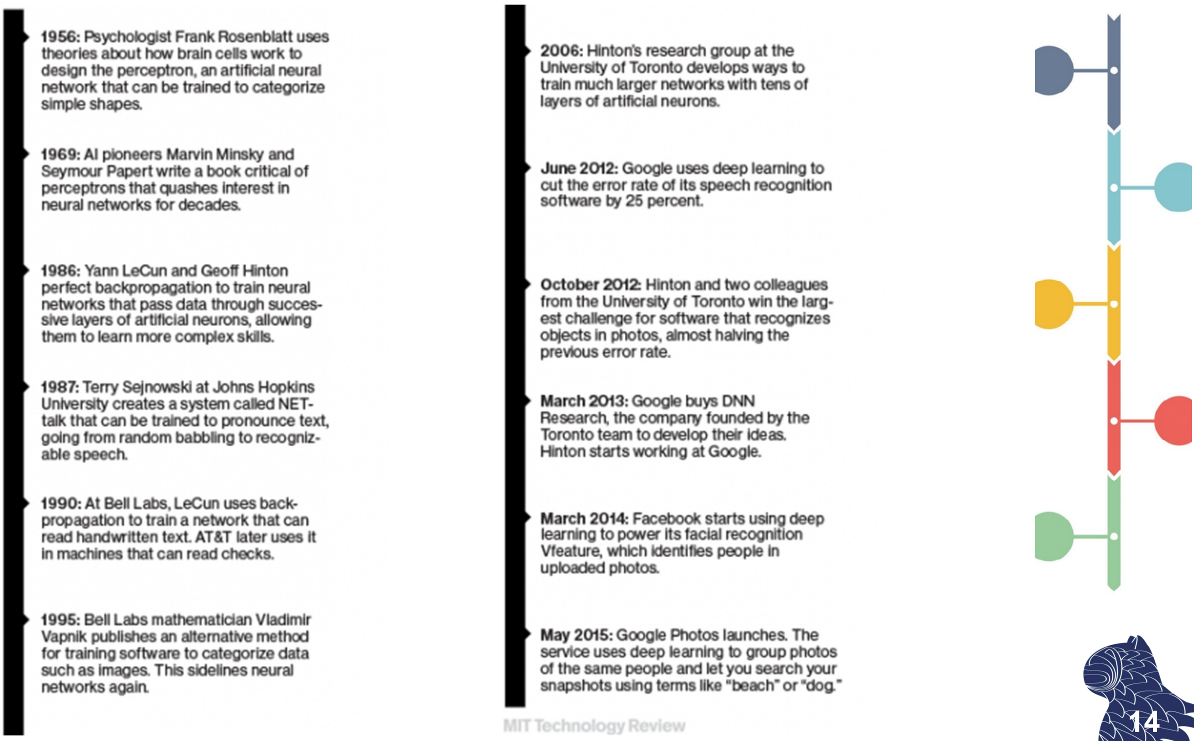
Redes neuronales artificiales

Modelo de neurona de McCulloch & Pitts

Ejemplo: Percepción fisiológica del calor y del frío



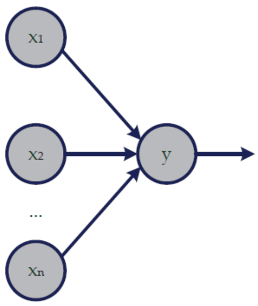
Historia

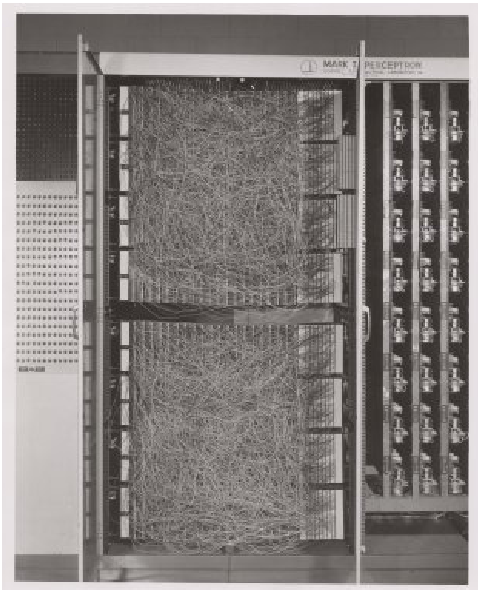


MIT Technology Review: "Teaching Machines to Understand Us", August 2015

Historia de las redes neuronales artificiales
El perceptrón

Primer algoritmo
de aprendizaje supervisado





Mark I Perceptron Machine
Primera implementación…

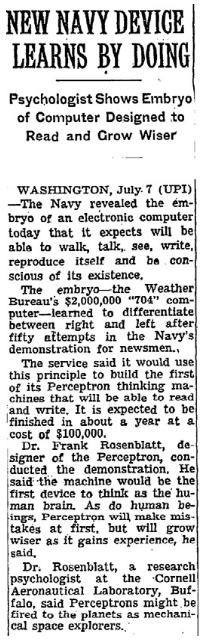
1957

Frank Rosenblatt:
"The Perceptron - A perceiving and recognizing automaton". Report 85-460-1,
Cornell Aeronautical Laboratory, 1957.

El perceptrón

En la prensa…

New York Times



NEW NAVY DEVICE LEARNS BY DOING

Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser

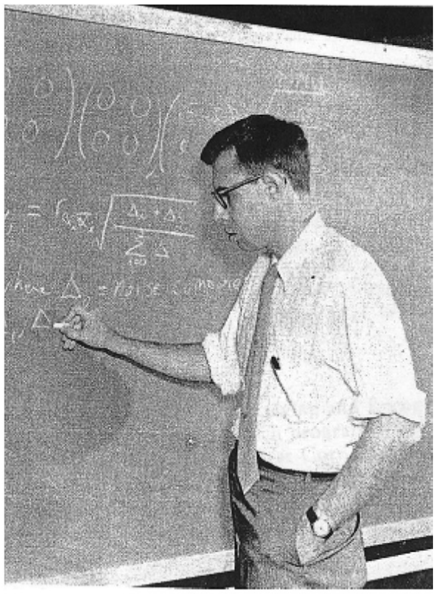
WASHINGTON, July 7 (UPI) —The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.

The embryo—the Weather Bureau's $2,000,000 "704" computer—learned to differentiate between right and left after fifty attempts in the Navy's demonstration for newsmen.

The service said it would use this principle to build the first of its Perceptron thinking machines that will be able to read and write. It is expected to be finished in about a year at a cost of $100,000.

Dr. Frank Rosenblatt, designer of the Perceptron, conducted the demonstration. He said the machine would be the first device to think as the human brain. As do human beings, Perceptron will make mistakes at first, but will grow wiser as it gains experience, he said.

Dr. Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said Perceptrons might be fired to the planets as mechanical space explorers.

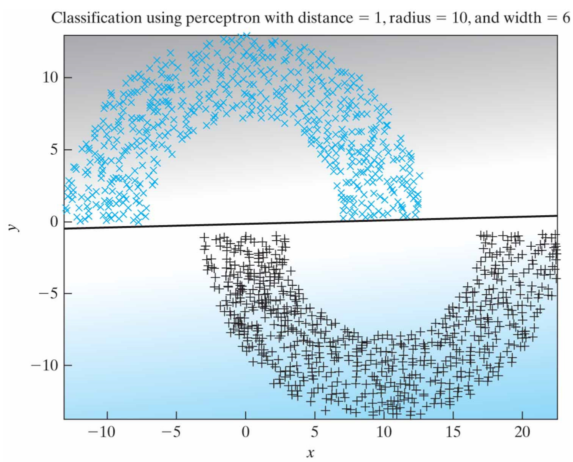


1958

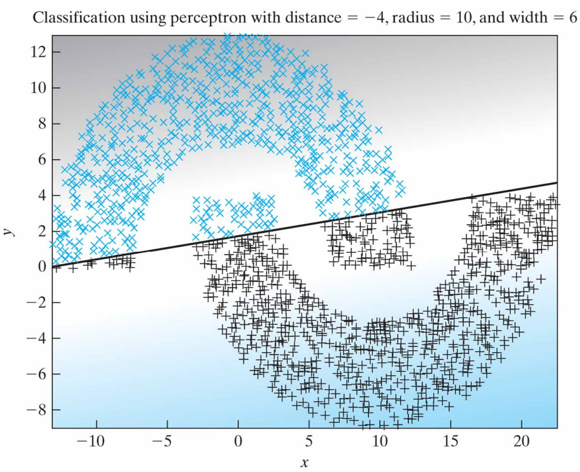
El perceptrón

En realidad…

Un clasificador lineal



(b) testing result



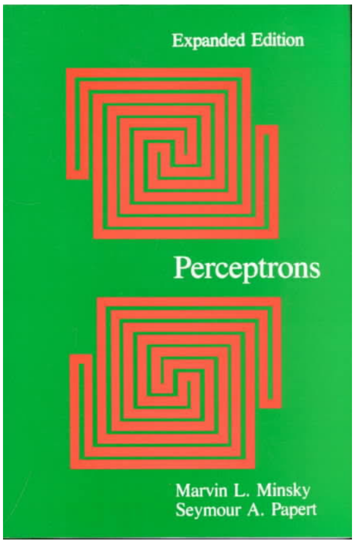
(b) testing result

[Haykin: "Neural Networks and Learning Machines", 3rd edition]

El perceptrón

Análisis de las capacidades
y limitaciones del perceptrón:

- Muchos pensaron que esas limitaciones se extendían a todos los modelos de redes neuronales, aunque no es así.
- Abandono de los modelos conexionistas.
- La investigación en redes neuronales casi desaparece.

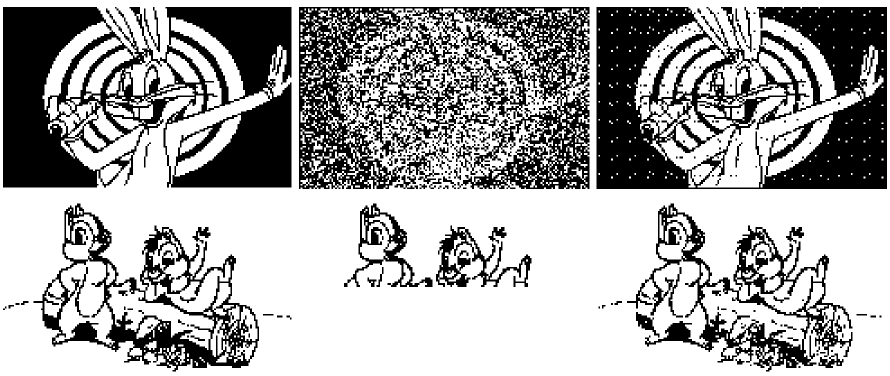


1969

Marvin Minsky & Seymour Papert:
"Perceptrons: An Introduction to Computational Geometry". MIT Press, expanded edition, 1987
ISBN 0262631113

Redes de Hopfield

Redes recurrentes
que funcionan como memorias asociativas



Original | Degraded | Reconstruction

1982

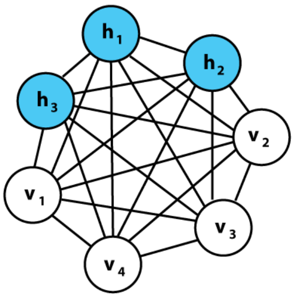
John J. Hopfield:
"Neural networks and physical systems with emergent collective computational abilities"
Proceedings of the National Academy of Sciences
PNAS 79(8):2554–2558, 1982

Historia de las redes neuronales artificiales

Máquinas de Boltzmann

Un contraejemplo:

Sí que se pueden entrenar redes con múltiples capas de neuronas.



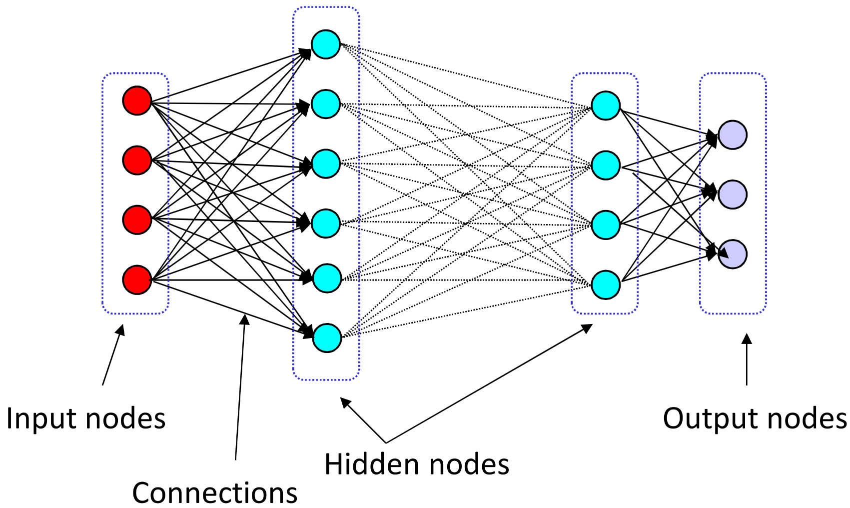
1985

David H. Ackley, Geoffrey E. Hinton & Terrence J. Sejnowski: "A Learning Algorithm for Boltzmann Machines", Cognitive Science 9(1):147–169, 1985. DOI 10.1207/s15516709cog0901_7

Historia de las redes neuronales artificiales

Backpropagation

Algoritmo de entrenamiento de redes multicapa



1986

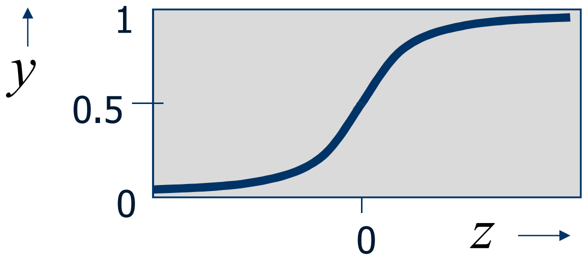
David E. Rumelhart, Geoffrey E. Hinton & Ronald J. Williams: "Learning representations by back-propagating errors" Nature 323(6088):533–536, 1986. DOI 10.1038/323533a0

Backpropagation

Modelo de neurona sigmoidal

$$z = \sum_i x_i w_i$$

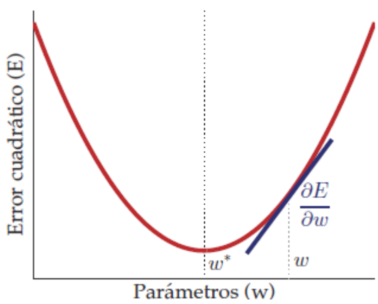
$$y = \frac{1}{1+e^{-z}}$$



1986

Backpropagation

Algoritmo de entrenamiento

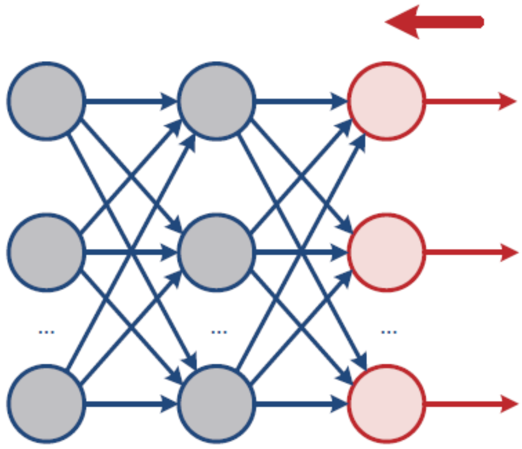


$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i}$$

1986

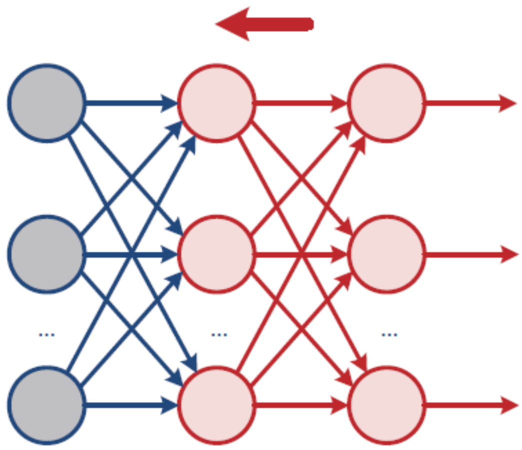
Backpropagation

Propagación de errores $\delta E / \delta y$



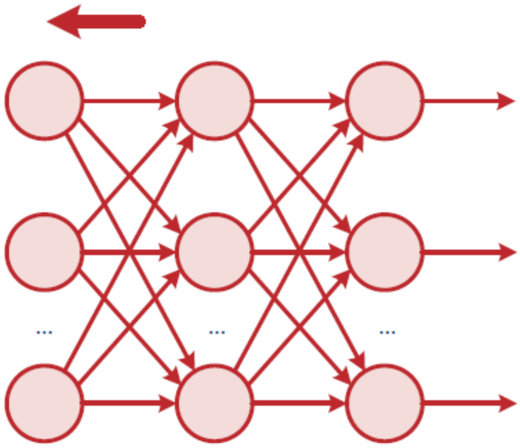
Backpropagation

Propagación de errores $\delta E / \delta y$



Backpropagation

Propagación de errores $\delta E/\delta y$

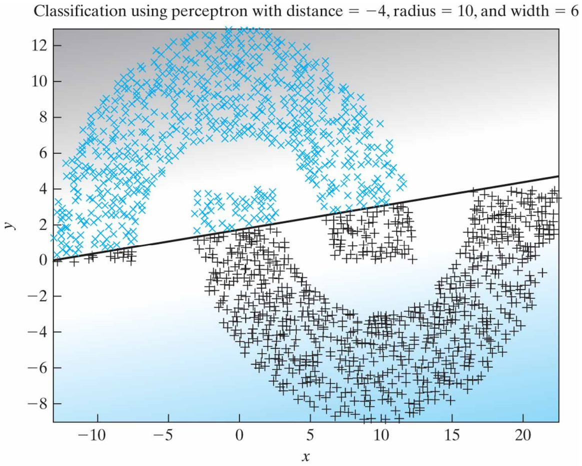


Historia de las redes neuronales artificiales

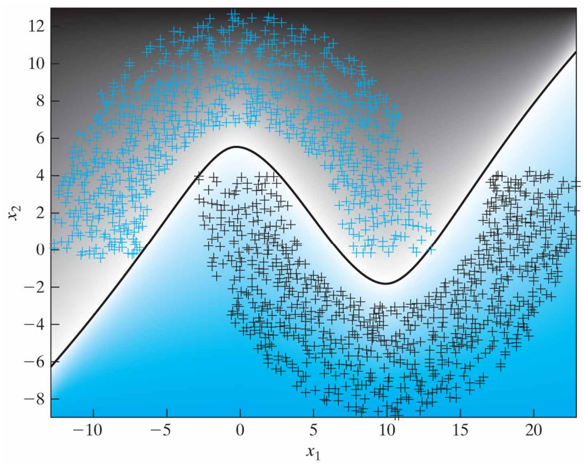
Backpropagation

El resultado…

Classification using perceptron with distance = −4, radius = 10, and width = 6



(b) testing result



(b) Testing result

Perceptrón **Red multicapa**

Historia de las redes neuronales artificiales
Backpropagation

Algoritmo redescubierto en múltiples ocasiones…

- Sistemas de control (años 60)
  Arthur E. Bryson, W.F. Denham & S.E. Dreyfus: "Optimal programming problems with inequality constraints. I: Necessary conditions for extremal solutions." AIAA J. 1(11):2544-2550, **1963**.
  Arthur E, Bryson & Yu-Chi Ho: "Applied optimal control: optimization, estimation, and control." Blaisdell Publishing Company / Xerox College Publishing, p. 481, **1969**.
- Diferenciación automática (años 70)
  Seppo Linnainmaa: The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors. Master's Thesis (in Finnish), University of Helsinki, 6–7, **1970**.
  Seppo Linnainmaa: "Taylor expansion of the accumulated rounding error". BIT Numerical Mathematics. 16(2):146–160, **1976**. DOI 10.1007/bf01931367.

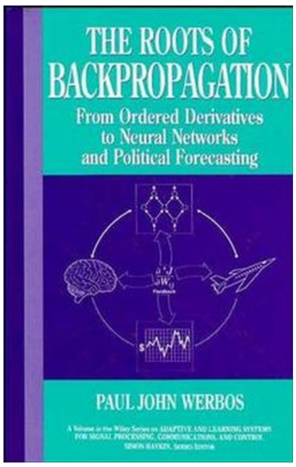
1986???

Historia de las redes neuronales artificiales
Backpropagation

Algoritmo redescubierto en múltiples ocasiones…

- Redes neuronales (1974!!!)
  Paul John Werbos: "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences." PhD thesis, Harvard University, **1974**.

Paul John Werbos:
"The Roots of Backpropagation:
From Ordered Derivatives
to Neural Networks and Political Forecasting."
John Wiley & Sons, Inc., 1994.
ISBN 0471598976



1986???

Backpropagation

Política & Publicaciones

Referencias bibliográficas del artículo de Nature

Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams*

Received 1 May; accepted 31 July 1986.

1. Rosenblatt, F. *Principles of Neurodynamics* (Spartan, Washington, DC, 1961).
2. Minsky, M. L. & Papert, S. *Perceptrons* (MIT, Cambridge, 1969).
3. Le Cun, Y. *Proc. Cognitiva* **85**, 599–604 (1985).
4. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 1: *Foundations* (eds Rumelhart, D. E. & McClelland, J. L.) 318–362 (MIT, Cambridge, 1986).

Backpropagation

Política & Publicaciones

Geoffrey Hinton interview

Neural Networks & Deep Learning

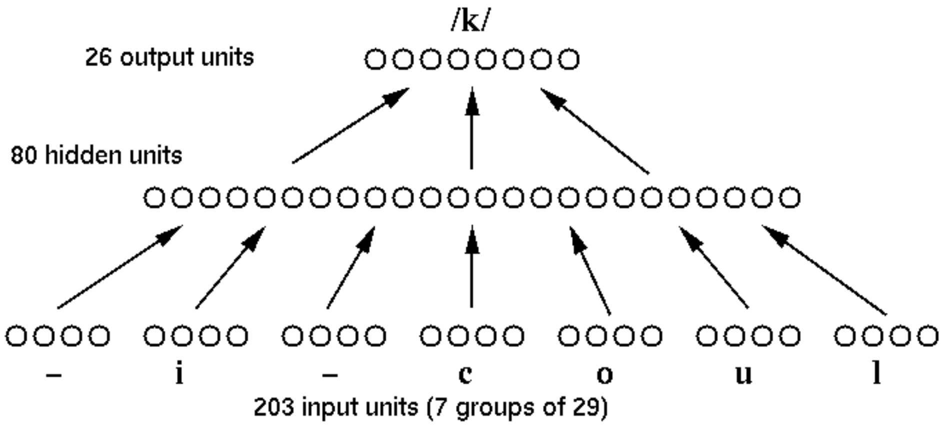
coursera

deeplearning.ai

"… we managed to get a paper into Nature in 1986. And **I did quite a lot of political work to get the paper accepted**. I figured out that one of the referees was probably going to be Stuart Sutherland, who was a well known psychologist in Britain. And I went to talk to him for a long time, and explained to him exactly what was going on. And he was very impressed by the fact that we showed that backprop could learn representations for words…"

Backpropagation

NETTalk
Síntesis de voz

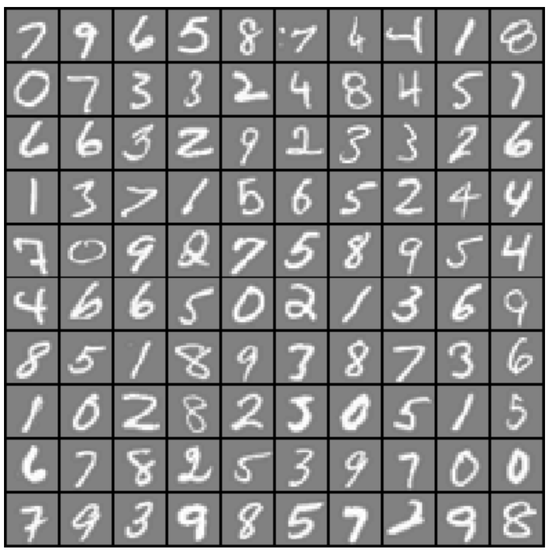


1986

Terrence J. Sejnowski & Charles Rosenberg: "NETtalk: a parallel network that learns to read aloud," Cognitive Science, 14, 179-211, 1986.

