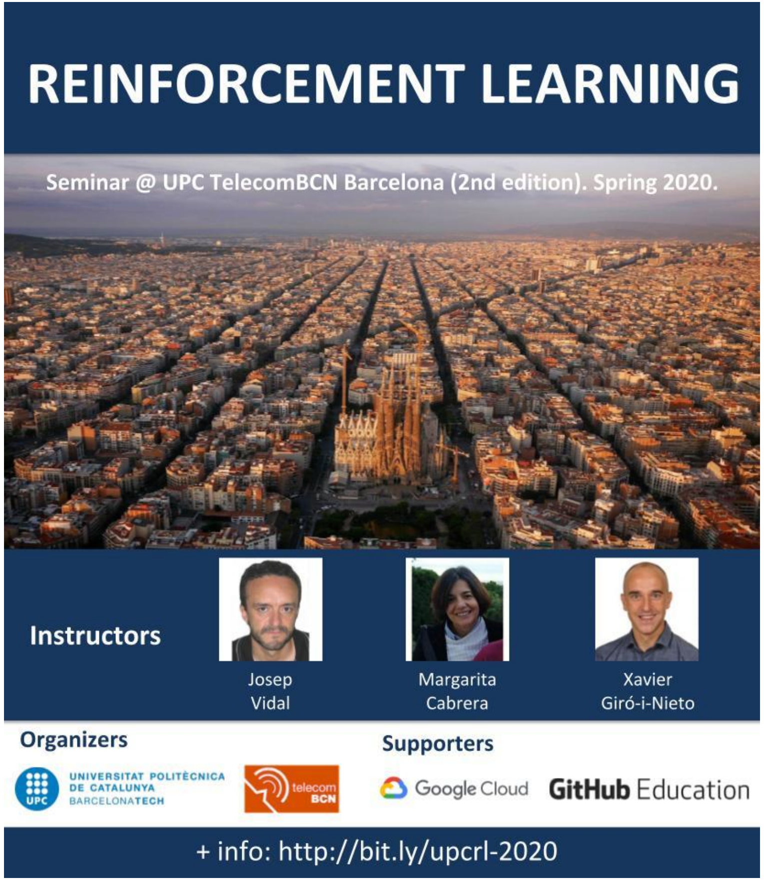


https://telecombcn-dl.github.io/mrl-2020/

MRL 2020 - Day 8 - Lecture 1

Deep Reinforcement Learning (DRL)

Xavier Giro-i-Nieto

@DocXavi
xavier.giro@upc.edu

Associate Professor
ETSETB TelecomBCN
Universitat Politècnica de Catalunya

Hola

Xavier Giro-i-Nieto

Associate Professor at Universitat Politècnica de Catalunya (UPC)

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Department of Signal Theory and Communications
Image Processing Group

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

IDEAI Center for Intelligent Data Science & Artificial Intelligence

telecom BCN

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
School of Professional & Executive Development

Master in Computer Vision Barcelona

BTS Barcelona Technology School

VRIJE UNIVERSITEIT BRUSSEL · SCIENTIA VINCERE TENEBRAS ·

SONY

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Research Group & Centers

Image Processing Group
Signal Theory and Communications Department
Universitat Politècnica de Catalunya. BARCELONATECH

- 11 faculty members
- 12 Phd students

https://imatge.upc.edu/

IDEAI (Intelligent Data Science and Artificial Intelligence)

- Center funded in 2017
- 60 researchers

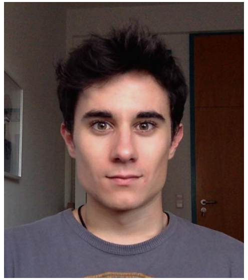
https://ideai.upc.edu/

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

- National computation center #1
- Supercomputer MareNostrum
- Emerging Technologies for Artificial Intelligence Group, directed by Prof. Jordi Torres.

https://www.bsc.es/

Acknowledgments



Víctor Campos
victor.campos@bsc.es

PhD Candidate

Barcelona Supercomputing Center

Míriam Bellver
miriam.bellver@bsc.edu

PhD Candidate
Barcelona Supercomputing Center

Reinforcement Learning (with extrinsic reward)

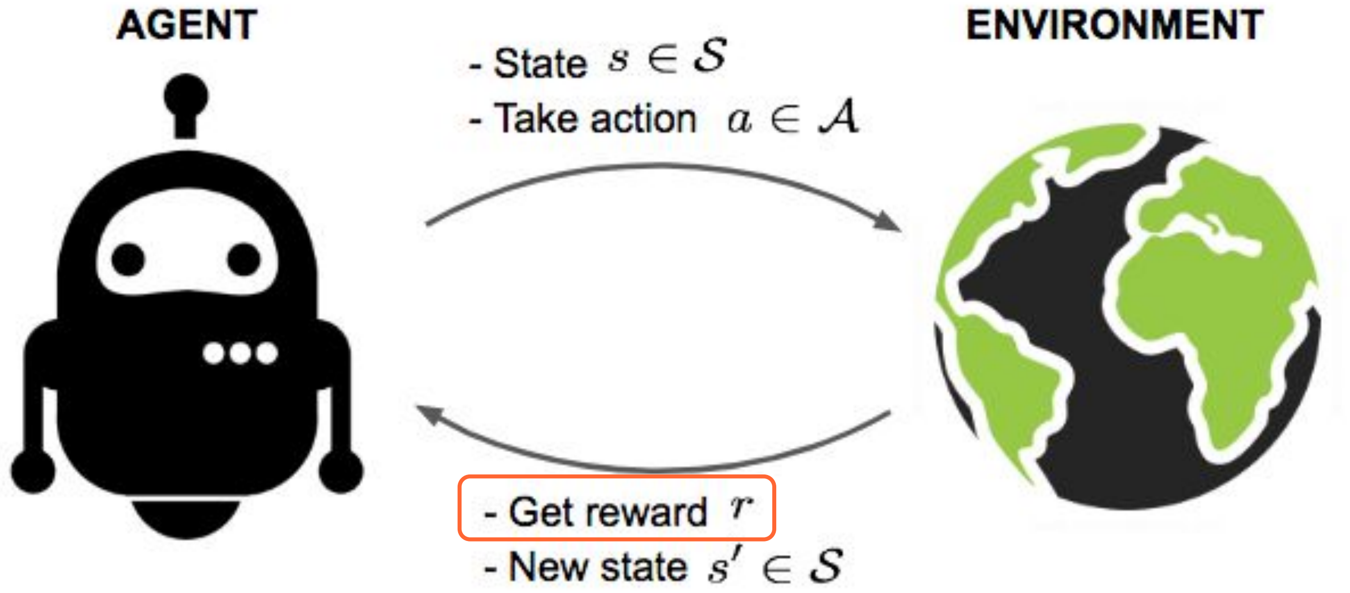
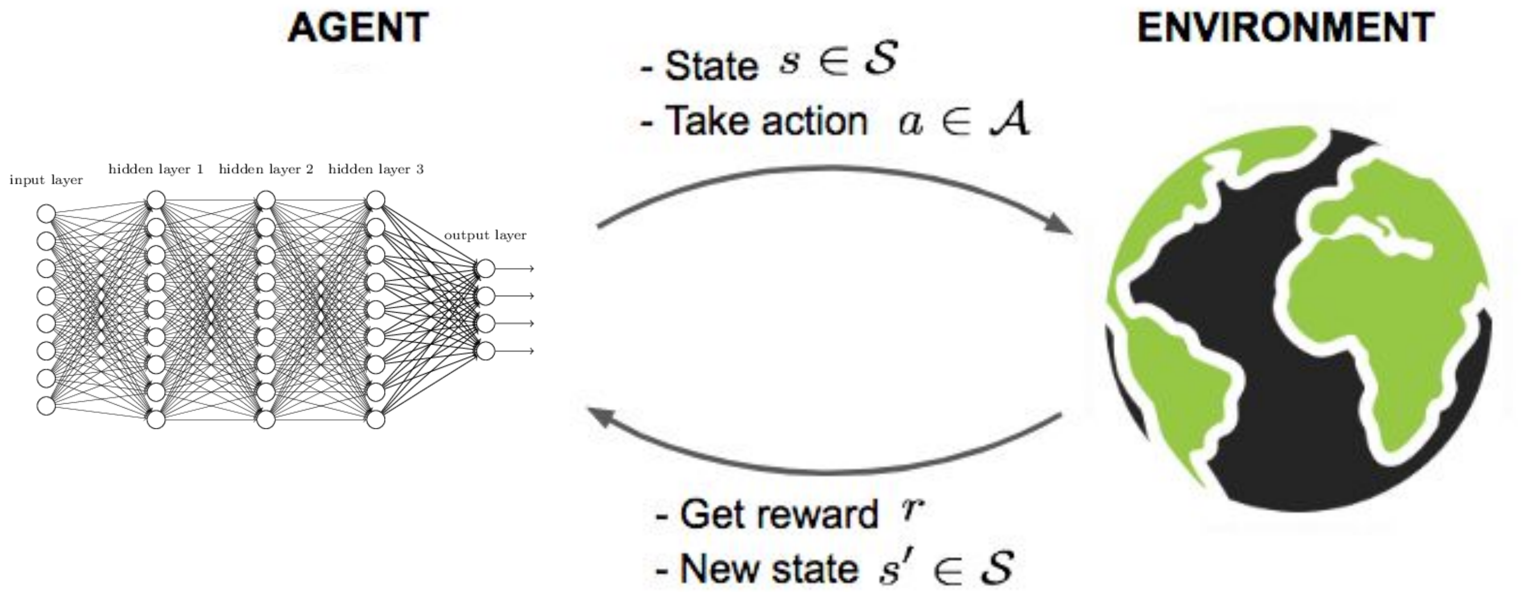


Figure: Lilian Weng, "A (Long) Peek into Reinforcement Learning" (2018)

Deep Reinforcement Learning (DRL)

Deep Reinforcement Learning (DRL) refers to agents controlled by deep neural networks.

