Universidad de Guadalajara

Segunda Feria de Posgrados

Clasificación de Imágenes con Deep Learning

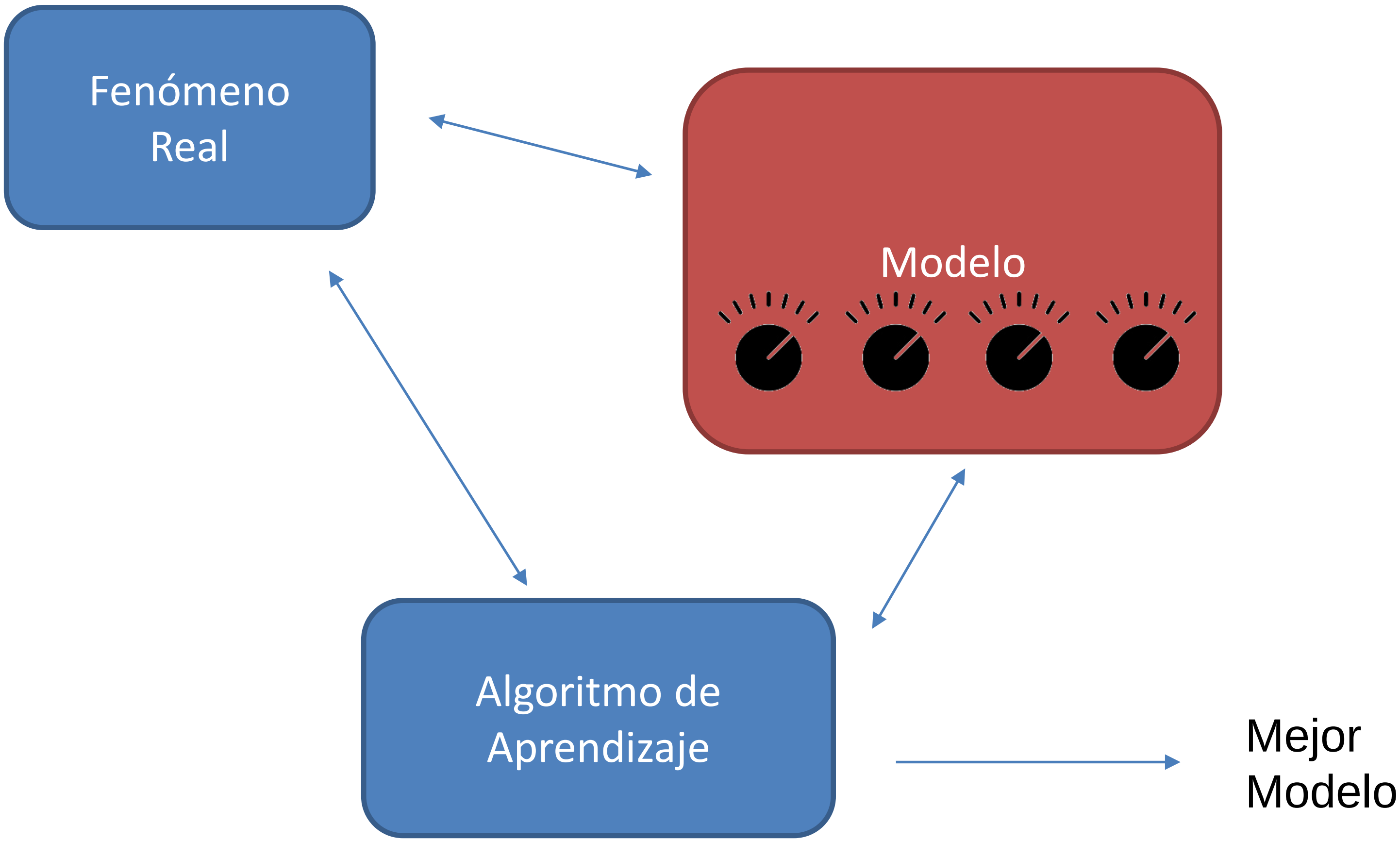
https://github.com/Dr-Carlos-Villasenor/Taller_CNN

Dr. Carlos Villaseñor

Contenido

- ¿Qué es Aprendizaje Automático?
- Introducción a redes neuronales densas
- Aprendizaje automático vs Aprendizaje profundo
- Introducción a redes neuronales convolucionales
- Transferencia de aprendizaje
- Búsqueda de hiperparámetros
- Poda neuronal

¿Qué es el aprendizaje automático?

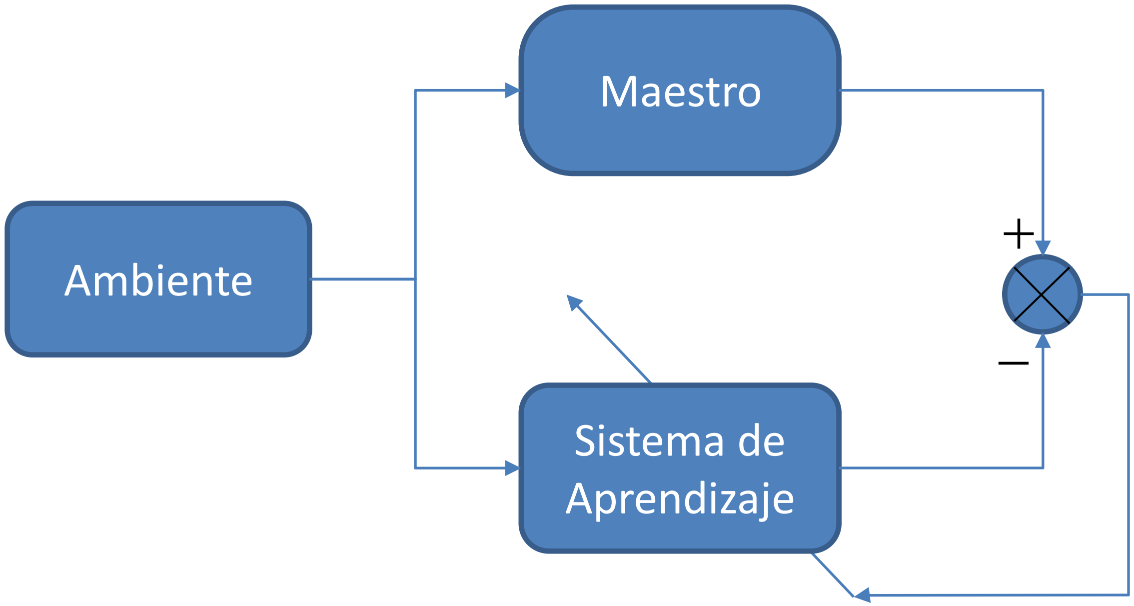


¿Qué es el aprendizaje?

- Según Tom Mitchell:

"Decimos que un programa de computadora aprende de la experiencia E con respecto a la tarea T y medida de desempeño D, si su desempeño D sobre la tarea T aumenta con la experiencia E"

Aprendizaje supervisado



Nota: Existen otros dos paradigmas llamados Aprendizaje no supervisado y aprendizaje reforzado, estos están fuera del alcance de esta clase

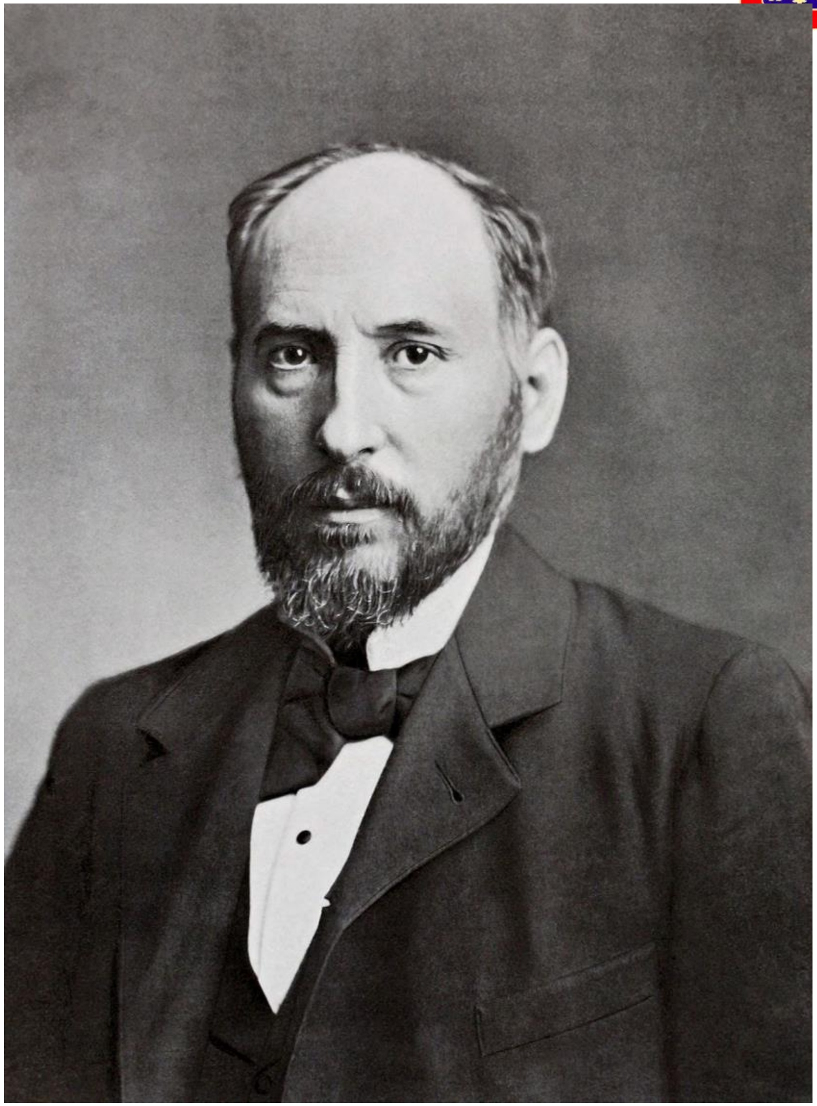
Problemas prototípicos

		Salida deseada	
		Discreta	Continua
Paradigma	Supervisado	Clasificación	Regresión
	No supervisado	Agrupamiento	Reducción de la dimensionalidad

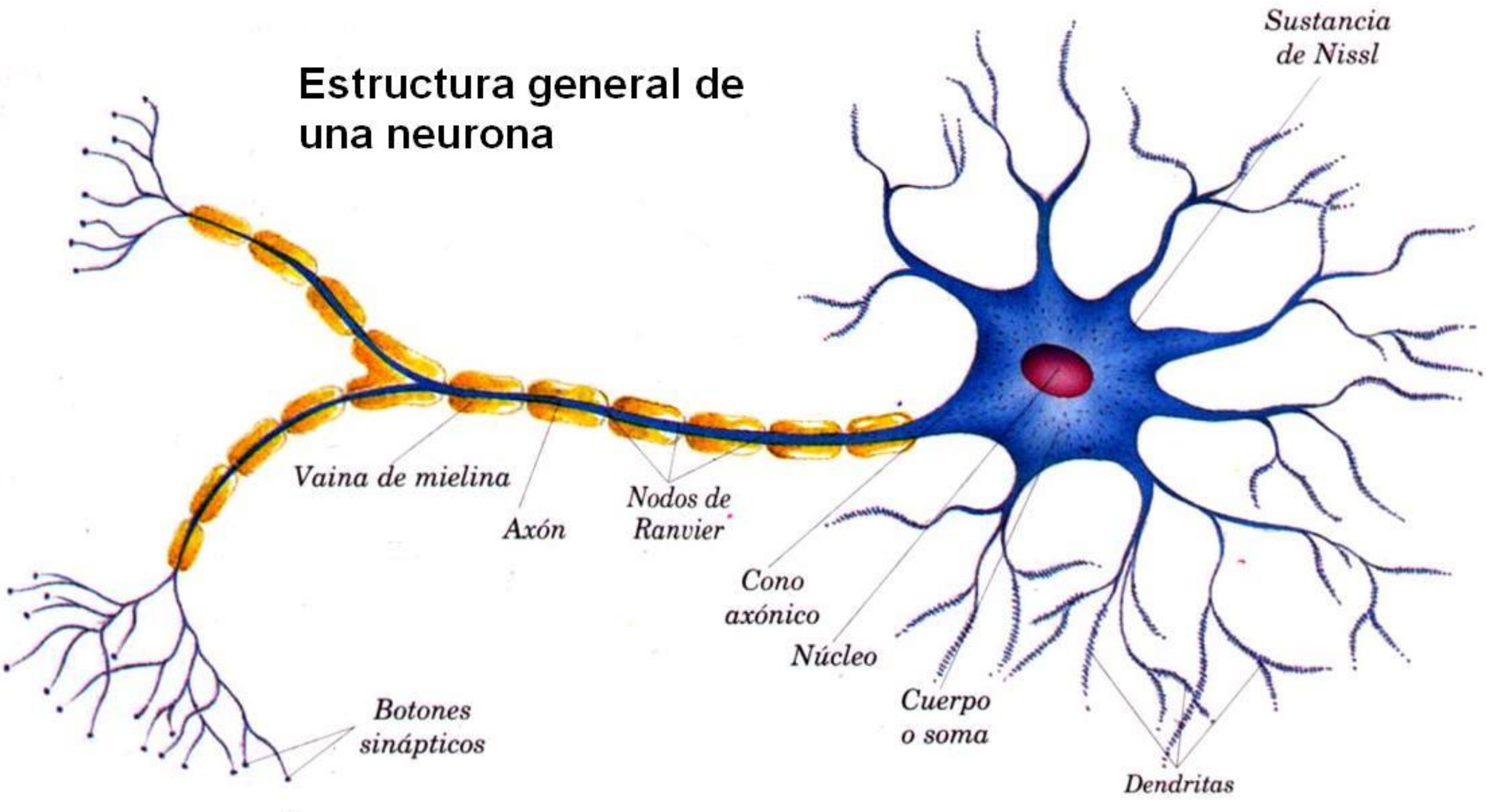
Introducción a las redes neuronales

- En 1906, Santiago Ramón y Cajal ganó el premio novel en medicina en reconocimiento de su trabajo sobre la estructura del sistema nervioso.

- Describió las neuronas como unidades de procesamiento de la información que se conectan y forman redes dinámicas para cumplir todas las funciones necesarias.

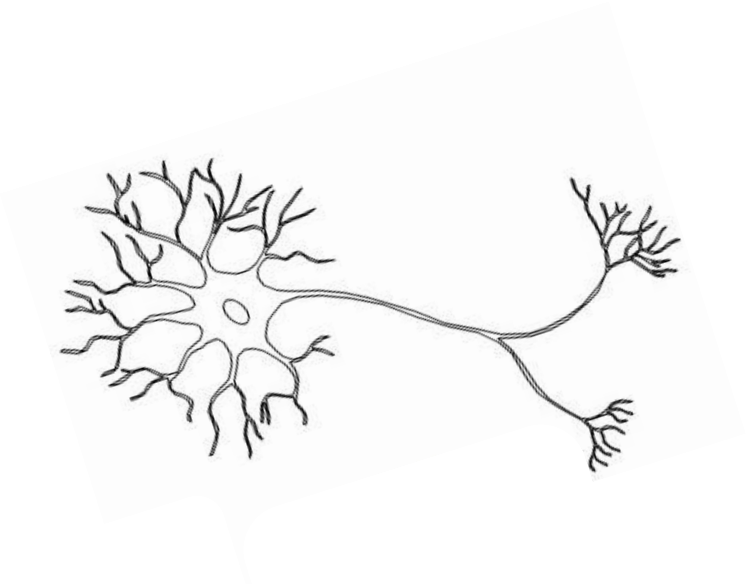


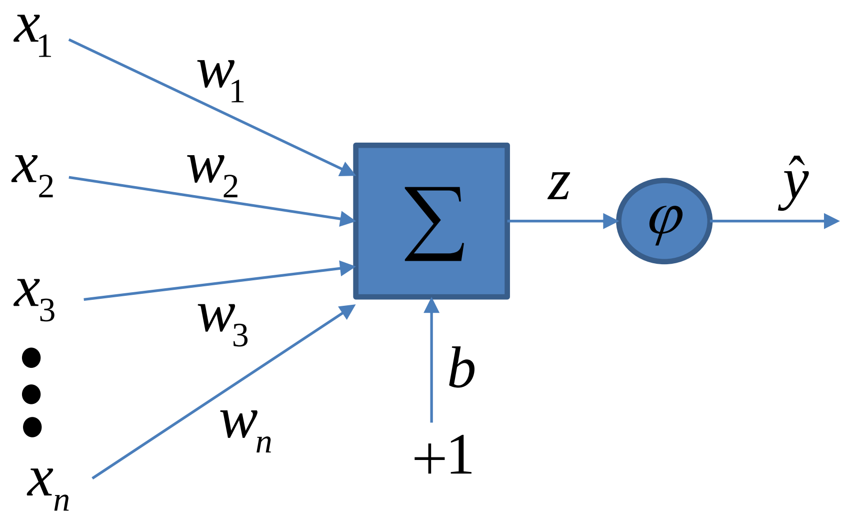
La neurona biológica



La neurona artificial

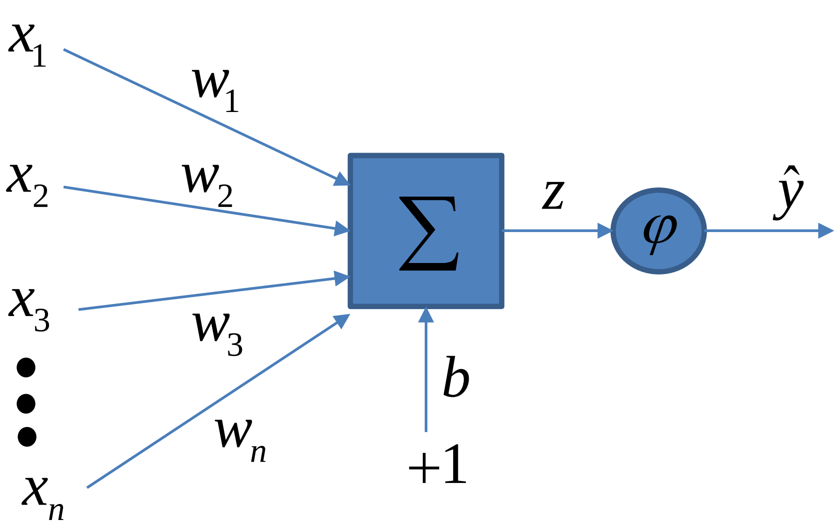
- En 1943, McCulluch y Pitts elaboraron el primer modelo matemático de una neurona





McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115-133.