Redes convolutivas

The MNIST database of handwritten digits
http://yann.lecun.com/exdb/mnist/

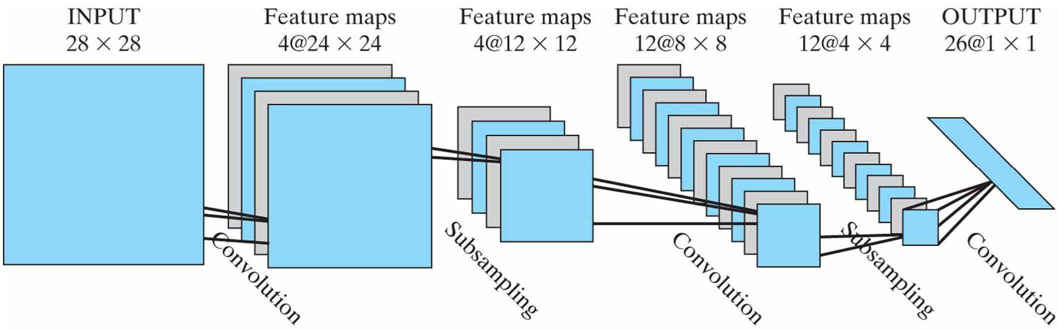


1990s

Redes convolutivas

LeNet

http://yann.lecun.com/exdb/lenet/

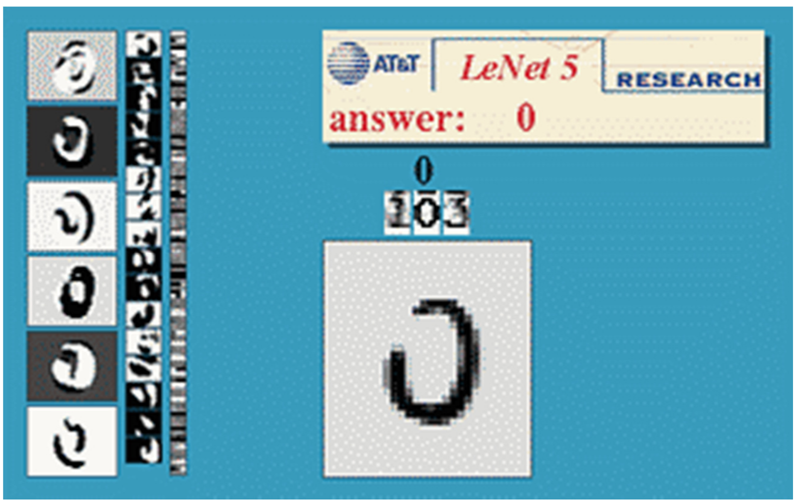


1990s

Redes convolutivas

LeNet

http://yann.lecun.com/exdb/lenet/

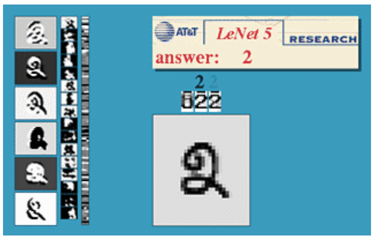


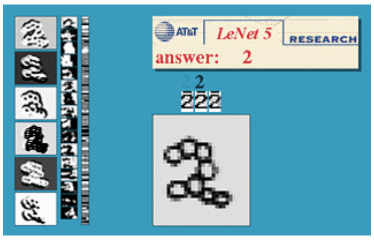
Redes convolutivas

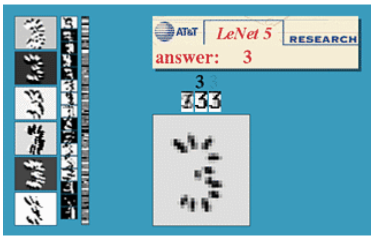
LeNet

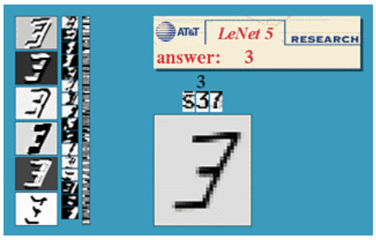
http://yann.lecun.com/exdb/lenet/

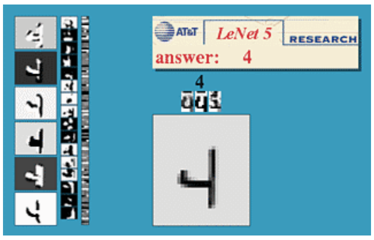
Ejemplos

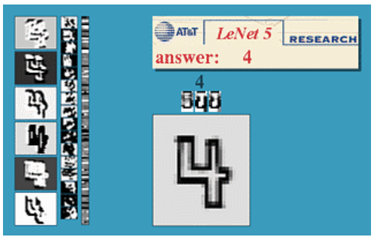


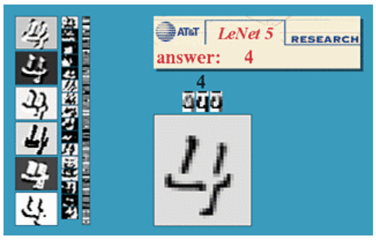












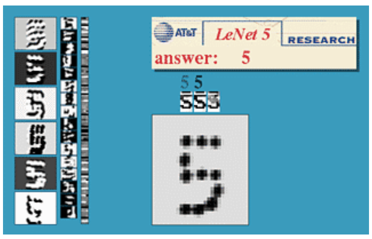
Historia de las redes neuronales artificiales

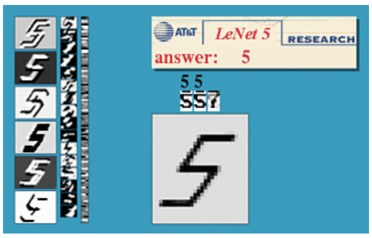
Redes convolutivas

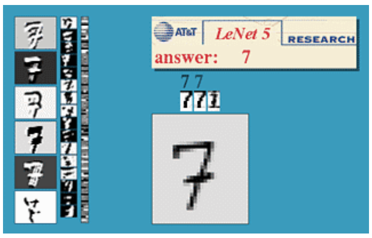
LeNet

http://yann.lecun.com/exdb/lenet/

Ejemplos





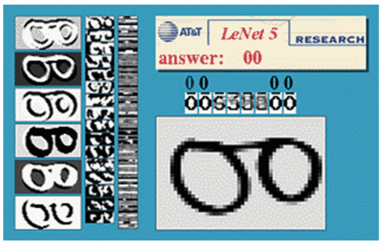


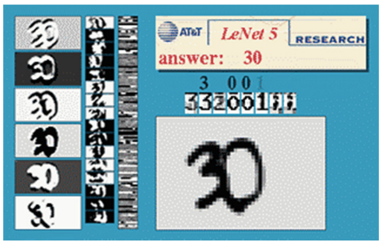
Redes convolutivas

LeNet

http://yann.lecun.com/exdb/lenet/

Variaciones en los datos de entrada



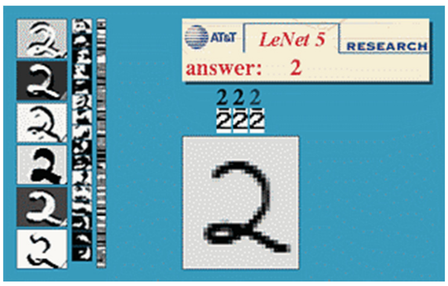


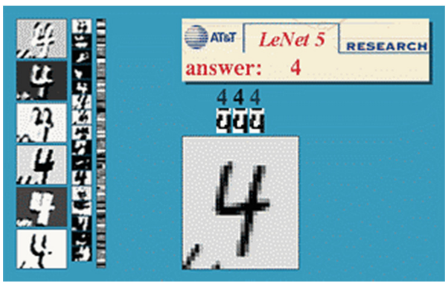
Redes convolutivas

LeNet

http://yann.lecun.com/exdb/lenet/

Robustez frente a la presencia de ruido en la imagen…





Redes convolutivas

LeNet

http://yann.lecun.com/exdb/lenet/

Casos curiosos

Redes convolutivas

Pooling operation used in convolutional neural networks is a big mistake, and the fact that it works so well is a disaster.

Geoffrey Hinton
Professor at University of Toronto
Google Brain Team Manager
Godfather of the MLP, Backpropagation and DNN

From Ask me anything on Reddit
https://www.reddit.com/r/MachineLearning/comments/2lmo0l/ama_geoffrey_hinton/

SVMs

Una apuesta…

AT&T Adaptive Systems Research Dept., Bell Labs

1. Jackel bets (one fancy dinner) that by March 14, 2000 people will understand
quantitatively why big neural nets working on large databases are not so bad.
(Understanding means that there will be clear conditions and bounds)

Vapnik bets (one fancy dinner) that Jackel is wrong.

But .. If Vapnik figures out the bounds and conditions, Vapnik still wins the bet.

****

2. Vapnik bets (one fancy dinner) that by March 14, 2005, no one in his right mind will use neural nets that are essentially like those used in 1995.

Jackel bets (one fancy dinner) that Vapnik is wrong

1995

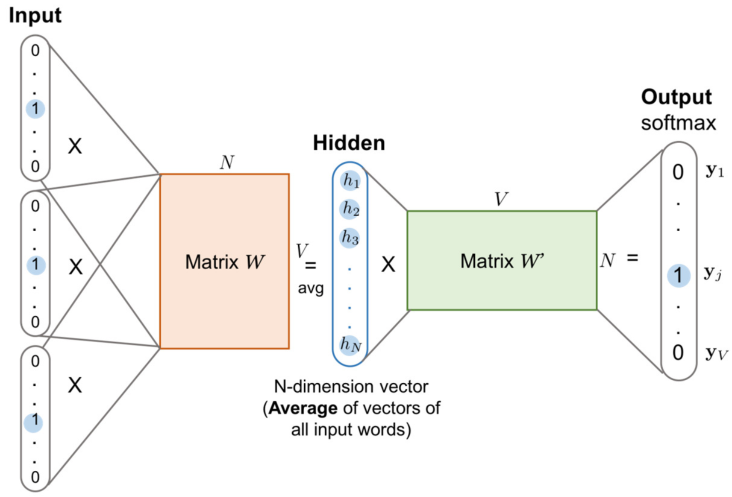
SVMs

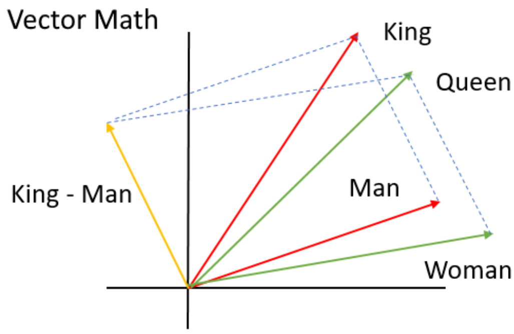
¿Por qué las SVMs nunca fueron una buena opción en IA?

Sólo son una reencarnación de los perceptrones…

- Expanden la entrada a una capa (enorme) de características **no adaptativas**.
- Sólo tienen una capa de pesos **adaptativos**.
- Disponen de un algoritmo eficiente para ajustar los pesos controlando el sobreaprendizaje
(una forma inteligente de seleccionar características y encontrar los pesos adecuados).

Word embeddings



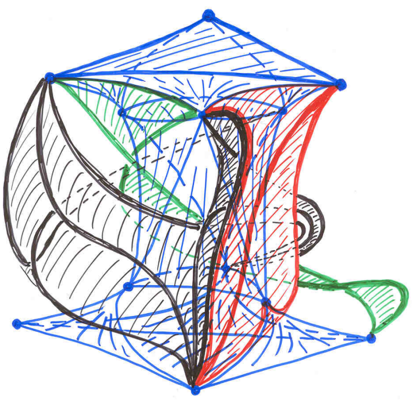


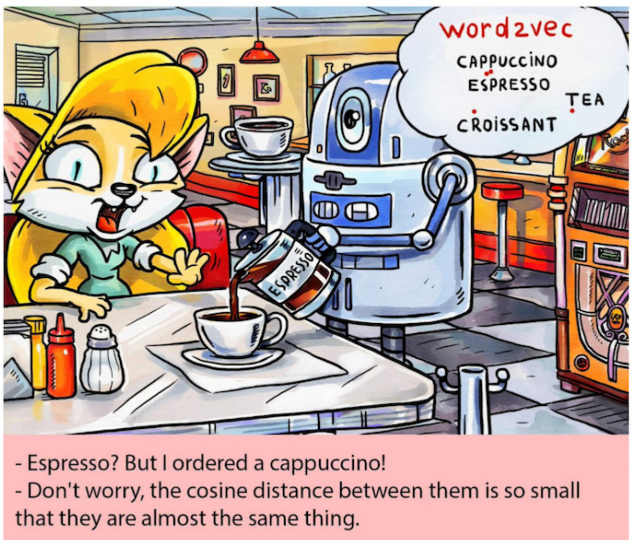
2000

Yoshua Bengio, Réjean Ducharme, Pascal Vincent
"A neural probabilistic language model."
NIPS 2000: 932-938
JMLR 3:1137-1155, 2003

Word embeddings

word2vec



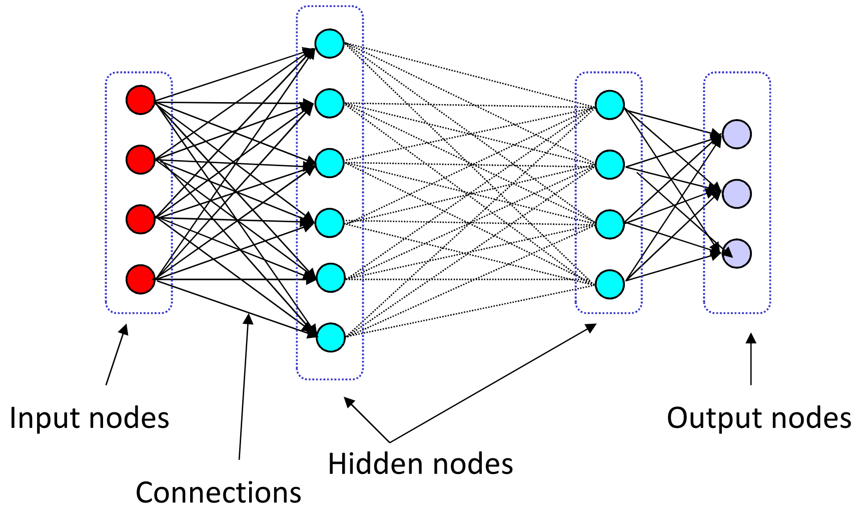


2000s

Tomas Mikolov et al.: "Efficient Estimation of Word Representations in Vector Space"
arXiv:1301.3781, 2013

Deep Learning

Backpropagation no funcionaba bien
con redes que tengan varias capas ocultas
(salvo en el caso de las redes convolutivas)…



Deep Learning

Algunos hechos hicieron que backpropagation no tuviera éxito en tareas en las que luego se ha demostrado útil:

- Capacidad de cálculo limitada.
- Disponibilidad de conjuntos de datos etiquetados.
- “Deep networks” demasiado pequeñas
  (e inicializadas de forma poco razonable).

Historia de las redes neuronales artificiales

Deep Learning

2006: The Deep Breakthrough

- Hinton, Osindero & Teh « A Fast Learning Algorithm for Deep Belief Nets », *Neural Computation*, 2006
- Bengio, Lamblin, Popovici, Larochelle « Greedy Layer-Wise Training of Deep Networks », *NIPS'2006*
- Ranzato, Poultney, Chopra, LeCun « Efficient Learning of Sparse Representations with an Energy-Based Model », *NIPS'2006*

2006

[Yoshua Bengio]

Historia de las redes neuronales artificiales

Deep Learning

Geoffrey Hinton
(University of Toronto & Google)

Yann LeCun
(AT&T Labs → NYU → Facebook)

Joshua Bengio
(University of Montréal & IBM Watson)

2018

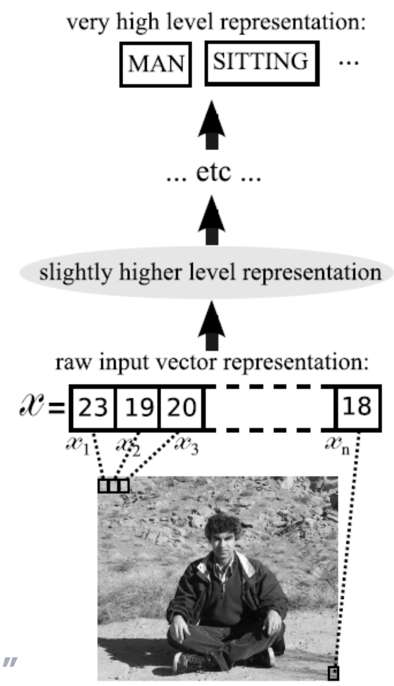
A.M. TURING AWARD

Historia de las redes neuronales artificiales
Deep Learning

Estadística	Inteligencia Artificial
Dimensionalidad baja (<100)	Dimensionalidad alta (>>100)
Mucho ruido en los datos	El ruido no es el mayor problema
Sin demasiada estructura en los datos (puede capturarse usando modelos simples)	Mucha estructura en los datos (demasiado complicada para modelos simples)
PRINCIPAL PROBLEMA	PRINCIPAL PROBLEMA
Separar estructura de ruido	Descubrir una forma de representar la estructura que se pueda aprender
TÉCNICAS	TÉCNICAS
SVM [Support Vector Machines]	Backpropagation

Historia de las redes neuronales artificiales
Deep Learning

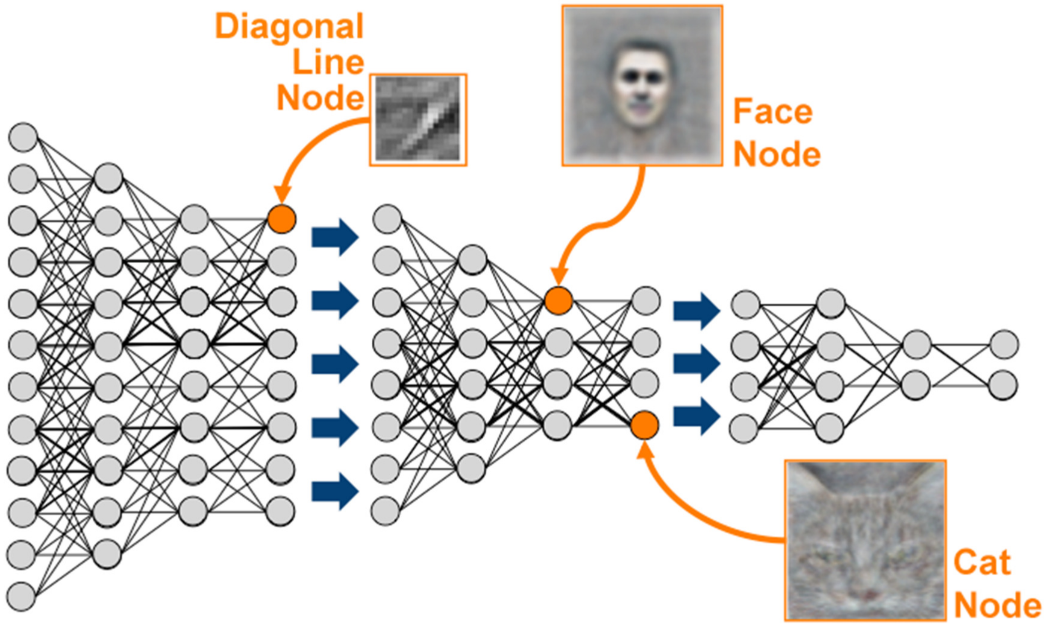
Motivación



Yoshua Bengio
"Learning Deep Architectures for AI"
2009

Deep Learning

Deep Learning as hierarchical feature representation



Deep Learning

¿Cuál era el problema de backpropagation?

- Requiere datos etiquetados,
  pero casi todos los datos disponibles no lo están.
- No resulta demasiado escalable:
  Demasiado lento en redes con múltiples capas ocultas.
- Se puede quedar atascado en óptimos locales
  (¿lejos de ser óptimos en "deep networks"?).

Deep Learning

Política & Publicaciones

Yann LeCun @ CVPR'2012

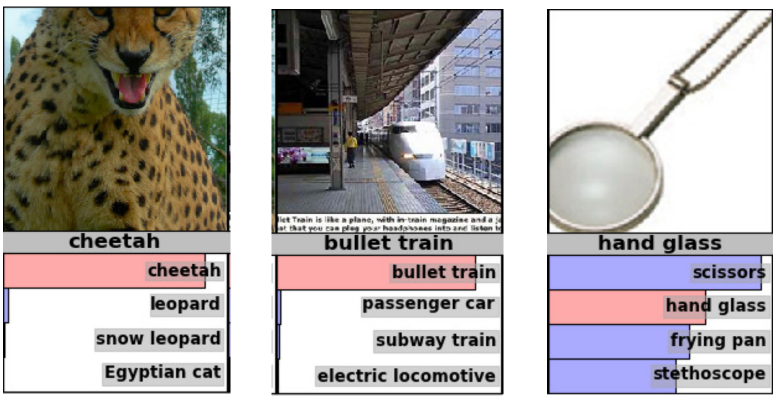
… the reviews [are] so ridiculous, that I don't know how to begin writing a rebuttal without insulting the reviewers … This time though, the reviewers were particularly clueless, or negatively biased, or both. I was very sure that this paper was going to get good reviews because: 1) it has two simple and generally applicable ideas for segmentation… 2) it uses no hand-crafted features… 3) it beats all published results on 3 standard datasets for scene parsing; 4) it's an order of magnitude faster than the competing methods.
If that is not enough to get good reviews, I just don't know what is."