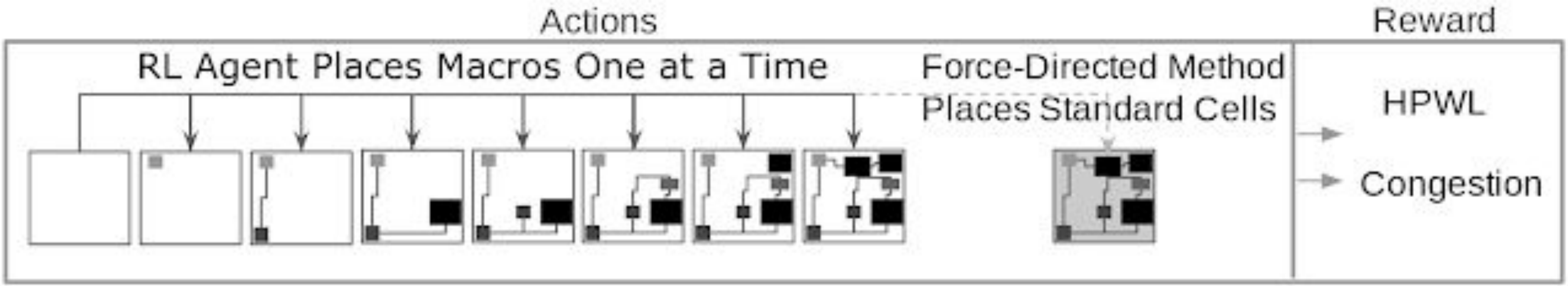


Wayve, "The first example of deep reinforcement learning on-board an autonomous car" (2018)

Covariant, "Bringing robots from lab to the real world" (2020)

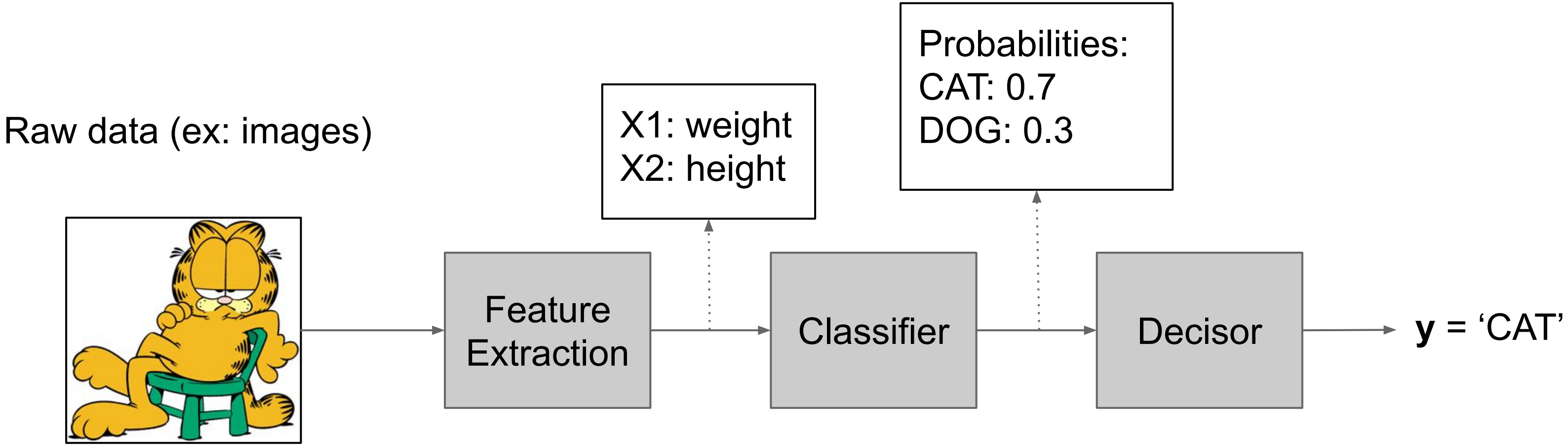
Deep Reinforcement Learning (DRL)

Hardware accelerator chip design.

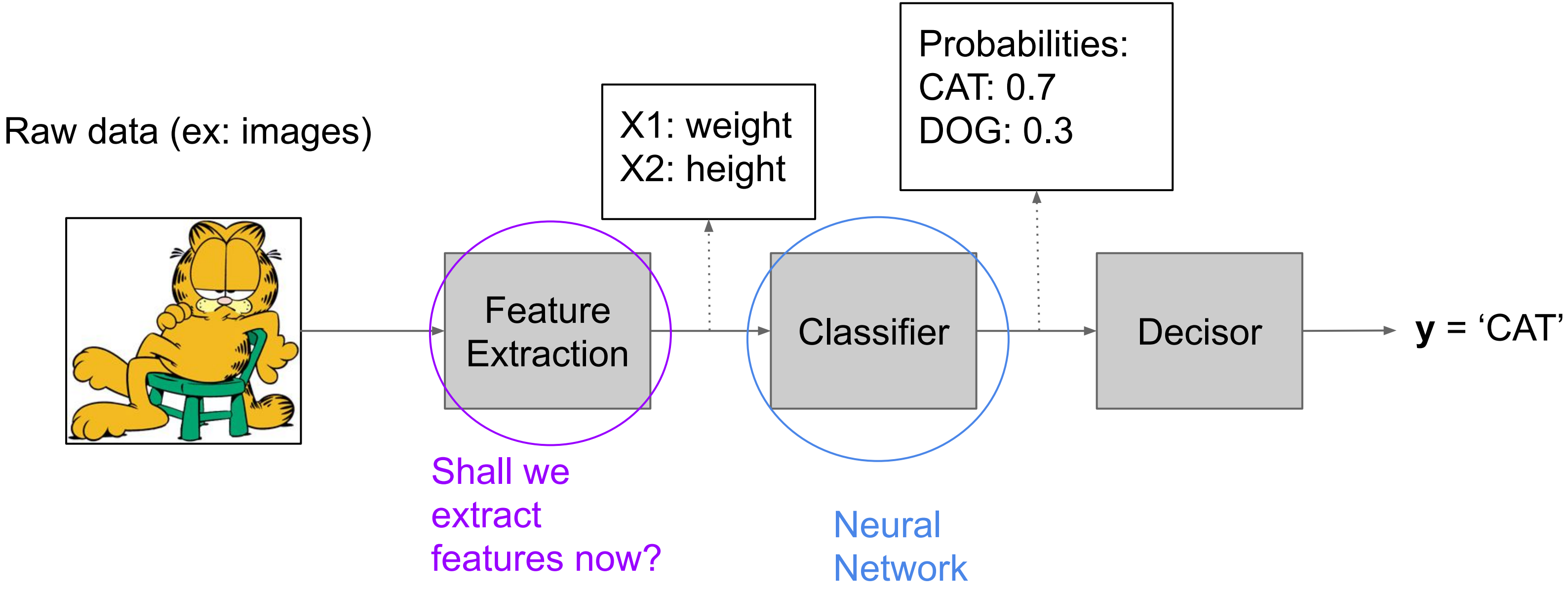


Azalia Mirhoseini et al, [“Chip Placement with Deep Reinforcement Learning”] arXiv 2020. [blog]

Classic Machine Learning classification pipeline



Classic Machine Learning classification pipeline

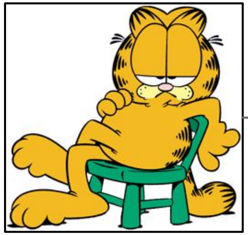


Slide credit: Santiago Pascual (UPC TelecomBCN 2019)

Deep Learning classification pipeline

Probabilities:
CAT: 0.7
DOG: 0.3

Raw data (ex: images)



Classifier

Decisor

y = ‘CAT’

We CAN inject the raw data, and features will be learned!!