La neurona artificial



$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$z = w^T x + b$$

$$\hat{y} = \varphi(z)$$

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115-133.

El Perceptrón

$$\text{Data:}\left\{\left(x^{(1)},y^{(1)}\right),\left(x^{(2)},y^{(2)}\right),\cdots,\left(x^{(p)},y^{(p)}\right)\right\}$$

$$x^{(i)}\in\mathbb{R}^{n},y^{(i)}\in\{0,1\}$$

$$\text{For e}\in\{1,2,\cdots,\text{epochs}\}$$

$$\text{For } i\ \in\{1,2,\cdots,p\}$$

$$\hat{y}=\varphi\left(w^{T}x^{(i)}+b\right)$$

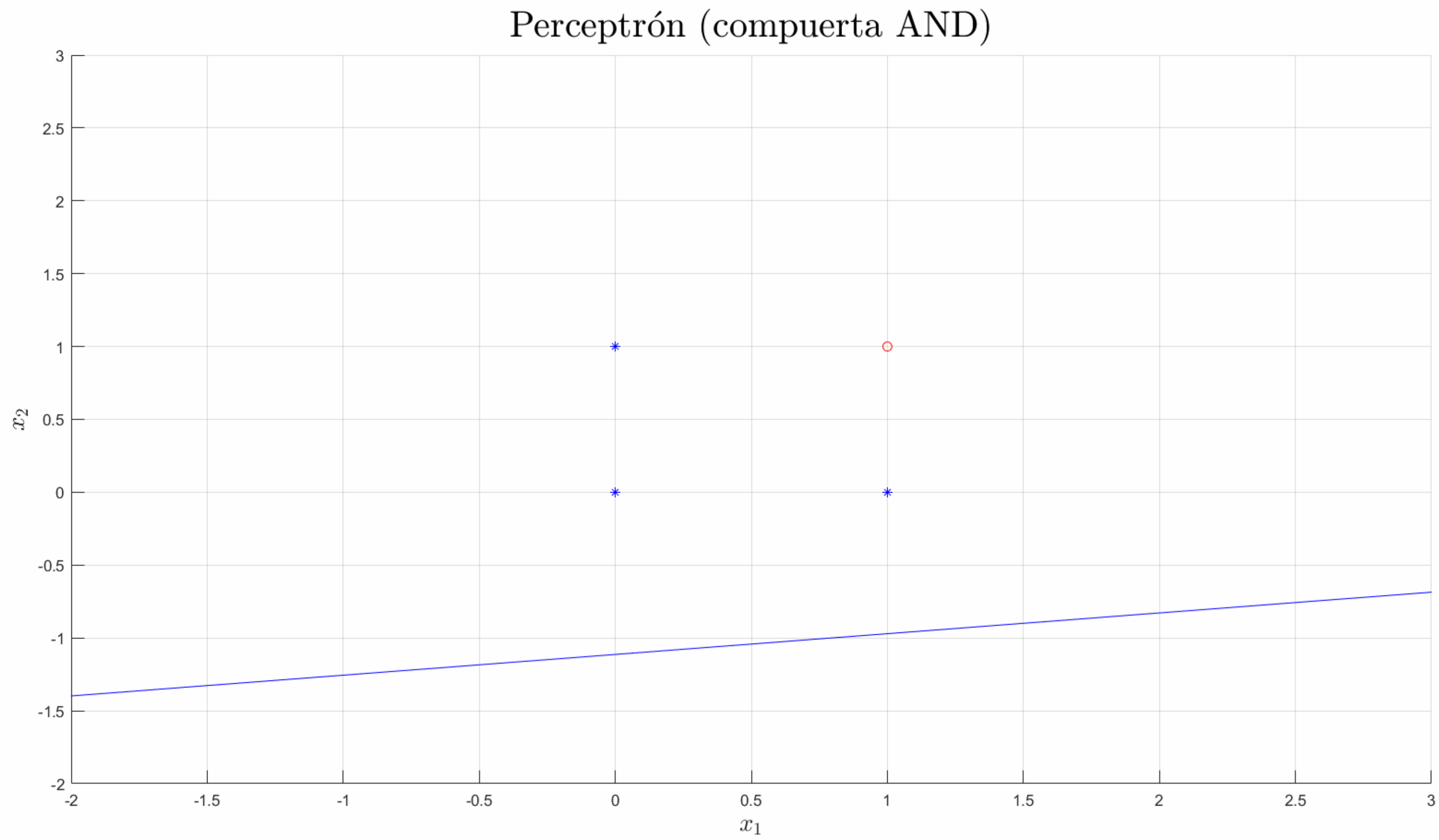
$$w\leftarrow w+\eta\left(y^{(i)}-\hat{y}\right)x^{(i)}$$

$$b\leftarrow b+\eta\left(y^{(i)}-\hat{y}\right)$$

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386.

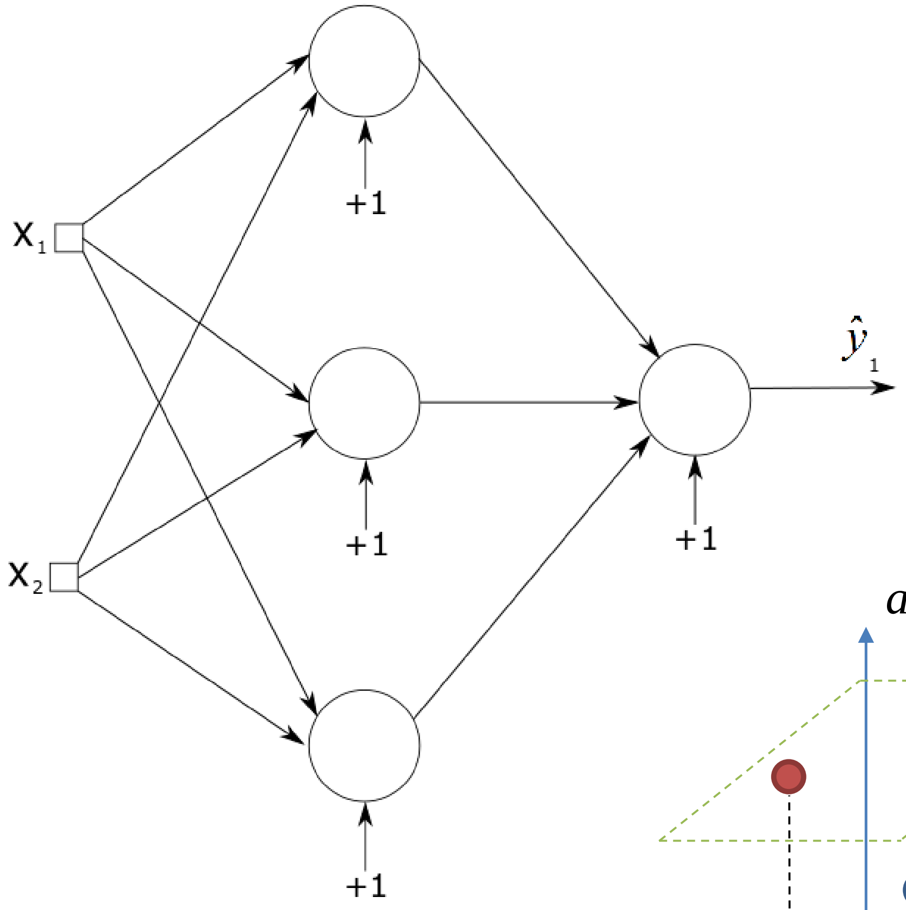
Demo 1

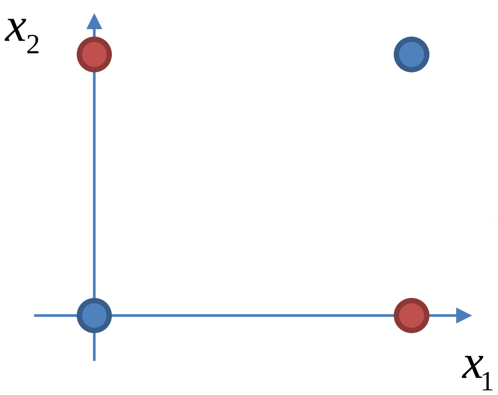
El Perceptrón

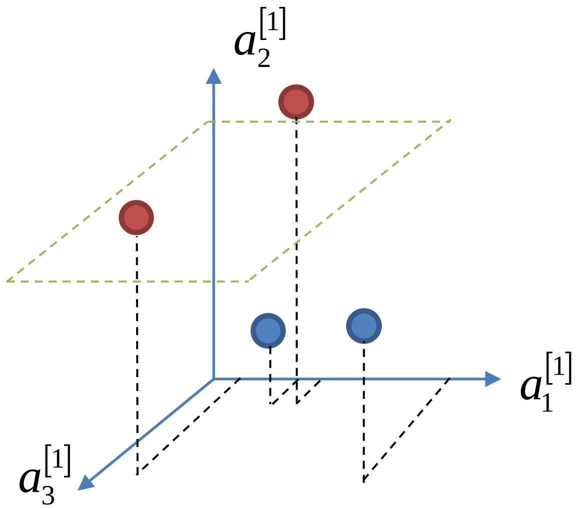


Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386.