Deep Learning

IMAGENET

Large Scale Visual Recognition Challenge

Reconocimiento de objetos reales en imágenes



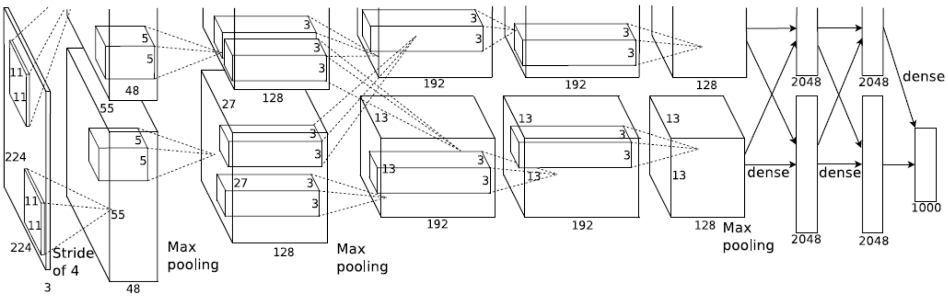
2012

Deep Learning

IMAGENET

AlexNet

Red neuronal diseñada por Alex Krizhevsky (NIPS 2012)



2012

Tasa de error

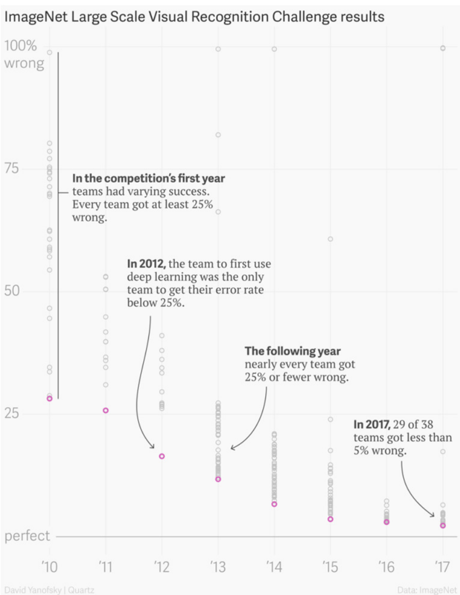
Clasificación de imágenes

16.4% vs. 25% (2010)

Deep Learning

IMAGENET

Large Scale Visual Recognition Challenge



Tasa de error

16.4% Alex Krizhevsky @ NIPS 2012

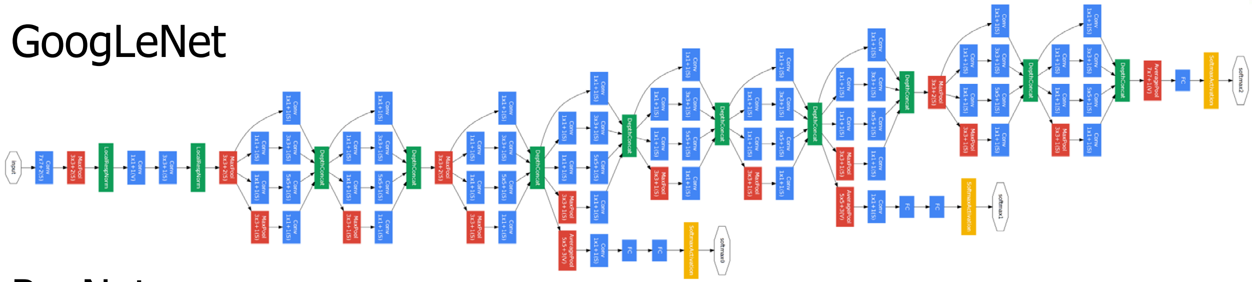
6.66% GoogLeNet @ ILSVRC'2014

4.94% PreLU-nets (MSR) @ 2015

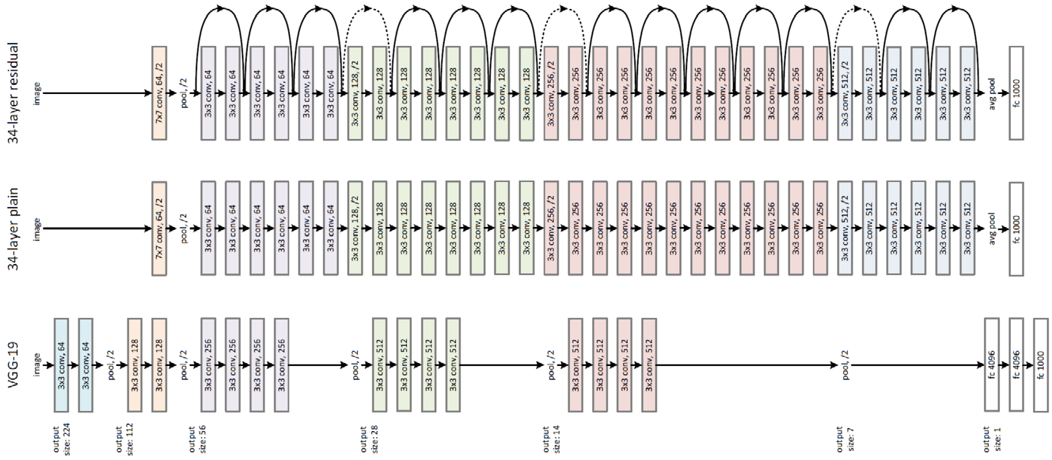
"Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification"
arXiv, 2015, http://arxiv.org/pdf/1502.01852v1.pdf

Deep Learning

GoogLeNet



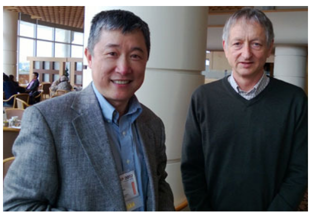
ResNets



Deep Learning

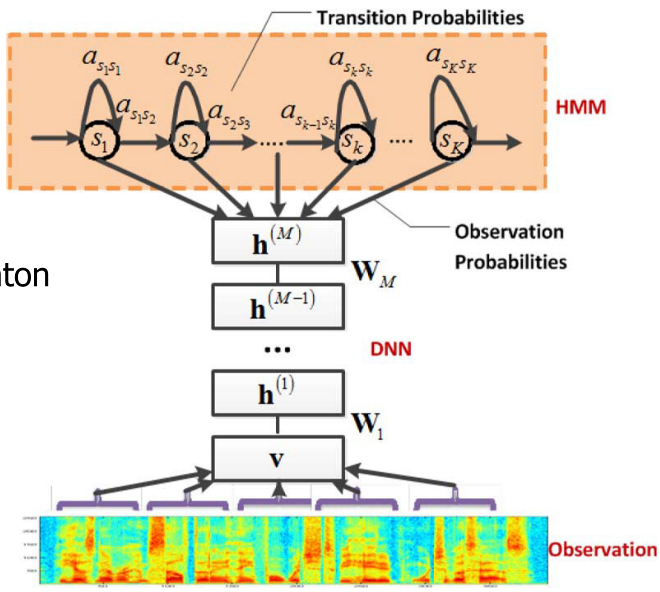
Reconocimiento de voz

Microsoft Research



Li Deng (MSR) & Geoff Hinton

Dong Yu (MSR)



Deep Learning

Reconocimiento de voz

Task	Hours of training data	Deep Neural Network	Gaussian Mixture Model	GMM with more data
Switchboard (Microsoft Research)	309	18.5%	27.4%	18.6% (2000 hrs)
English broadcast news (IBM)	50	17.5%	18.8%	
Google Voice Search (Android 4.1)	5,870	12.3% (and falling)		16.0% (>>5,870 hrs)

Microsoft Research

IBM

Google

2012

Geoffrey Hinton, Li Deng, Dong Yu et al.:
"Deep Neural Networks
for Acoustic Modeling in Speech Recognition"
IEEE Signal Processing Magazine, 2012

Deep Learning

Traducción simultánea

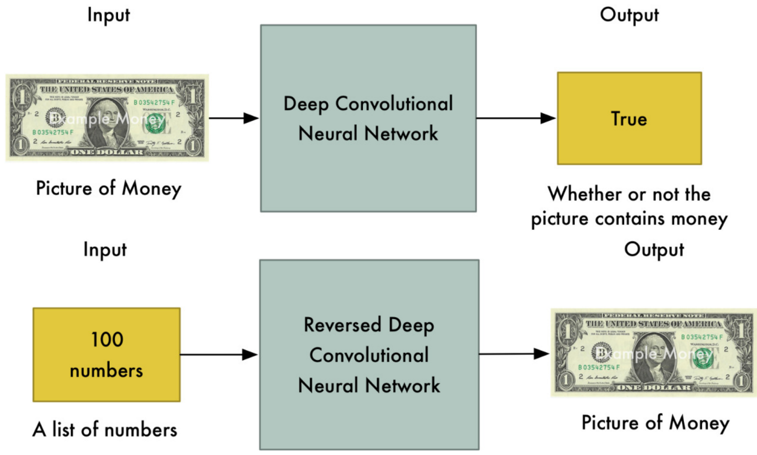
2012

Deep Learning

GANs [Generative Adversarial Networks]

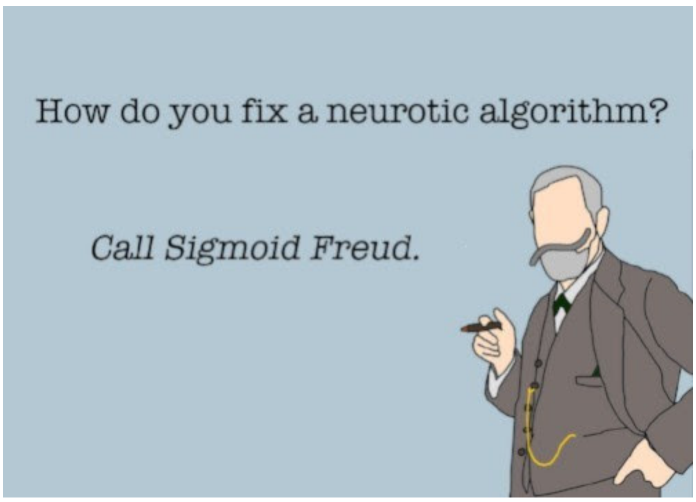
Discriminador

Generador



2014

Técnicas de deep learning



Técnicas de Deep Learning

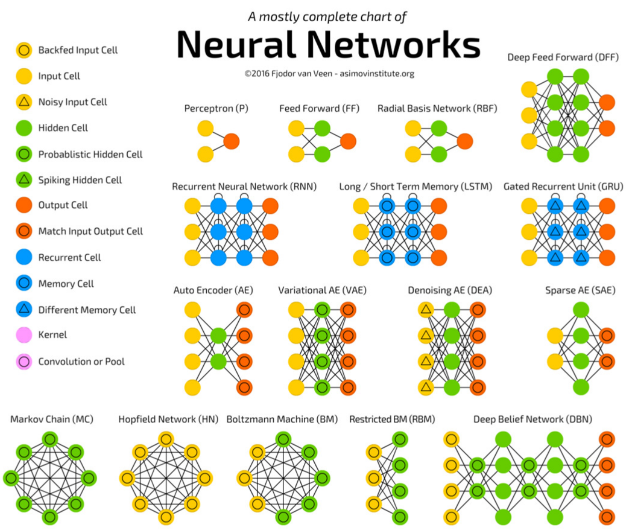
Pero resulta que, después de todo,
la solución era muy simple:

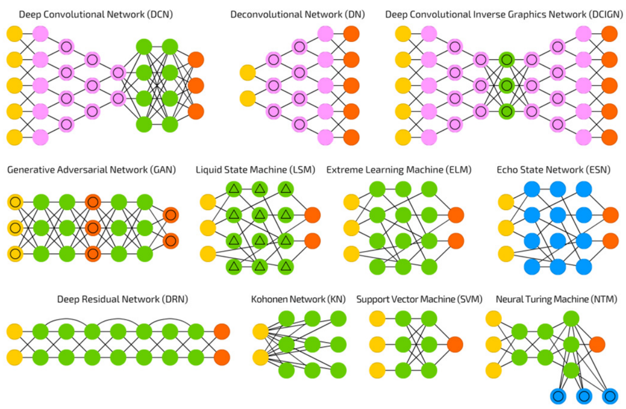
- Modelos paramétricos suficientemente grandes (redes neuronales con muchos pesos ajustables).
- Conjuntos de entrenamiento suficientemente grandes para entrenar las redes usando el gradiente descendente.

Richard Feynman, sobre el Universo:
It's not complicated, it's just a lot of it

Técnicas de Deep Learning

El zoo de las redes neuronales

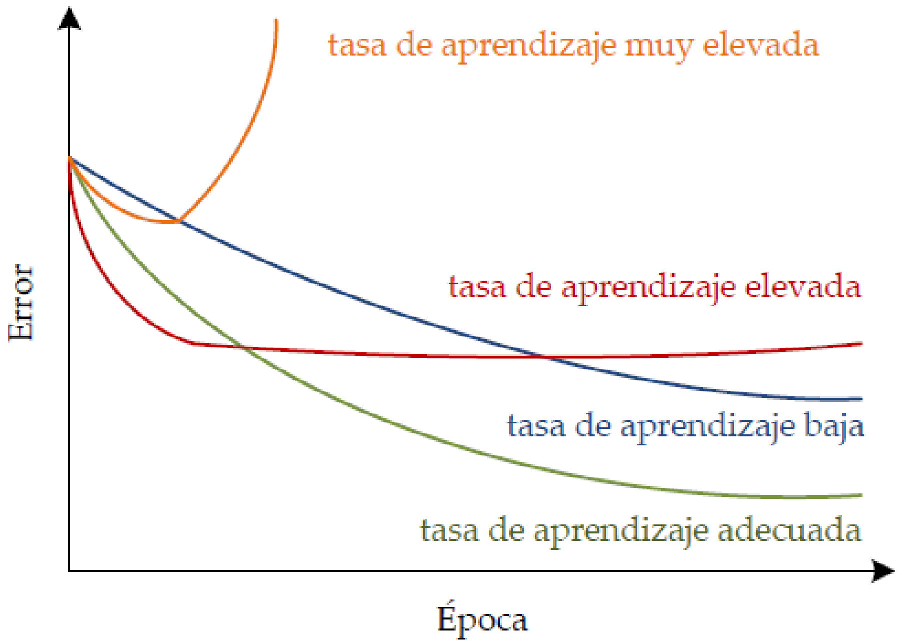




Técnicas de Deep Learning

Algoritmos de optimización

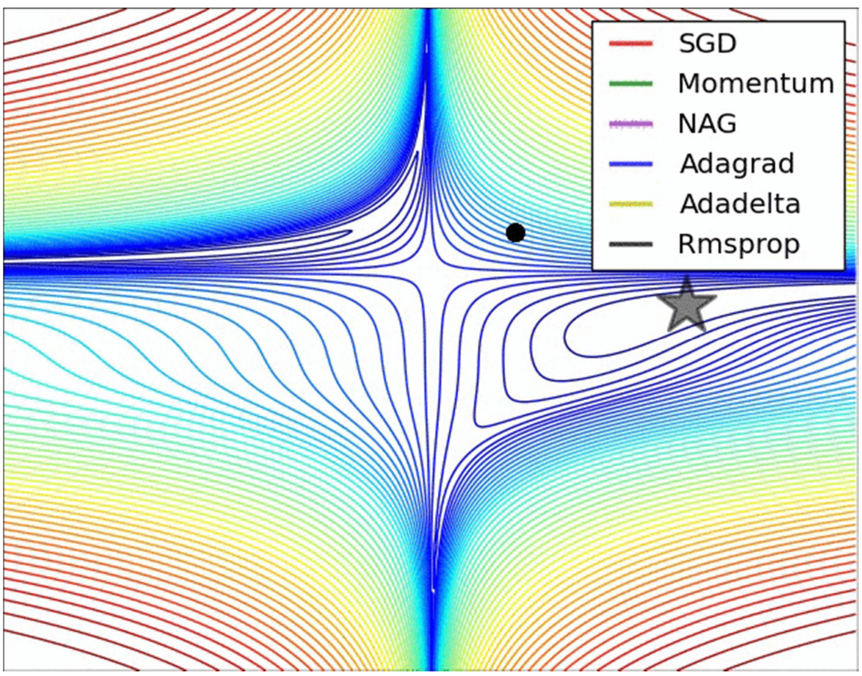
$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i}$$



Técnicas de Deep Learning

Algoritmos de optimización

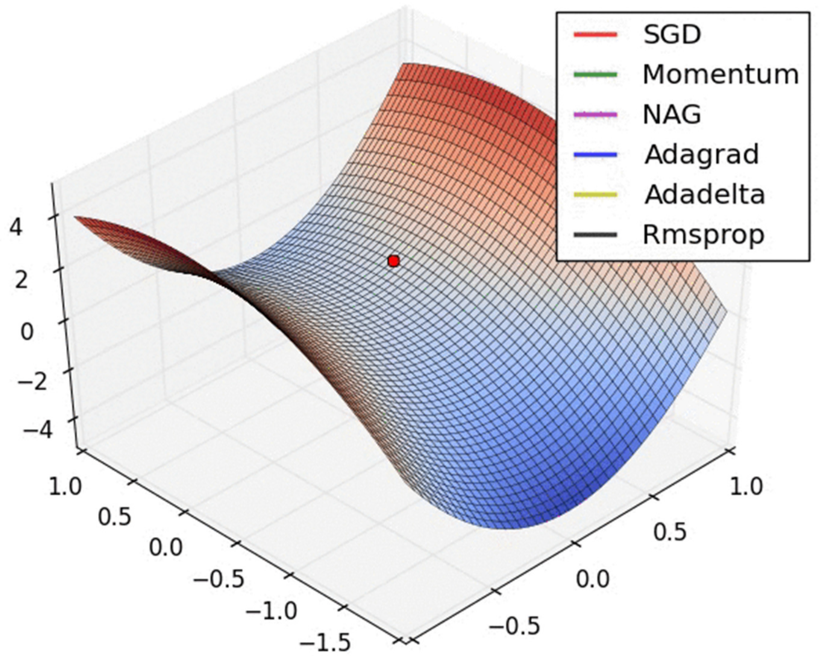
SGD [Stochastic Gradient Descent]



Alec Radford
https://twitter.com/alecrad

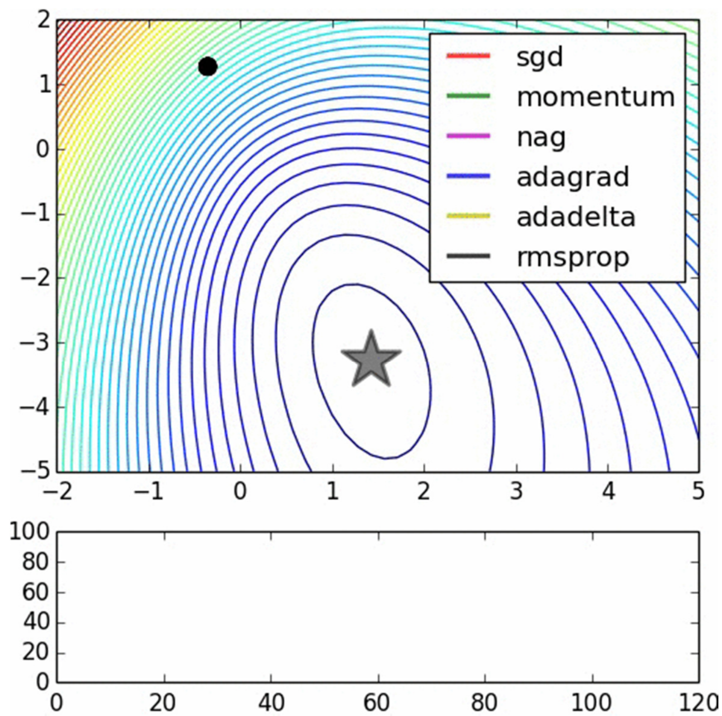
Algoritmos de optimización

SGD [Stochastic Gradient Descent] @ saddle point



Alec Radford
https://twitter.com/alecrad

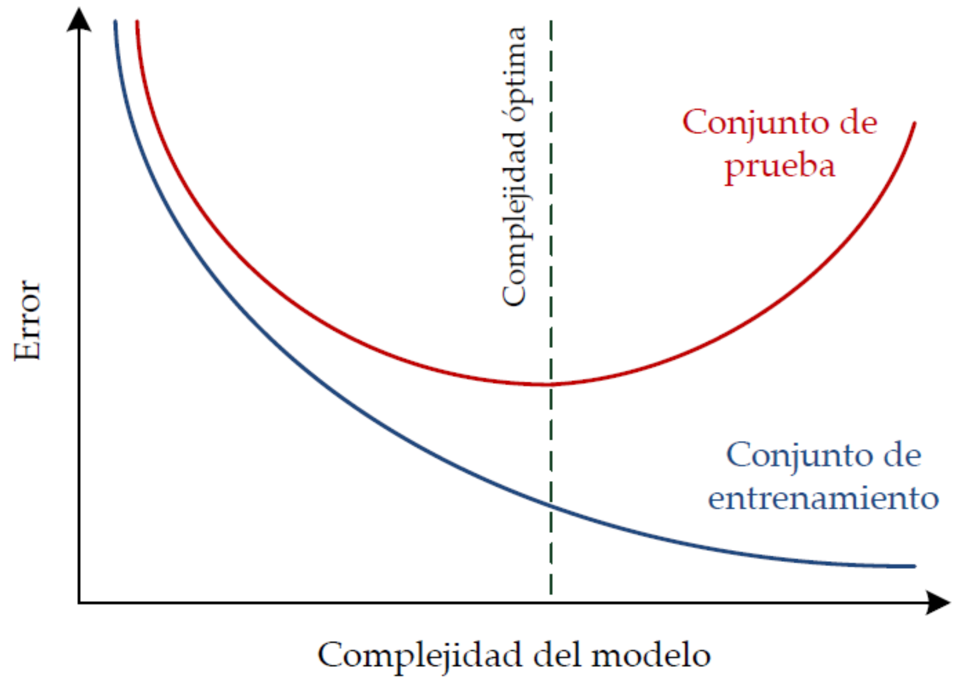
Algoritmos de optimización



Alec Radford
https://twitter.com/alecrad

Regularización

El problema del sobreaprendizaje



Regularización

Estrategias para evitar el sobreaprendizaje:

- Obtener más datos (la mejor opción si tenemos capacidad para entrenar la red usando más datos).

"You can't keep adjusting the data to prove that you would be the best Valentine's date for Scarlett Johansson."

- Ajustar los parámetros de la red para que tenga la capacidad adecuada (suficiente para identificar las regularidades en los datos, pero no demasiada para ajustarse a las espúreas, suponiendo que sean más débiles que las auténticas).

Regularización

Capacidad de la red: Topología

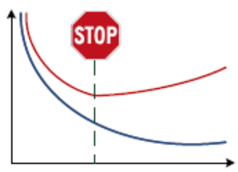
Algunas formas de limitar la capacidad de la red actuando sobre su topología:

- **Arquitectura de la red**:
  Se limita el número de capas ocultas y/o el número de unidades por capa.

- **Weight sharing**:
  Se reduce el número de parámetros de la red haciendo que distintas neuronas compartan los mismos pesos (p.ej. redes convolutivas).

Regularización

Capacidad de la red: Entrenamiento



Algunas formas de limitar la capacidad de la red actuando sobre su algoritmo de entrenamiento:

- **Early stopping**: Se comienza a entrenar la red con pesos pequeños y se para el entrenamiento antes de que sobreaprenda.

- **Weight decay**: Se penalizan los pesos grandes en función de sus valores al cuadrado (penalización L2) o absolutos (penalización L1).

- **Ruido**: Se añade ruido a los pesos o actividades de las neuronas de la red que se está entrenando.

Regularización

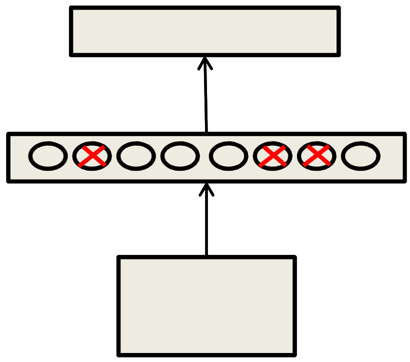
Una tercera estrategia: Combinar modelos

Model averaging” (a.k.a. “ensembles”):
Muchos modelos diferentes con distintos parámetros o el mismo tipo de modelo utilizando distintos subconjuntos del conjunto de entrenamiento [bagging].

Dropout

¡las conspiraciones complejas no son robustas!

-- Geoff Hinton.



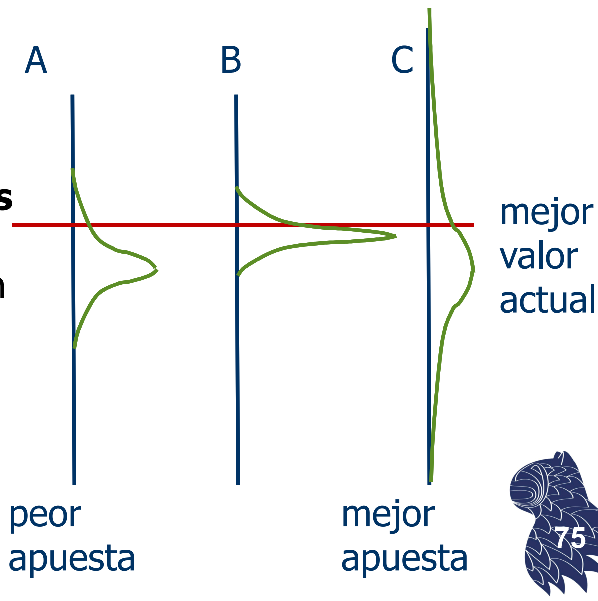
Ajuste de hiperparámetros

Una de las principales dificultades prácticas del uso de redes neuronales es la destreza que requiere establecer todos sus parámetros (“arte” más que ciencia)

p.ej.

Modelo de procesos gaussianos
A partir de la mejor configuración conocida, se elige una combinación de hiperparámetros tal que la mejora esperada sea grande (sin preocuparse por la posibilidad de empeorar).