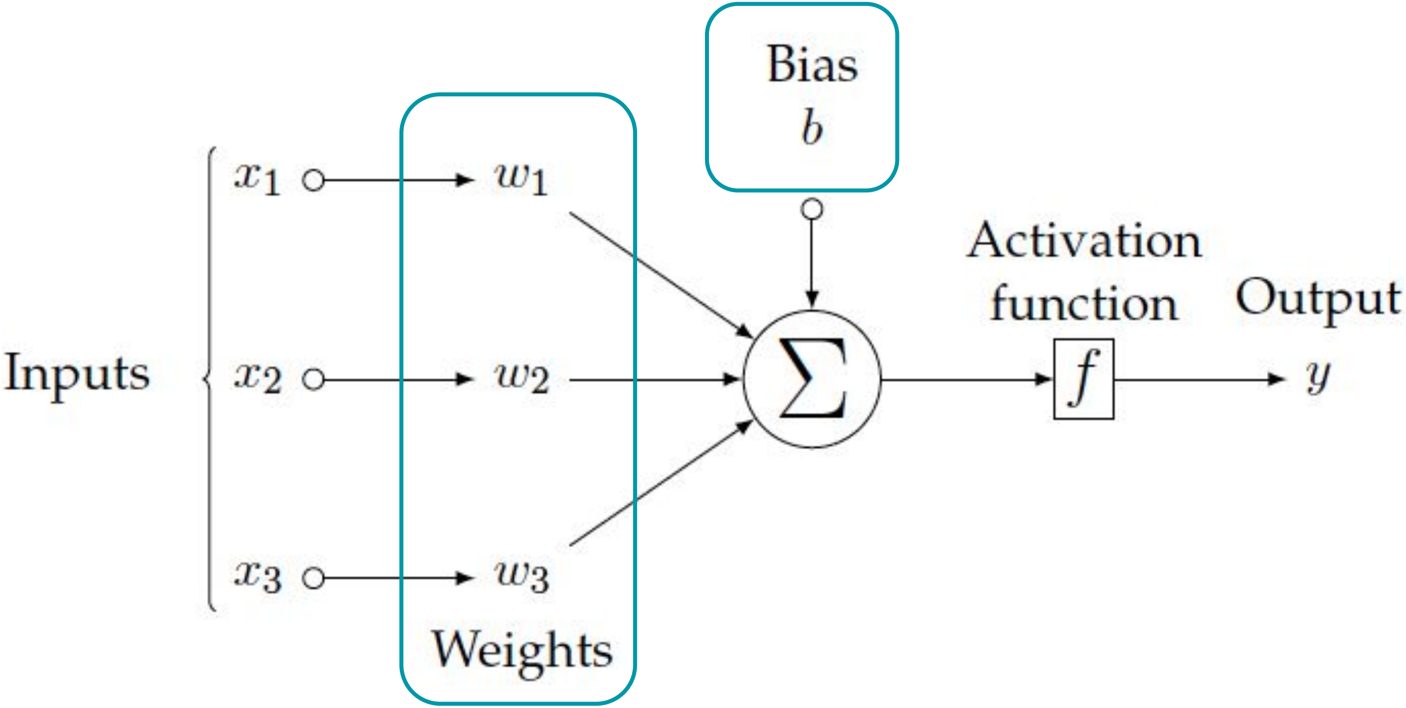
Neural Network

End to End concept

Slide credit: Santiago Pascual (UPC TelecomBCN 2019)

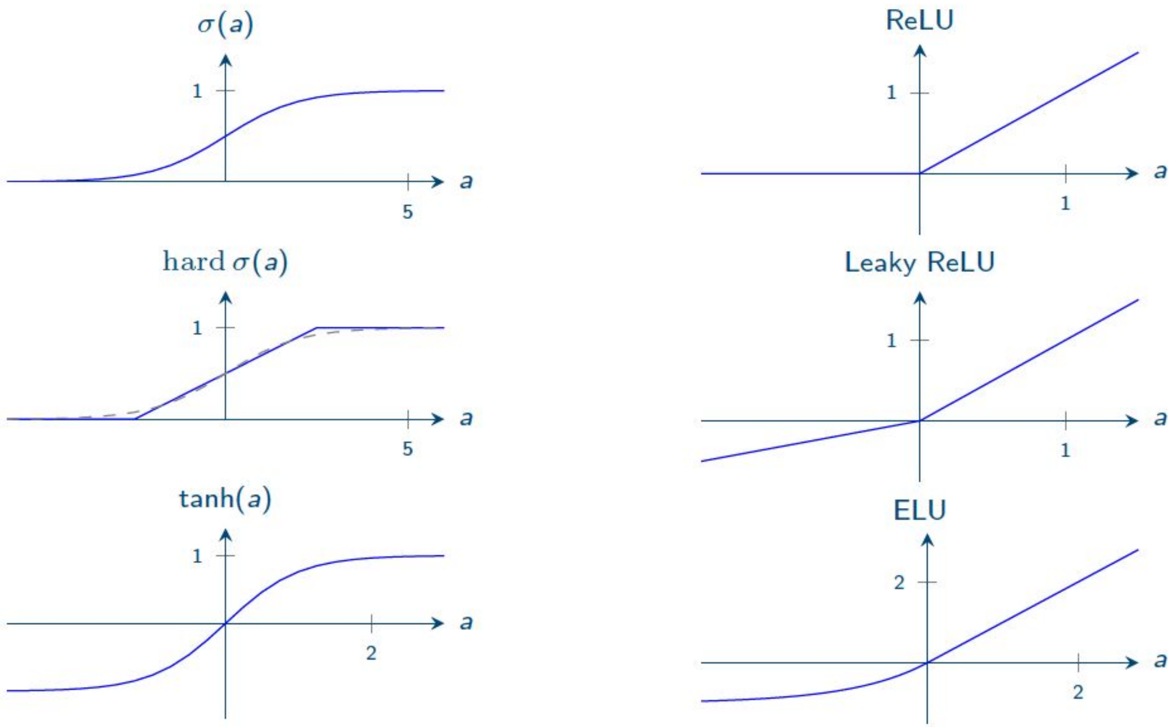
DL basic unit: The Perceptron

Weights and bias are the parameters that define the behavior. They must be estimated during training.



DL basic unit: The Perceptron

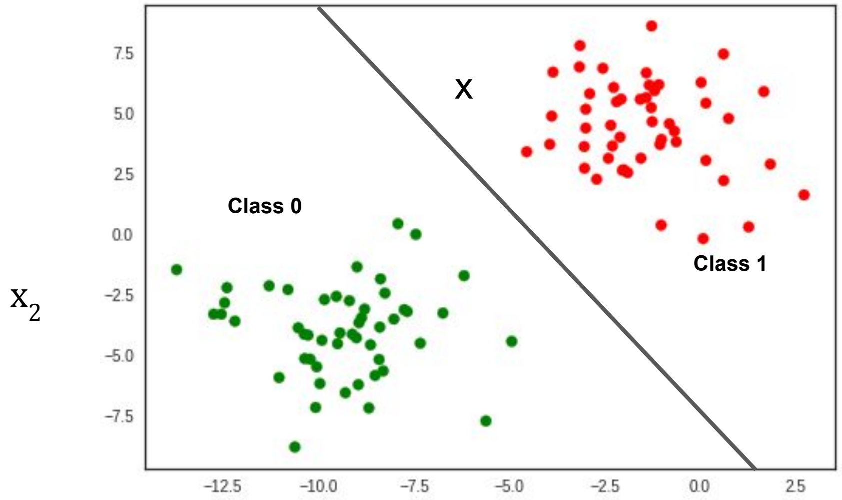
Multiple options as activation functions f(·):



DL basic unit: The Perceptron

A single perceptron can only define linear decision boundaries.

2D input space data

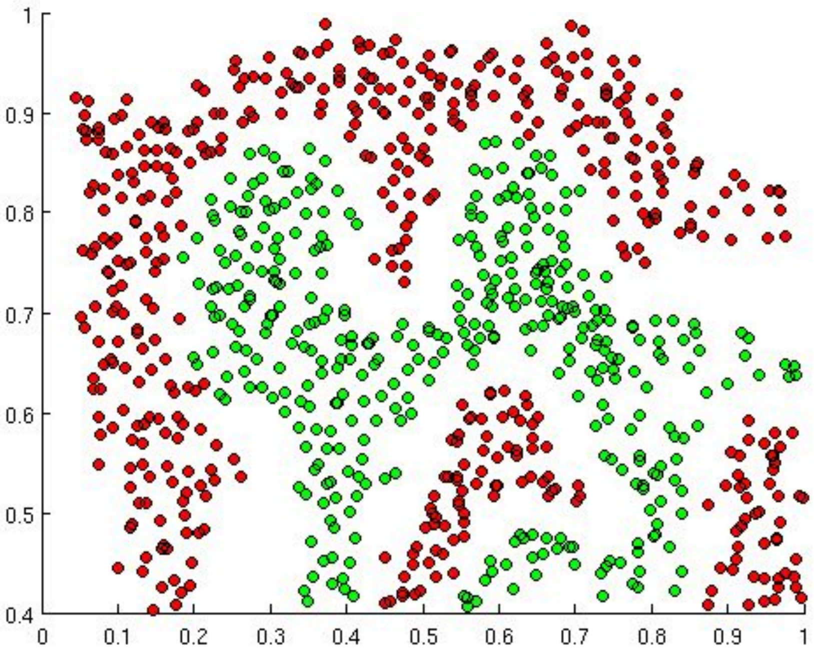


$$f(x) = \begin{cases} 1 & \text{if } w \cdot x + b > 0 \\ 0 & \text{otherwise} \end{cases}$$