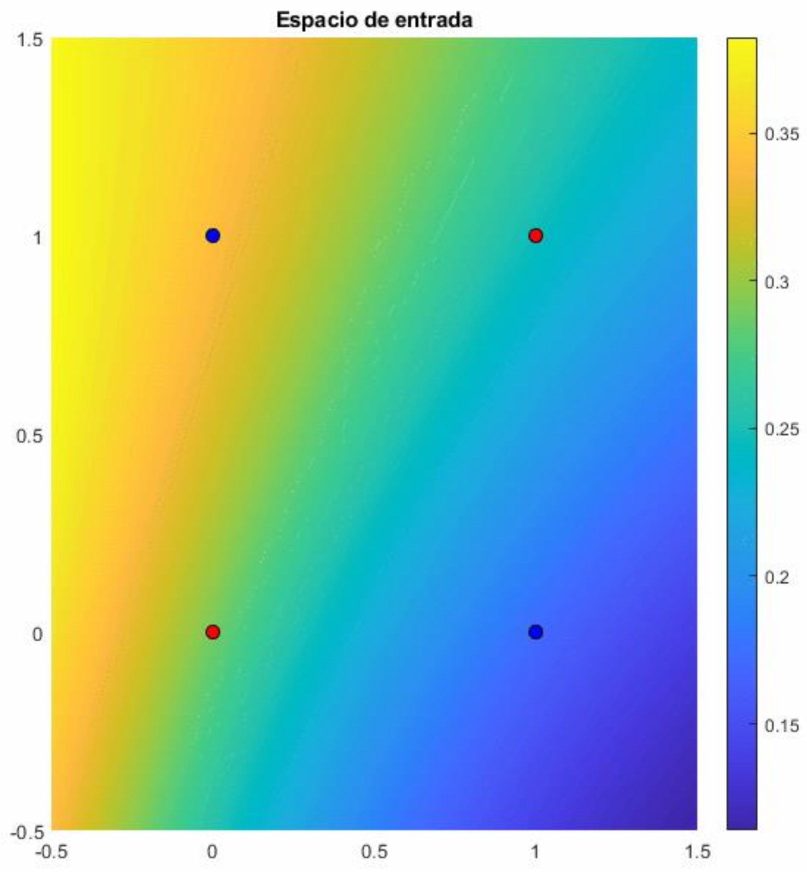
Red neuronal densa

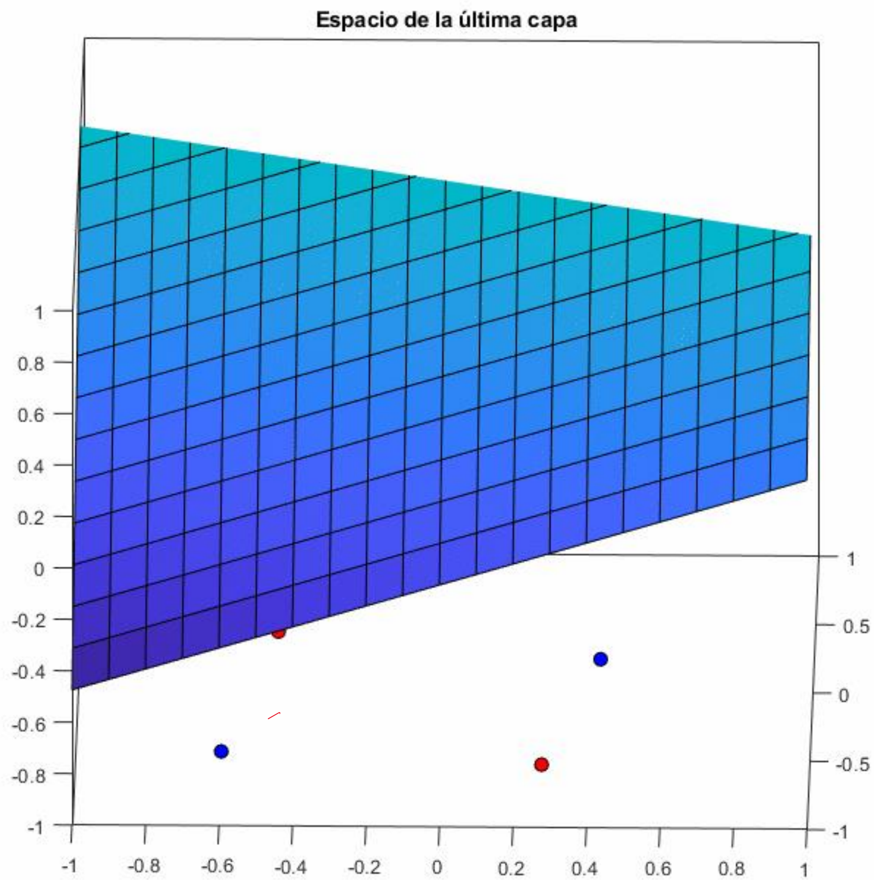






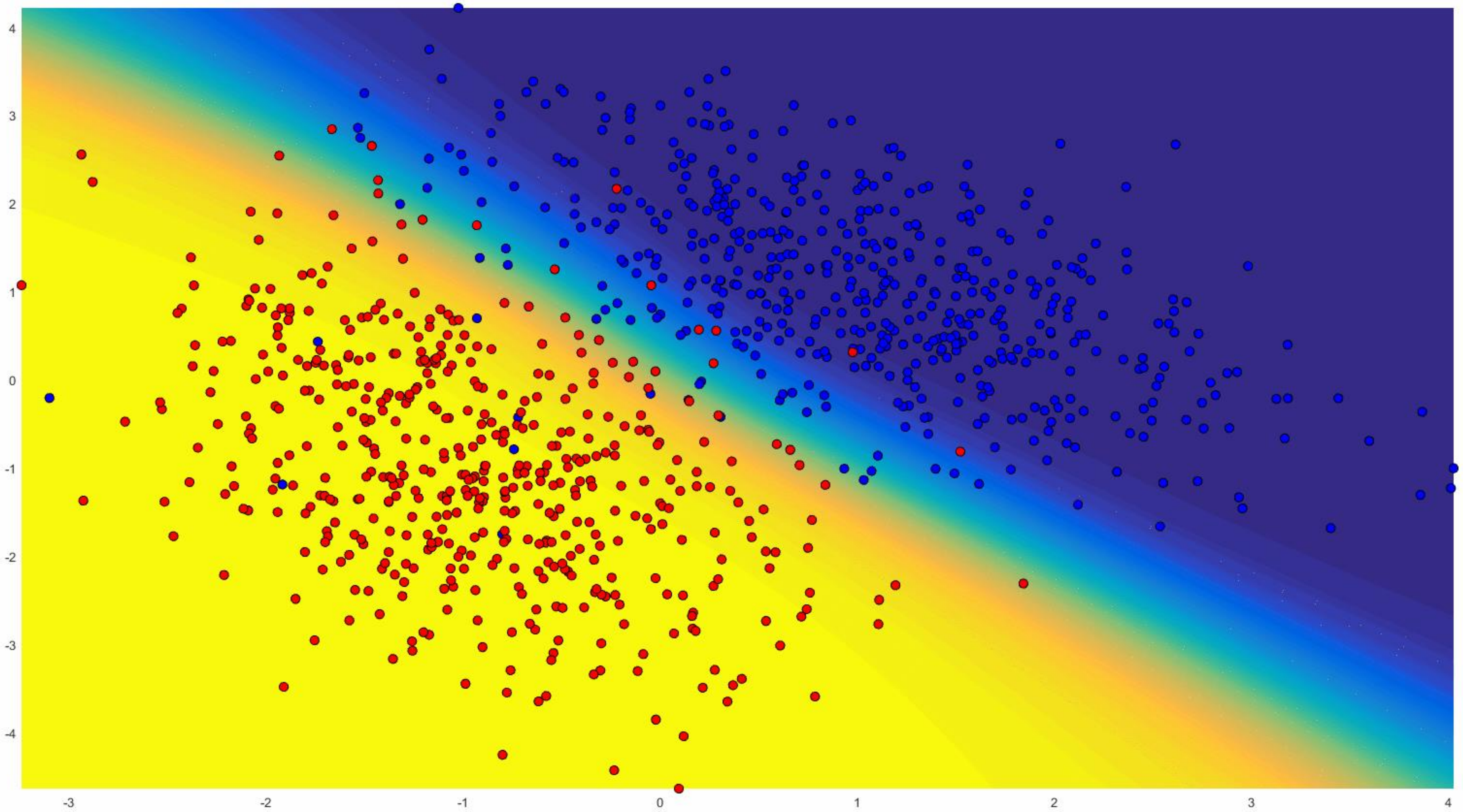
Red neuronal densa



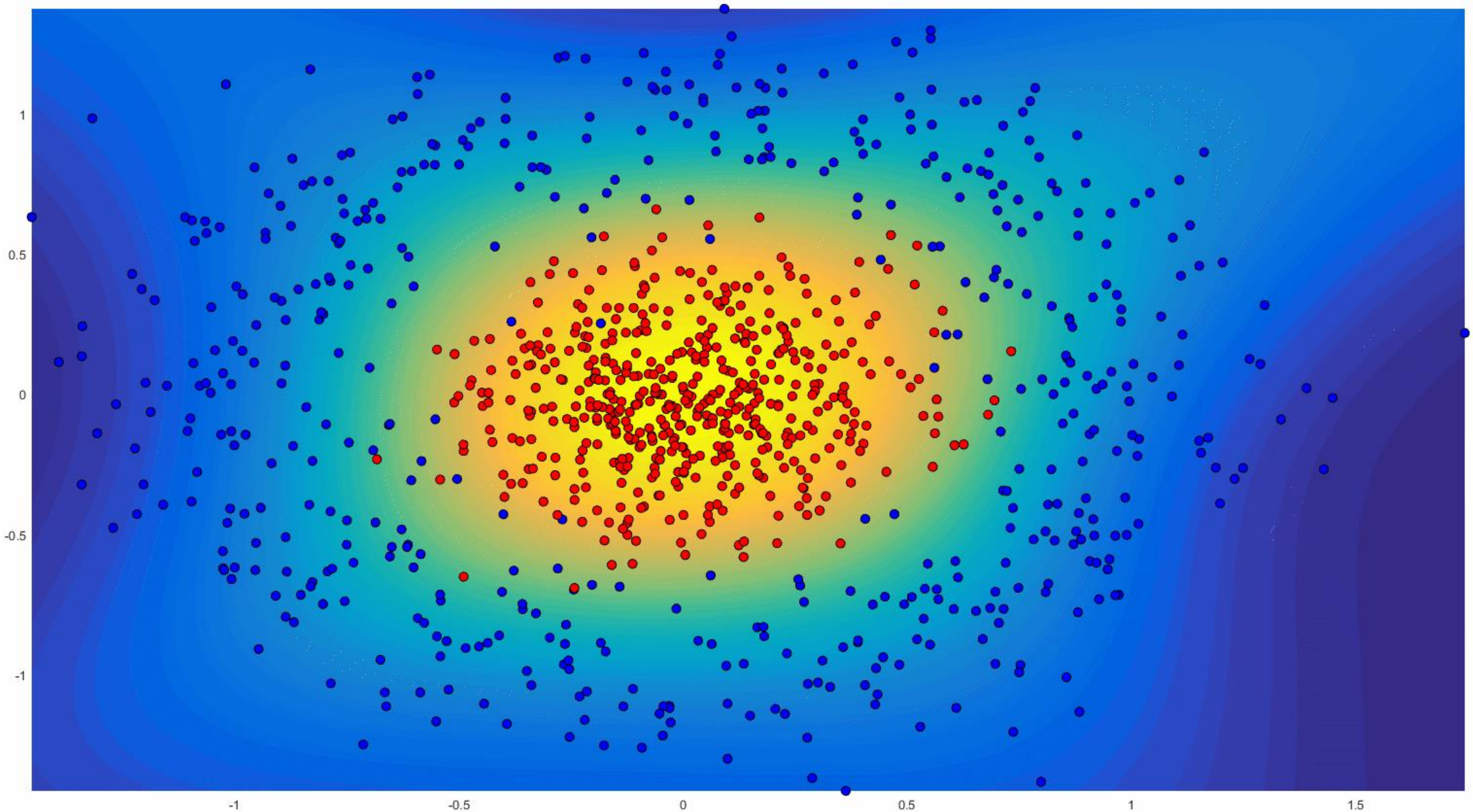


Demo 2

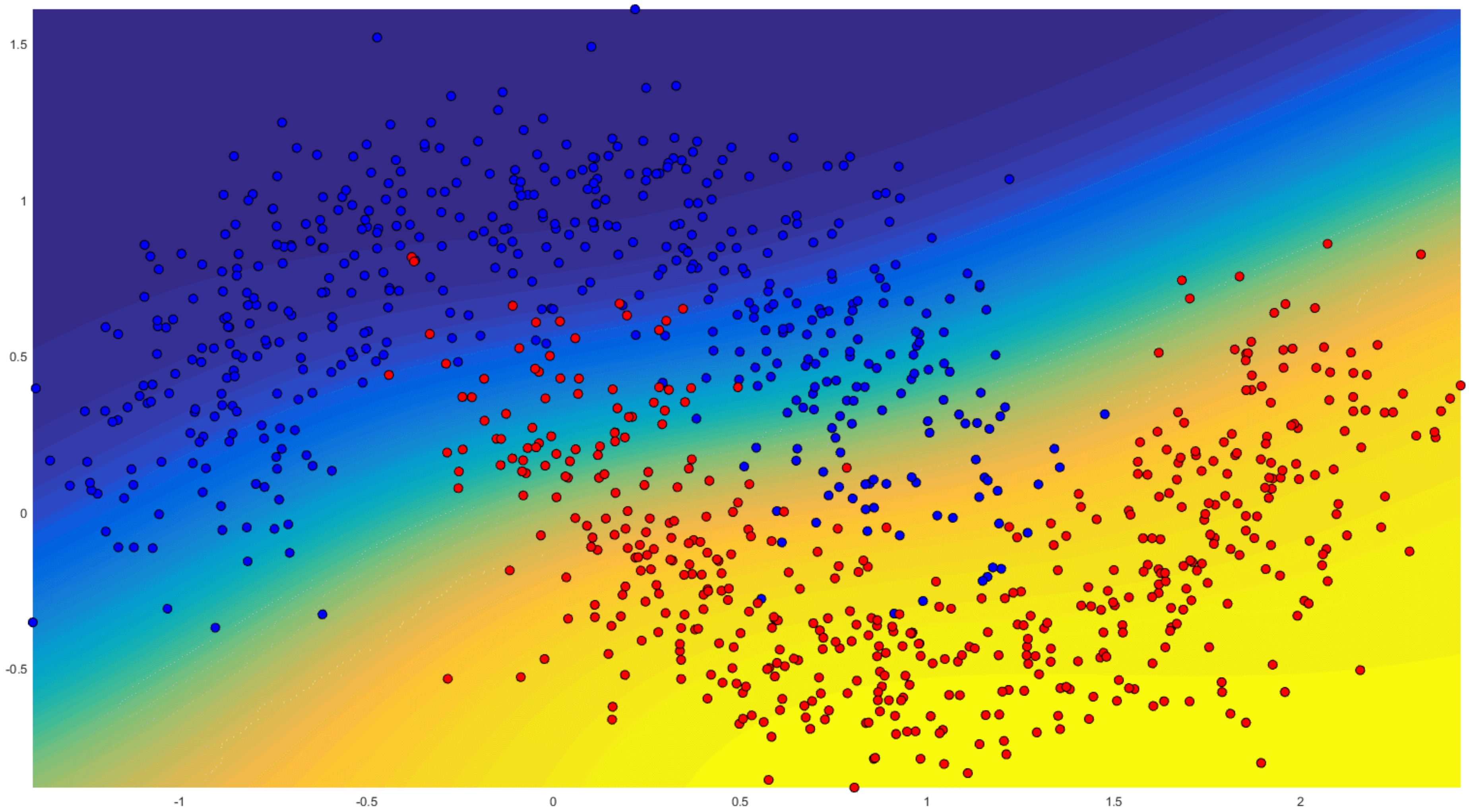
Backpropagation



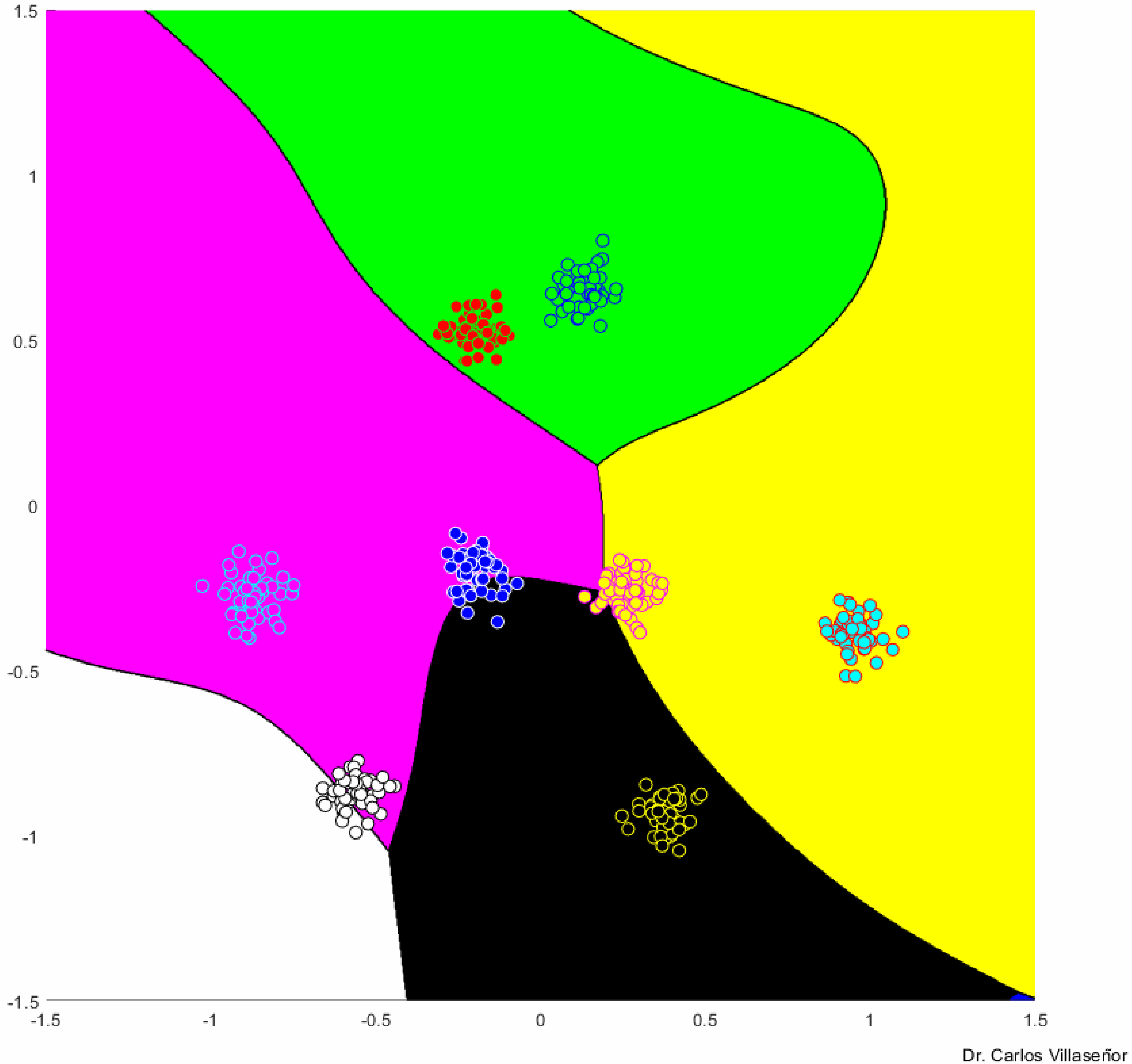
Backpropagation



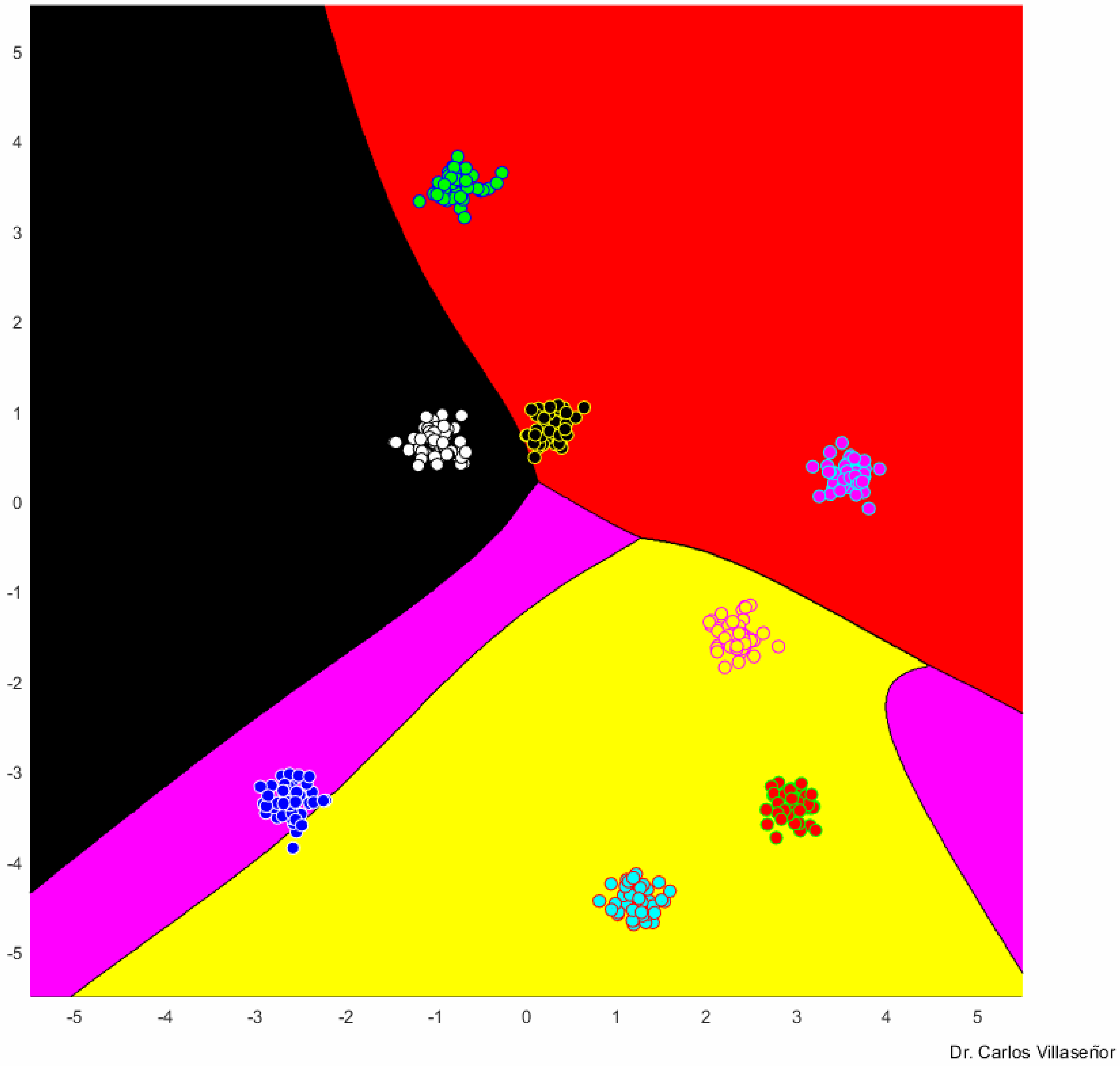
Backpropagation