Snoek, Larochelle & Adams
NIPS 2012



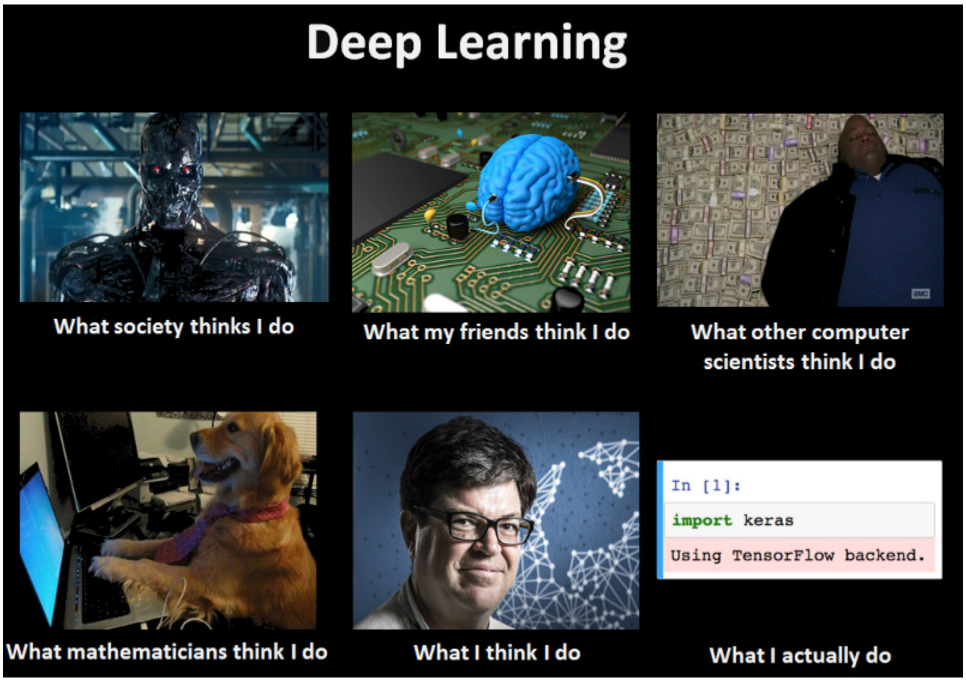
Ajuste de hiperparámetros

AutoML
Aprendizaje automático [Machine Learning]

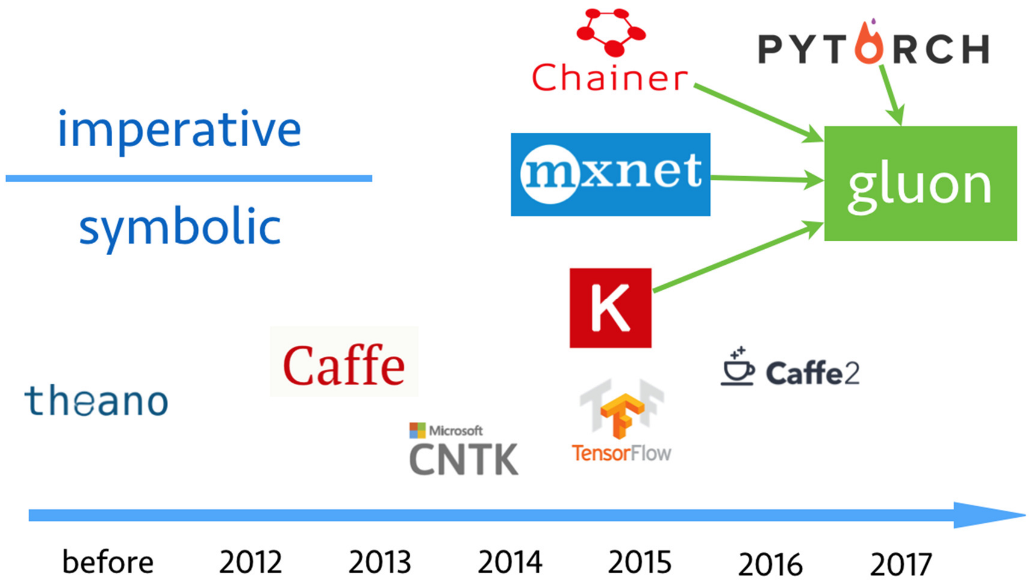
- Mucho mejor que ir haciendo pruebas manualmente (no es el tipo de tarea que los humanos hacemos bien).

- Evita sesgos psicológicos no deseados:
método menos propenso a funcionar mejor
con el método que nos gusta y peor con el que no
(las personas no podemos evitarlo ;-)

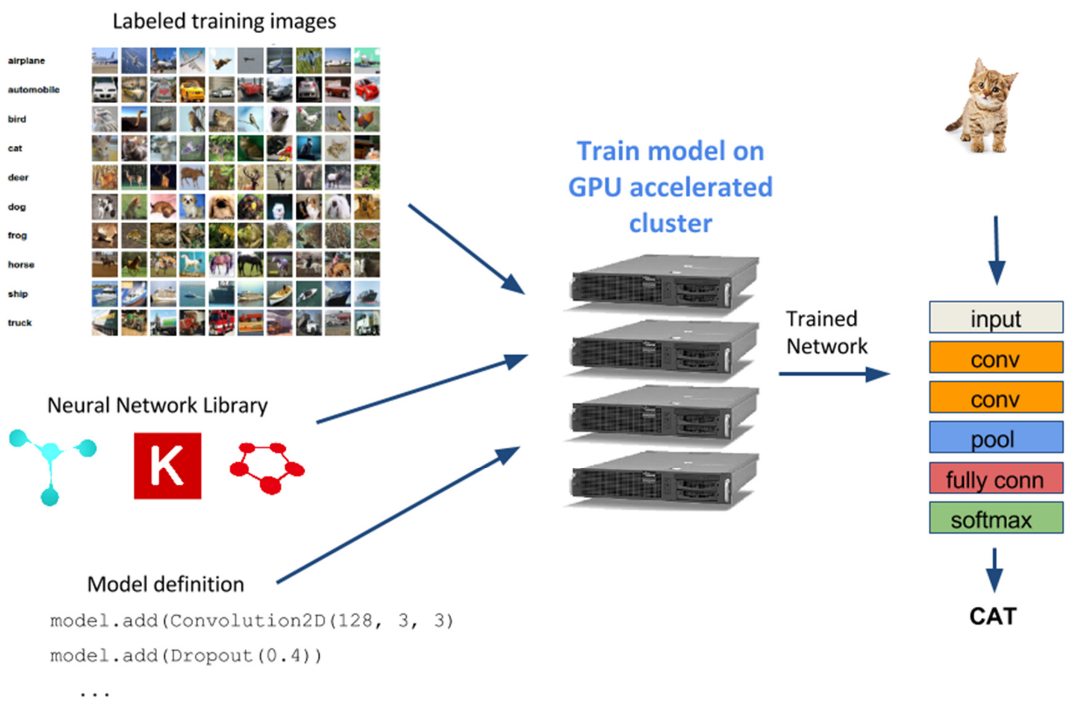
En la práctica



En la práctica
Software para deep learning



En la práctica
Software para deep learning



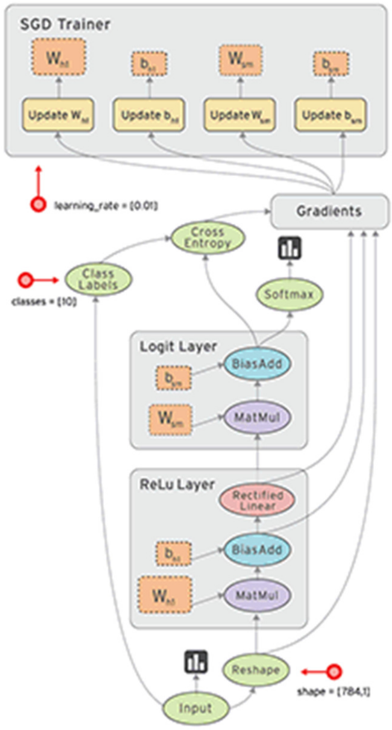
En la práctica

Software para deep learning

Google TensorFlow

https://www.tensorflow.org/

Licencia Apache 2.0



Data flow graph

En la práctica

My favorite definition of Deep Learning is matrix multiplication, a lot of matrix multiplication...

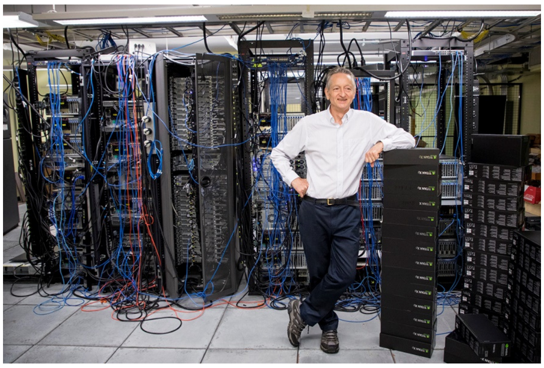
Barbara Fusinska

Machine Learning Programmer at Microsoft

From Keynote Networks are like onions: Practical Deep Learning with TensorFlow
https://www.youtube.com/watch?v=95IV8DoWRwI

En la práctica

En deep learning, todo son vectores…



The Vector Institute for AI

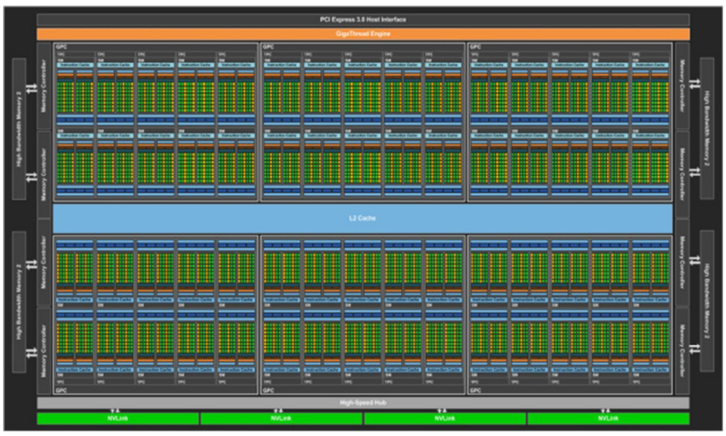
University of Toronto

http://vectorinstitute.ai/

En la práctica
Hardware para deep learning

GPU [procesador SIMD]: Data-level parallelism

NVIDIA Pascal SIMD Processor

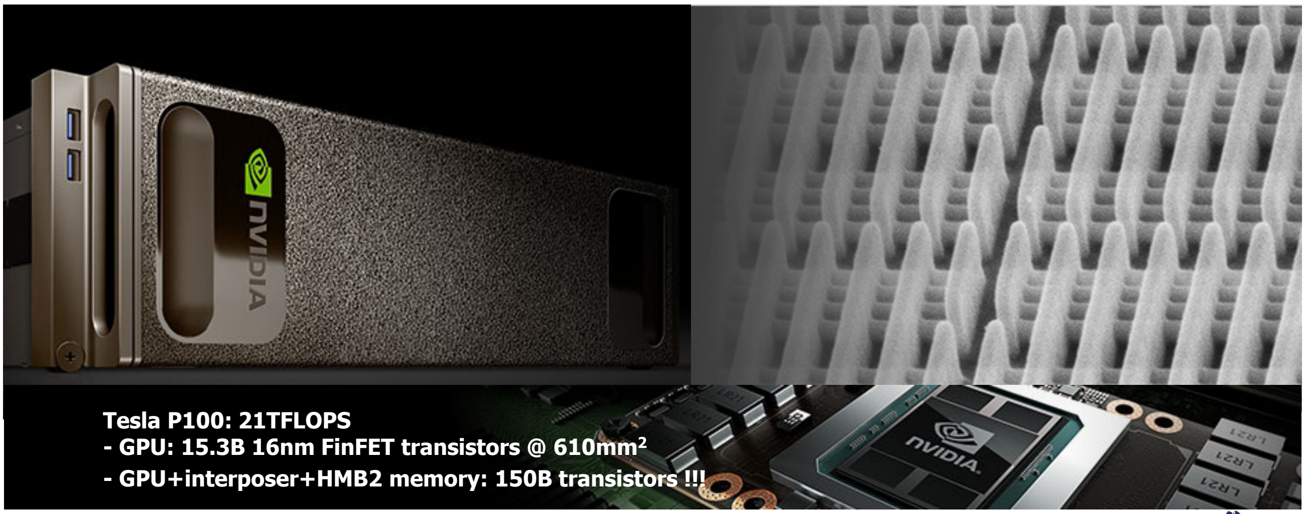


NVIDIA Pascal P100 GPU

Hardware para deep learning

NVIDIA DGX-1 deep learning supercomputer $ 129 000

8x GPUs NVIDIA Tesla P100, 28672 CUDA cores



DGX-1: 8xP100, 512 GB DDR4, 4x1.92TB SSD, 170 TFLOPS, 60kg, 3200W

A $2 Billion Chip to Accelerate Artificial Intelligence, MIT Technology Review, April 2016

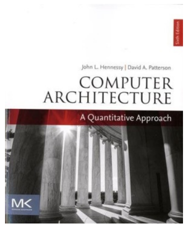
https://www.technologyreview.com/s/601195/a-2-billion-chip-to-accelerate-artificial-intelligence/

http://www.nvidia.com/object/deep-learning-system.html

Hardware para deep learning

DSAs [Domain-Specific Architectures]

- **ASICs** [Application-Specific Integrated Circuits]
- Circuitos reconfigurables,
  p.ej. **FPGAs** [Field-Programmable Gate Arrays]



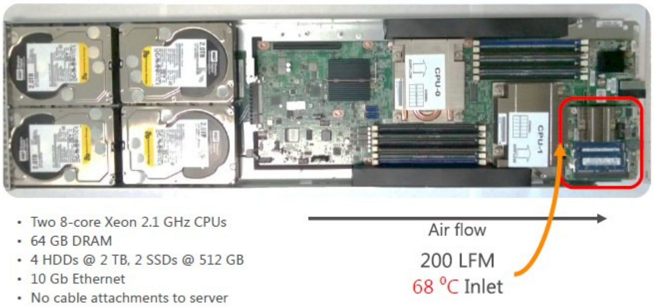
Diseño energéticamente eficiente

- Memorias dedicadas [scratchpad] gestionadas por el programador en lugar de memorias caché.
- Recursos ahorrados en la microarquitectura dedicados a más unidades funcionales o más memoria.
- Paralelismo ajustado al dominio de aplicación (SIMD).
- Reducción de la precisión (8-, 16-bit),
  para aprovechar mejor el ancho de banda de memoria.
- DSL [domain-specific language] para portar código a la DSA, p.ej. Halide (visión) o TensorFlow (DNNs).

Microsoft Research Catapult

FPGAs [Field Programmable Gate Arrays]

- Menor consumo de energía: 25W
- Menor ancho de banda: 11GB/s
  (datos en la memoria DDR3 de la propia FPGA para evitar PCIe)



CPU-FPGA minimalista:
FPGA Altera Stratix V @ Open CloudServer

Toward Accelerating Deep Learning at Scale Using Specialized Logic
HOTCHIPS'2015: A Symposium on High Performance Chips, August 2015

Microsoft Research Project BrainWave

FPGAs [Field Programmable Gate Arrays]

- DNNs as "hardware microservices"

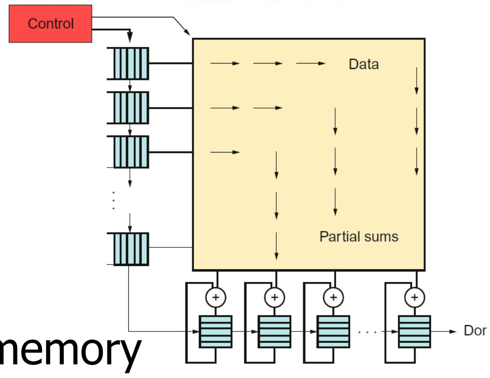
FPGA Intel Stratix 10 (e.g. GRU @ 39.5TFLOPS)

Accelerating Persistent Neural Networks at Datacenter Scale
HOTCHIPS'2017 & NIPS'2017

Hardware para deep learning

TPU [Tensor Processing Unit]

Google, 2015-

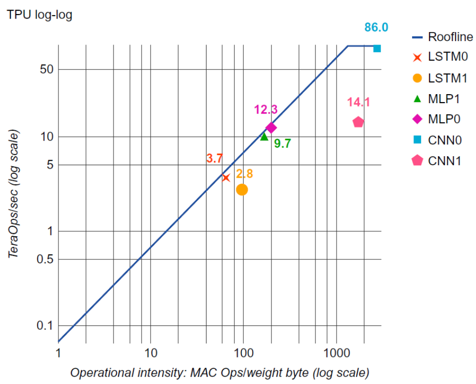


- 65,536 (256x256) 8-bit
  ALU Matrix Multiply Unit
- Large software-managed on-chip memory
- Single-threaded, deterministic execution model
- Coprocessor on the PCIe I/O bus
- The host server sends instructions over the PCIe bus directly to the TPU for it to execute,
  rather than having the TPU fetch the instructions
  (closer to FPUs than to GPUs).

Hardware para deep learning

TPU Applications

Name	LOC	DNN layers: FC	Conv	Element	Pool	Total	Weights	TPU Ops/Weight	% deployed TPUs 2016
MLP0	100	5				5	20M	200	61%
MLP1	1000	4				4	5M	168	
LSTM0	1000	24		34		58	52M	64	29%
LSTM1	1500	37		19		56	34M	96	
CNN0	1000		16			16	8M	2888	5%
CNN1	1000	4	72		13	89	100M	1750	



6 applications, 95% workload @ Google

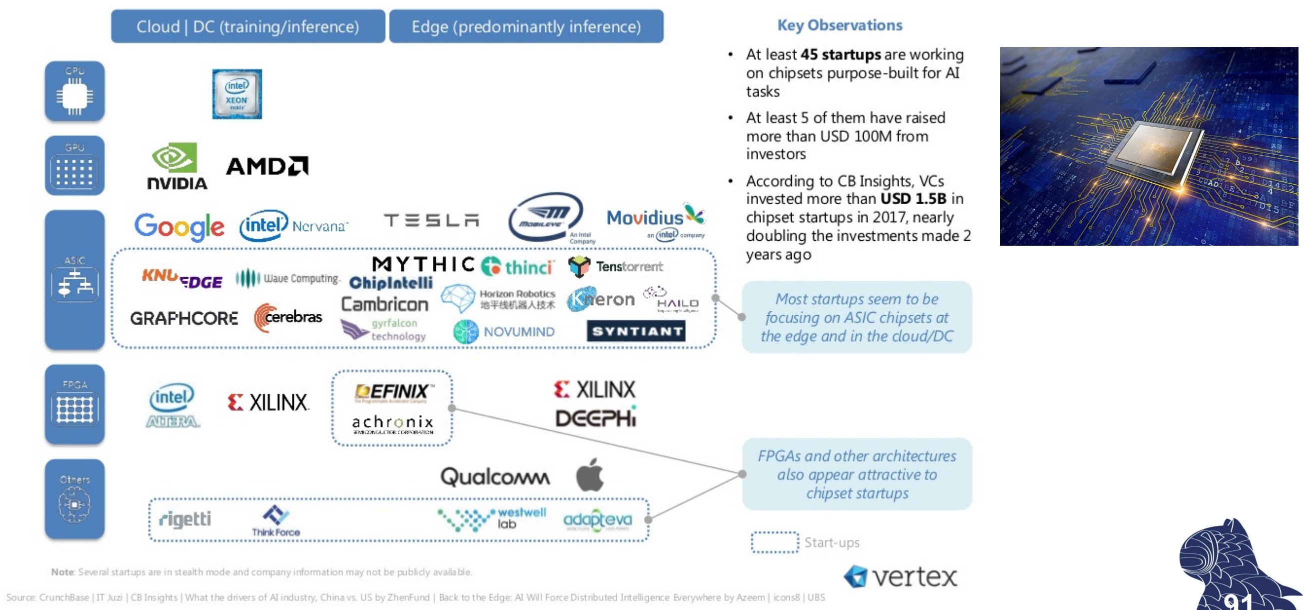
- MLP: RankBrain
- LSTM: Google Neural Translator
- CNN: DeepMind AlphaGo

En la práctica
Hardware para deep learning

CPU Intel i7	GPU NVIDIA	FPGA Catapult	ASIC TPU
SIMD extensions (MMX, SSE, AVX)	Streaming multiprocessors	3926 18-bit PEs (reconfigurable)	256x256 Matrix Multiply Unit
1D SIMD		2D SIMD	
Multithreading		Pipelined systolic array	
Out-of-order execution	Multiprocessing	Programmable controller	Simple execution of instructions
32 & 64-bit FP	32 & 64-bit FP	18-bit integers	8-bit integers
x86 ISA (C, Java, Python…)	**PTX** (CUDA, OpenCL)	**RTL** (Verilog, VHDL)	**Reduced CISC** ISA (API via DSLs: Halide, TensorFlow)

En la práctica
Hardware para deep learning

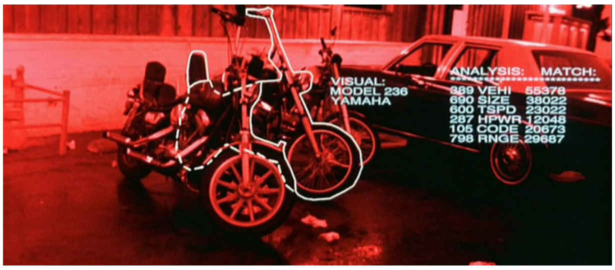
Docenas de empresas desarrollan chips para IA…



Aplicaciones del Deep Learning

Existen problemas para los que es extremadamente difícil desarrollar manualmente un programa de ordenador que los resuelva.

Ejemplo: Visión artificial



[Terminator, 1984]

Aplicaciones del Deep Learning

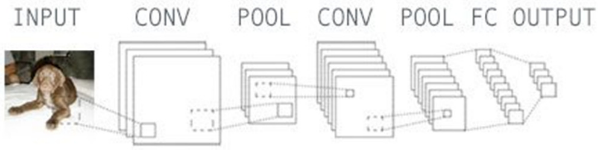
De hecho, el reconocimiento de objetos…

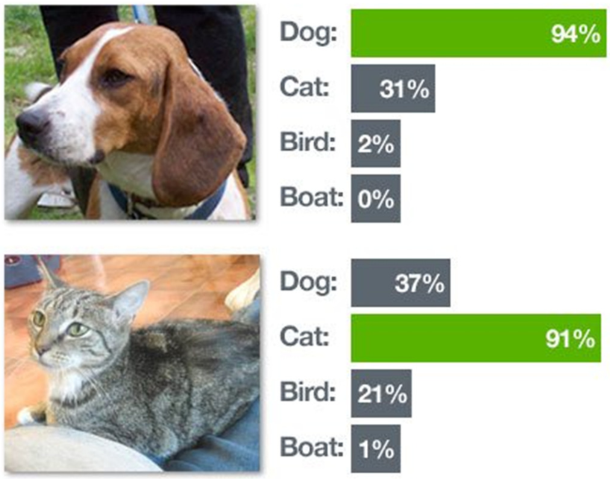
- Ni siquiera sabemos cómo se hace realmente en nuestro cerebro (por lo que difícilmente podremos diseñar un algoritmo que haga exactamente lo mismo).