DL basic unit: The Perceptron (limitation)

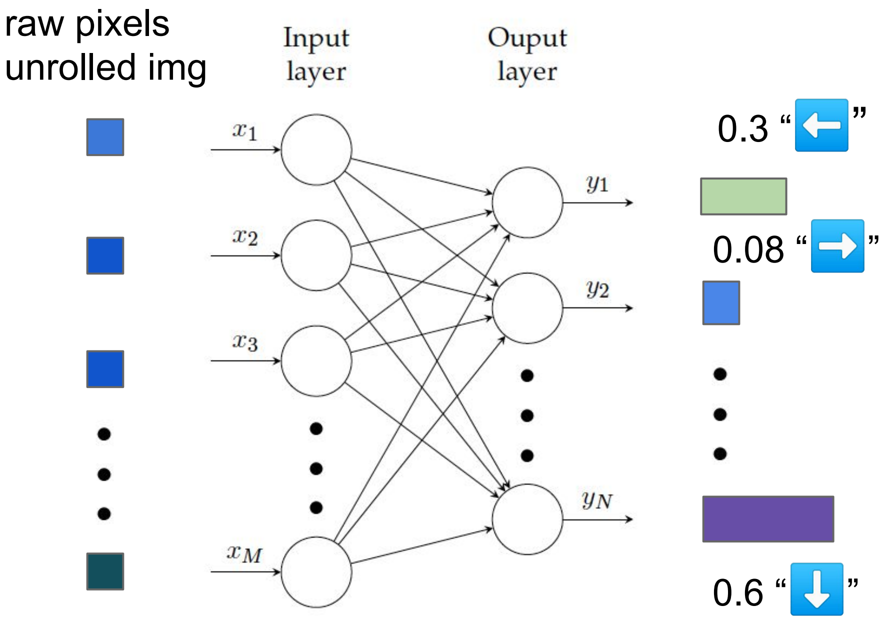
Real world data often needs a non-linear decision boundary

- Images
- Audio
- Text
- **Value functions**
- **Policies**

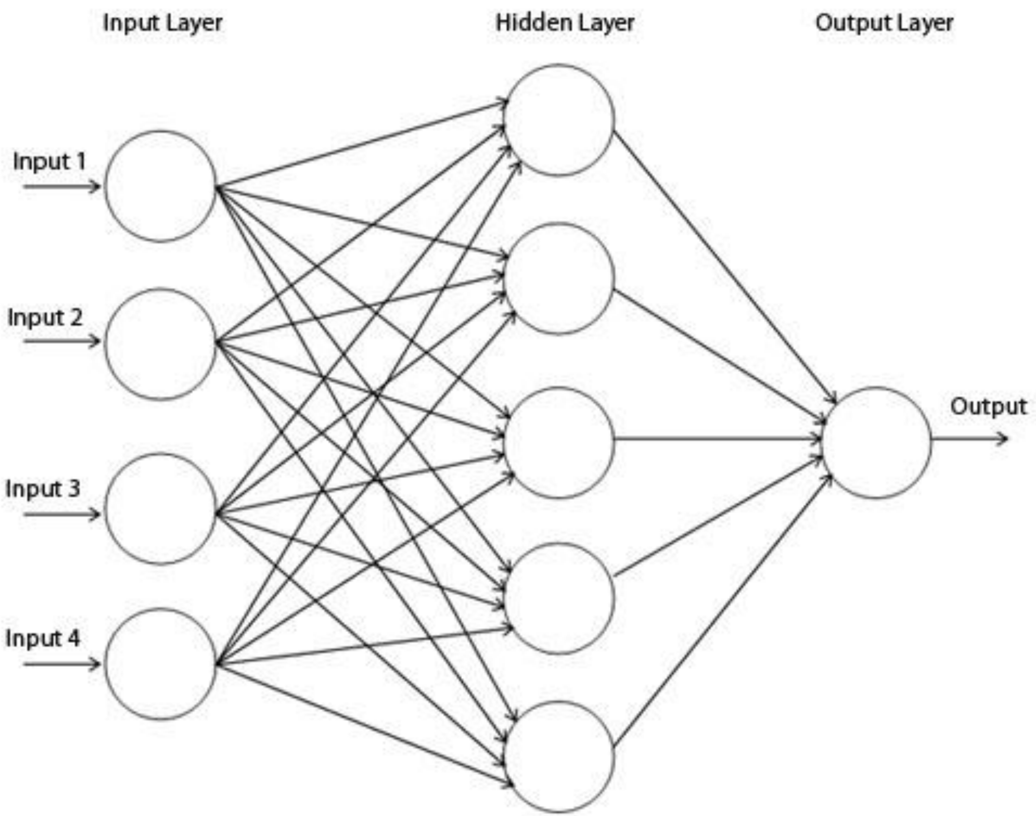


A Single Layer of N Perceptrons

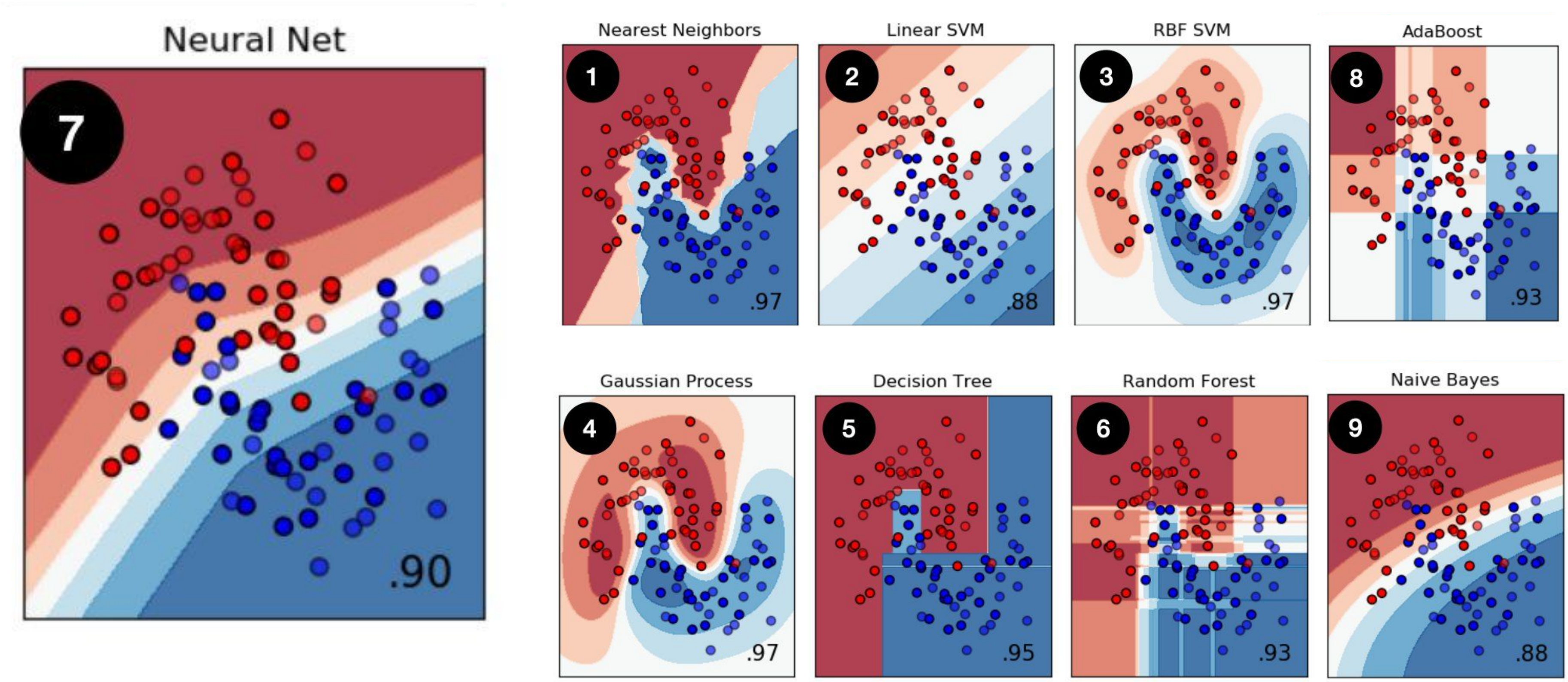
A multiclass classification problem can be solved by assigning a perceptron for each class and choosing the **maximum** output.