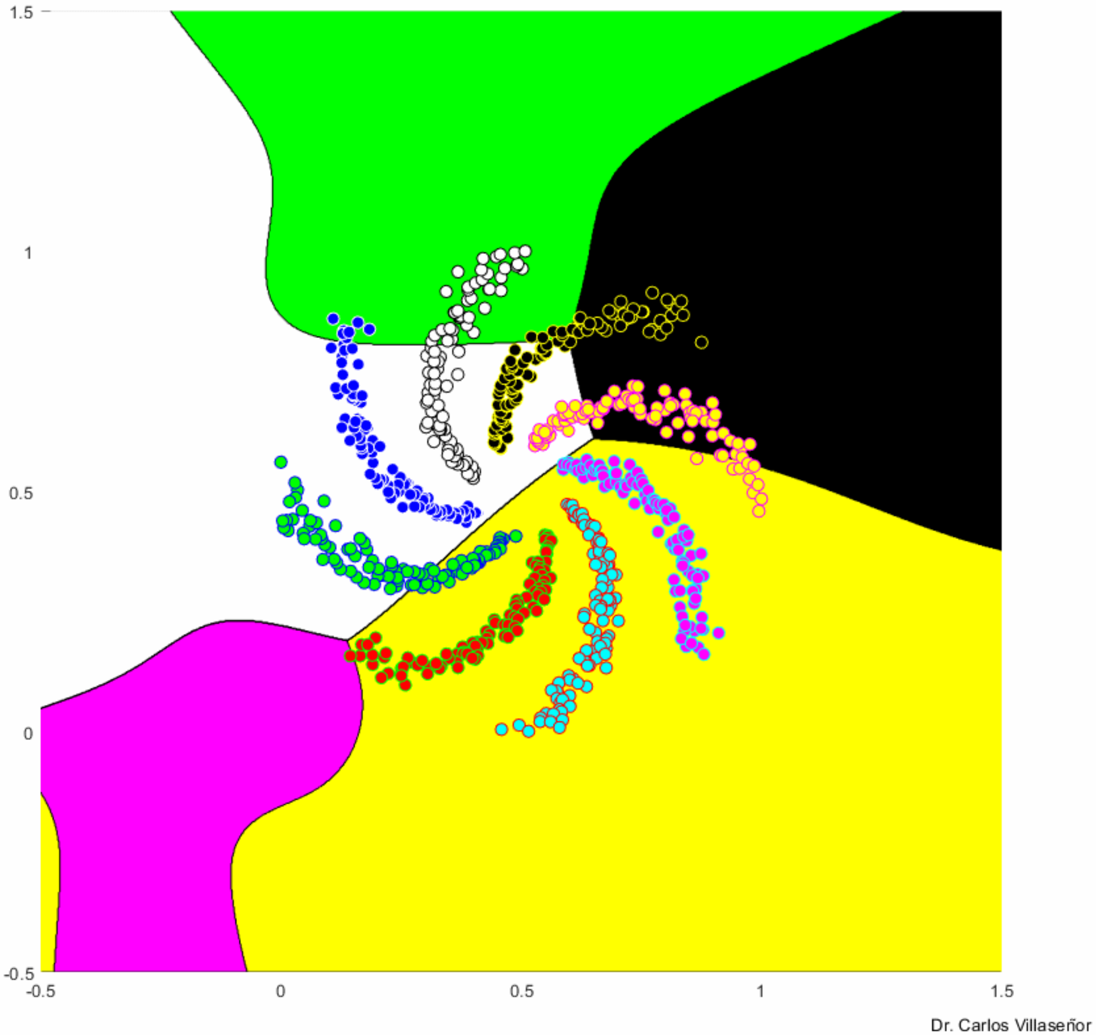
MLP with Softmax

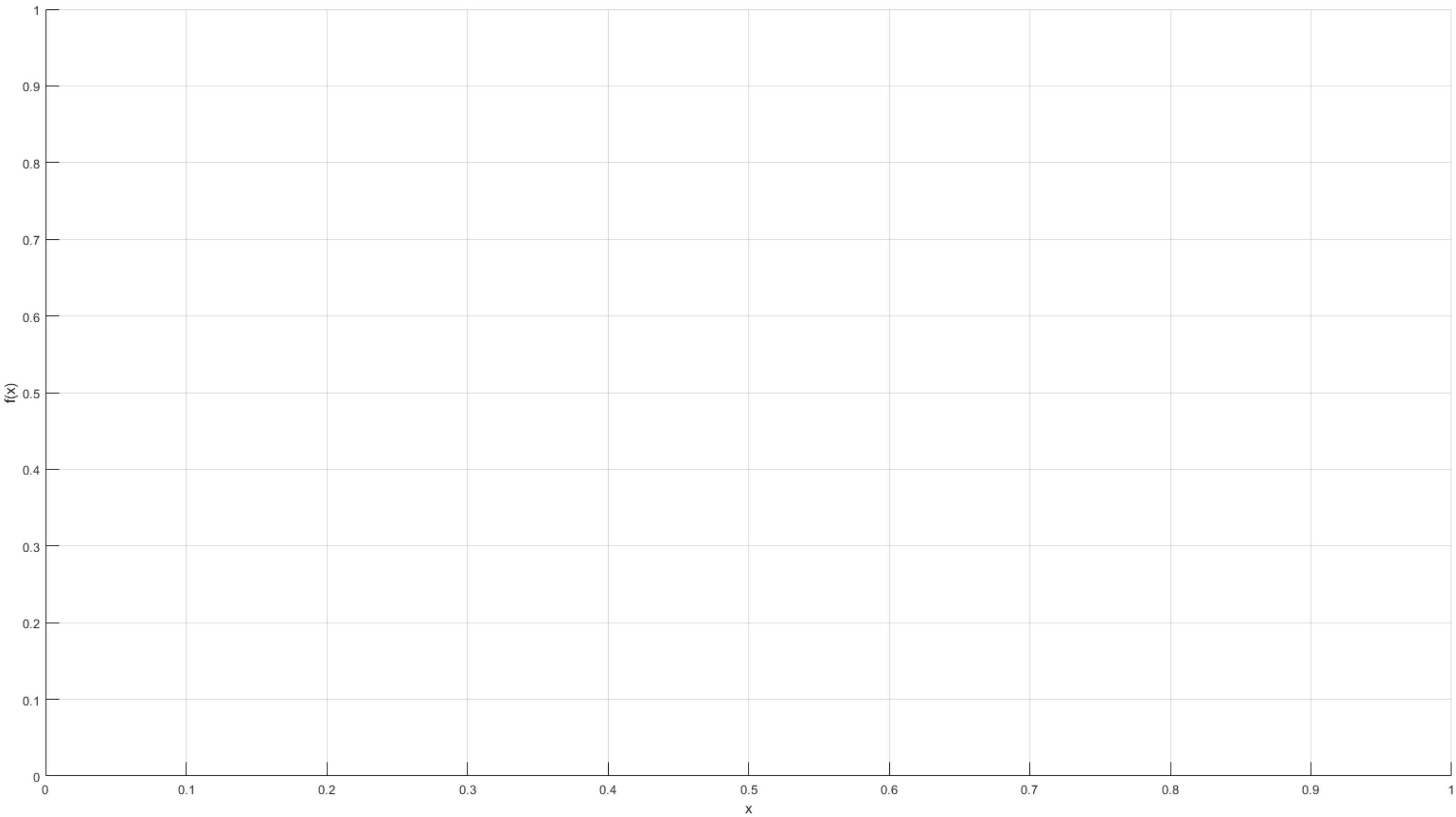


MLP with Softmax

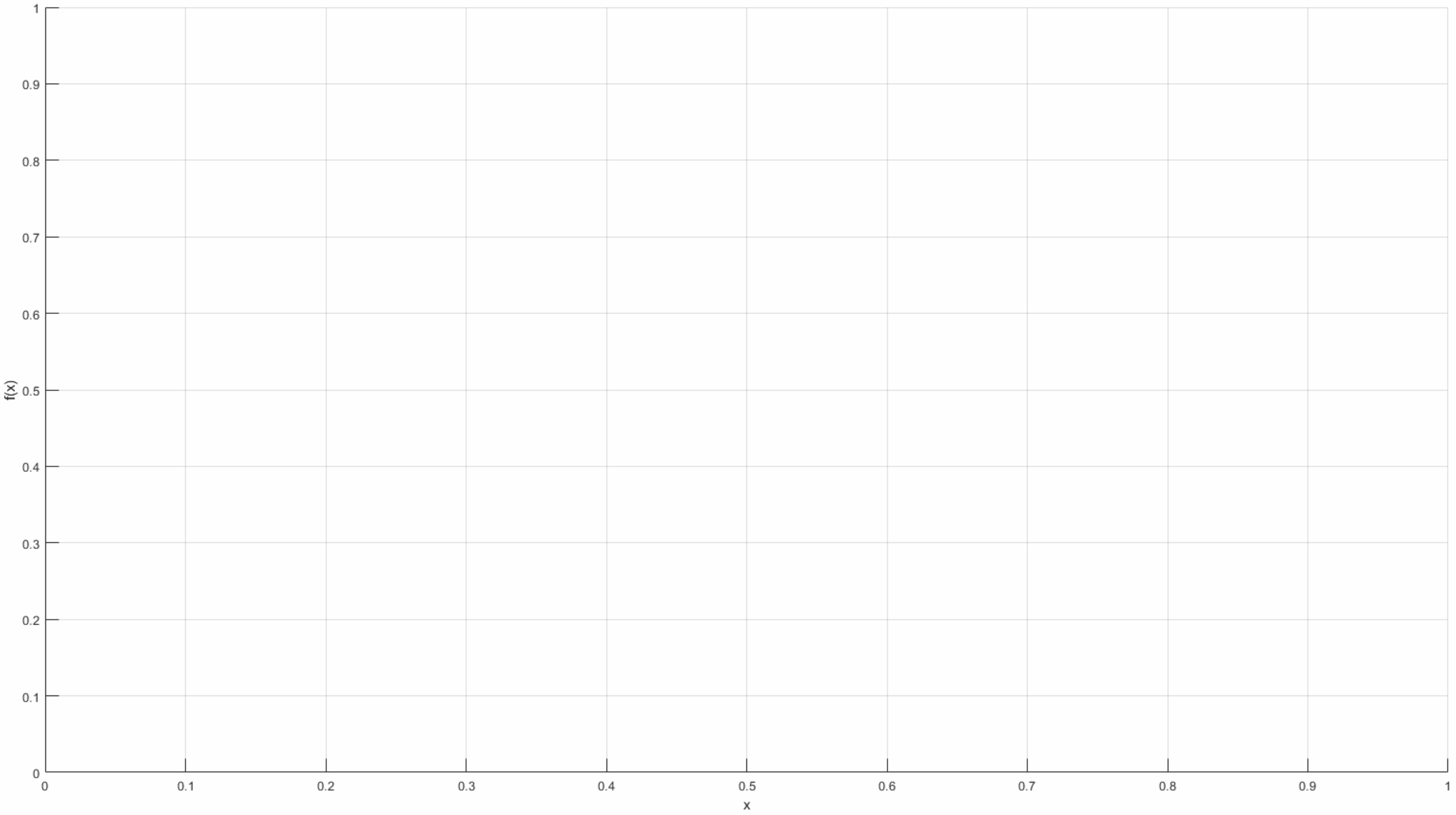


MLP with Softmax

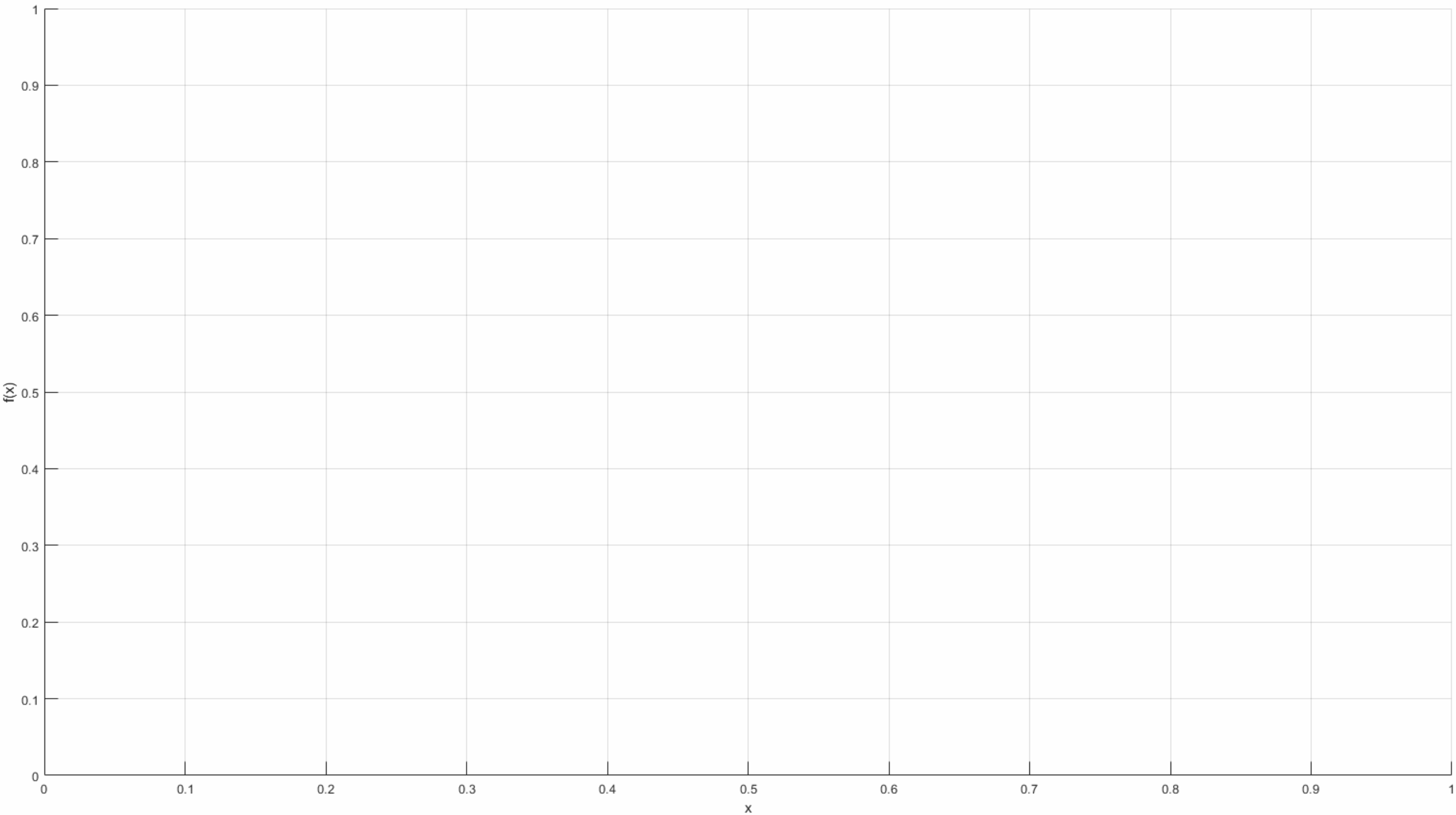




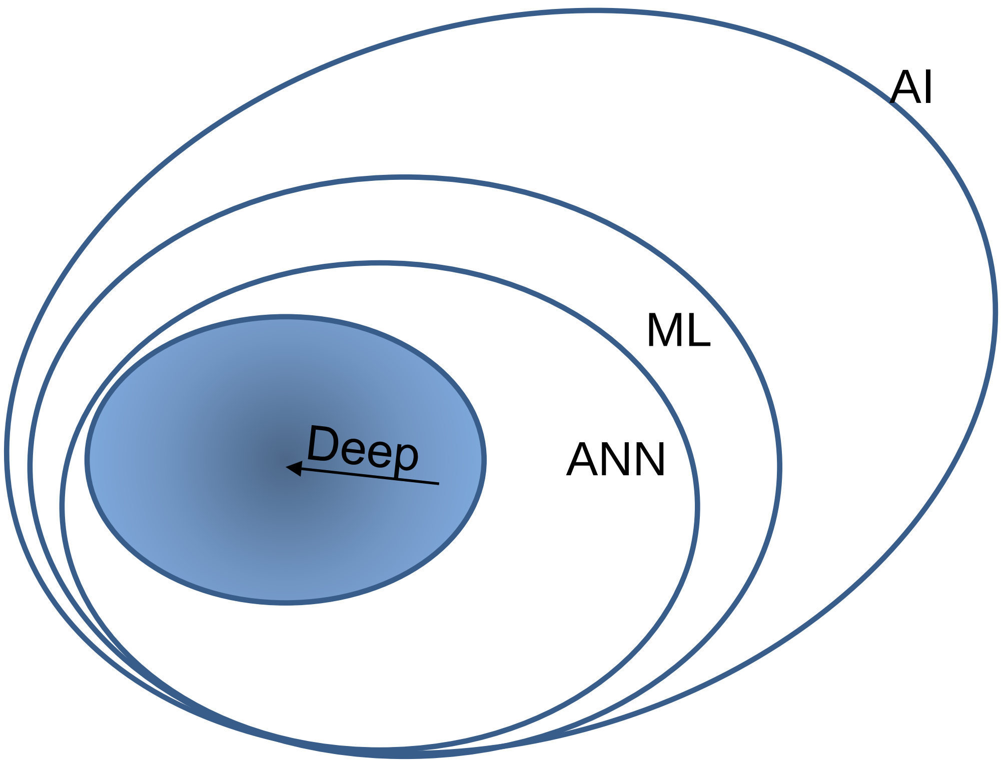
f(x)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
f(x)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
x