- Incluso aunque tuviésemos una idea más precisa de cómo se hace en nuestro cerebro, el programa necesario podría ser tremendamente complicado :-(

Aplicaciones del Deep Learning

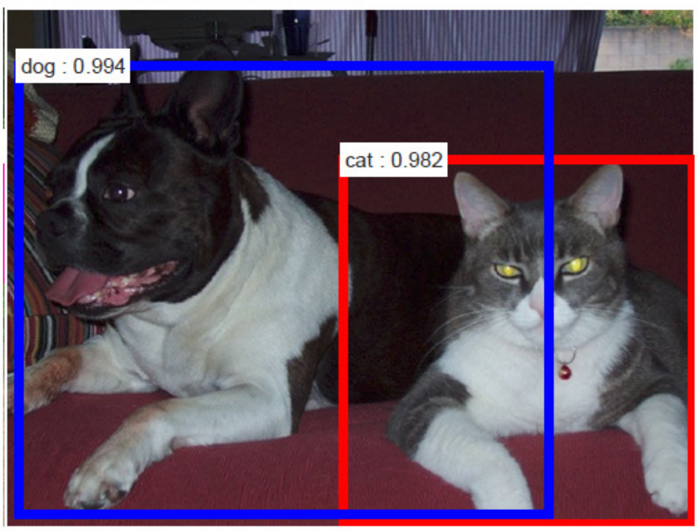
Clasificación de imágenes





Aplicaciones del Deep Learning

Detección de objetos



Detección de objetos usando redes convolutivas, 2017

Aplicaciones del Deep Learning

Detección de objetos

CVPR'2018

Aplicaciones del Deep Learning

Segmentación de imágenes

Input image

Aplicaciones del Deep Learning

Vehículos autónomos

Aplicaciones del Deep Learning

Vehículos autónomos

2005 DARPA Grand Challenge

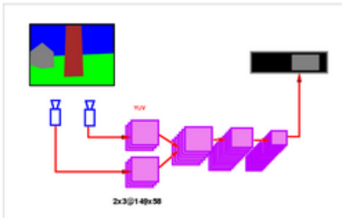
Aplicaciones del Deep Learning

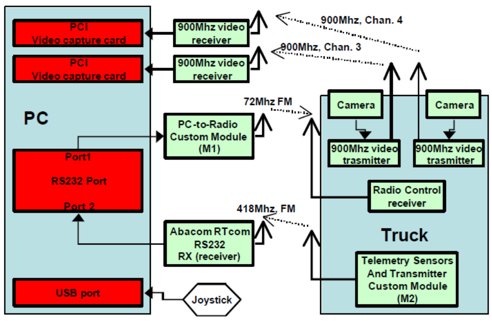
Vehículos autónomos

Autonomous Land Vehicle In a Neural Network (ALVINN)
NAVigational LABoratory II (NAVLAB II)
Control de dirección de un vehículo, CMU Ph.D. thesis, 1992

Aplicaciones del Deep Learning

Vehículos autónomos

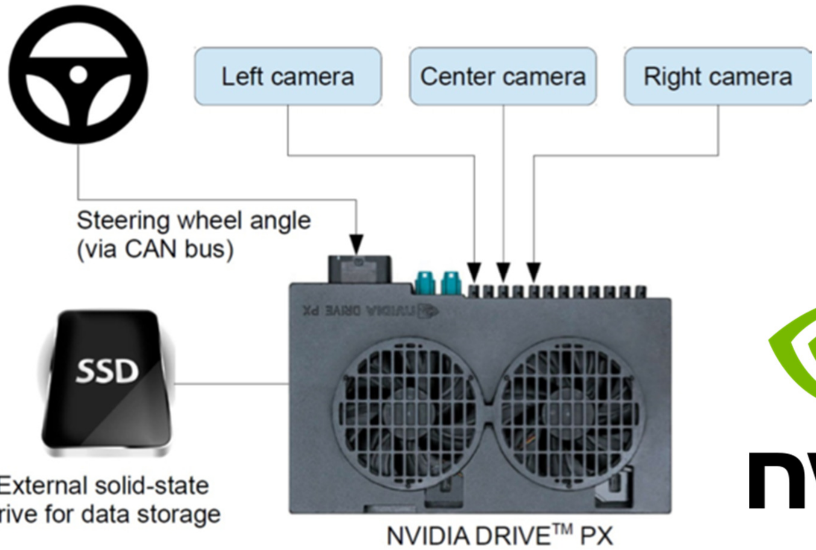




DAVE, 2004
Autonomous Off-Road Vehicle Control using End-to-End Learning
NYU Courant Institute / CBLL [Computational & Biological Learning Lab]
http://www.cs.nyu.edu/~yann/research/dave/

Aplicaciones del Deep Learning

Vehículos autónomos



DAVE2, after DARPA Autonomous Vehicle (DAVE) project
NVIDIA, 2016. http://arxiv.org/abs/1604.07316
Completamente autónomo con sólo 100 horas de entrenamiento!!!

Aplicaciones del Deep Learning

Aprendizaje automático
/ Inteligencia Computacional
/ Redes neuronales artificiales [deep learning]

- En vez de diseñar un algoritmo que resuelva el problema, recopilamos un montón de datos (ejemplos).
- Diseñamos un algoritmo que aprenda de esos datos y cree el programa necesario para resolver el problema.

Aplicaciones del Deep Learning

La solución basada en deep learning

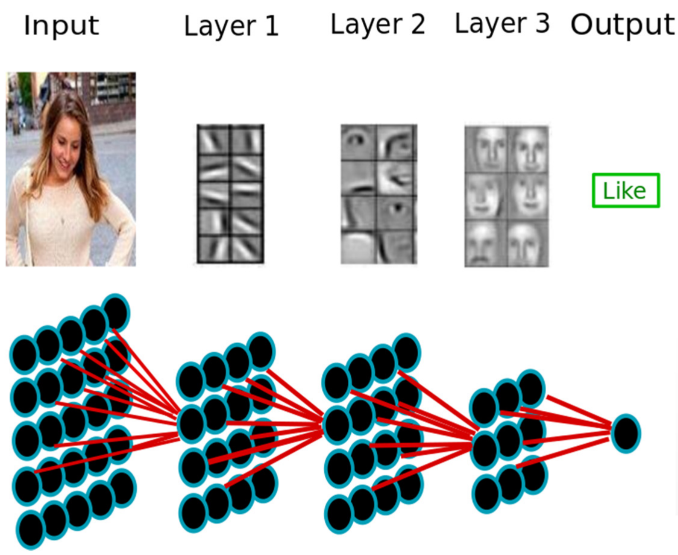
- El programa generado automáticamente no tiene por qué parecerse a un programa implementado manualmente (en el caso de las redes neuronales, puede contener millones de números reales).
- Si tenemos éxito, el programa funcionará bien para nuevos ejemplos, aunque sean diferentes a los que utilizamos para su entrenamiento.
- Si los datos cambian, el programa puede cambiar entrenándolo de nuevo.

Aplicaciones del Deep Learning

"Sweetheart, my neural net predicts that you and I are 98.9% compatible. Will you be my Valentine?"

Aplicaciones del Deep Learning

Citas por Internet, e.g. Tinder & OKCupid.com



Harm De Vries & Jason Yosinski:
Can deep learning help you find the perfect match?
ICML'2015 Deep Learning Workshop

Aplicaciones del Deep Learning

Detección de tiburones

SharkSpotter, Australia, 2017

Aplicaciones del Deep Learning

Traducción automática de señales

Google app

Otavio Good (Google Translate):

How Google Translate squeezes deep learning onto a phone

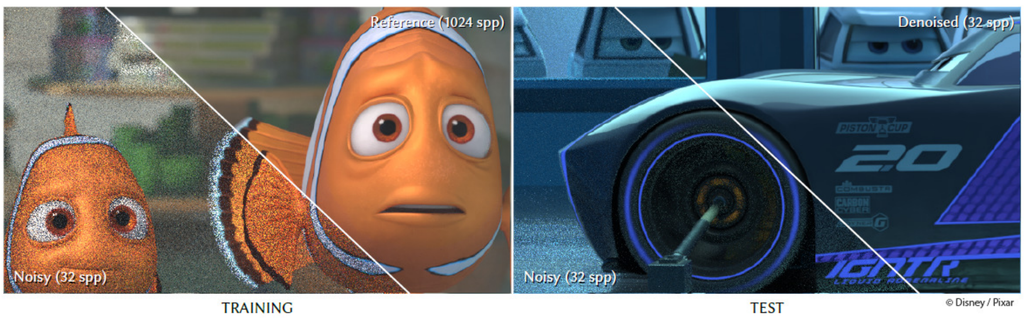
http://googleresearch.blogspot.com.es/2015/07/how-google-translate-squeezes-deep.html

Aplicaciones del Deep Learning

Síntesis de imágenes

Eliminación de ruido

@ UCSB, Disney & Pixar



SIGGRAPH 2017

http://cvc.ucsb.edu/graphics/Papers/SIGGRAPH2017_KPCN/

Aplicaciones del Deep Learning

Retoque fotográfico

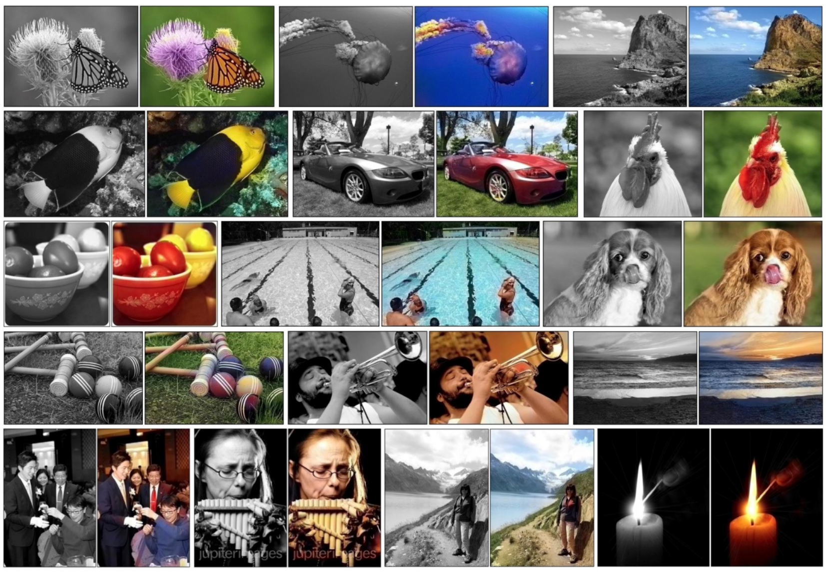
Figure 6: Original (top) vs. enhanced (bottom) images for iPhone 6, HTC One M9 and Huawei P9 cameras.

WESPE: Weakly Supervised Photo Enhancer for Digital Cameras. CVPR 2018. https://arxiv.org/abs/1709.01118

Aplicaciones del Deep Learning

Síntesis de imágenes

Coloreado de fotografías



ECCV 2016 http://richzhang.github.io/colorization/

Aplicaciones del Deep Learning

Síntesis de imágenes

Coloreado de fotografías interactivo

SIGGRAPH 2017

https://richzhang.github.io/ideepcolor/

Aplicaciones del Deep Learning

Síntesis de imágenes: "Inceptionism"

Usando una red ya entrenada para reconocer objetos…

"El grito"
Edvard Munch
… visto por una red neuronal

http://deepdreamgenerator.com

https://github.com/google/deepdream

Aplicaciones del Deep Learning

Síntesis de imágenes

Transferencia de estilos

A

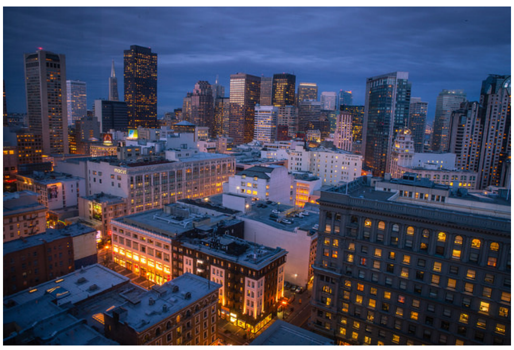
D

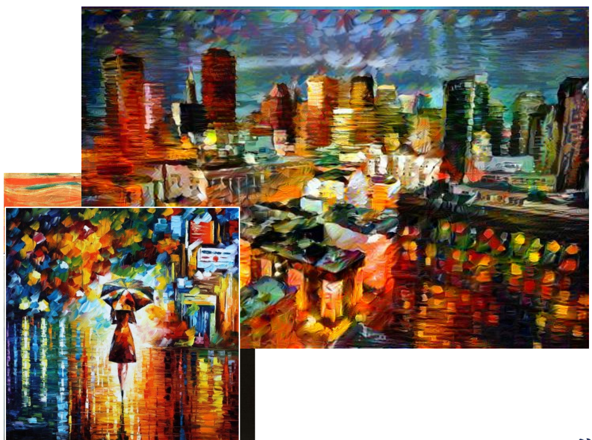
Leon A. Gatys, Alexander S. Ecker & Matthias Bethge:
A Neural Algorithm of Artistic Style
arXiv, 2015. http://arxiv.org/abs/1508.06576

Aplicaciones del Deep Learning

Síntesis de imágenes

Transferencia de estilos





Aplicaciones del Deep Learning

Transferencia de estilos

Aplicaciones del Deep Learning

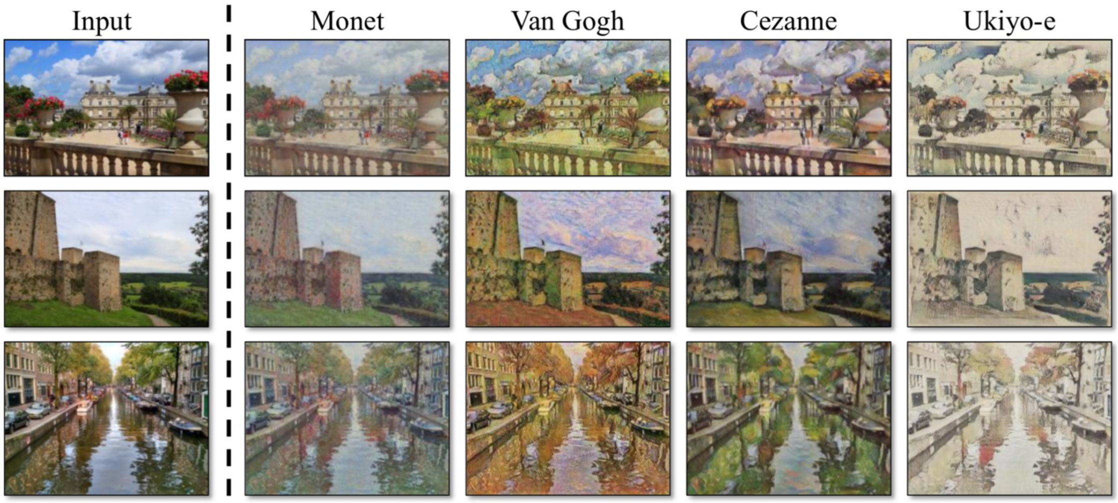
Síntesis de imágenes

https://thispersondoesnotexist.com/

StyleGAN https://arxiv.org/abs/1812.04948 CVPR'2019

Aplicaciones del Deep Learning

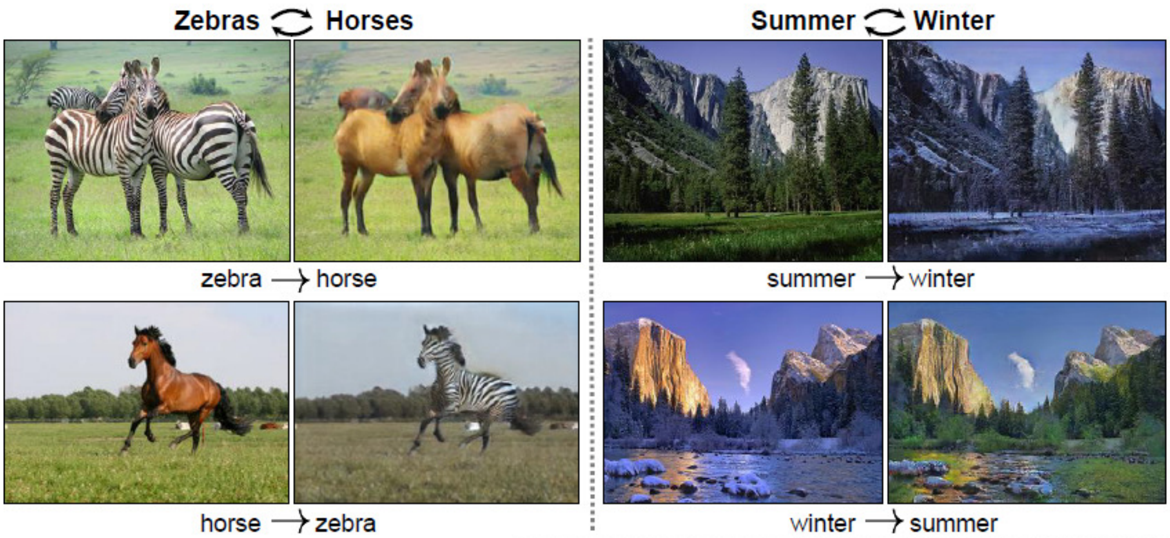
"Traducción de imágenes"



CycleGAN Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, ICCV'2017

Aplicaciones del Deep Learning

"Traducción de imágenes"



CycleGAN Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, ICCV'2017

Aplicaciones del Deep Learning

"Traducción de imágenes"

CycleGAN Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, ICCV'2017

Aplicaciones del Deep Learning

Unsupervised Image-to-Image Translation Networks, NIPS'2017

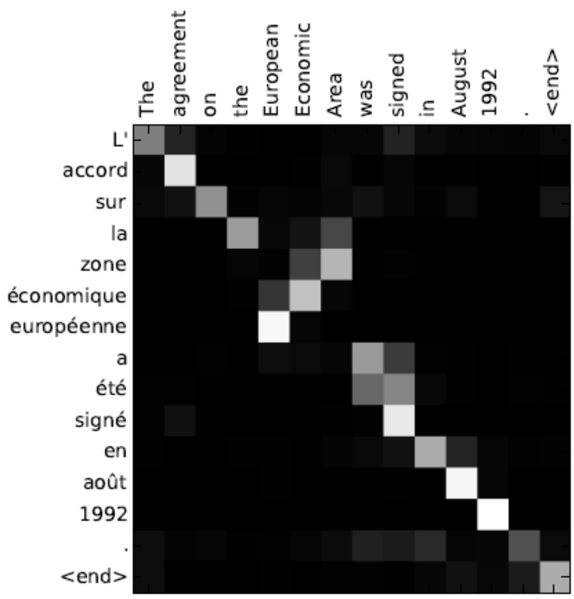
Aplicaciones del Deep Learning

"You sketch, the AI paints"

GauGAN, NVIDIA, CVPR'2019

Aplicaciones del Deep Learning

Traducción automática [machine translation]

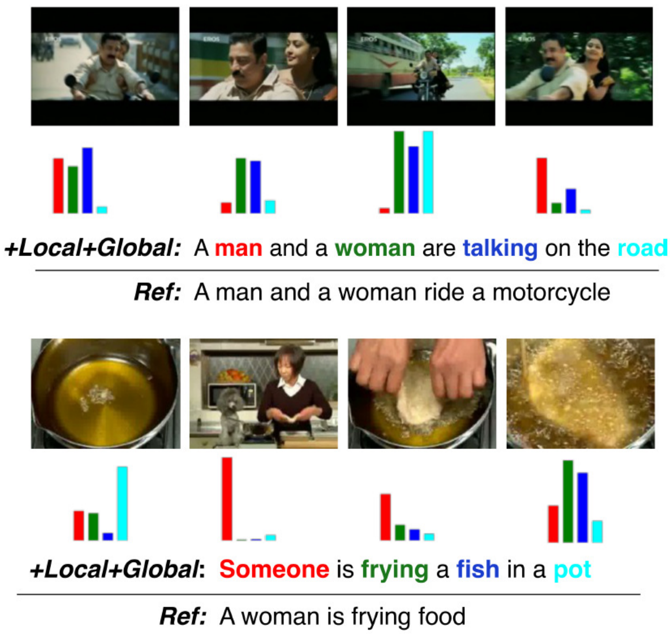


Aplicaciones del Deep Learning

Descripción textual de imágenes [image captioning]

Aplicaciones del Deep Learning

Descripción textual de vídeos [video clip description]



Youtube2Text

Aplicaciones del Deep Learning

Videojuegos (Atari 2600)

Google DeepMind

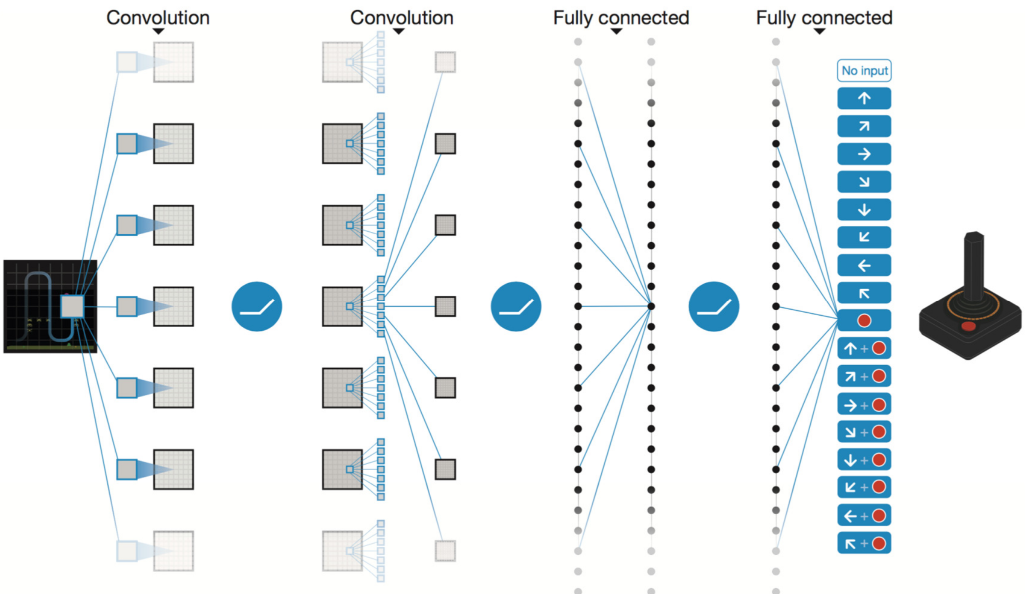
"Google AI beats humans at more classic arcade games than ever before"

http://arxiv.org/pdf/1509.06461v1.pdf (September 2015)

Aplicaciones del Deep Learning

Videojuegos

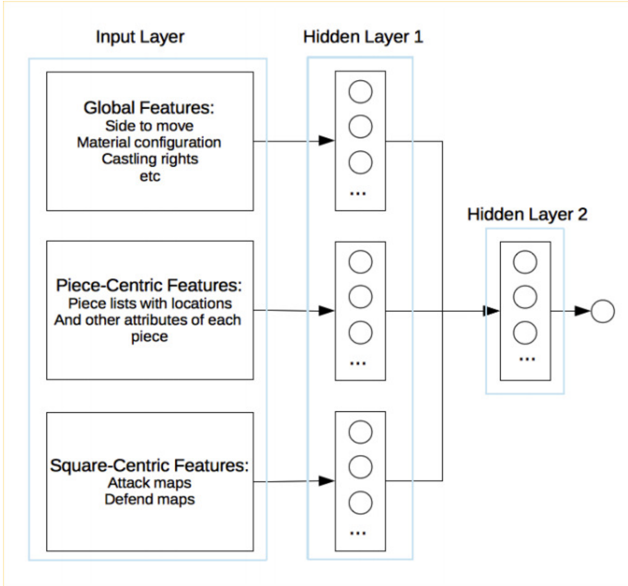
Deep Q Learning (Nature, 2015)



Aplicaciones del Deep Learning

Juegos

Ajedrez



Matthew Lai (Imperial College London):
"Giraffe: Using Deep Reinforcement Learning to Play Chess"
http://arxiv.org/abs/1509.01549 (September 2015)

Aplicaciones del Deep Learning

Juegos: Poker

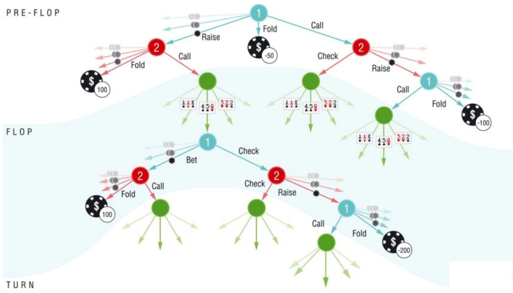
DeepStack

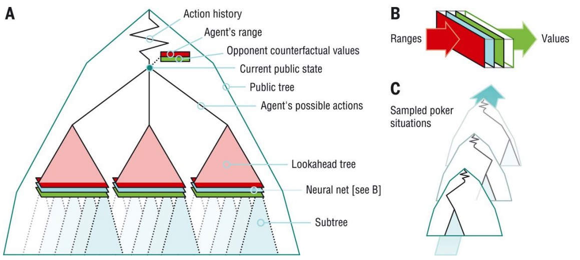
https://www.deepstack.ai/

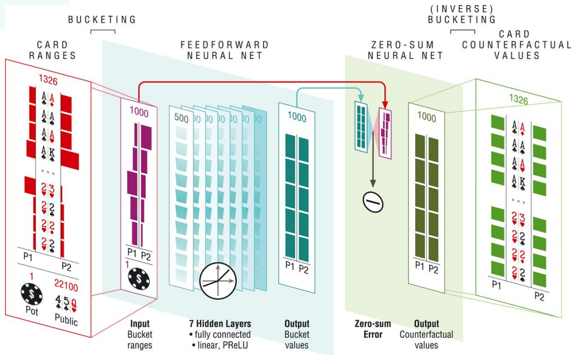
Aplicaciones

Juegos

Poker





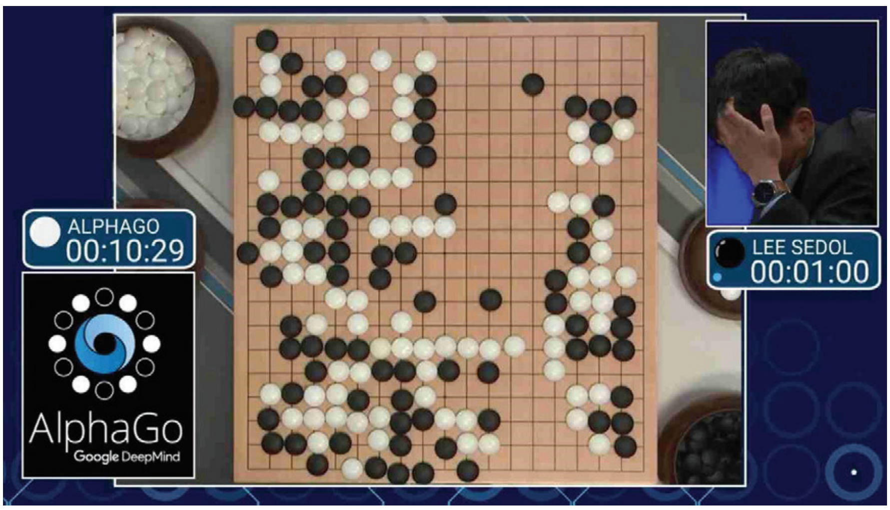


DeepStack: Expert-level artificial intelligence in heads-up no-limit poker
Science, Vol. 356, Issue 6337, pp. 508-513, 5 May 2017
DOI: 10.1126/science.aam6960

Aplicaciones del Deep Learning

Juegos: Go

AlphaGo

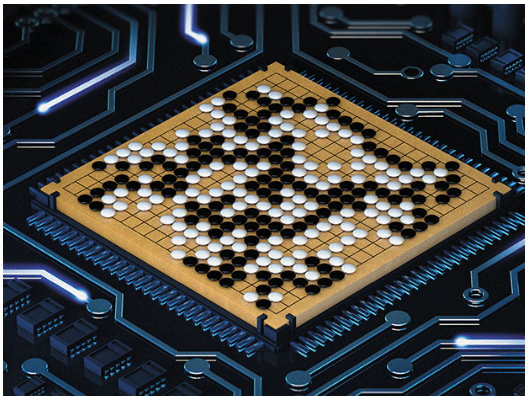


https://deepmind.com/research/alphago/

Aplicaciones

Juegos

Go: Los campeones humanos se negaban a jugar contra ordenadores porque eran demasiado malos (b>300)…



Octubre 2015, Londres:
AlphaGo (Google DeepMind) vence al campeón europeo Fan Hui [2-dan], 5-0.