Neural Network (single hidden layer)



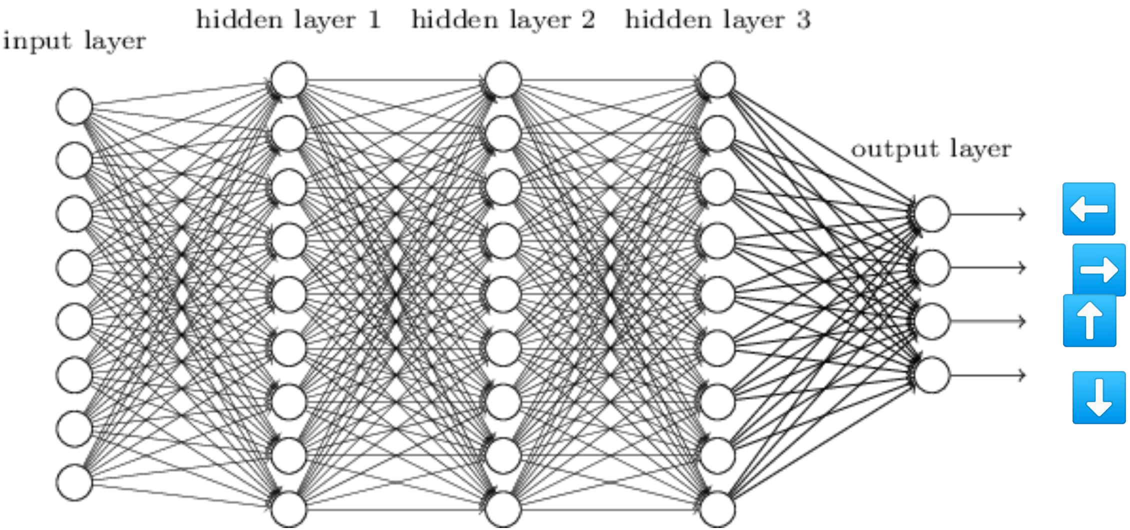
Neural Network (single hidden layer)



Source: [scikit-learn]

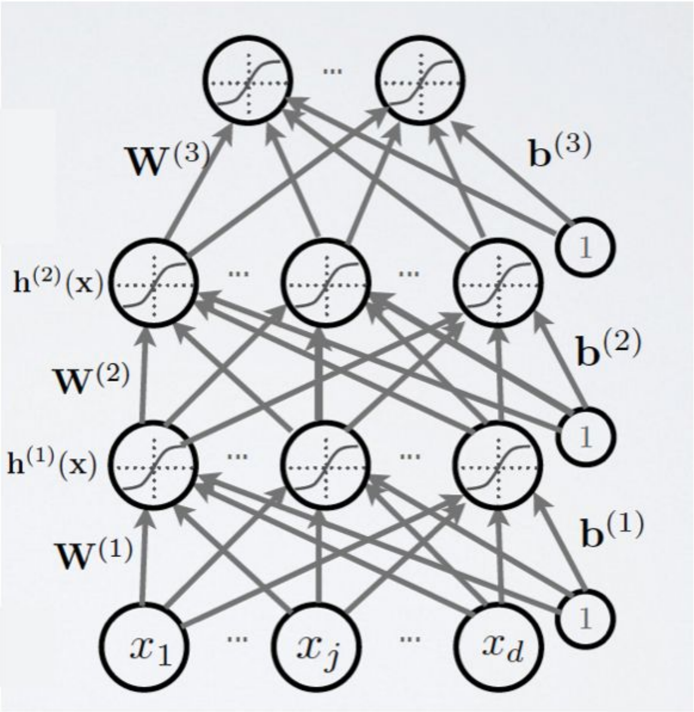
Multilayer Perceptron (MLP)

In practice, **deep neural networks** nets can usually represent more complex functions with less total neurons (and therefore, less parameters)



Multilayer Perceptron (MLP)

Feed Forward



<u>Figure</u>: Hugo Larochelle

OUTPUT (y)

Feed-forward Weights (W_i)

$$y = f(W_3 \cdot h_2 + b_3)$$

Hidden States h_1 & h_2

$$h_2 = f(W_2 \cdot h_1 + b_2)$$

$$h_1 = f(W_1 \cdot x + b_1)$$

INPUT (x)

ELSEVIER

Neural Networks
Volume 2, Issue 5, 1989, Pages 359-366

Original contribution

Multilayer feedforward networks are universal approximators

Kurt Hornik, Maxwell Stinchcombe, Halbert White [1]

⊞ Show more

https://doi.org/10.1016/0893-6080(89)90020-8

Abstract

This paper rigorously establishes that standard multilayer feedforward networks with as few as one hidden layer using arbitrary squashing functions are capable of approximating any Borel measurable function from one finite dimensional space to another to any desired degree of accuracy, provided sufficiently many hidden units are available. In this sense, multilayer feedforward networks are a class of universal approximators.

ELSEVIER

Neural Networks
Volume 2, Issue 5, 1989, Pages 359-366

Original contribution

Multilayer feedforward networks are universal approximators

Kurt Hornik, Maxwell Stinchcombe, Halbert White [1]

⊞ Show more

https://doi.org/10.1016/0893-6080(89)90020-8

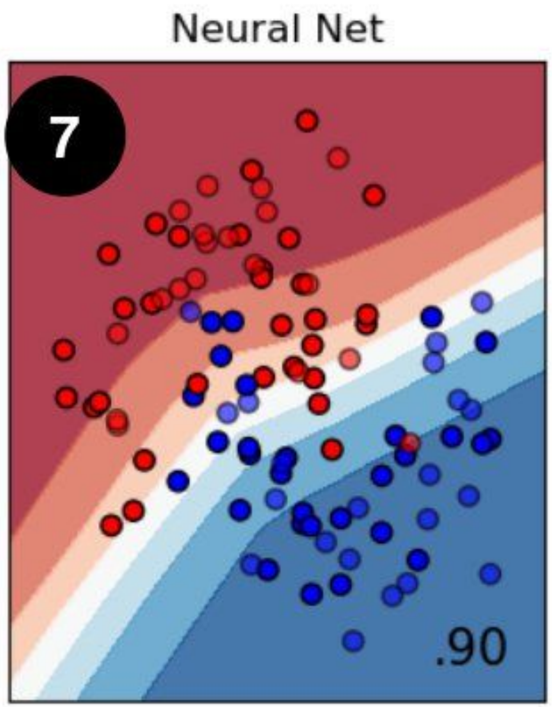
Get rights and content

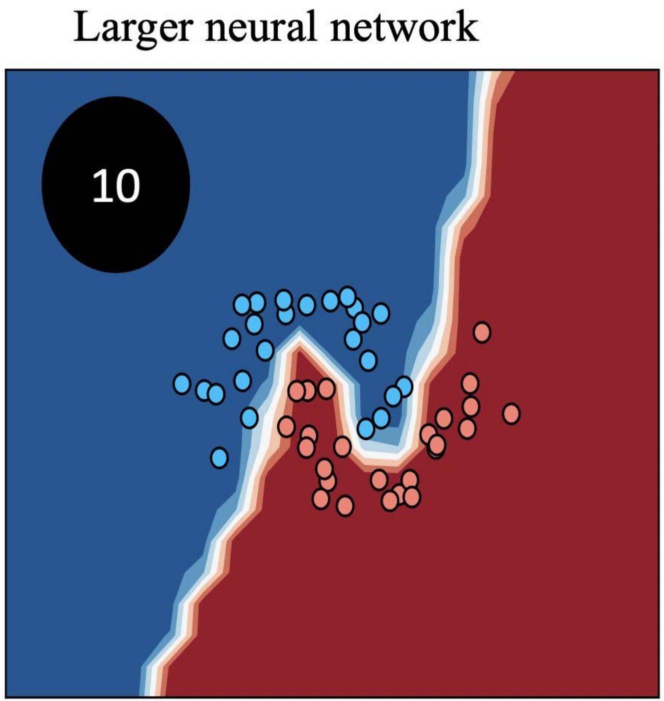
Abstract

This paper rigorously establishes that standard multilayer feedforward networks with as few as one hidden layer using arbitrary squashing functions are capable of approximating any Borel measurable function from one finite dimensional space to another to any desired degree of accuracy, provided sufficiently many hidden units are available. In this sense, multilayer feedforward networks are a class of universal approximators.

- Needs a “finite number of hidden neurons”: finite may be extremely large
- How to find the parameters (weights, biases) of these neurons ?