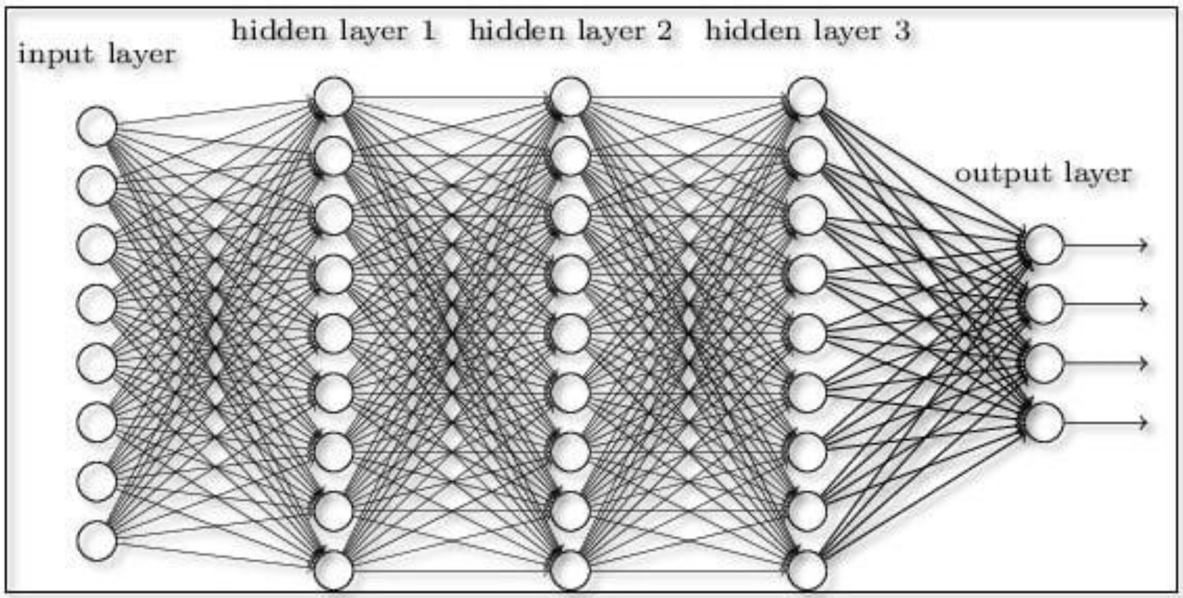

f(x)
x

Aprendizaje prodfungo



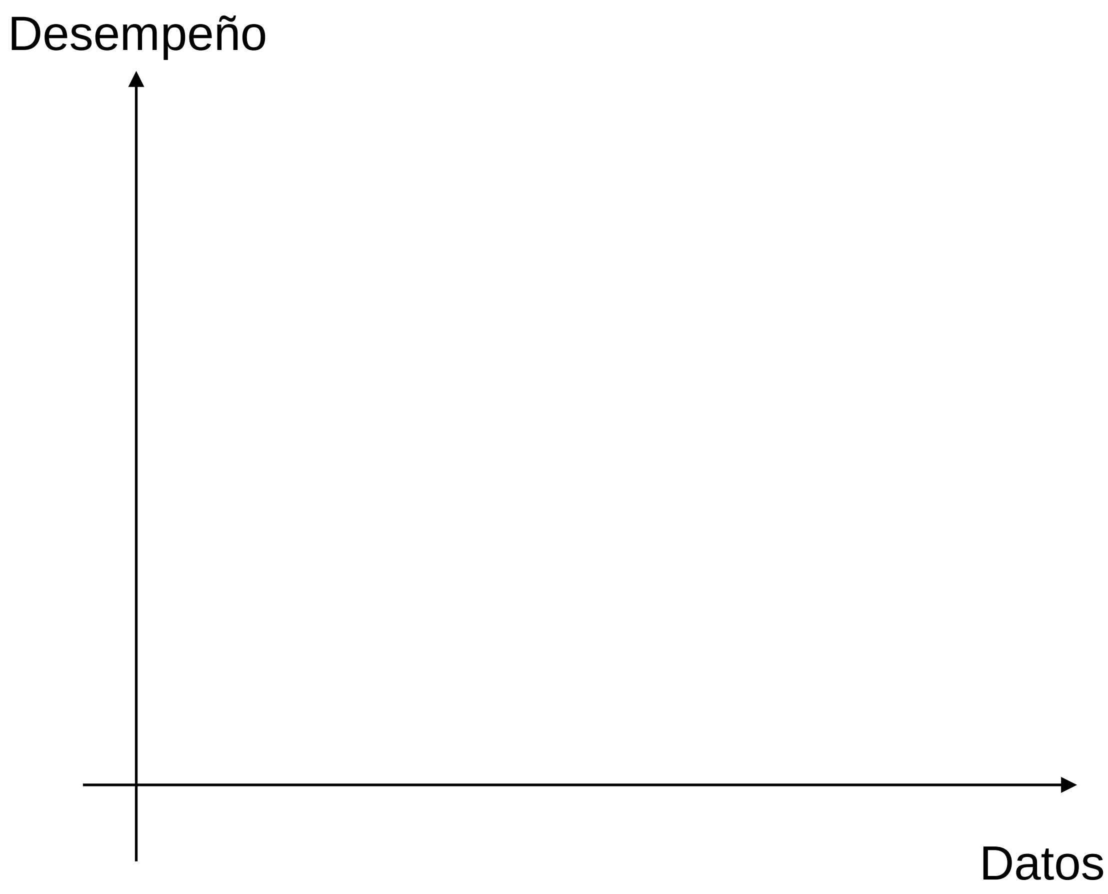
Deep Learning

Profunda (Deep)

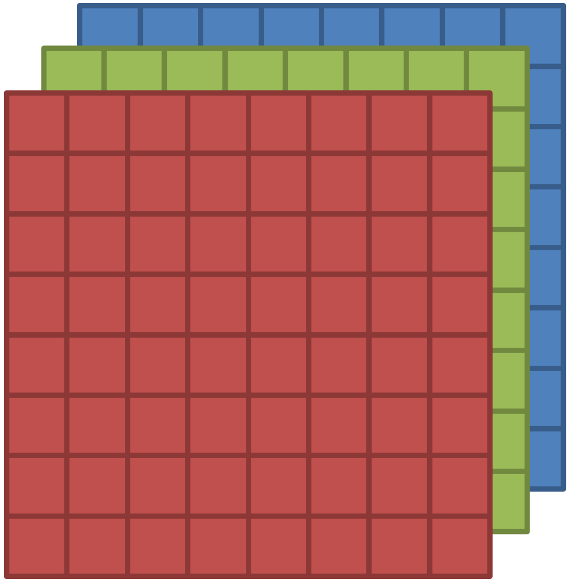


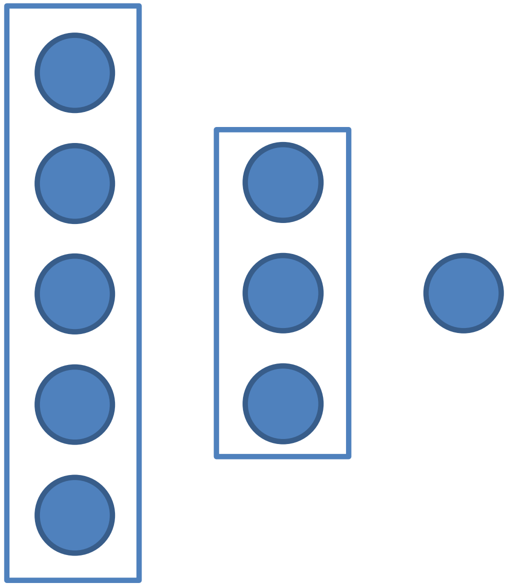
Superficial (Shallow)

Deep Learning vs Machine Learning



Visión con una Red Neuronal





Si / No

Para una imagen de 626x417 el número de pixeles que hay es de 261,042, multiplicando por 3 capas de color, obtenemos 783,126. Si en la primer capa oculta hay 5 neuronas, el total de parámetros en la primer capa es de 3,915,635

Red Convolucional

En 1989, Yann LeCun presentó una red que entrenaba filtros convolucionales con back-propagation.

Convoluciones

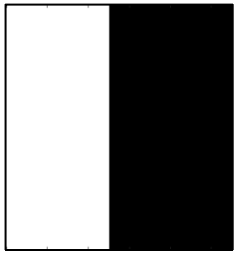
$$
\begin{bmatrix} 3 & 0 & 1 & 2 & 7 & 4 \\ 1 & 5 & 8 & 9 & 3 & 1 \\ 2 & 7 & 2 & 5 & 1 & 3 \\ 0 & 1 & 3 & 1 & 7 & 8 \\ 4 & 2 & 1 & 6 & 2 & 8 \\ 2 & 4 & 5 & 2 & 3 & 9 \end{bmatrix} * \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -12 & -8 & 5 & 16 \\ -10 & 0 & 3 & 5 \\ -3 & -2 & -8 & -14 \\ 0 & -6 & -4 & -18 \end{bmatrix}
$$

$$6 \times 6 \qquad 3 \times 3 \qquad 4 \times 4$$

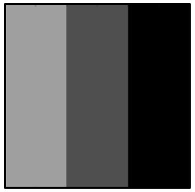
$$n_h \times n_w * f \times f \rightarrow n_h - f + 1 \times n_w - f_2 + 1$$