Marzo de 2016, Seúl: $1M
AlphaGo (Google DeepMind) vence a Lee Sedol [9-dan], 4-1.

https://en.wikipedia.org/wiki/AlphaGo

AlphaGo

Aplicaciones del Deep Learning

Juegos

AlphaGo Zero

AlphaZero: Una máquina se ense…

https://elpais.com/elpais/2018/12/05/ciencia/1544007034_265553.html

REDES NEURONALES & DEEP LEARNING

amazon.es

REDES NEURONALES & DEEP LEARNING

TAPA BLANDA

(7)

€41,59 prime

ESP | AME | BRA | CAT | ENG

NEWSLETTER SUSCRÍBETE

EL PAÍS

Materia

ALPHAZERO ›

Una máquina se enseña a sí misma a ganar en todo

DeepMind desarrolla una inteligencia artificial invencible en los tres juegos de tablero más complejos gracias a un algoritmo que compite sin instrucciones

188

JAVIER SALAS

7 DIC 2018 - 16:29 CET

https://elpais.com/elpais/2018/12/05/ciencia/1544007034_265553.html

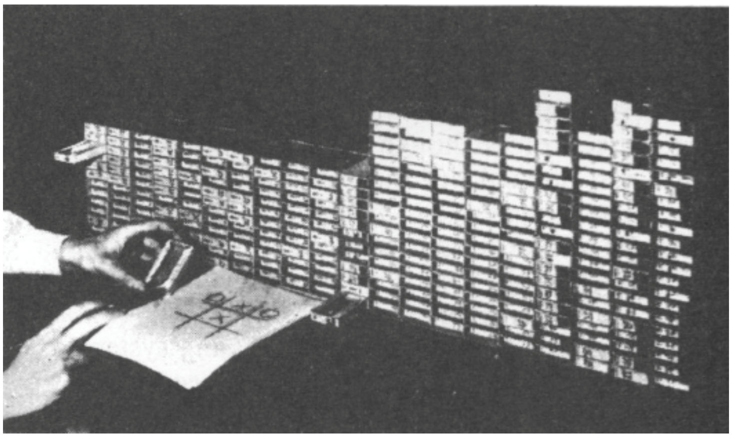
Aplicaciones del Deep Learning

Juegos: Go

Really???

The Matchbox Machine

1961



MENACE
Matchbox Educable Noughts And Crosses Engine

Donald Michie:
"Experiments on the mechanization of game-learning
Part I. Characterization of the model and its parameters"
The Computer Journal, 6(3):232–236, November 1963,
The British Computer Society, https://doi.org/10.1093/comjnl/6.3.232

Limitaciones del Deep Learning

Limitaciones

Cualquier cosa que requiera razonar, planificar a largo plazo o manipular datos de forma algorítmica está fuera del alcance de las técnicas actuales de deep learning.

p.ej. Ordenar un conjunto de datos puede ser extremadamente difícil usando una red neuronal.

Limitaciones del Deep Learning

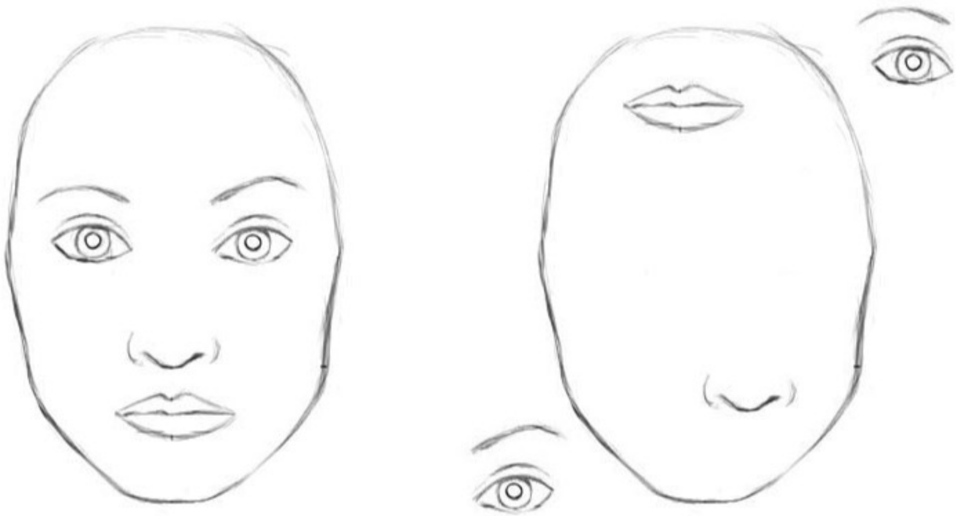
Limitaciones

Falta de compresión (en sentido humano) …

The boy is holding a baseball bat.

Limitaciones del Deep Learning

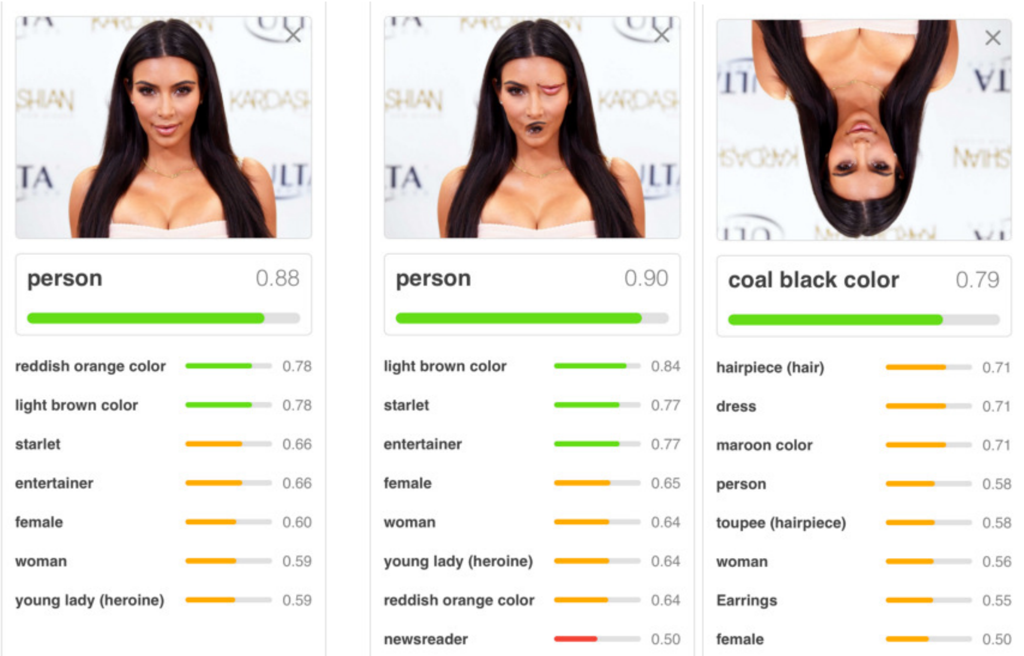
Las redes convolutivas [CNNs] funcionan muy bien en la práctica, pero…



… para una CNN, ambas imágenes son similares ☹

Limitaciones del Deep Learning

"Convolutional neural networks are doomed"
—Geoffrey Hinton

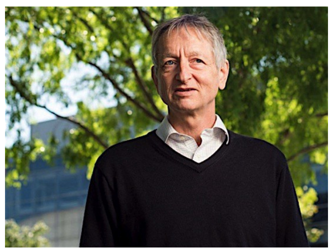


Limitaciones del Deep Learning

- Las redes convolutivas detectan características, pero no su colocación relativa (traslación & rotación).

- Las redes convolutivas ignoran las posiciones relativas utilizando "pooling", un apaño que funciona sorprendentemente bien en la práctica:

"The pooling operation used in convolutional neural networks is a big mistake and the fact that it works so well is a disaster." – Geoffrey Hinton



Limitaciones del Deep Learning

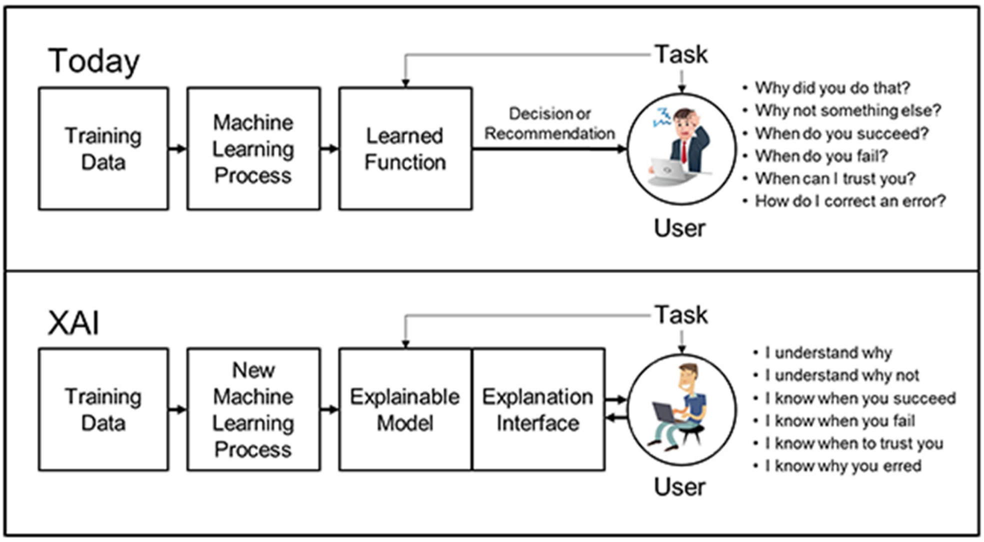
Problema clave de las redes convolutivas

La representación interna de una red convolutiva no tiene en cuenta las relaciones espaciales entre objetos, ni la jerarquía existente entre objetos simples y los objetos compuestos de los que forman parte.

Limitaciones del Deep Learning

Limitaciones

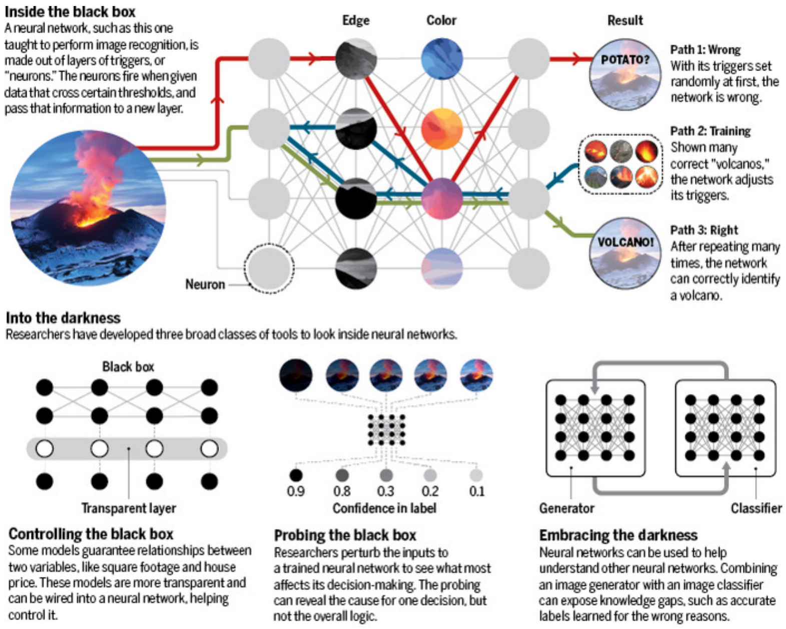
Falta de interpretabilidad:
Redes neuronales como cajas negras



Limitaciones del Deep Learning

Limitaciones

Falta de interpretabilidad



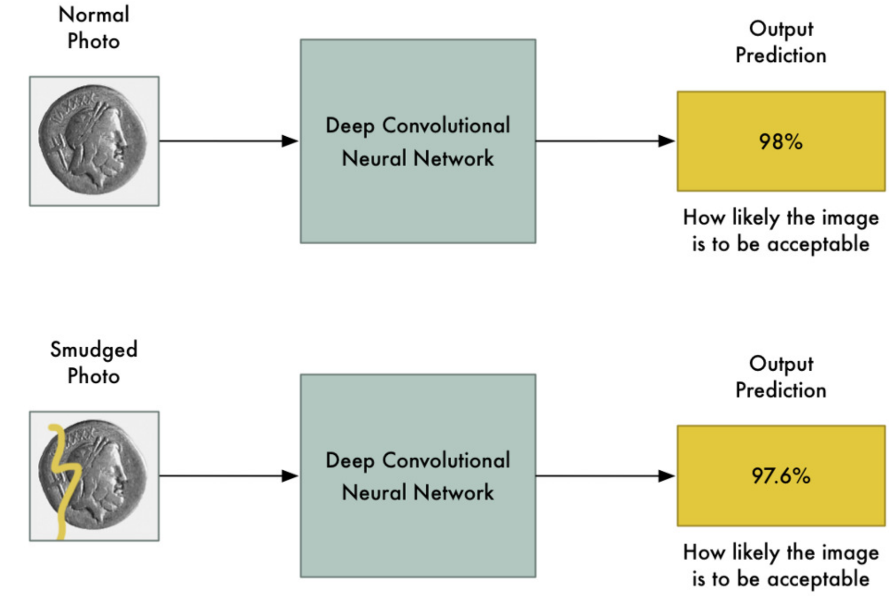
Mecanismos de atención

Limitaciones

Mecanismos de atención en la descripción textual de imágenes [image captioning]

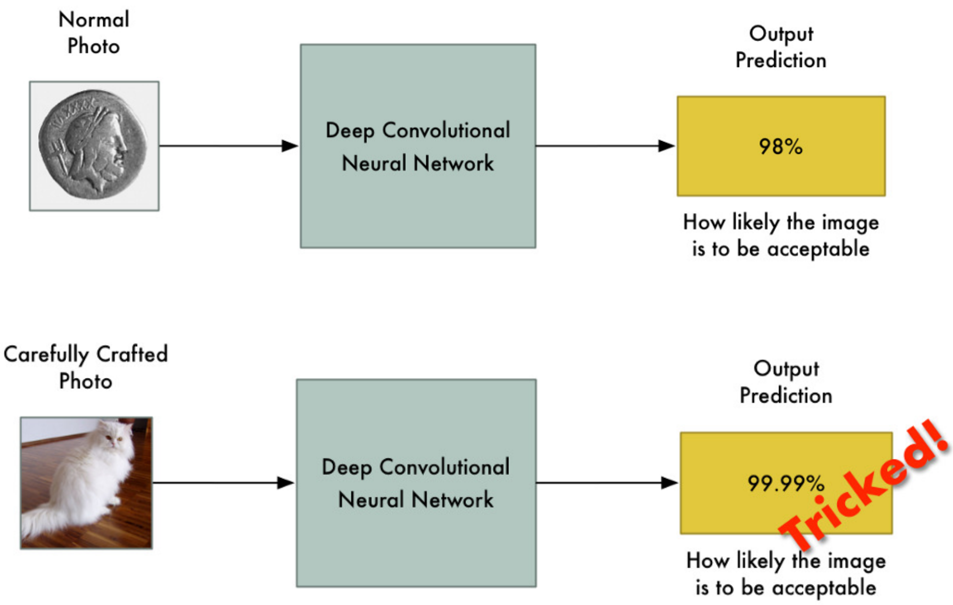
Limitaciones del Deep Learning

Lo deseable…



Limitaciones del Deep Learning

Lo que puede pasar…

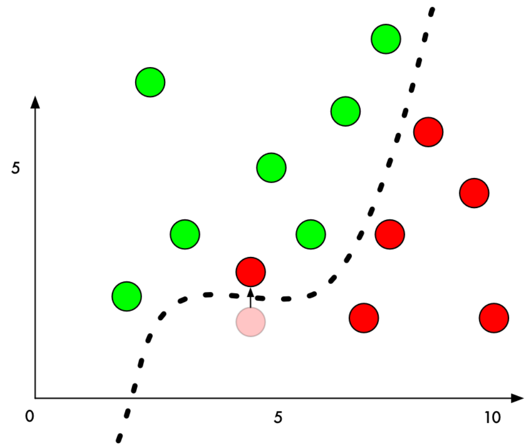


Limitaciones del Deep Learning

Ejemplos diseñados por un adversario
(o cómo engañar fácilmente a una red neuronal)

Si conocemos la red, podemos saber exactamente cómo modificar mínimamente la entrada para confundir a la red neuronal…

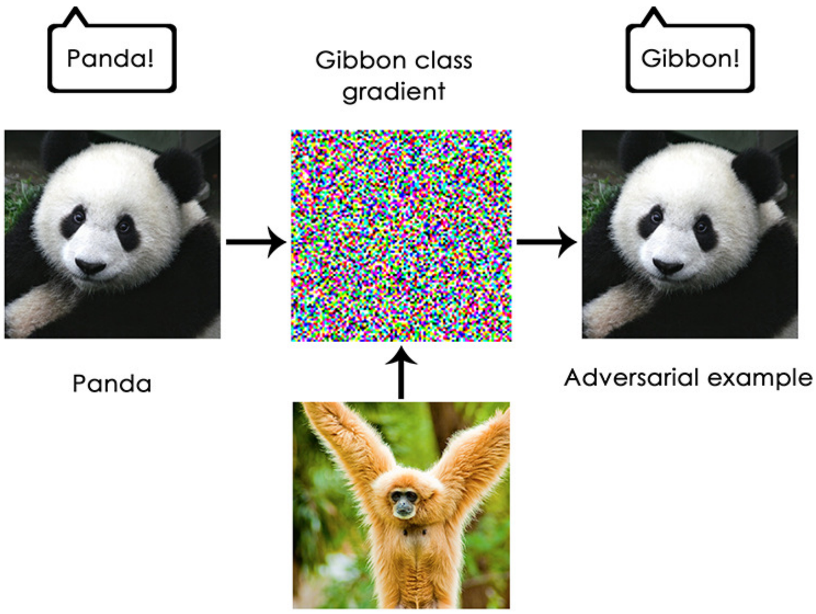
… en la dirección del gradiente !!!



Limitaciones del Deep Learning

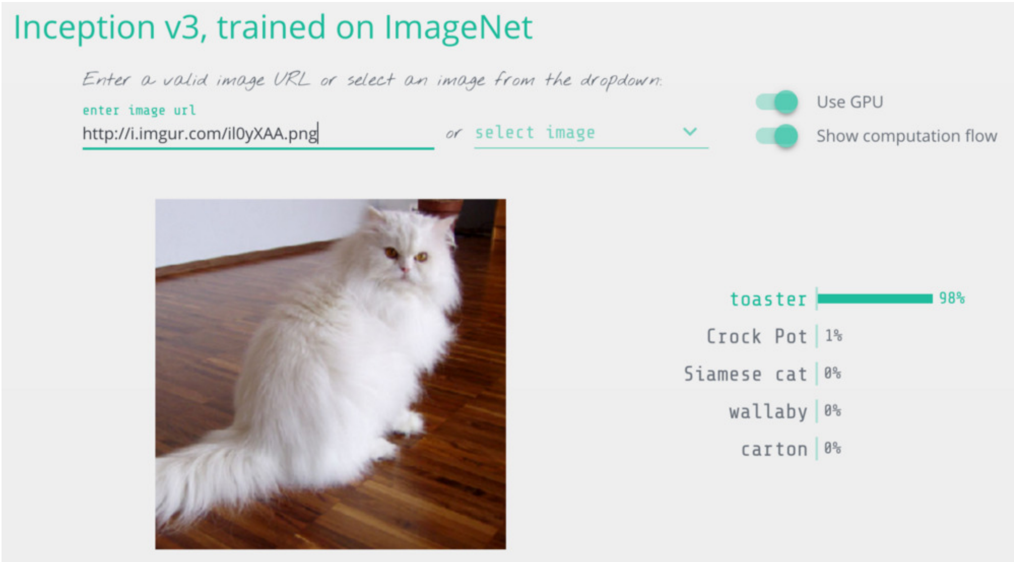
Limitaciones

Situaciones con adversario [adversarial examples]



Limitaciones del Deep Learning

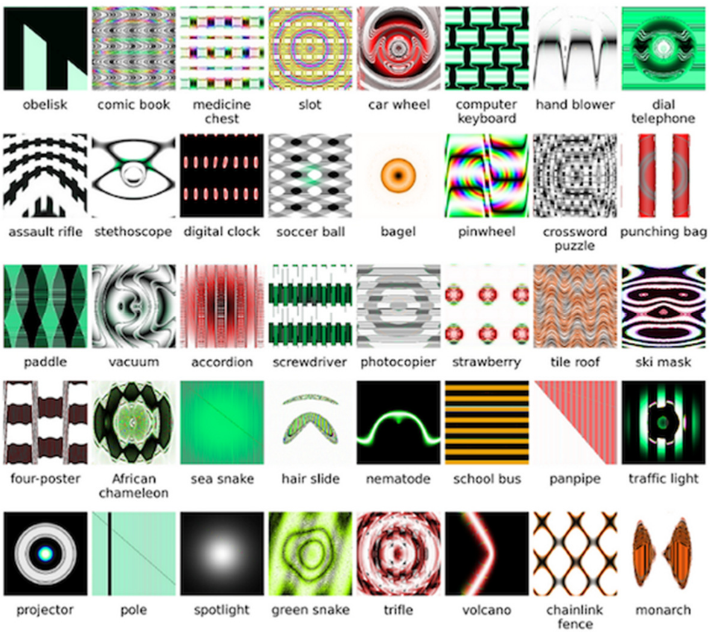
**Ejemplos diseñados por un adversario
(o cómo engañar fácilmente a una red neuronal)**