Deep Neural Networks (DNN)

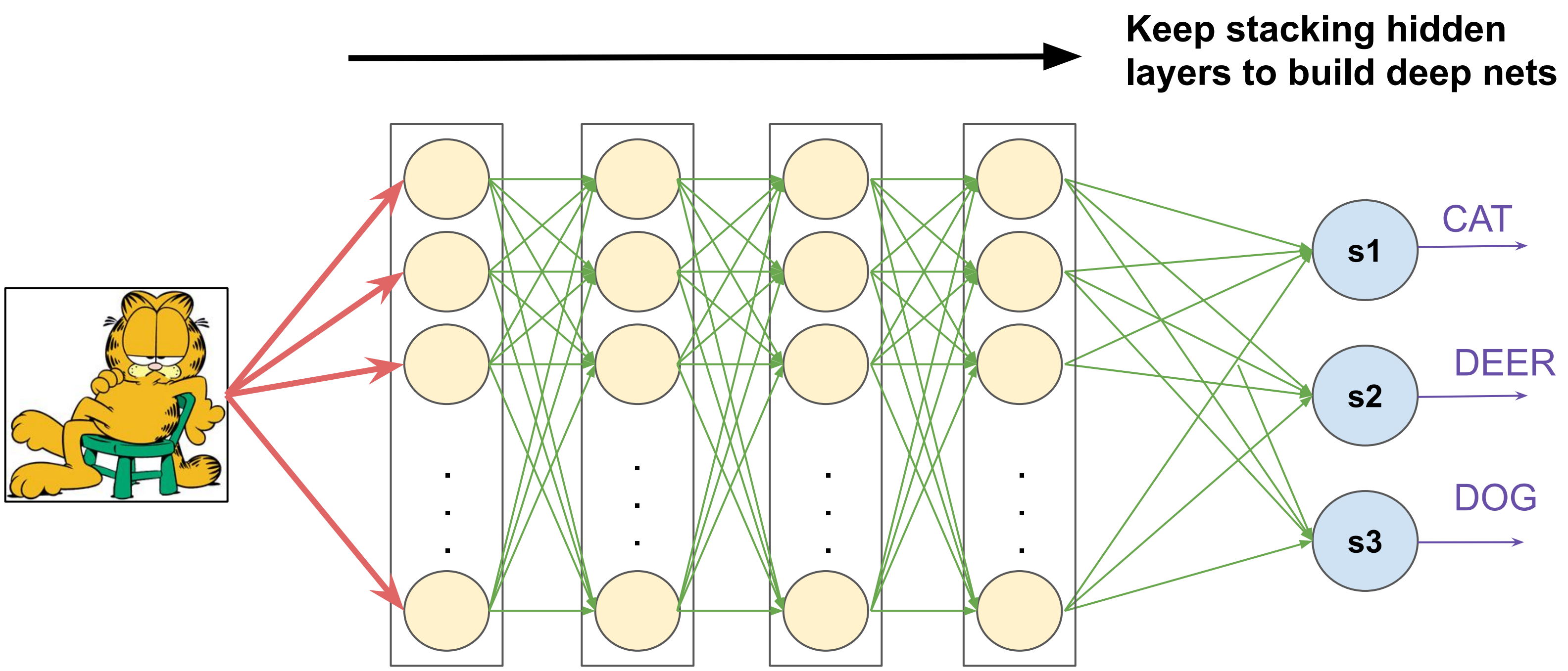




Source: scikit-learn

Source: Tim Pearce

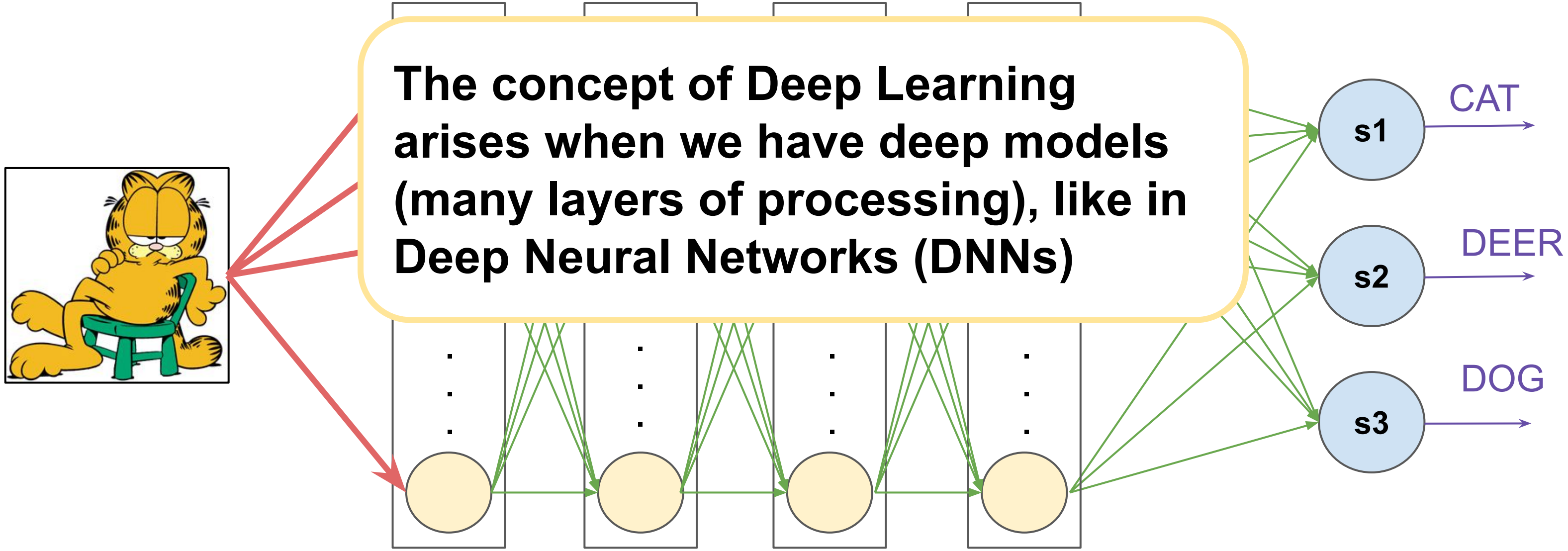
Deep Neural Networks (DNN)



Slide credit: Santiago Pascual (UPC TelecomBCN 2019)

Deep Neural Networks (DNN)

Keep stacking hidden layers to build deep nets



The concept of Deep Learning arises when we have deep models (many layers of processing), like in Deep Neural Networks (DNNs)

Slide credit: Santiago Pascual (UPC TelecomBCN 2019)