Convoluciones

9	9	9	0	0	0
9	9	9	0	0	0
9	9	9	0	0	0
9	9	9	0	0	0
9	9	9	0	0	0
9	9	9	0	0	0



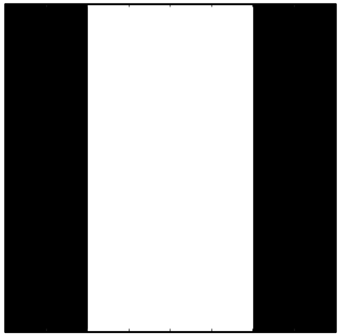
*

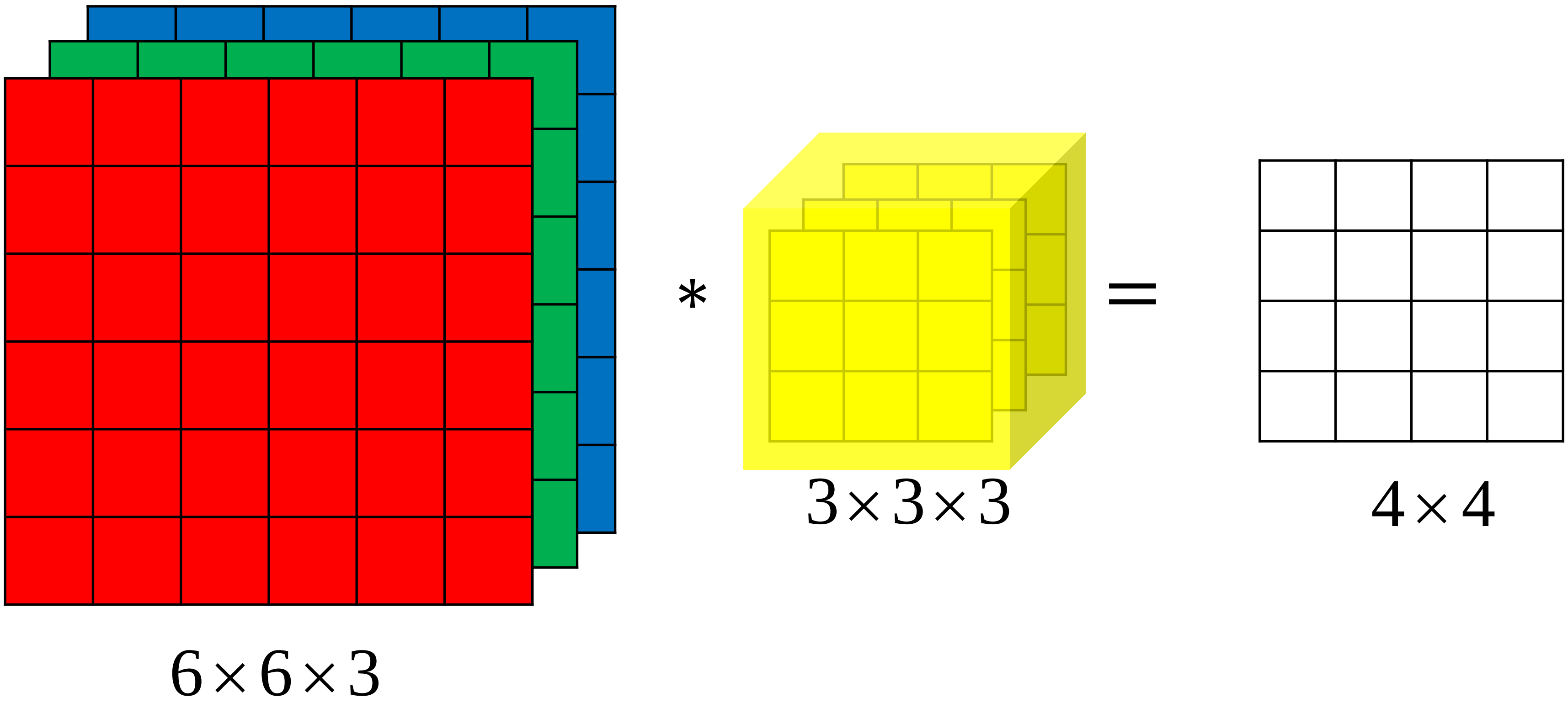


1	0	-1
1	0	-1
1	0	-1

=

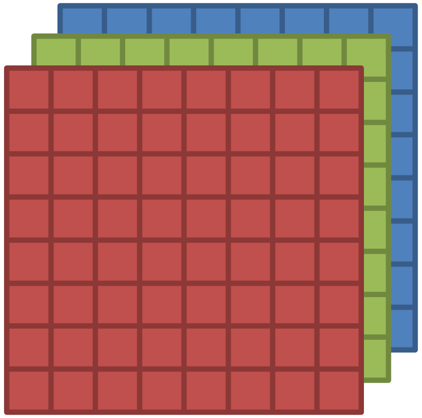
0	27	27	0
0	27	27	0
0	27	27	0
0	27	27	0



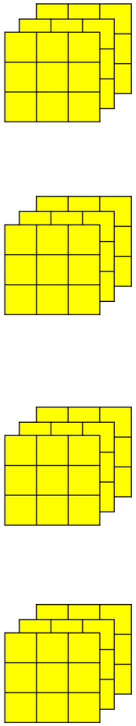


$6 \times 6 \times 3$
*
$3 \times 3 \times 3$
=
4×4

Múltiples filtros



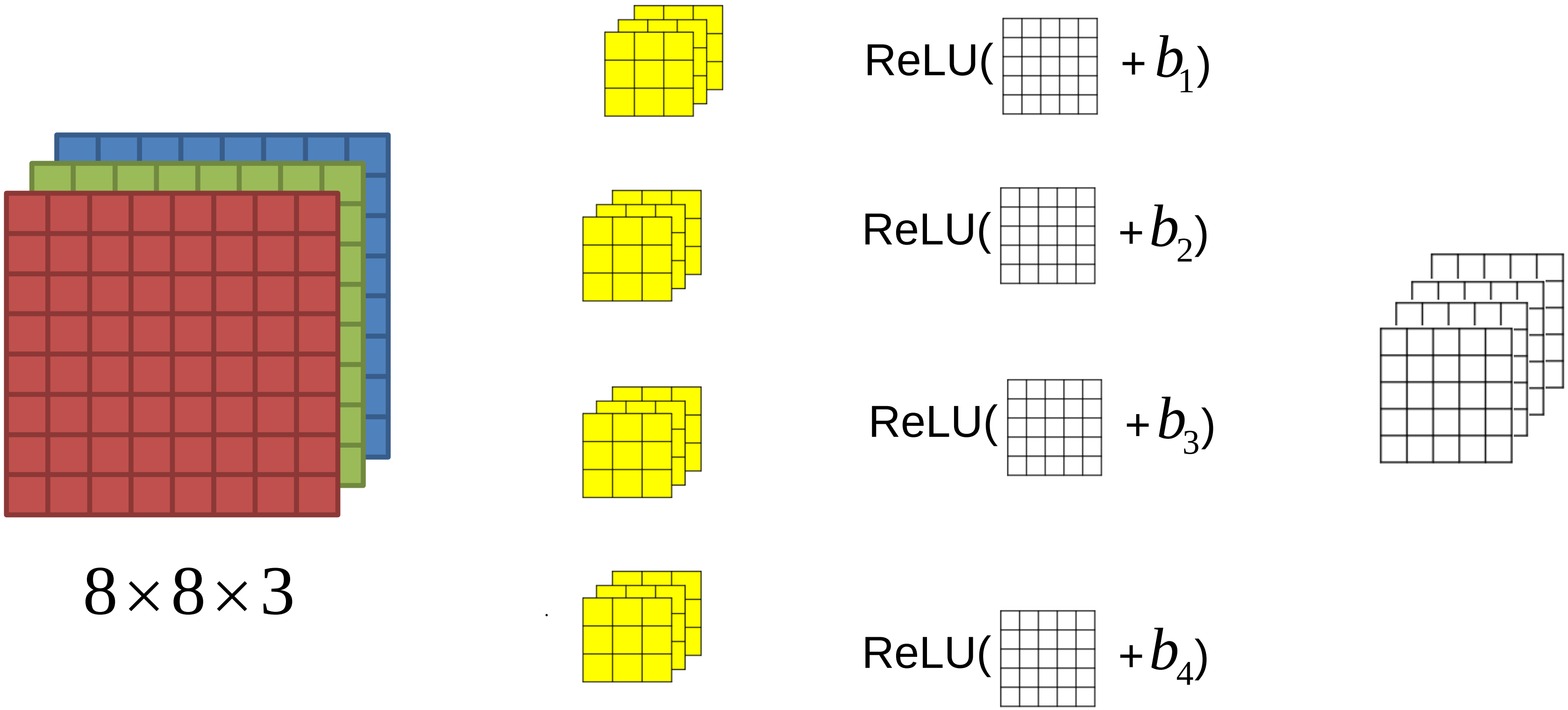
$8 \times 8 \times 3$



$3 \times 3 \times 3$

Filtros=4

Capa convolucional



$8 \times 8 \times 3$

Capa convolucional

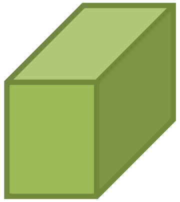
$$n_h \times n_w \times n_c$$

$$f \times f \times n_c$$

$$s$$

$$p$$

$$n_f$$

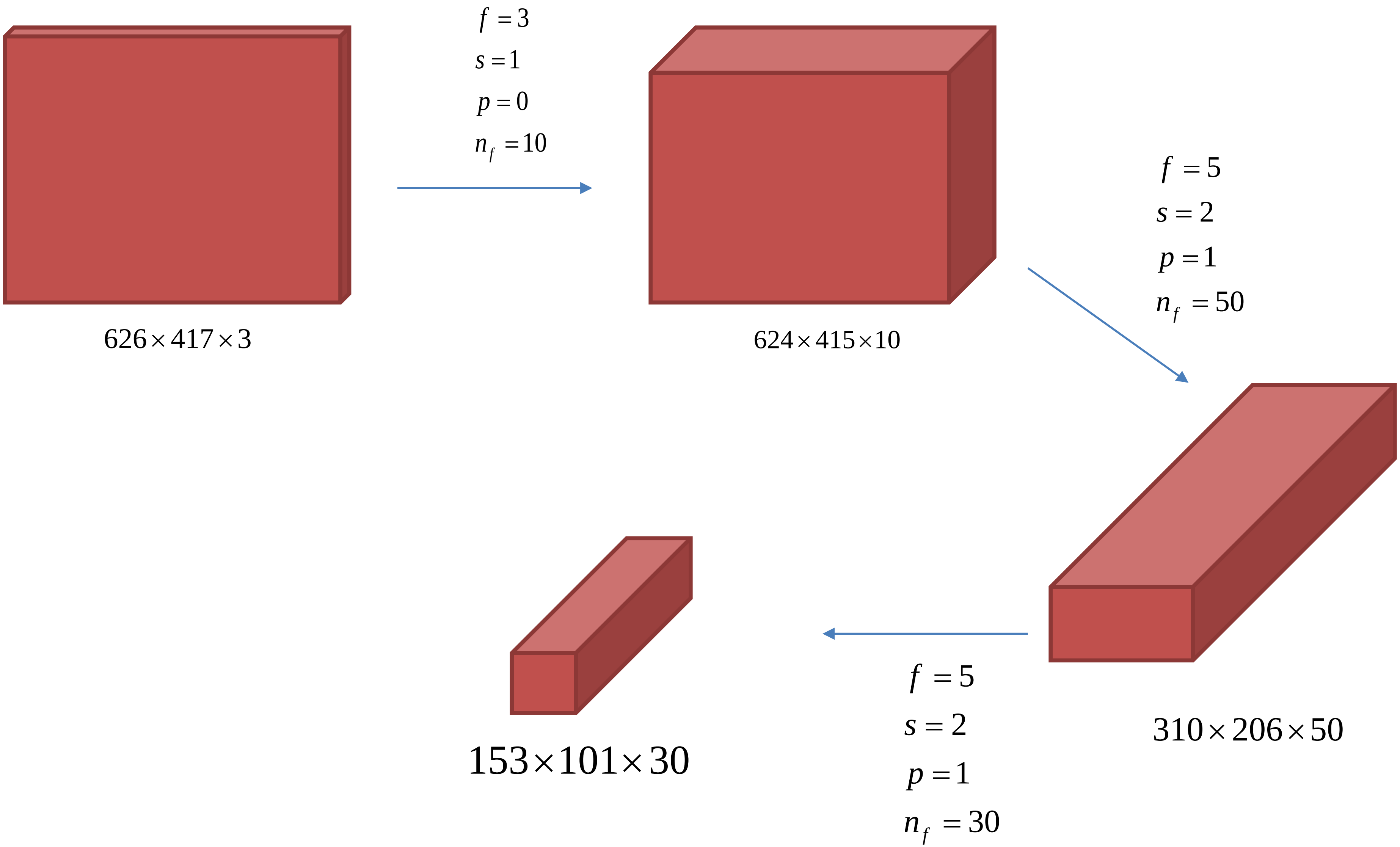


$$\left\lfloor \frac{n_h + 2p - f}{s} + 1 \right\rfloor \times \left\lfloor \frac{n_w + 2p - f}{s} + 1 \right\rfloor \times n_f$$

Conteo de parámetros

Para una imagen RGB de 626x417 y si suponemos 5 filtros convolucionales de (3, 3) se tienen los siguientes parámetros:

$$5(3\times 3\times 3)+5=140$$


$f = 3$
$s = 1$
$p = 0$
$n_f = 10$
$626 \times 417 \times 3$
$624 \times 415 \times 10$
$f = 5$
$s = 2$
$p = 1$
$n_f = 50$
$310 \times 206 \times 50$
$f = 5$
$s = 2$
$p = 1$
$n_f = 30$
$153 \times 101 \times 30$

Pooling

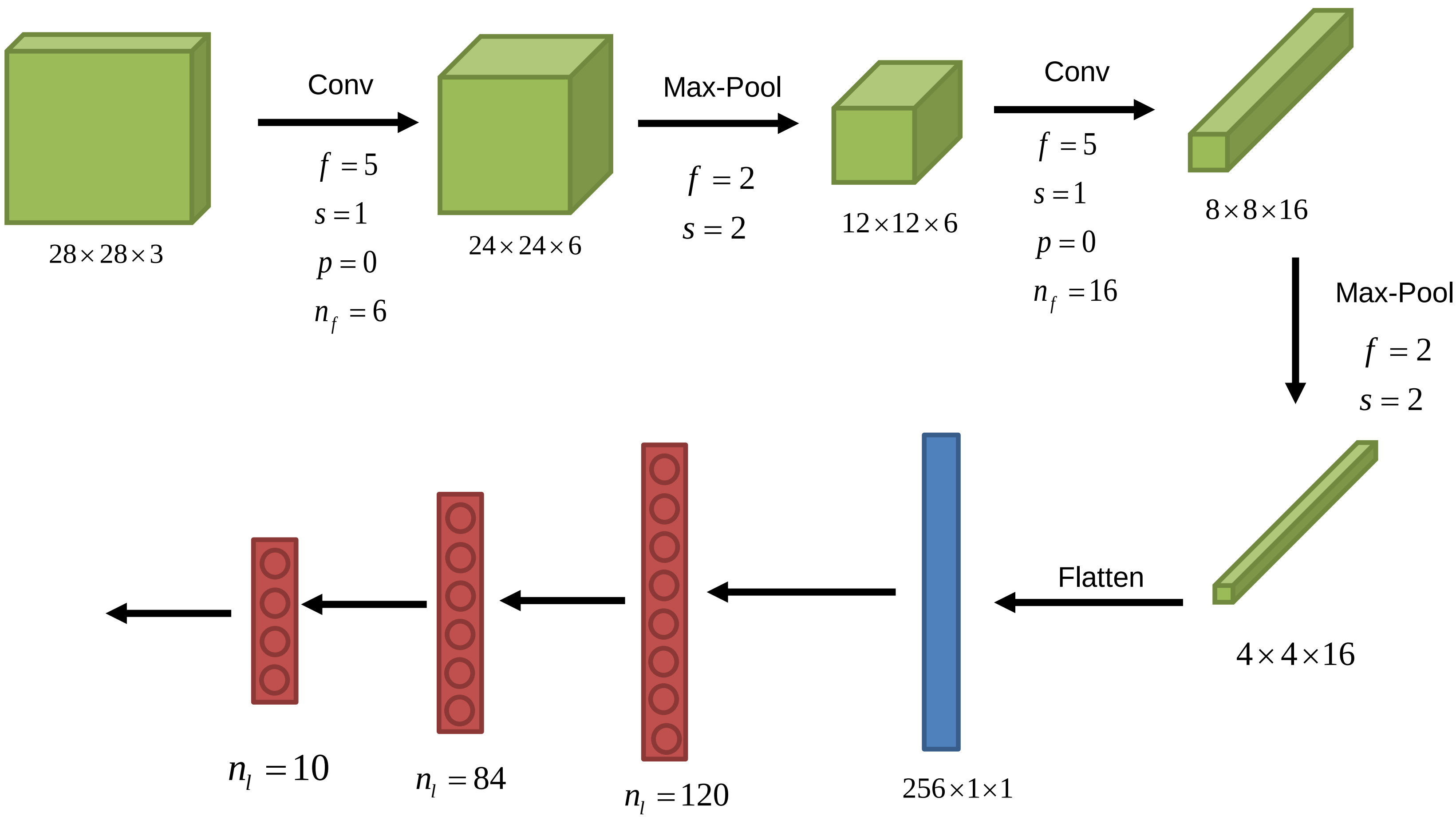
Max-Pooling:

5	3	2	6
7	4	2	1
5	7	8	1
9	5	4	1

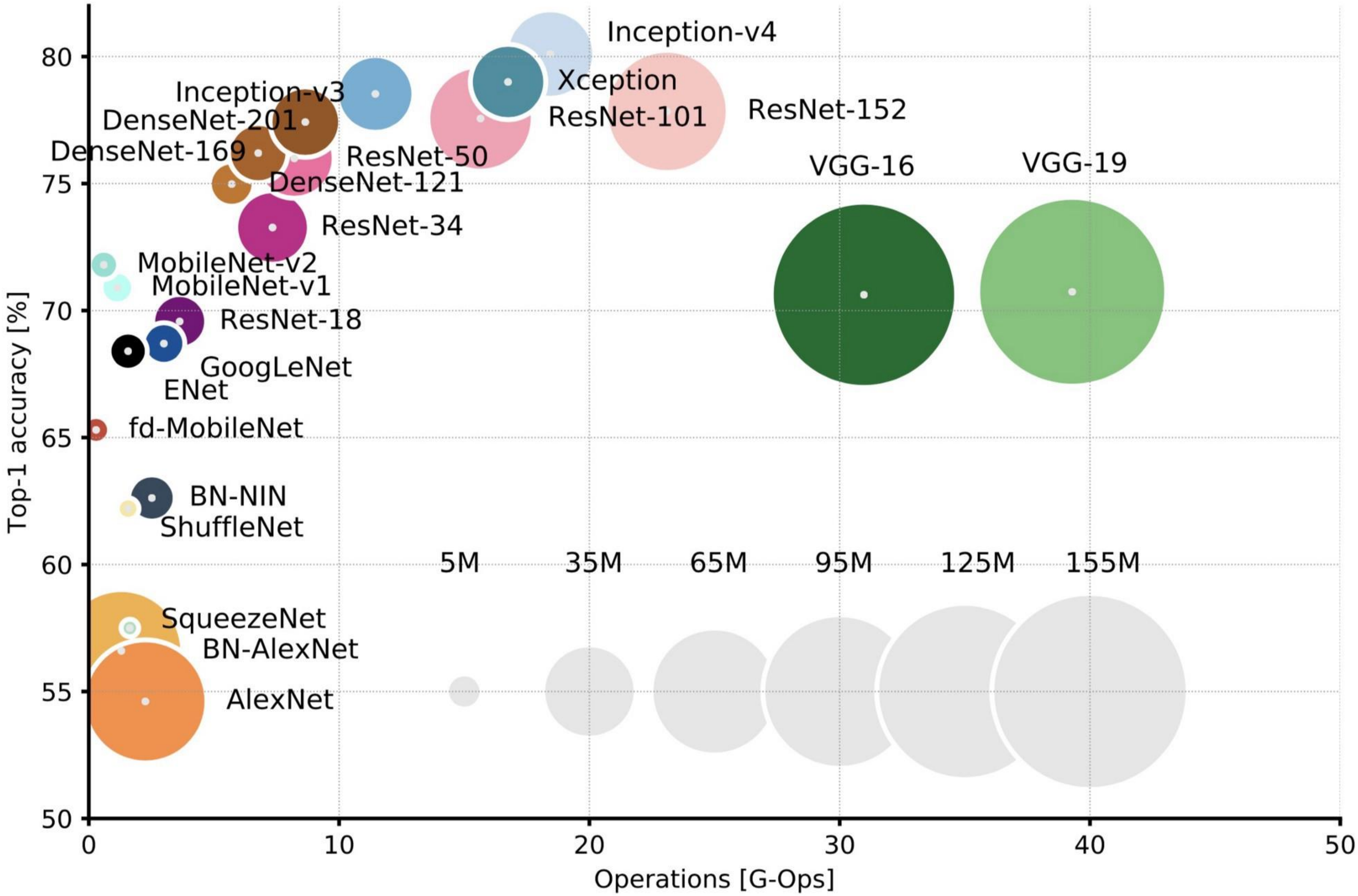
$$\xrightarrow[s=2]{f=2}$$

7	6
9	8

Red Neuronal Convolucional (CNN)