Limitaciones del Deep Learning

Limitaciones

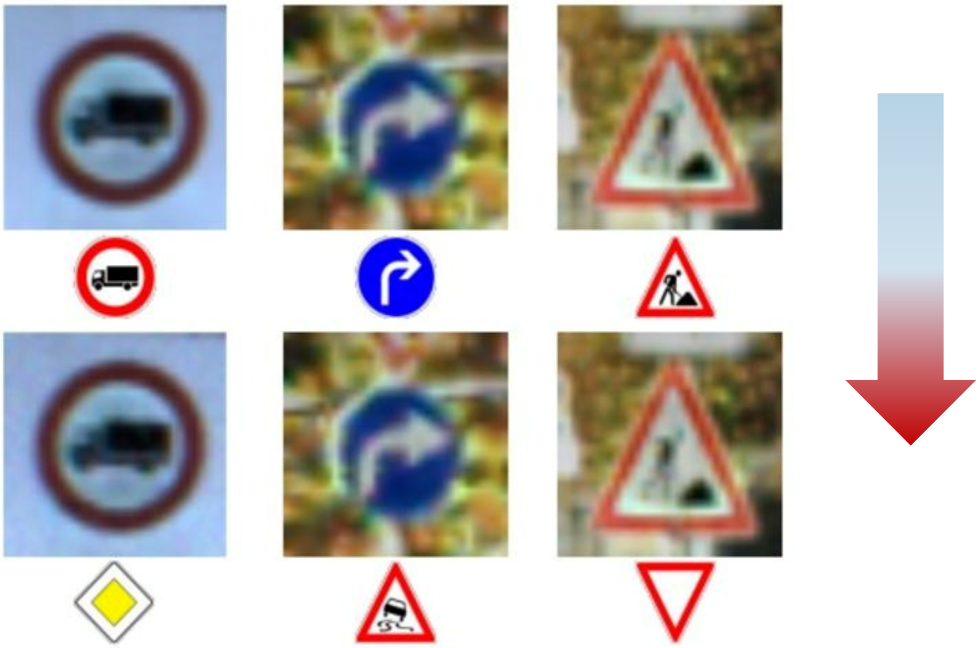
Situaciones con adversario [adversarial examples]



Limitaciones del Deep Learning

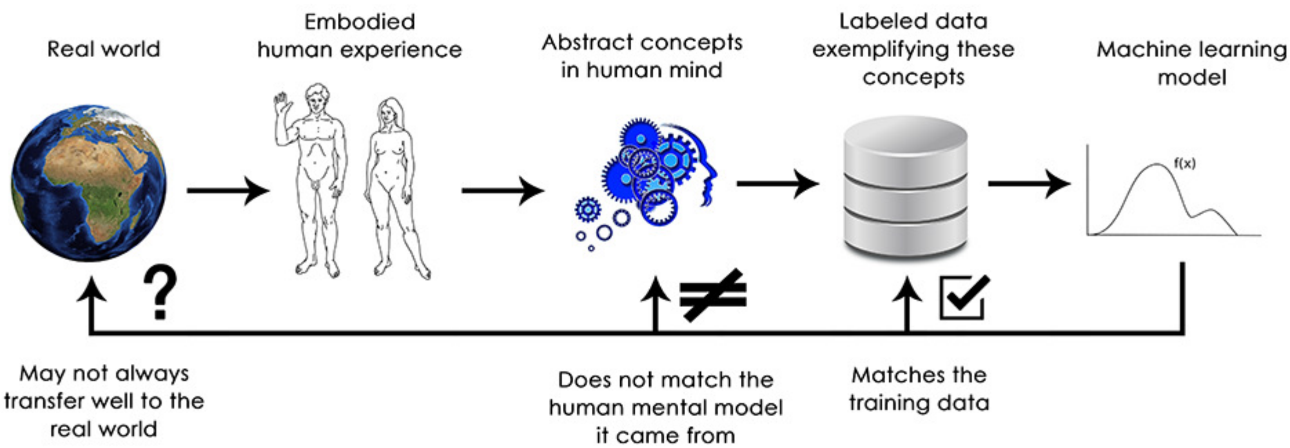
Limitaciones

Situaciones con adversario [adversarial examples]



Limitaciones del Deep Learning

Limitaciones

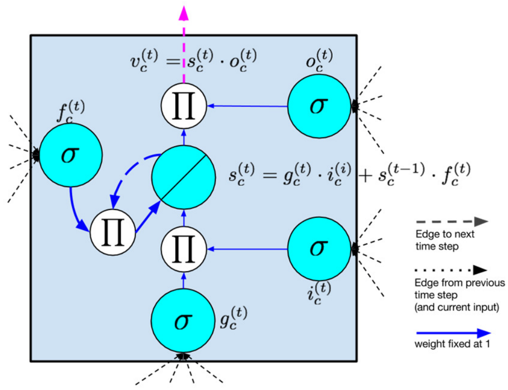


"Never fall into the trap of believing that neural networks understand the task they perform"

https://blog.keras.io/the-limitations-of-deep-learning.html

Deep Learning

Para algunos investigadores, que intentan obtener garantías teóricas basadas en resultados matemáticos: “[deep learning] might seem to be a regression” ;-)



LSTM (red recurrente)

En la práctica, los algoritmos con las mejores propiedades teóricas no son siempre los que mejor funcionan (sin restar importancia al estudio de las propiedades de los algoritmos de aprendizaje).

Deep Learning

Las técnicas heurísticas tienen éxito gracias a la disponibilidad de grandes conjuntos de datos (en los que el riesgo de sobreaprendizaje es menor) y la capacidad de cálculo de los sistemas actuales.

La validación con conjuntos de datos de prueba independientes ofrece una estimación de su comportamiento esperado en situaciones reales (los análisis teóricos se centran en el peor caso).

Few Things Are Guaranteed

When attainable, theoretical guarantees are beautiful. They reflect clear thinking and provide deep insight to the structure of a problem. Given a working algorithm, a theory which explains its performance deepens understanding and provides a basis for further intuition. Given the absence of a working algorithm, theory offers a path of attack.

However, there is also beauty in the idea that well-founded intuitions paired with rigorous empirical study can yield consistently functioning systems that outperform better-understood models, and sometimes even humans at many important tasks. Empiricism offers a path forward for applications where formal analysis is stifled, and potentially opens new directions that might eventually admit deeper theoretical understanding in the future.

Zachary Lipton:
“Deep Learning and the Triumph of Empiricism”
KDnuggets, July 2015

They are neither neural nor networks!

They are chains of differentiable, parameterized geometric functions, trained with gradient descent (obtained via chain rule)

A small set of high school level ideas put together

Francois Chollet
Creator of Keras
Google AI Researcher

https://twitter.com/fchollet/status/951906139632840704

Deep Learning

The only real success of deep learning so far has been the ability to map space X to space Y using a continuous geometric transform, given large amounts of human-annotated data. Doing this well is a game-changer for essentially every industry, but it is still a very long way from human-level AI.

… machine learning models could be defined as "learnable programs"; currently we can only learn programs that belong to a very narrow and specific subset of all possible programs. But what if we could learn *any* program, in a modular and reusable way?

-- François Chollet: "The limitations of deep learning"
https://blog.keras.io/the-limitations-of-deep-learning.html

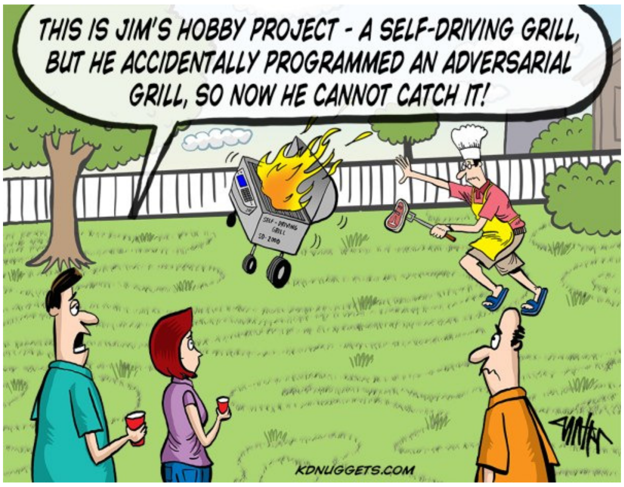
Deep Learning

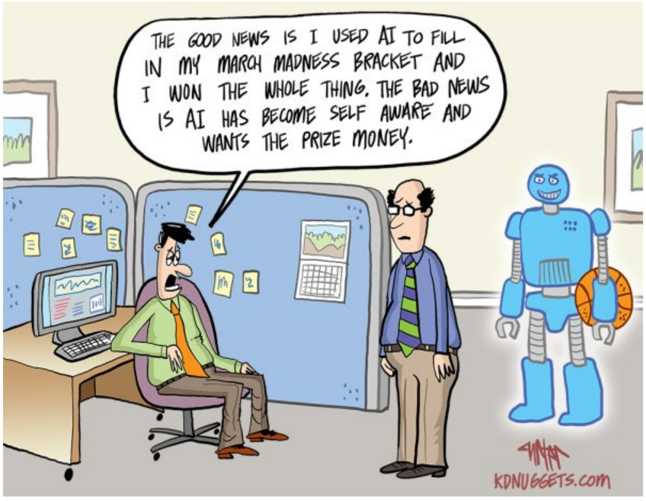
With all due respect
to the brilliant Geoff Hinton,
thought is not a vector,
and AI is not a problem in statistics.

-- Oren Etzioni

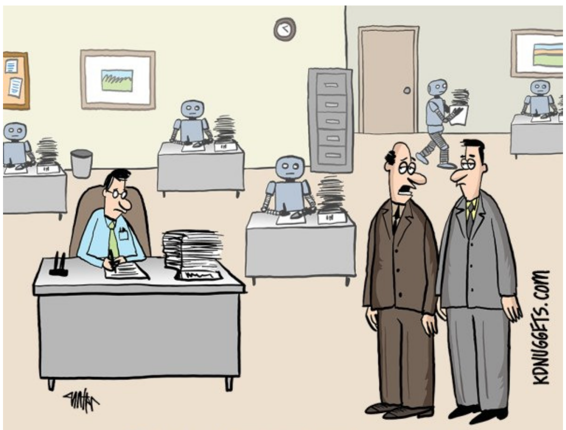
Shortcomings of Deep Learning
https://www.kdnuggets.com/2016/11/shortcomings-deep-learning.html

Deep Learning



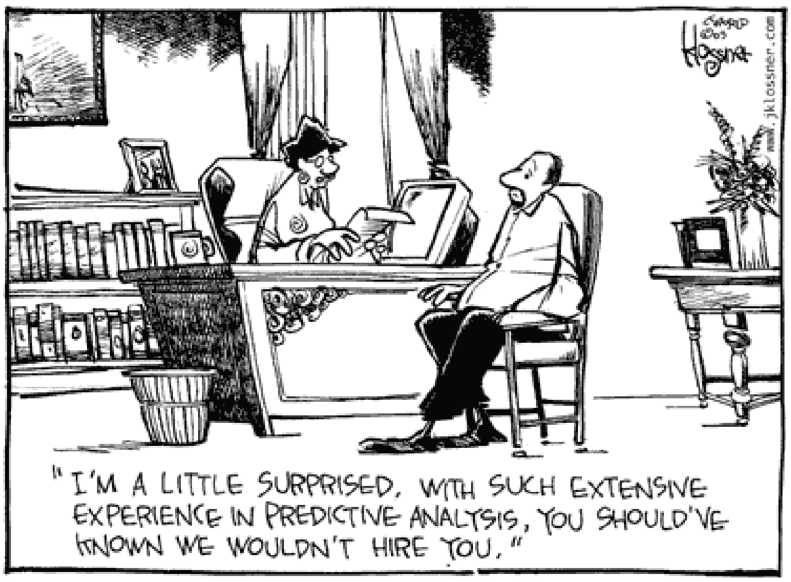


Deep Learning

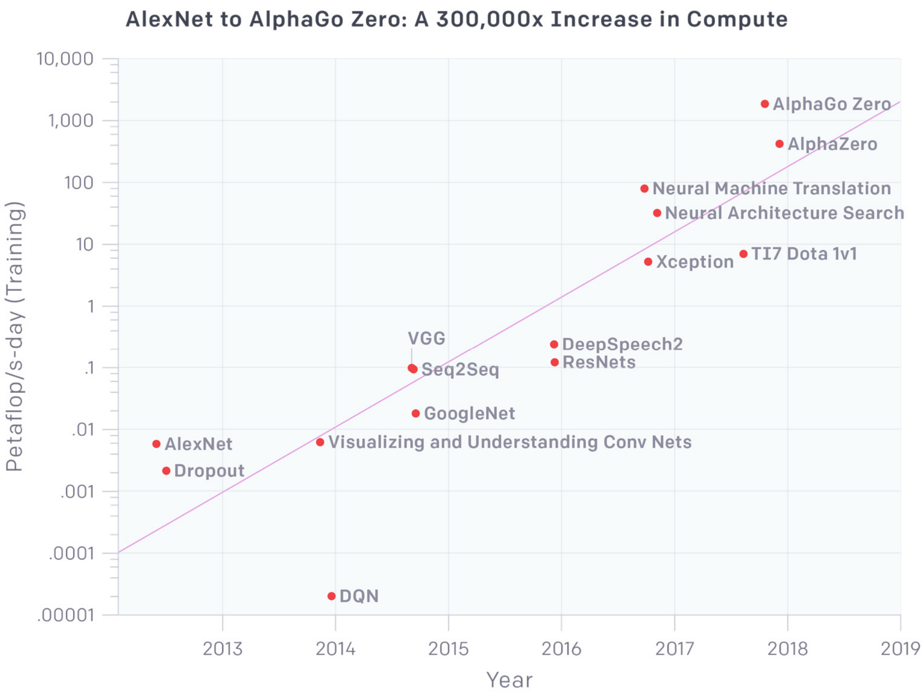


"The robots have taken over almost all jobs, but we still need Jenkins to cheat on our taxes"

Deep Learning

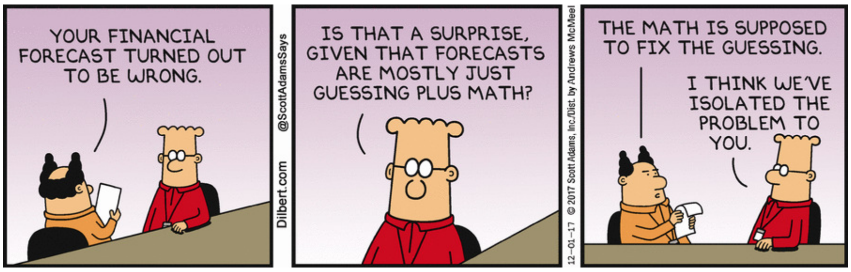


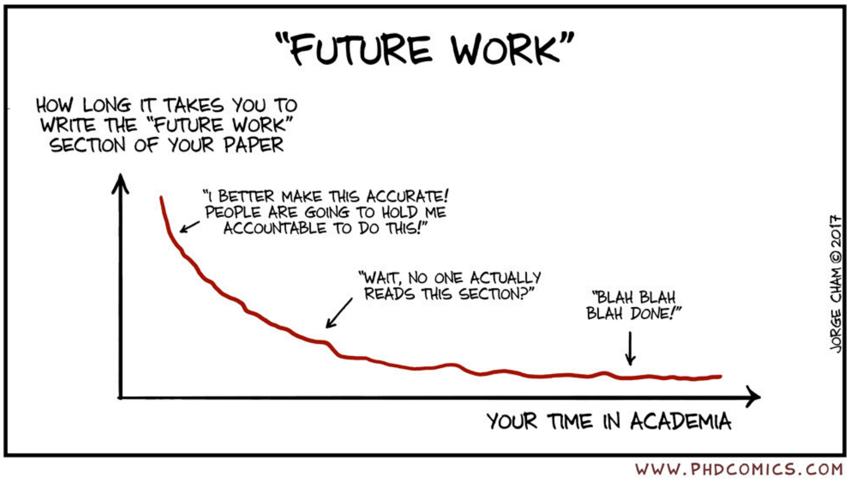
El futuro…



https://openai.com/blog/ai-and-compute/

El futuro…





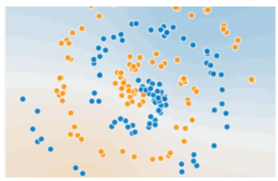
Demos

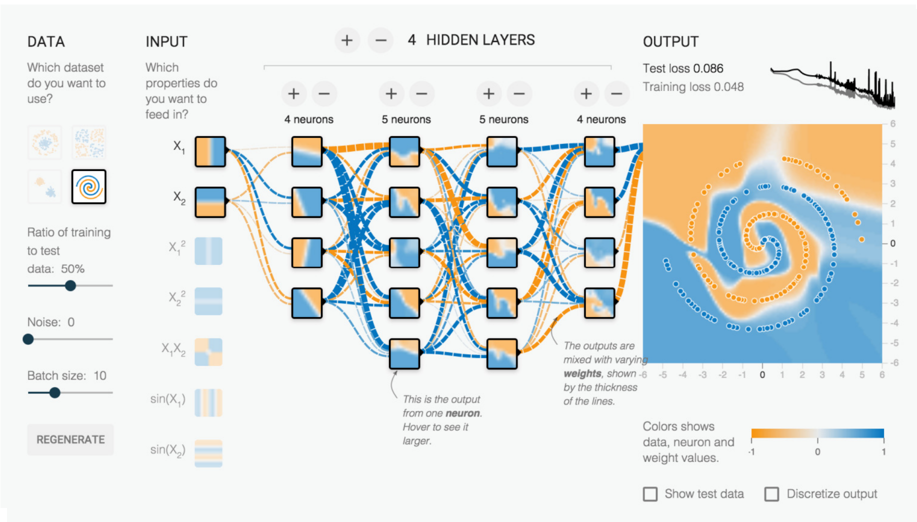
Para jugar un poco…

http://playground.tensorflow.org/

http://ml4a.github.io/demos/

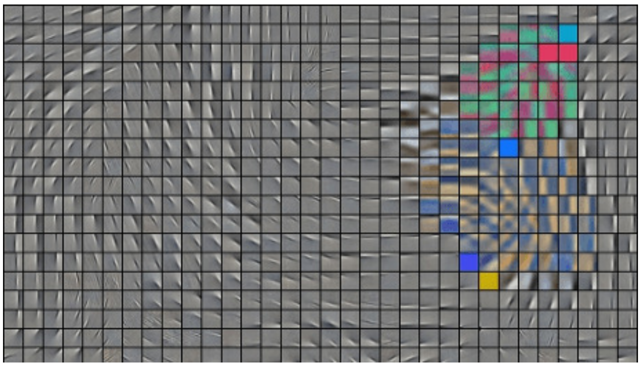
http://demos.algorithmia.com/classify-places/





Cursos

Neural Networks for Machine Learning
by Geoffrey Hinton
(University of Toronto & Google)
https://www.coursera.org/course/neuralnets



Cursos

Deep Learning Specialization
by Andrew Ng, 2017

- Neural Networks and Deep Learning
- Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
- Structuring Machine Learning Projects
- Convolutional Neural Networks
- Sequence Models

https://www.coursera.org/specializations/deep-learning

Cursos & Tutoriales

- **Deep Learning Tutorial**
  Andrew Ng et al. (Stanford University)
  http://ufldl.stanford.edu/tutorial/

- **Deep Learning: Methods and Applications**
  Li Deng & Dong Yu (Microsoft Research)
  http://research.microsoft.com/apps/pubs/default.aspx?id=209355

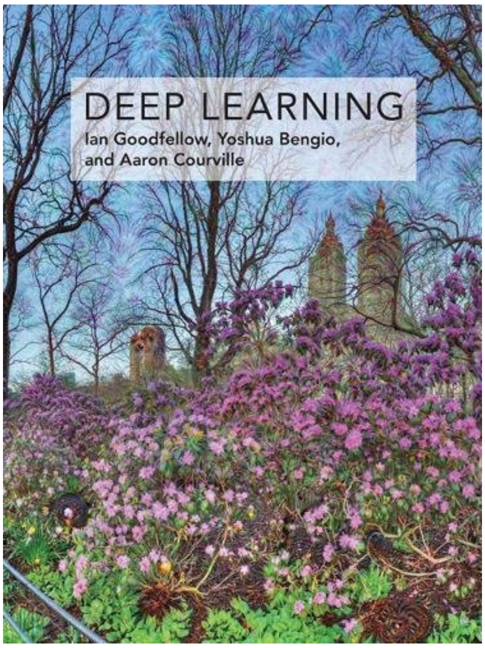
- **Deep Learning for Natural Language Processing**
  Richard Socher et al. (Stanford University CS224d)
  http://cs224d.stanford.edu/

- **Convolutional Neural Networks for Visual Recognition**
  Andrej Karpathy (Stanford University CS231n)
  http://cs231n.github.io/

Bibliografía

Lecturas recomendadas

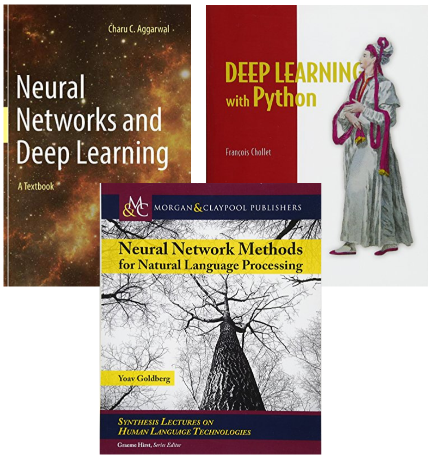
Ian Goodfellow,
Yoshua Bengio
& Aaron Courville:
Deep Learning
MIT Press, 2016
ISBN 0262035618



http://www.deeplearningbook.org

Bibliografía

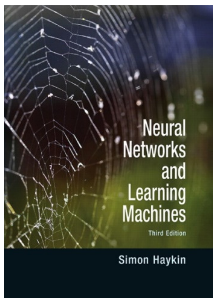
Lecturas complementarias

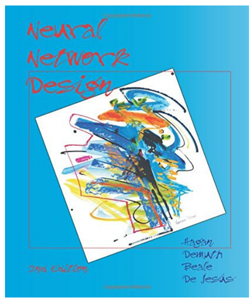


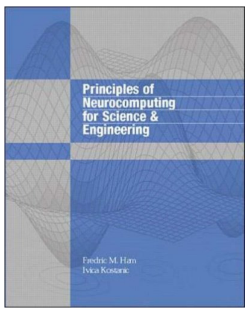
- Charu C. Aggarwal:
  Neural Networks and Deep Learning: A Textbook. Springer, 2018
  ISBN 3319944622
  http://link.springer.com/978-3-319-94463-0
- François Chollet:
  Deep Learning with Python
  Manning Publications, 2018
  ISBN 1617294438
  https://www.manning.com/books/deep-learning-with-python
- Yoav Goldberg:
  Neural Network Methods in Natural Language Processing
  Morgan & Claypool Publishers,, 2017
  ISBN 1627052984
  https://doi.org/10.2200/S00762ED1V01Y201703HLT037

Bibliografía

Lecturas complementarias





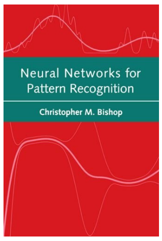


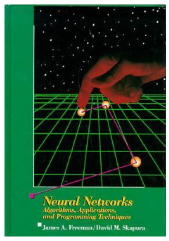
- Simon Haykin:
  Neural Networks and Learning Machines
  Prentice Hall, 3rd edition, 2008
  ISBN 0131471392
- Martin T. Hagan, Howard B. Demuth, Mark H. Beale & Orlando de Jesús:
  Neural Network Design
  Martin Hagan, 2nd edition, 2014
  ISBN 0971732116
  http://hagan.okstate.edu/NNDesign.pdf
- Fredric M. Ham & Ivica Kostanic:
  Principles of Neurocomputing for Science and Engineering
  McGraw-Hill Higher Education, 2000
  ISBN 0070259666

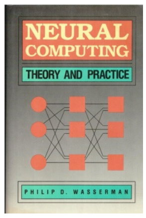
Bibliografía complementaria

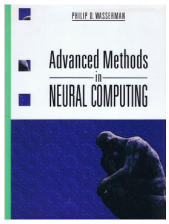
Redes neuronales artificiales

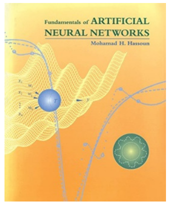
- Christopher M. Bishop:
  Neural Networks for Pattern Recognition
  Oxford University Press, 1996. ISBN 0198538642
- James A. Freeman & David M. Skapura:
  Neural Networks: Algorithms, Applications, and Programming Techniques
  Addison-Wesley, 1991. ISBN 0201513765
- Mohamad Hassoun:
  Fundamentals of Artificial Neural Networks
  MIT Press, 2003. ISBN 0262514672
- Philip D. Wasserman:
  Neural Computing: Theory and Practice,
  Van Nostrand Reinhold, 1989. ISBN 0442207433
  Advanced Methods in Neural Computing
  Van Nostrand Reinhold, 1993. ISBN 0442004613