Deep (Hierarchical) Data Representations

Image **Speech**

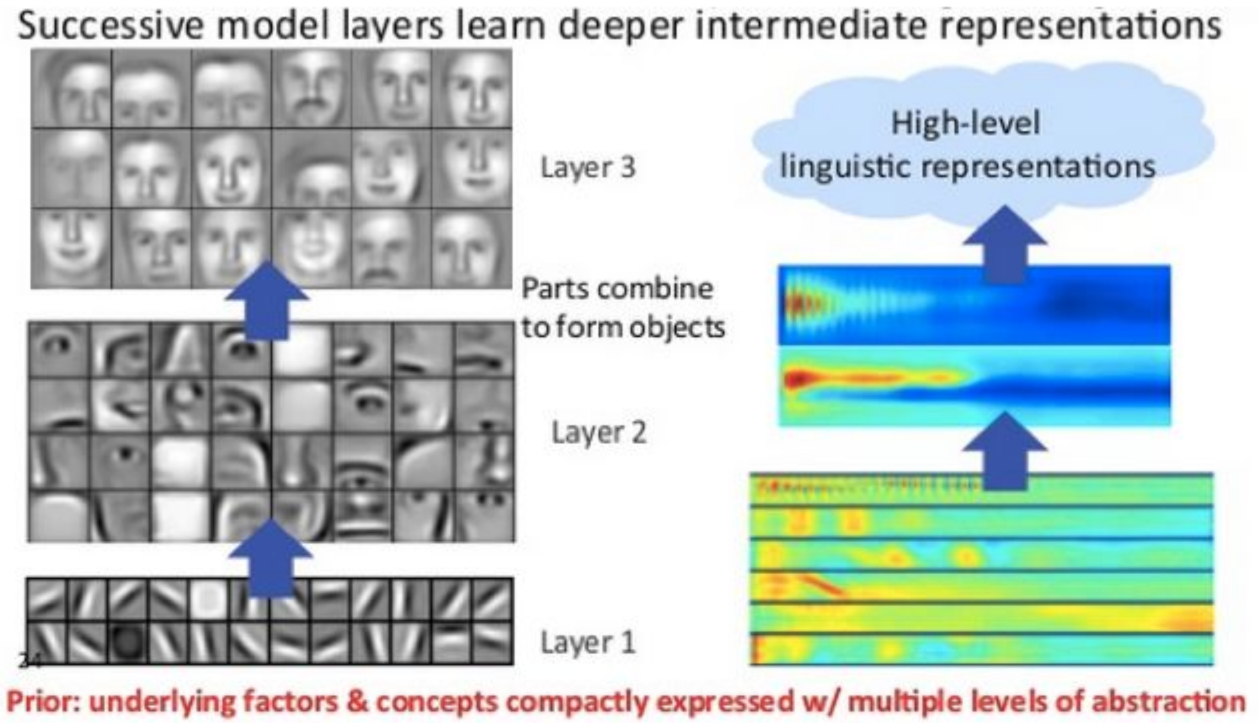


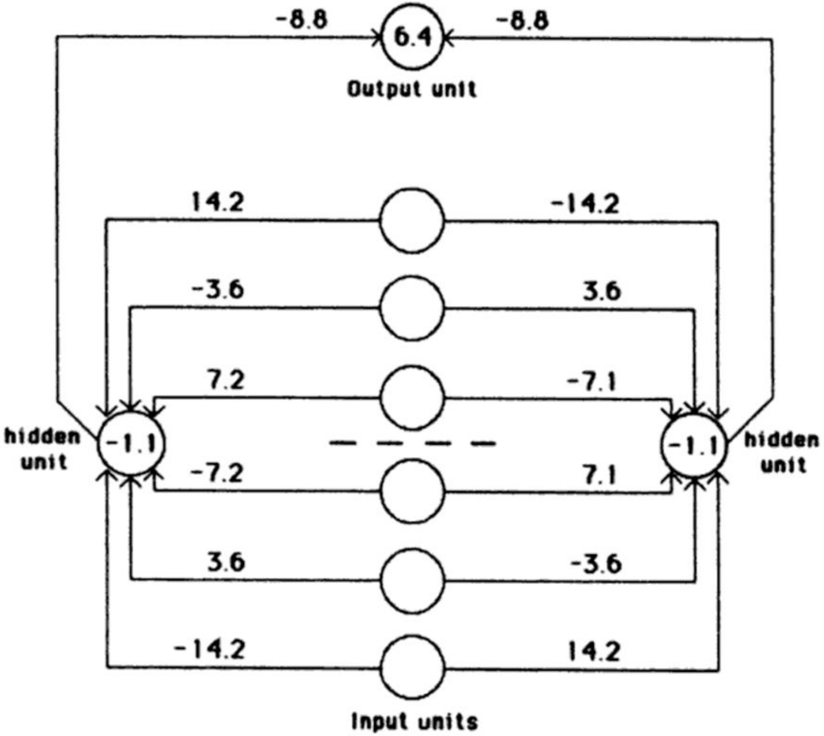
Figure ref

Slide credit: Santiago Pascual (UPC TelecomBCN 2019)

How to estimate the parameters ?

Training a neural network with the **back-propagation** algorithm.

Geoff Hinton after writing the paper on backprop in 1986

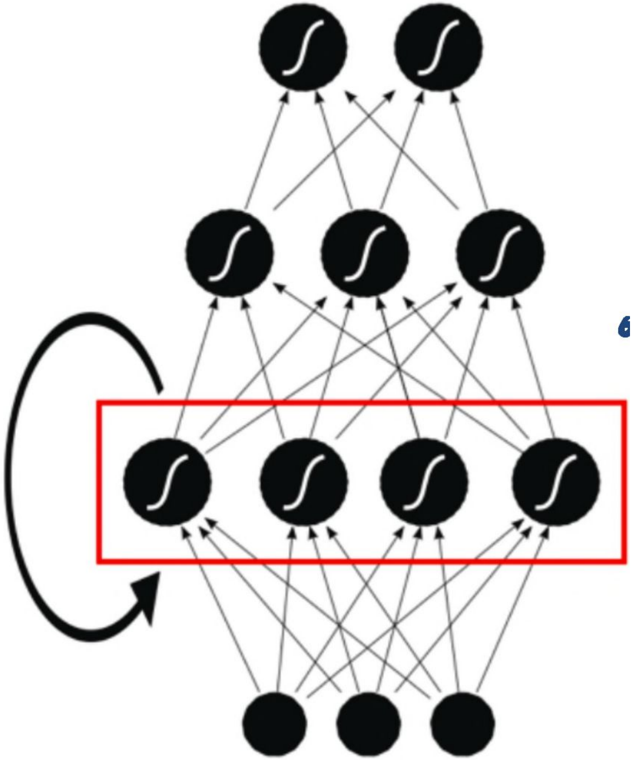


Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. ["Learning representations by back-propagating errors."]() Nature 1986.

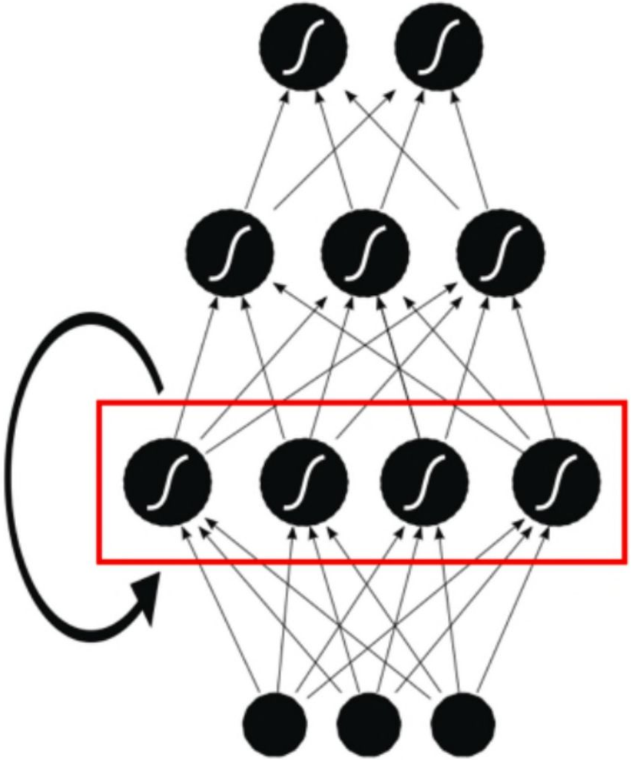
How to learn a memory unit ?

Recurrent layer (RNN)

The hidden layers and the output depend from previous states of the hidden layers



How to learn a memory unit ?



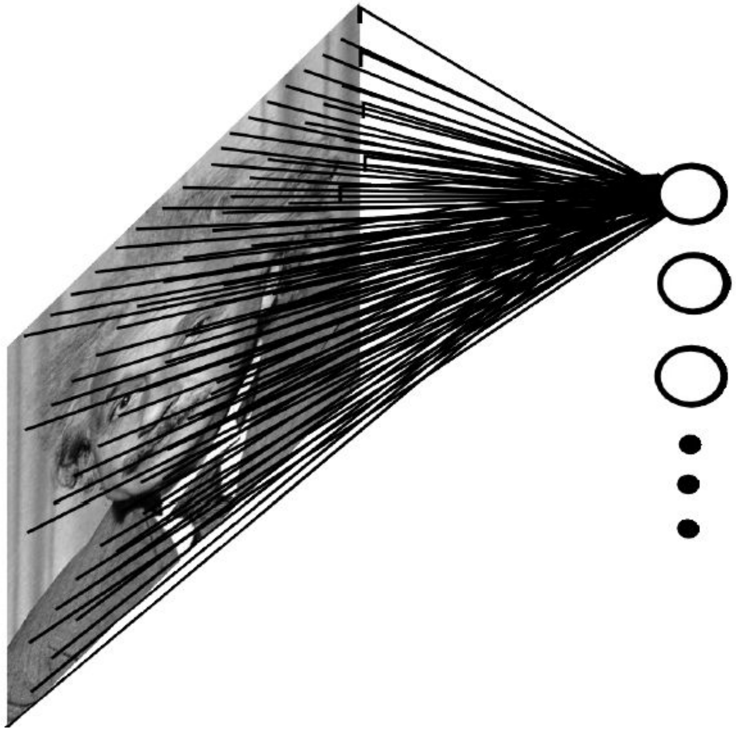
Feed-forward
Weights (W)

$$h_t = f(\boxed{W} \cdot x_t + \boxed{U} \cdot h_{t-1} + b)$$

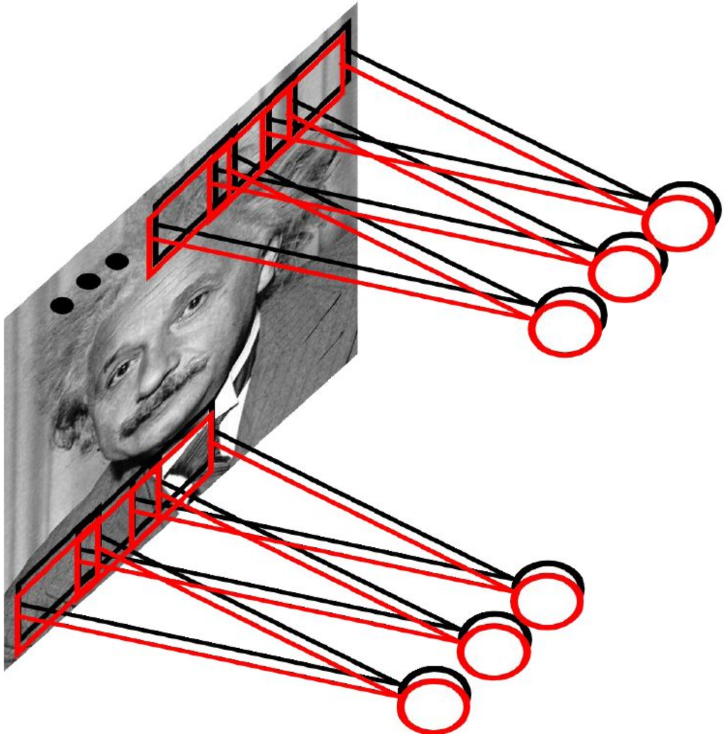
Recurrent
Weights (U)

How to reuse neurons ?