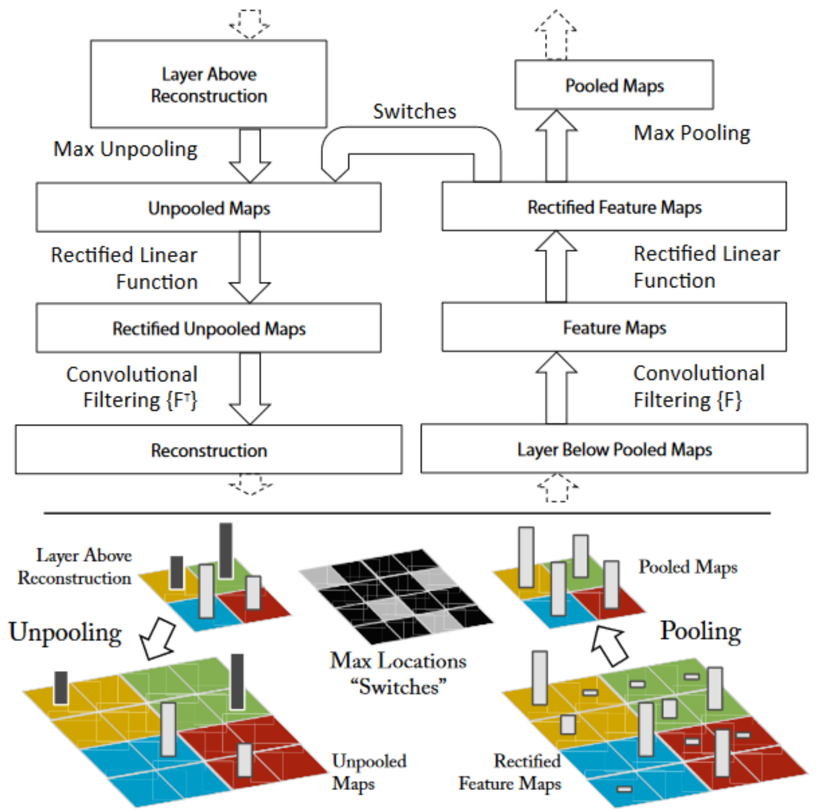
Demo 3


Inception-v4
Xception
Inception-v3
DenseNet-201
ResNet-101
ResNet-152
DenseNet-169
ResNet-50
DenseNet-121
VGG-16
VGG-19
ResNet-34
MobileNet-v2
MobileNet-v1
ResNet-18
GoogLeNet
ENet
fd-MobileNet
BN-NIN
ShuffleNet
5M
35M
65M
95M
125M
155M
SqueezeNet
BN-AlexNet
AlexNet
Top-1 accuracy [%]
80
75
70
65
60
55
50
0
10
20
30
40
50
Operations [G-Ops]

Aprendizaje de Transferencia

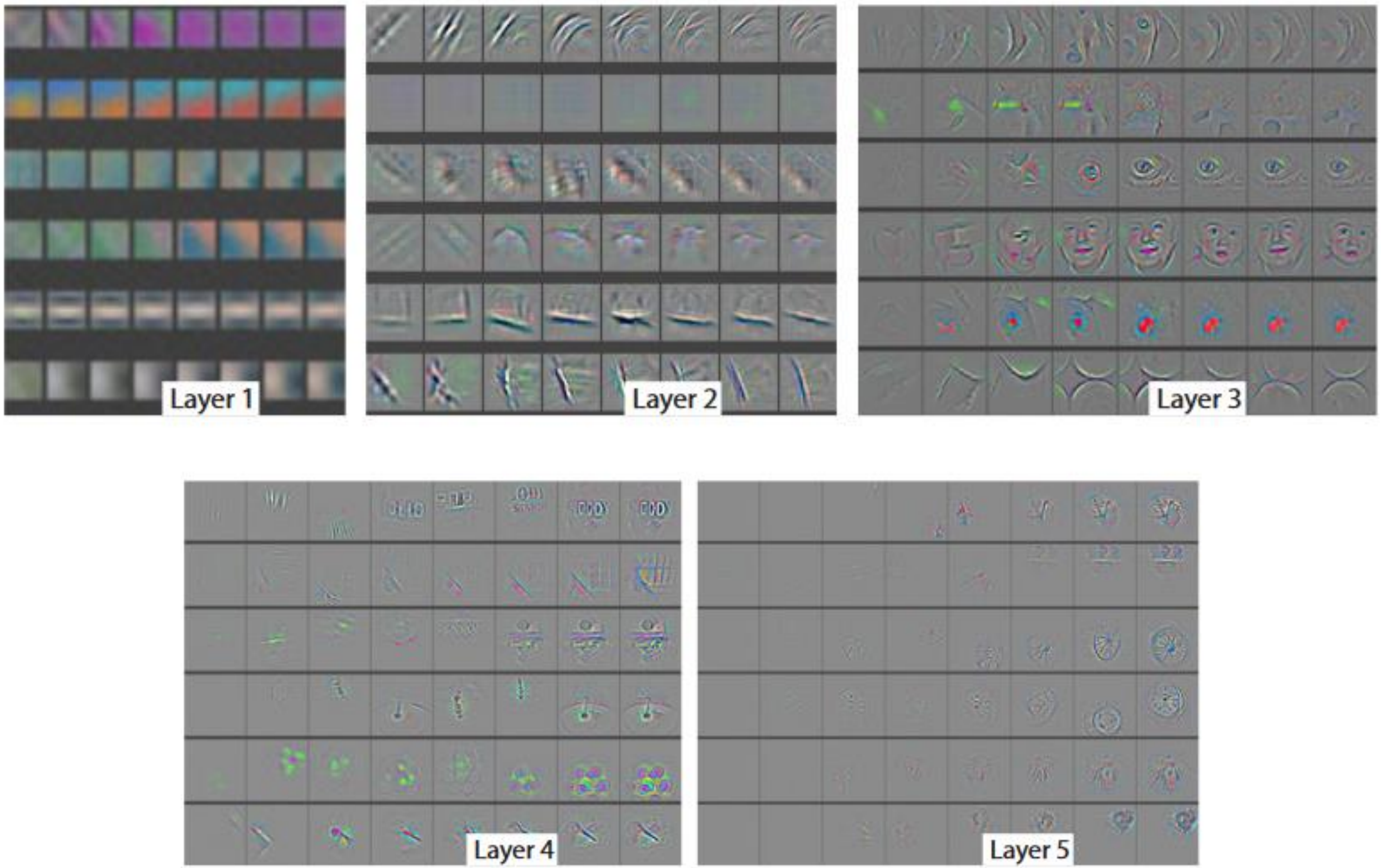
- Dificultades de Deep Learning
  - Entrenar una red neuronal profunda puede durar semanas y requiere múltiples GPUs.
  - Se necesita una cantidad enorme de datos para lograr un buen desempeño en una arquitectura profunda.

Aprendizaje de Transferencia

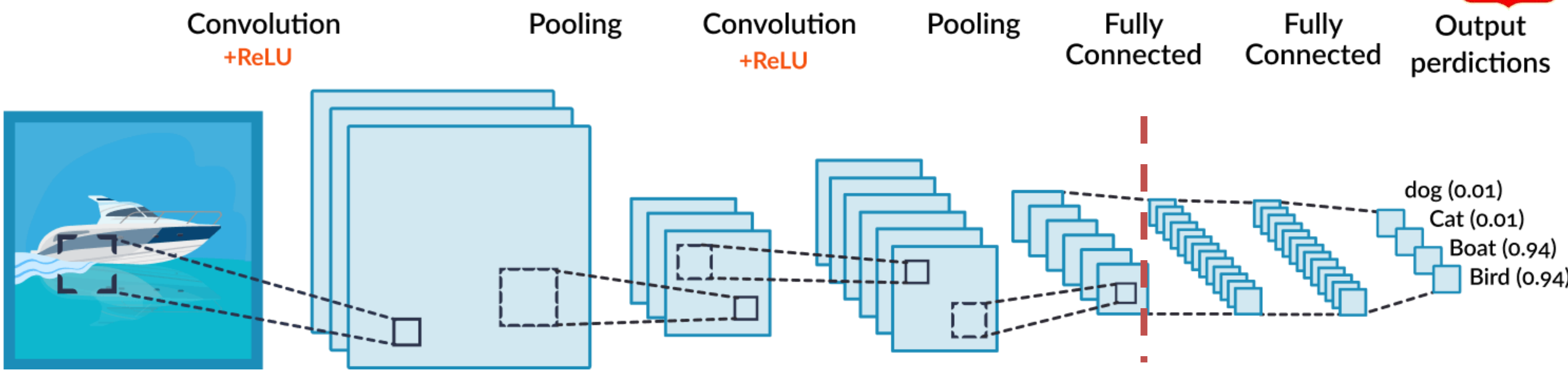


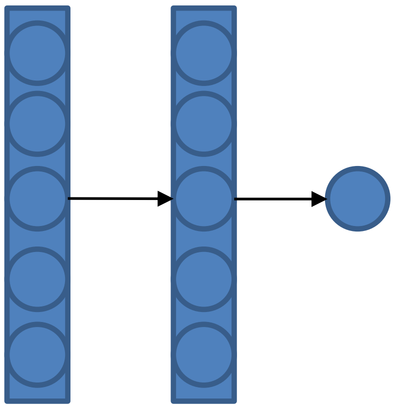
Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

Aprendizaje de Transferencia



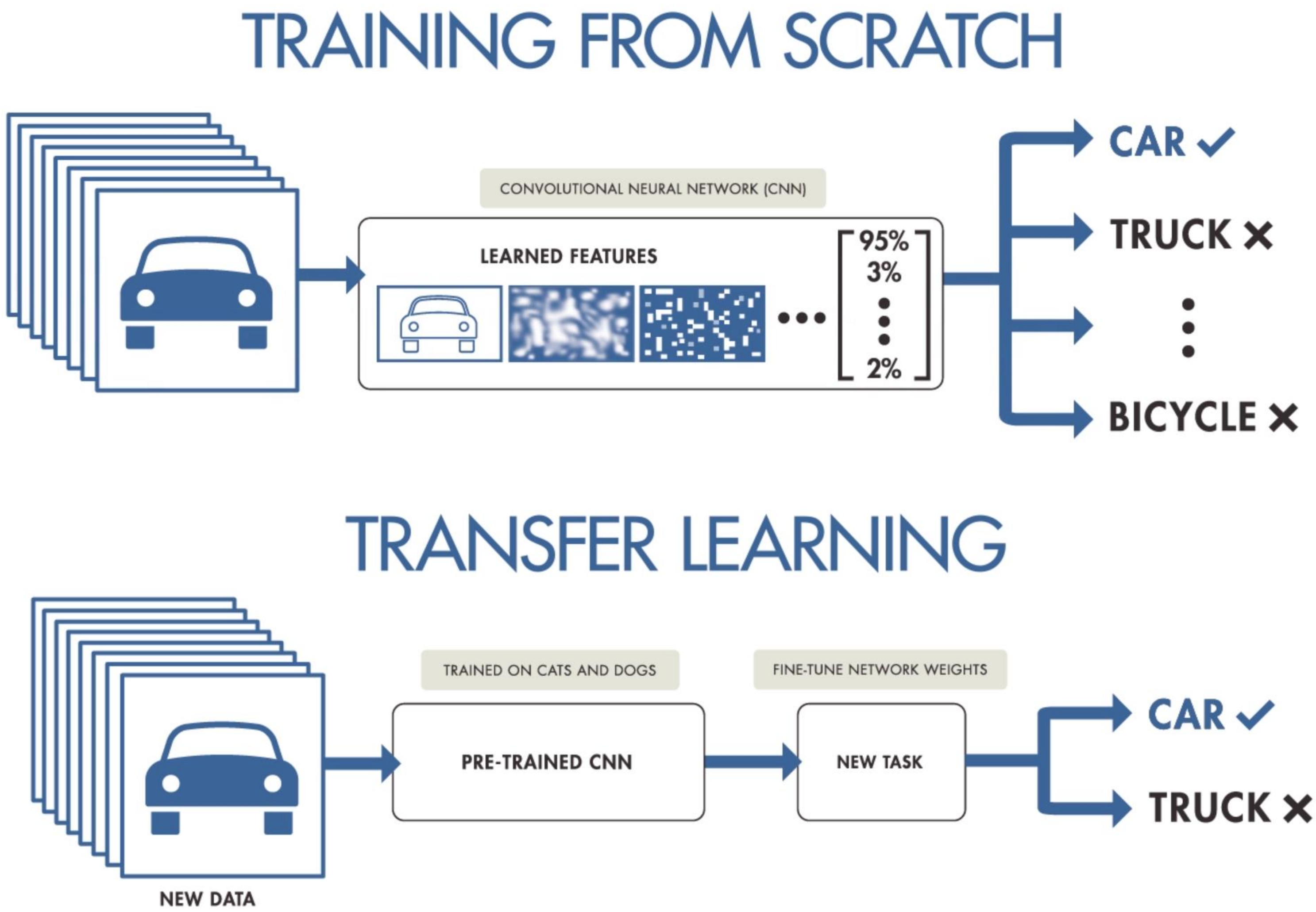
Aprendizaje de Transferencia





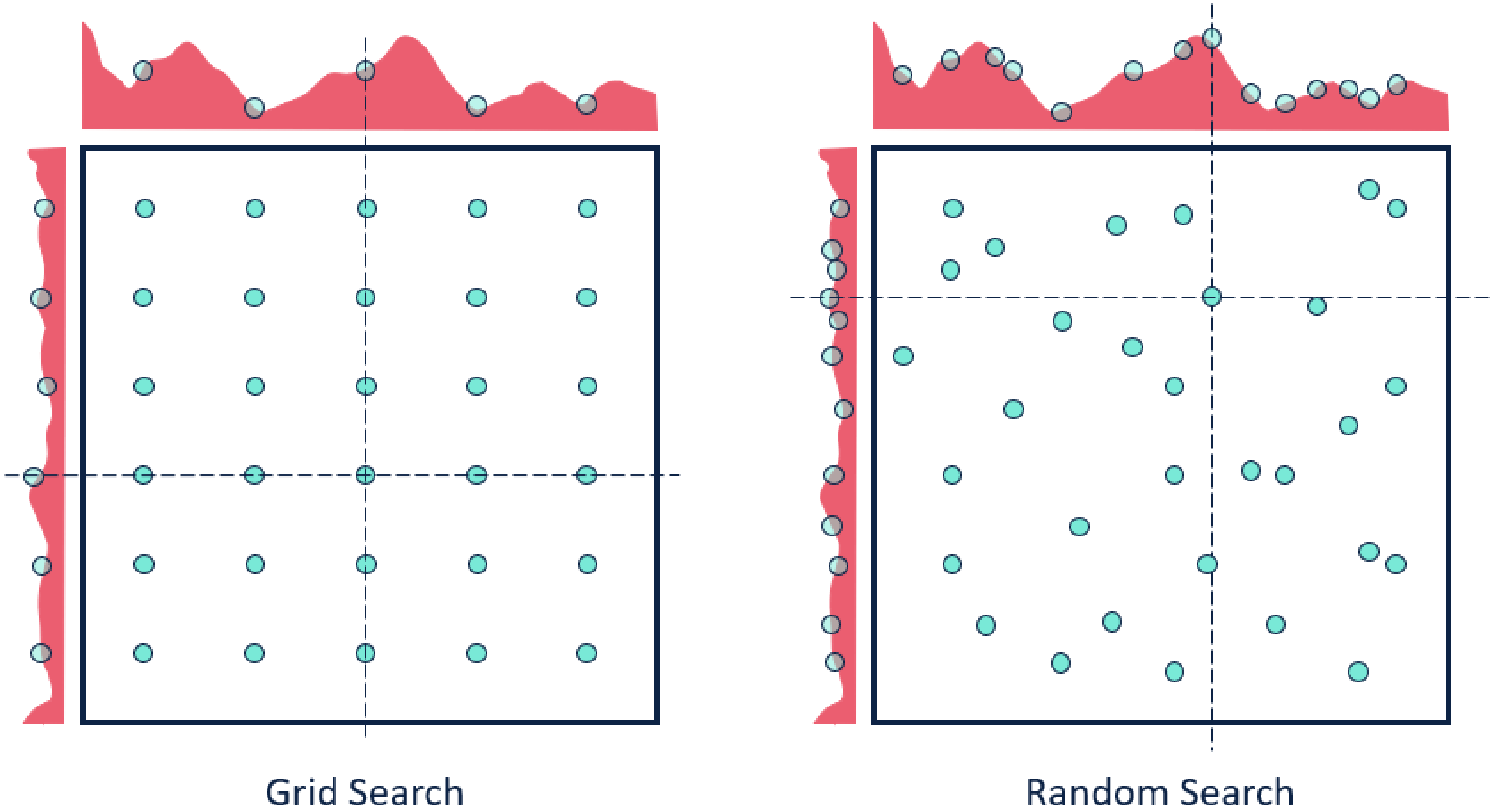
https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-tutorial-basic-advanced/