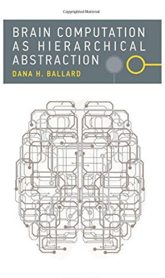


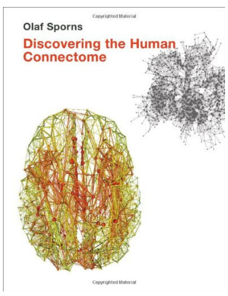


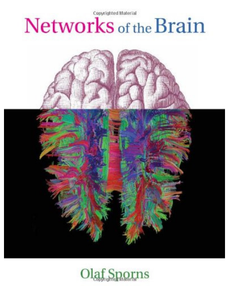
Bibliografía complementaria

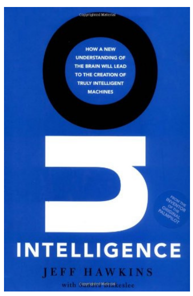
"wetware"

- Dana H. Ballard: **Brain Computation as Hierarchical Abstraction.** MIT Press, 2015. ISBN 0262028611
- Olaf Sporns: **Discovering the Human Connectome.** MIT Press, 2012. ISBN 0262017903
- Olaf Sporns: **Networks of the Brain**. MIT Press, 2010. ISBN 0262014696
- Jeff Hawkins: **On Intelligence.** Times Books, 2004. ISBN 0805074562





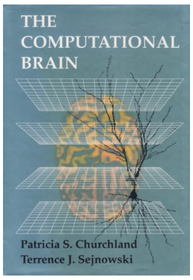


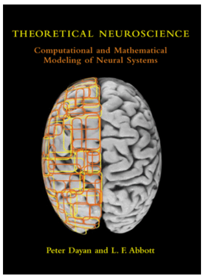


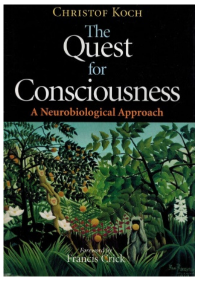
Bibliografía complementaria

Computational Neuroscience

- Patricia S. Churchland & Terrence J. Sejnowski:
  The Computational Brain. MIT Press, 1992. ISBN 0262031884
- Peter Dayan & L.F. Abbott: **Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.** MIT Press, 2001. ISBN 0262041995.
- Christof Koch: **The Quest for Consciousness: A Neurobiological Approach**.
  Roberts & Company Publishers, 2004. ISBN 0974707708



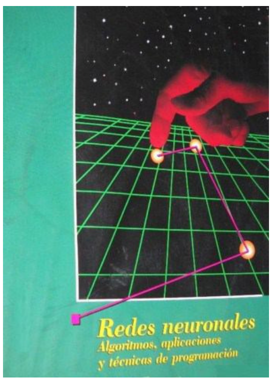




Bibliografía

Bibliografía en castellano

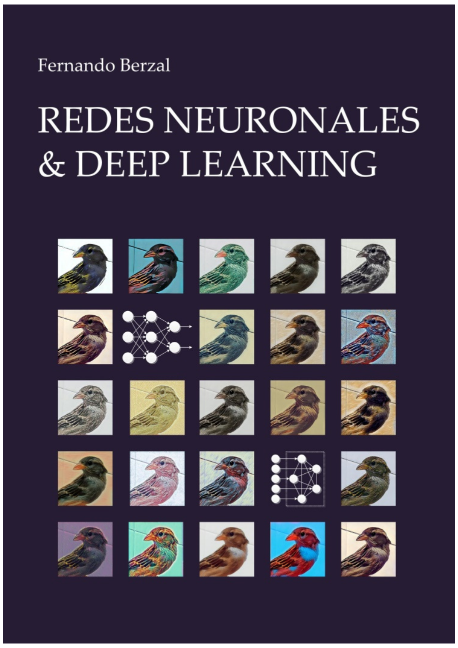
- James A. Freeman & David M. Skapura:
  Redes Neuronales: Algoritmos, aplicaciones y técnicas de programación
  Addison-Wesley / Díaz de Santos, 1993.
  ISBN 020160115X



Bibliografía

Bibliografía en castellano

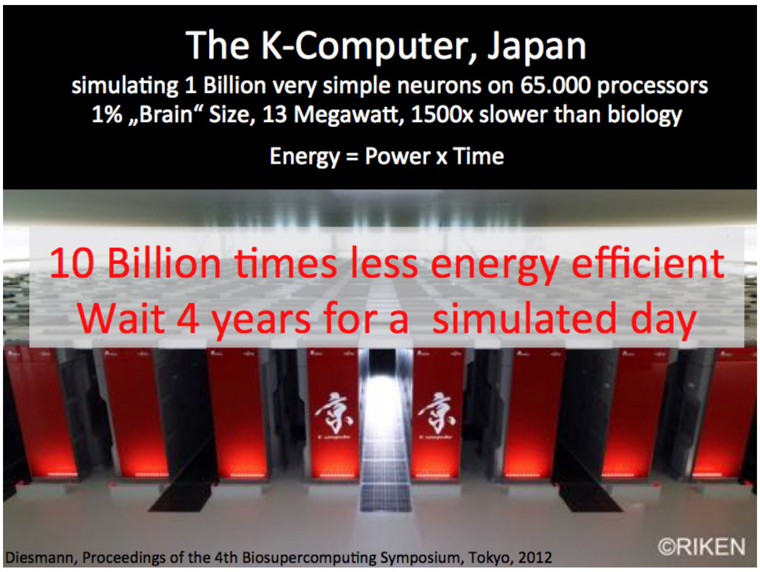
- Fernando Berzal:
  Redes Neuronales
  & Deep Learning
  Edición independiente, 2018
  ISBN 1-7312-6538-7 (b&n)
  ISBN 1-7313-1433-7 (color)
  http://deep-learning.ikor.org



Otros modelos de redes

Arquitecturas basadas en el cerebro

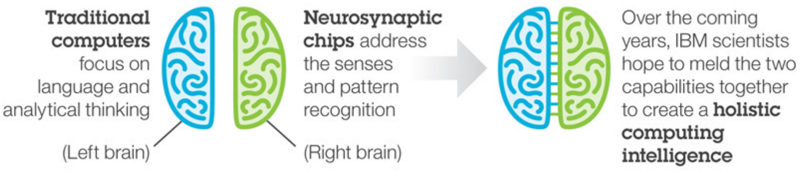
Simulación (muy ineficiente) ➔ "Neuromorphic Computing"

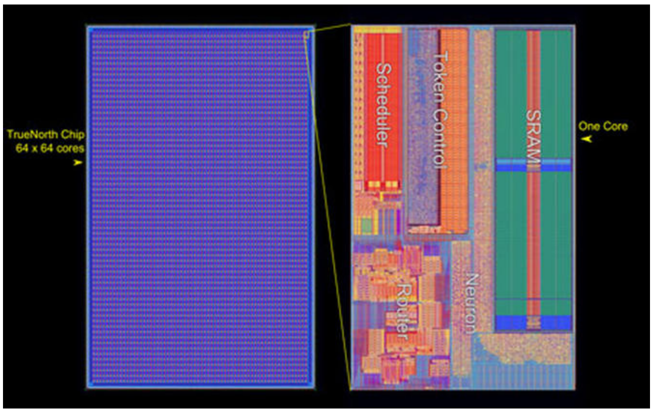


Otros modelos de redes

Arquitecturas basadas en el cerebro

IBM TrueNorth Brain-inspired Computer



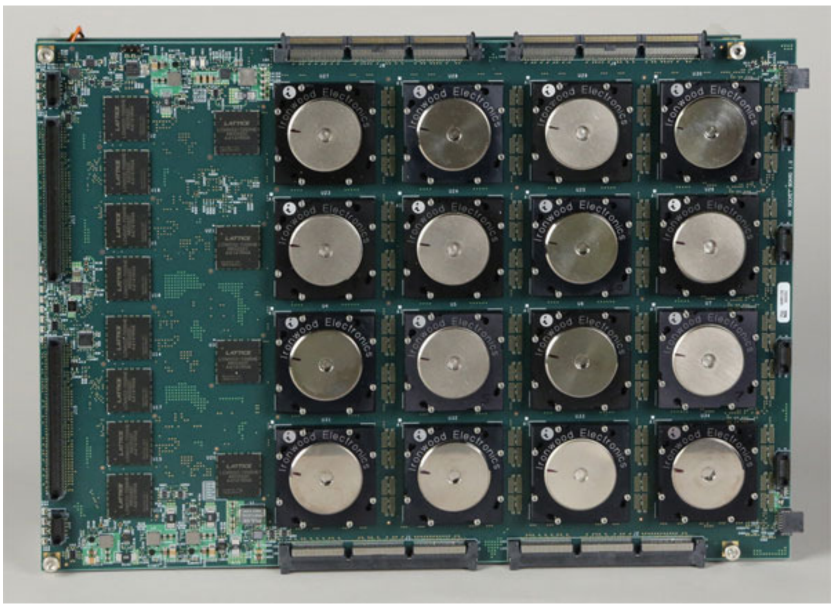


4096 cores
1M neurons
256M synapses
5.4B transistors
CMOS
70mW

Otros modelos de redes

Arquitecturas basadas en el cerebro

IBM TrueNorth Brain-inspired Computer

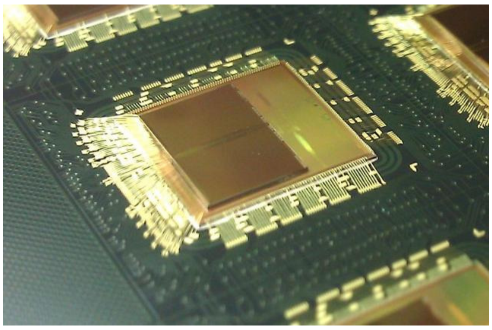


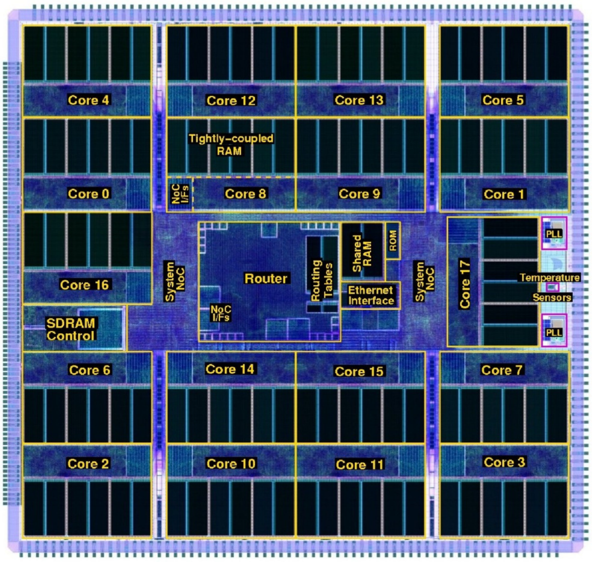
Synapse 16
16M neurons
4B synapses

Otros modelos de redes

Arquitecturas basadas en el cerebro

SpiNNaker project (UK)



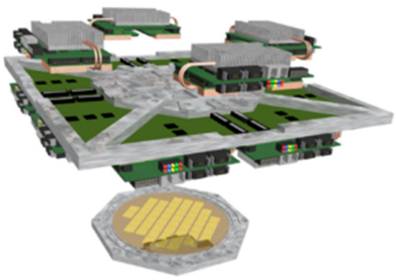


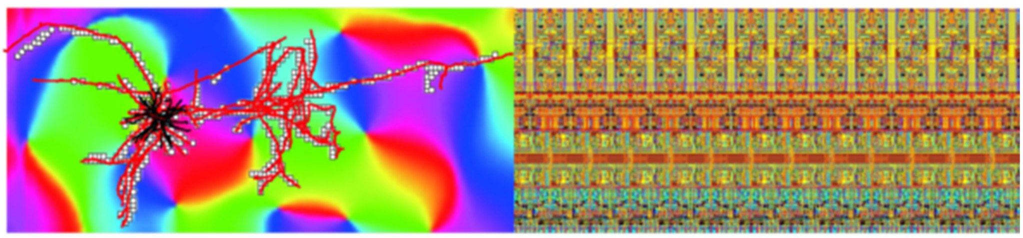
Globally Asynchronous Locally Synchronous (GALS) chip:
18 ARM968 processor nodes + 128MB Mobile DDR SDRAM
http://apt.cs.manchester.ac.uk/projects/SpiNNaker/project/

Otros modelos de redes

Arquitecturas basadas en el cerebro

BrainScaleS (Germany)





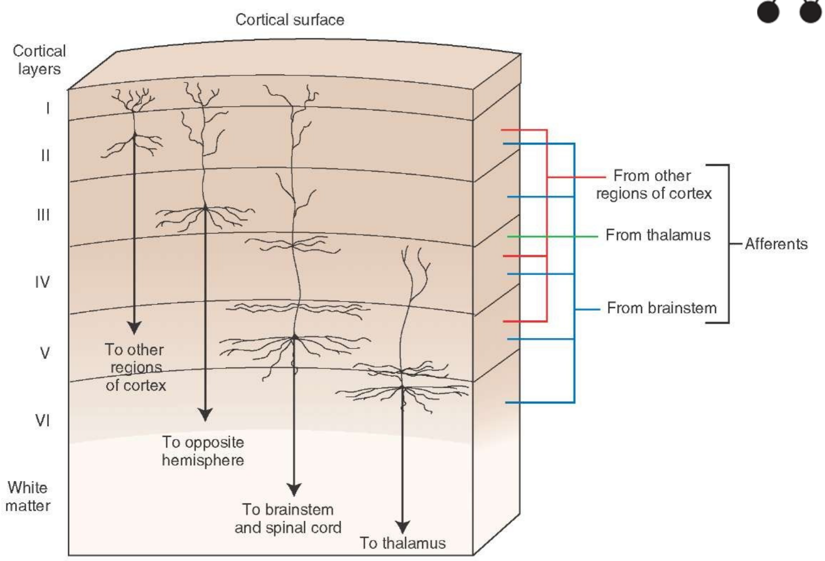
Mixed CMOS signals
https://brainscales.kip.uni-heidelberg.de/

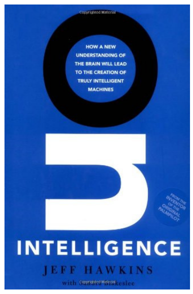
Otros modelos de redes

Arquitecturas basadas en el cerebro

HTM [Hierarchical Temporal Memory]

Numenta



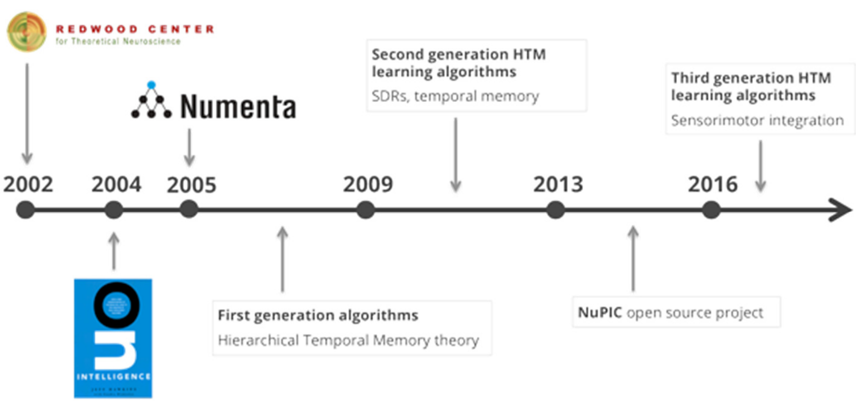


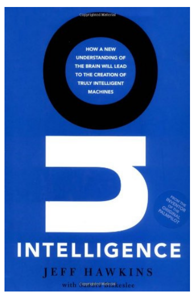
Otros modelos de redes

Arquitecturas basadas en el cerebro

HTM [Hierarchical Temporal Memory]

Numenta





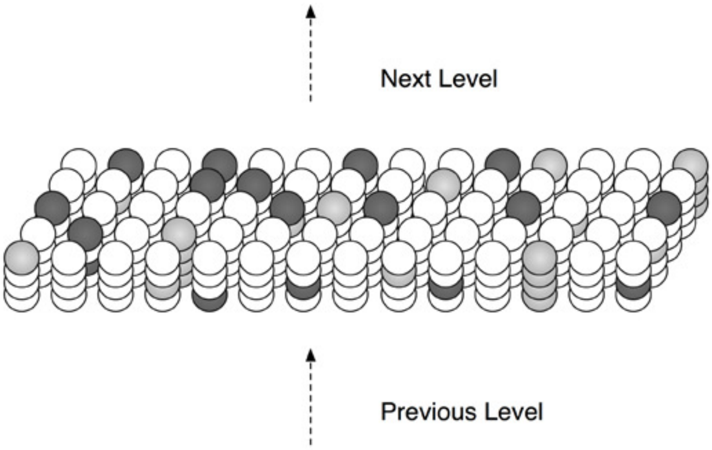
http://numenta.com/

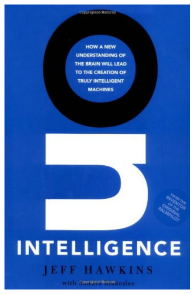
Otros modelos de redes

Arquitecturas basadas en el cerebro

HTM [Hierarchical Temporal Memory]

Numenta





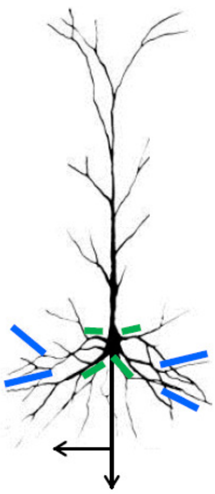
http://numenta.com/

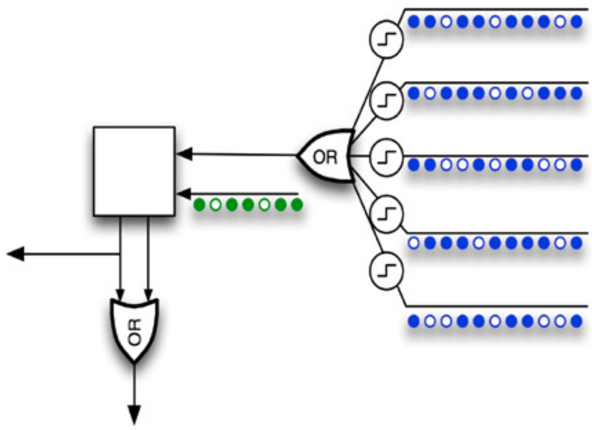
Otros modelos de redes

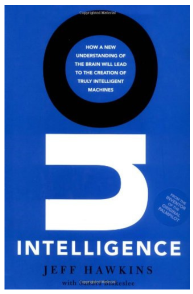
Arquitecturas basadas en el cerebro

HTM [Hierarchical Temporal Memory]

Numenta







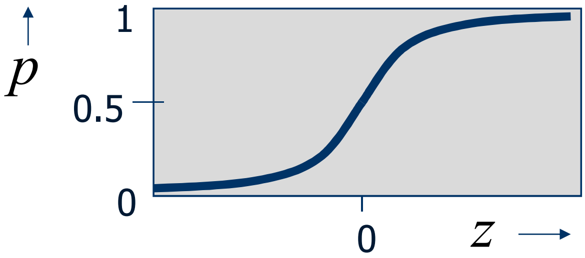
http://numenta.com/

Otros modelos de redes

Neuronas binarias estocásticas

$$z = \sum_i x_i w_i$$

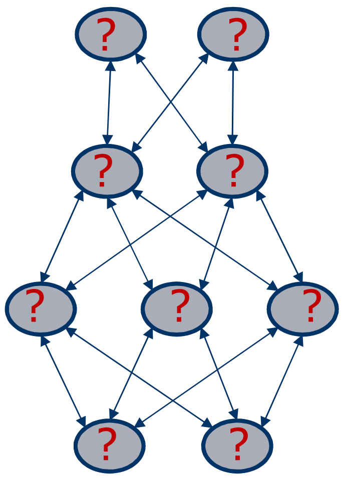
$$p = \frac{1}{1+e^{-z}}$$



Las mismas ecuaciones que las neuronas sigmoidales, si bien su salida se interpreta como una probabilidad (de producir un spike en una pequeña ventana de tiempo).

Otros modelos de redes

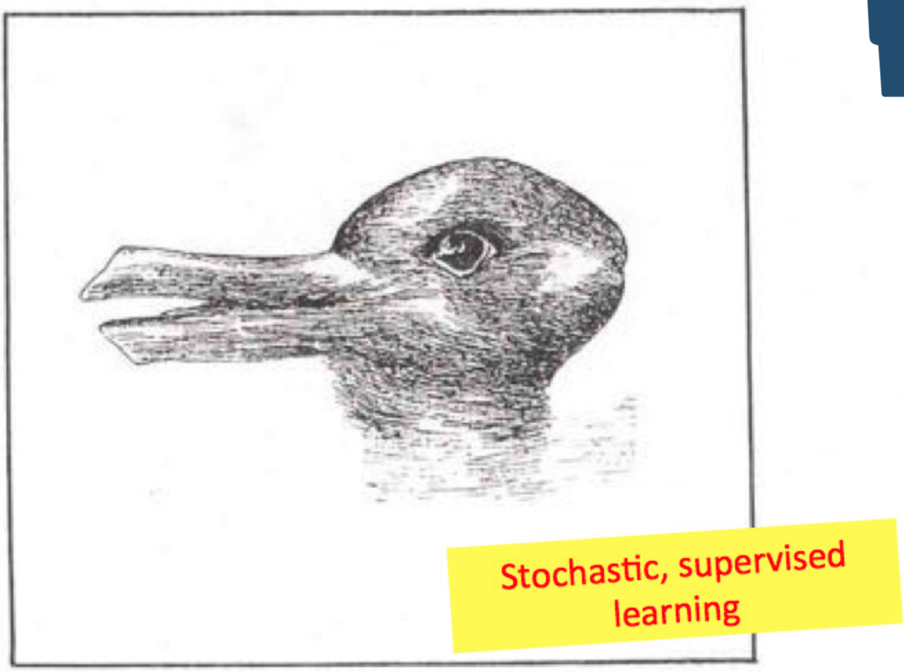
Máquinas de Boltzmann



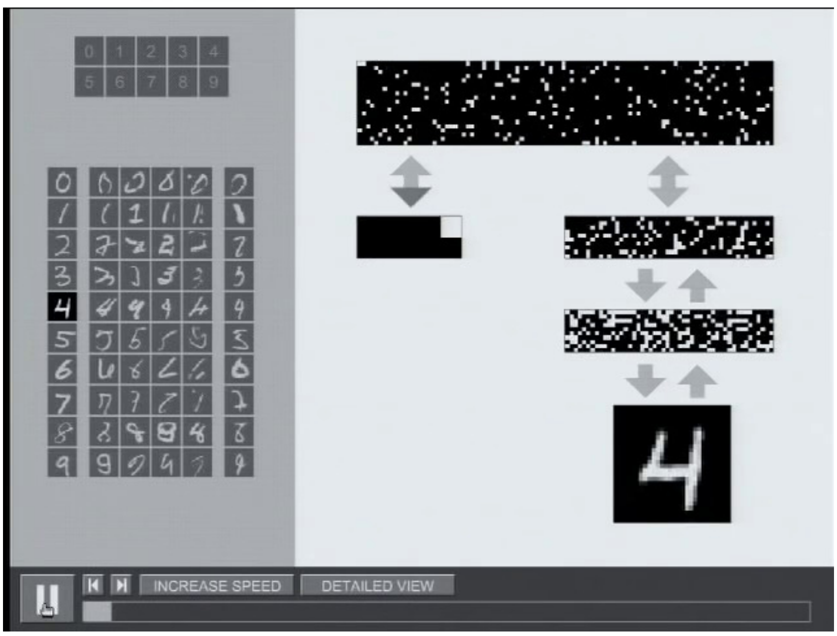
- En una máquina de Boltzmann, las actualizaciones estocásticas de las distintas unidades deben ser secuenciales.
- Existe una arquitectura que admite actualizaciones paralelas alternas mucho más eficientes: **DBM [Deep Boltzmann Machine]**
  - Sin conexiones entre unidades de una misma capa.
  - Sin conexiones entre capas no adyacentes.

Otros modelos de redes

Máquinas de Boltzmann



Otros modelos de redes



Geoffrey Hinton: "The Next Generation of Neural Networks"
Google Tech Talks, 2007
https://www.youtube.com/watch?v=AyzOUbkUf3M

Otros modelos de redes

Geoff Hinton doesn't need to make hidden units. They hide by themselves when he approaches.

Geoff Hinton doesn't disagree with you, he contrastively diverges

Deep Belief Nets actually believe deeply in Geoff Hinton.

Yann LeCun: "Geoff Hinton facts"
http://yann.lecun.com/ex/fun/index.html