Fully Connected layer (FC)

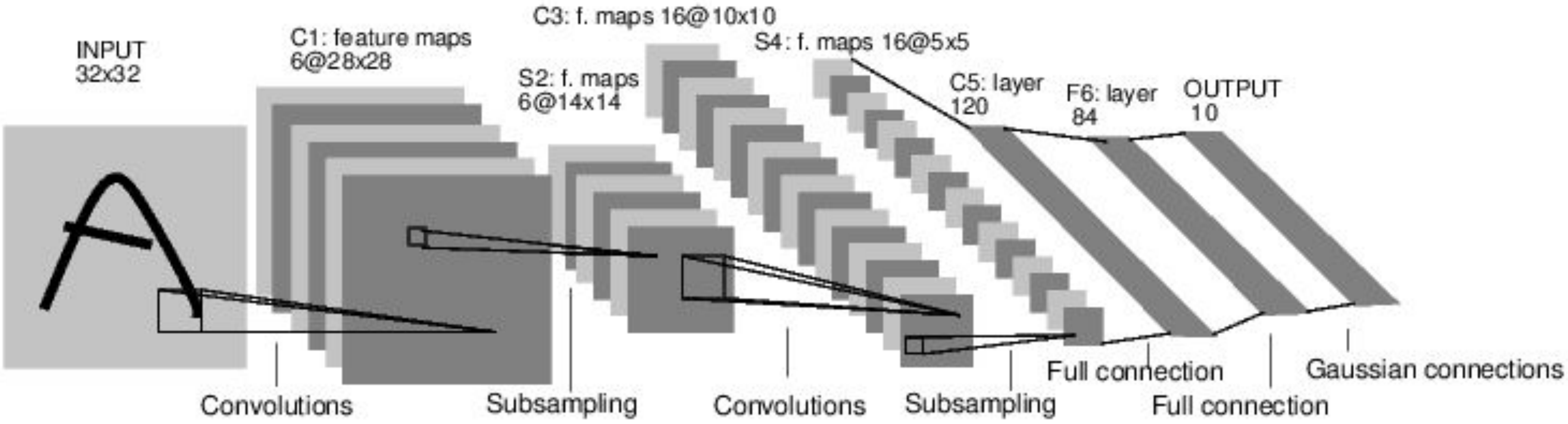


Convolutional layer (Conv)



Figures: Ranzatto

Convolutional Neural Network (CNN)



#CNN #LeNet-5 LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). [Gradient-based learning applied to document recognition](). *Proceedings of the IEEE*, *86*(11), 2278-2324.

Many other researchers have also contributed to the field as, for example, those pointed out by LSTM author Jürgen Schmidhuber in [“Deep Learning Conspiracy”](#).

Le Cake

Le Cake

A Pieter Abbeel i 7 més els agrada

Yann LeCun
@ylecun

Pierre Pieter Abbeel gave a seminar at NYU today.
He showed up at the Turing reception and was given the RL piece of "Le Cake" (and lots of cherries).

Tradueix el tuit

1:10 · 13/4/19 · Facebook

Association for Computing Machinery
@TheOfficialACM

David Silver, a Professor at University College London (@ucl) and a Principal Research Scientist at @DeepMind, will receive the 2019 ACM Prize in Computing for breakthrough advances in computer game-playing: bit.ly/3bEpsVd #ACMPrize

Tradueix el tuit

6:00 a. m. · 1 d’abr. de 2020 · Social Media Publisher App

#DQN Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. "Playing atari with deep reinforcement learning." NIPS Deep Learning Workshop (2013).

Artificial intelligence (AI)

Google buys UK artificial intelligence startup Deepmind for £400m

Google makes its biggest EU purchase yet with the technology that aims to make computers think like humans

Samuel Gibbs

Monday 27 January 2014 13.23 GMT

AlphaGo

Google DeepMind

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. and Dieleman, S., 2016. [Mastering the game of Go with deep neural networks and tree search](). *Nature*, *529* (7587), pp.484-489

Greg Kohs, “AlphaGo” (2017)

When you move on to deep learning

Learn more

- MSc course [2017] [2018] [2019]
- BSc course [2018] [2019] [2020]

- 1st edition (2016)
- 2nd edition (2017)
- 3rd edition (2018)
- 4th edition (2019)

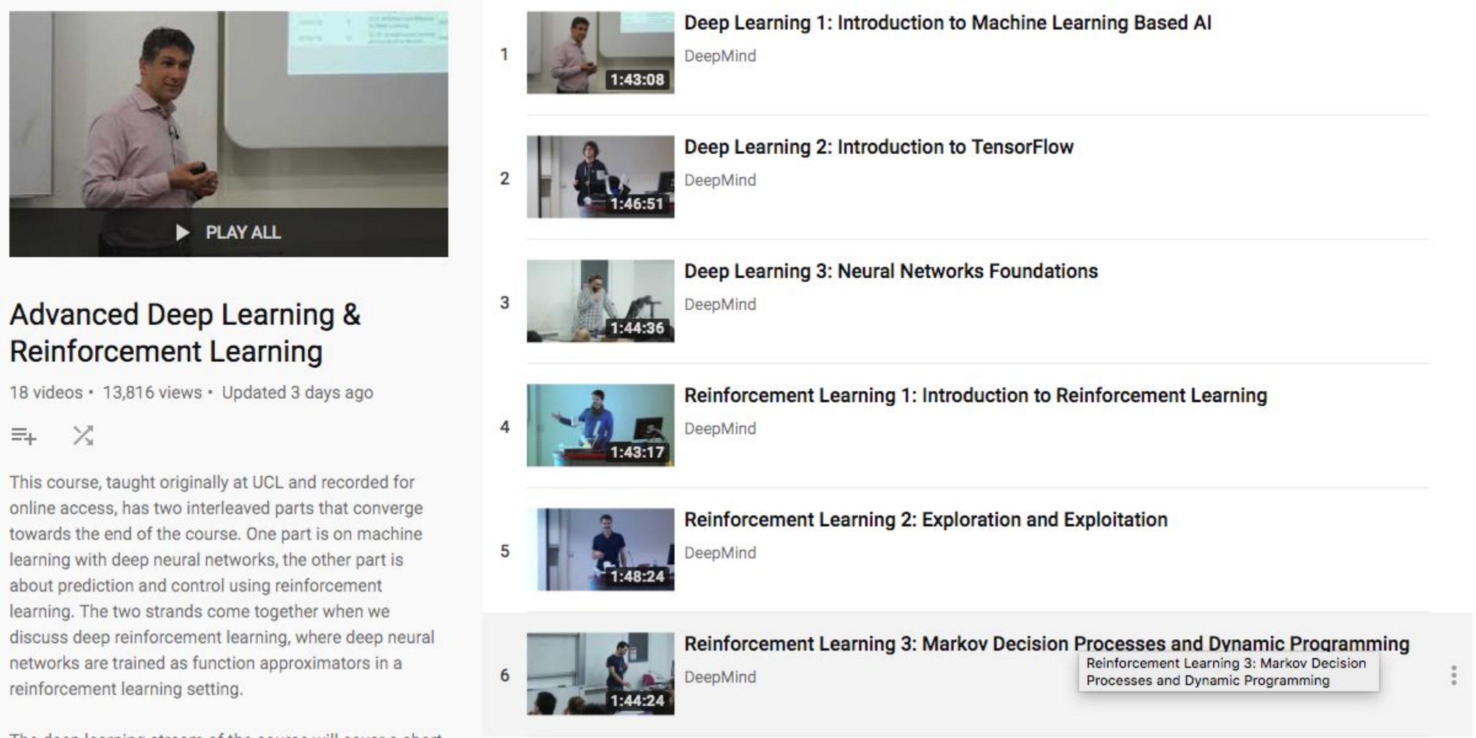
- 1st edition (2017)
- 2nd edition (2018)
- 3rd edition - NLP (2019)

Learn more

David Silver, UCL COMP050, [Reinforcement Learning](#) (2015)

Learn more

UCL & Deepmind: [“Advanced Deep Learning & Reinforcement Learning”]() (2018) [[slides]()]



Learn more

Pieter Abbeel and John Schulman, [CS 294-112 Deep Reinforcement Learning](), Berkeley.

Slides: [“Reinforcement Learning - Policy Optimization”]() OpenAI / UC Berkeley (2017)

Learn more

[OpenAI Spinning Up in Deep RL]

OpenAI