Aprendizaje de Transferencia

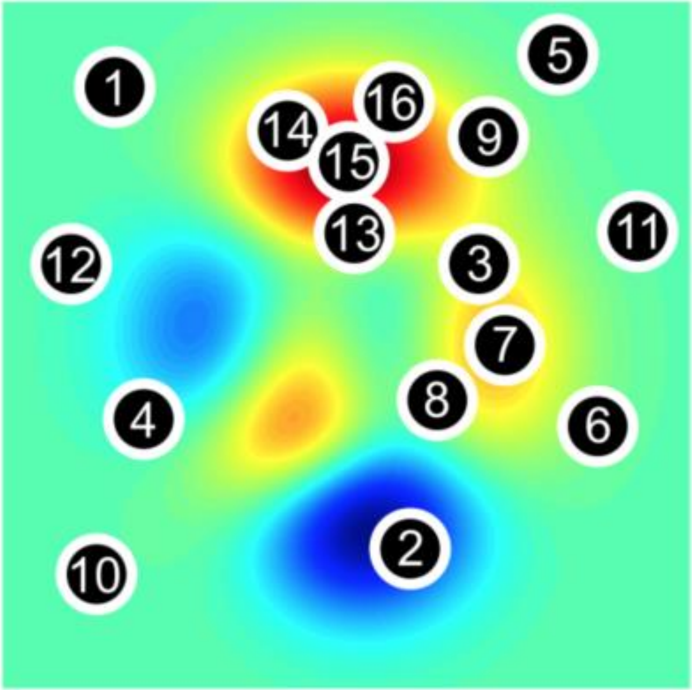


Demo 4

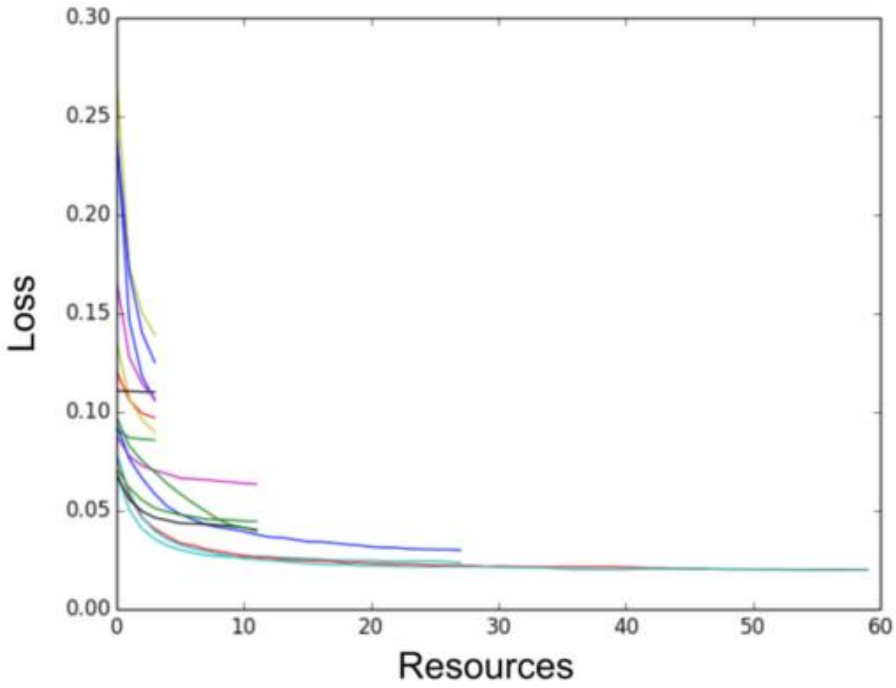
Búsqueda de hiperparámetros



Hyperband



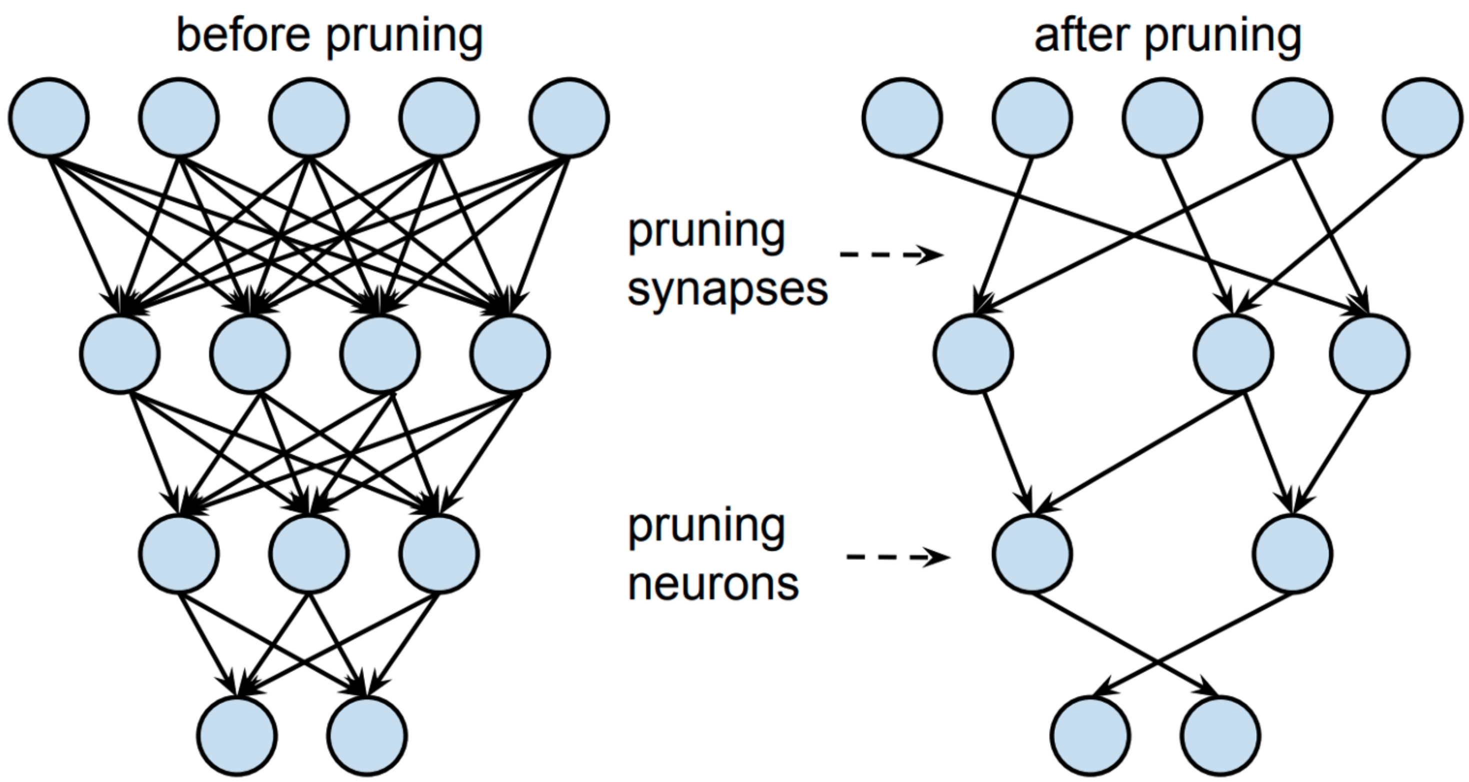
(a) Configuration Selection



(b) Configuration Evaluation

Demo 5

¿Para qué podar la red artificial ?



Poda neuronal

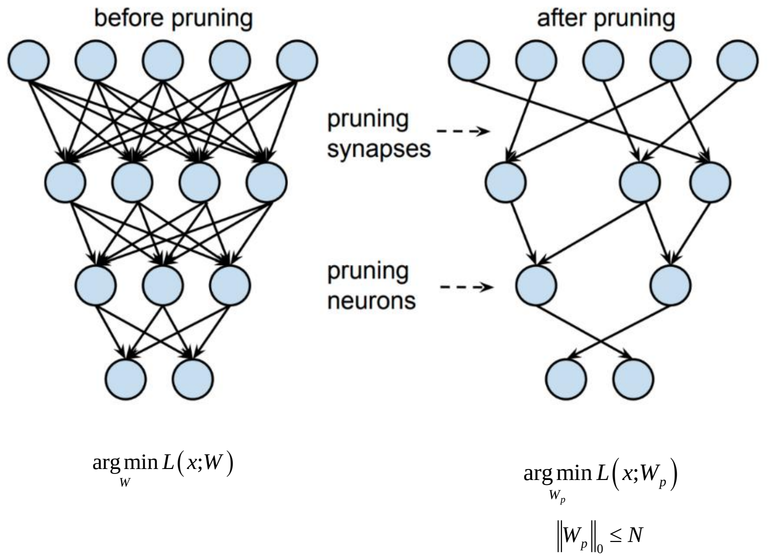
- Normalmente tenemos

$$\arg\min_{W} L(x;W)$$

cuándo hacemos poda suponemos

$$\arg\min_{W_p} L(x;W_p) \quad \|W_p\|_0 \leq N$$

- Características del podado
  1. Granularidad
  2. Criterio de poda
  3. Radio de poda
  4. Entrenamiento fino de red podada

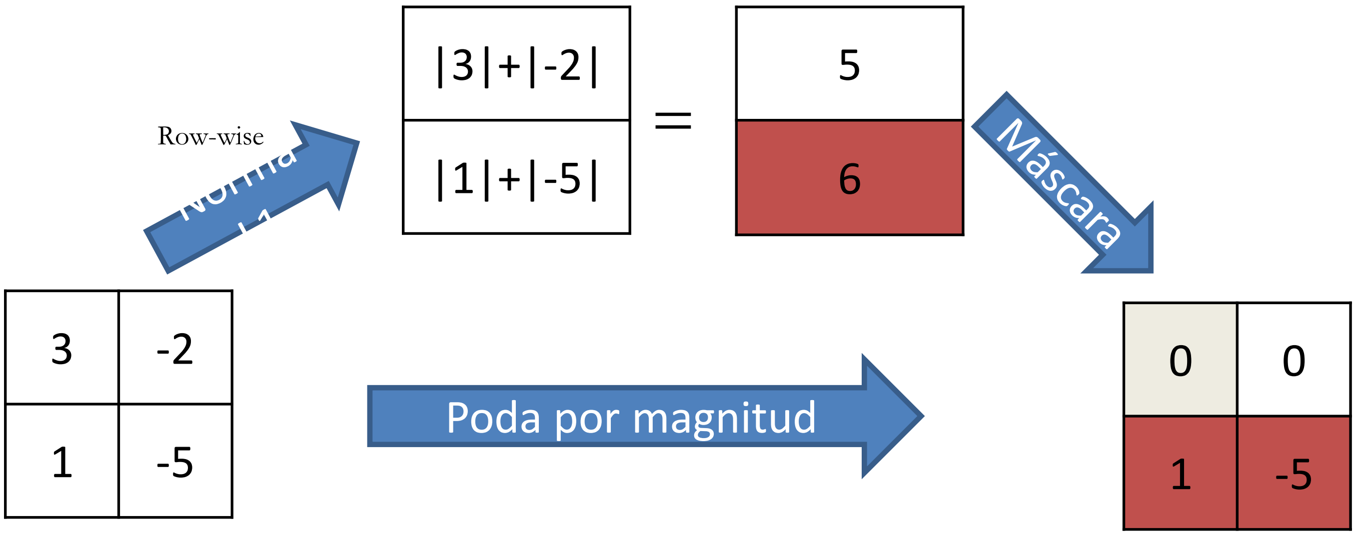


Criterio de poda

Existen muchos criterios de poda, a continuación veremos los más importantes

- Basados en la magnitud (Magnitud-based prunning)

$$\text{importancia} = |w|$$



Pruning Scheduling

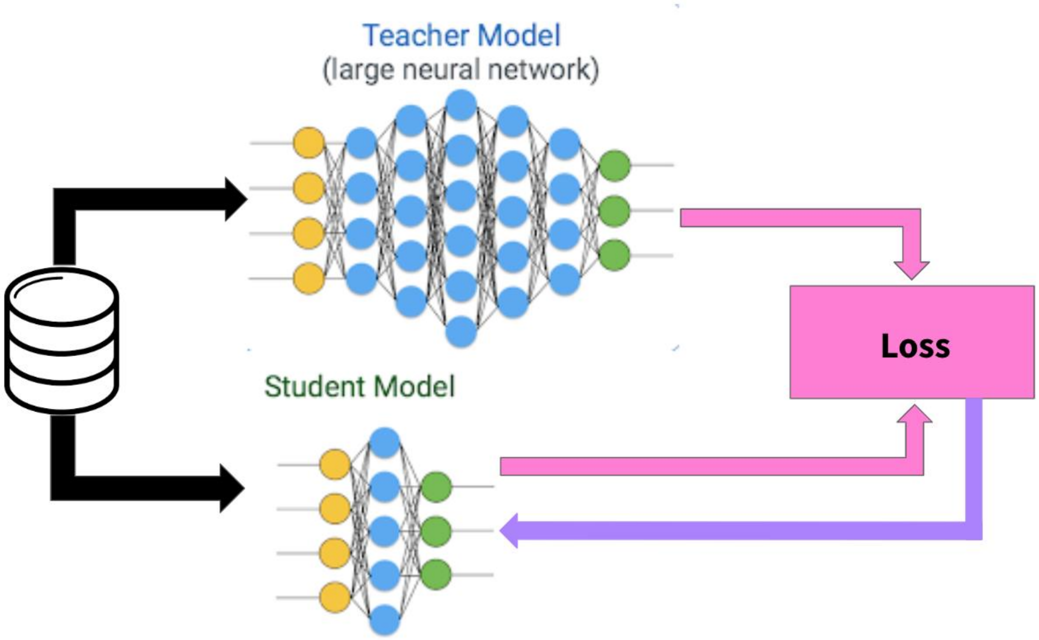
- Existen dos métodos populares
  - ConstantSparsity
    - La dispersión se mantiene constante durante e entrenamiento

Args
target_sparsity
begin_step
end_step
frequency

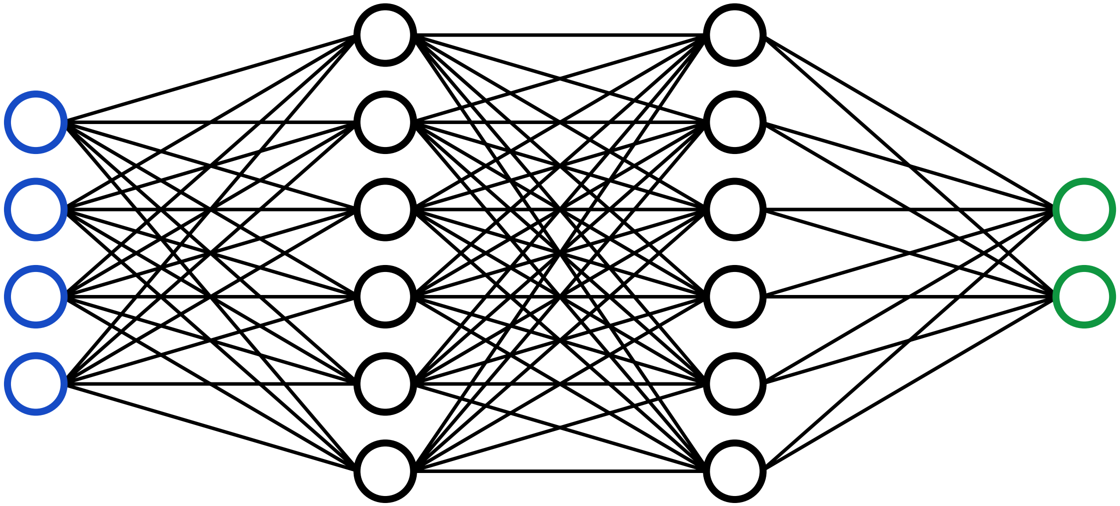
  - PolynomialDecay
    - Las dispersión va aumentando junto con el ent o

Args
initial_sparsity
final_sparsity
begin_step
end_step
power
frequency

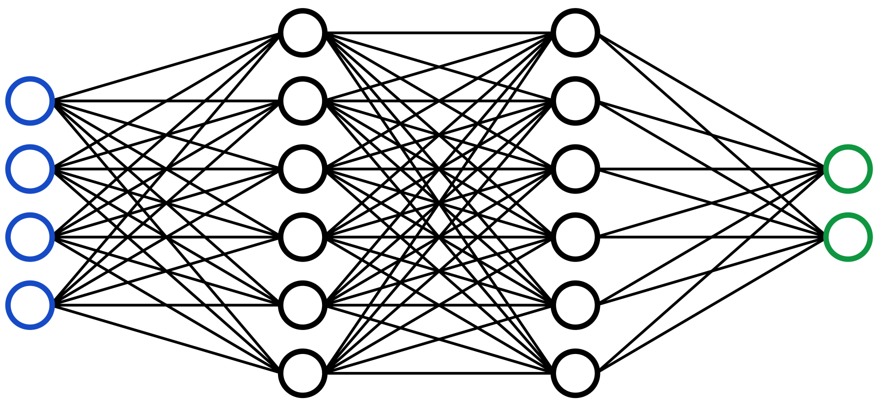
Destilación del conocimiento



Red evaluada en otros sistemas numéricos



Cuantización de una red



Demo 6

Fin del curso