

Parte: 2

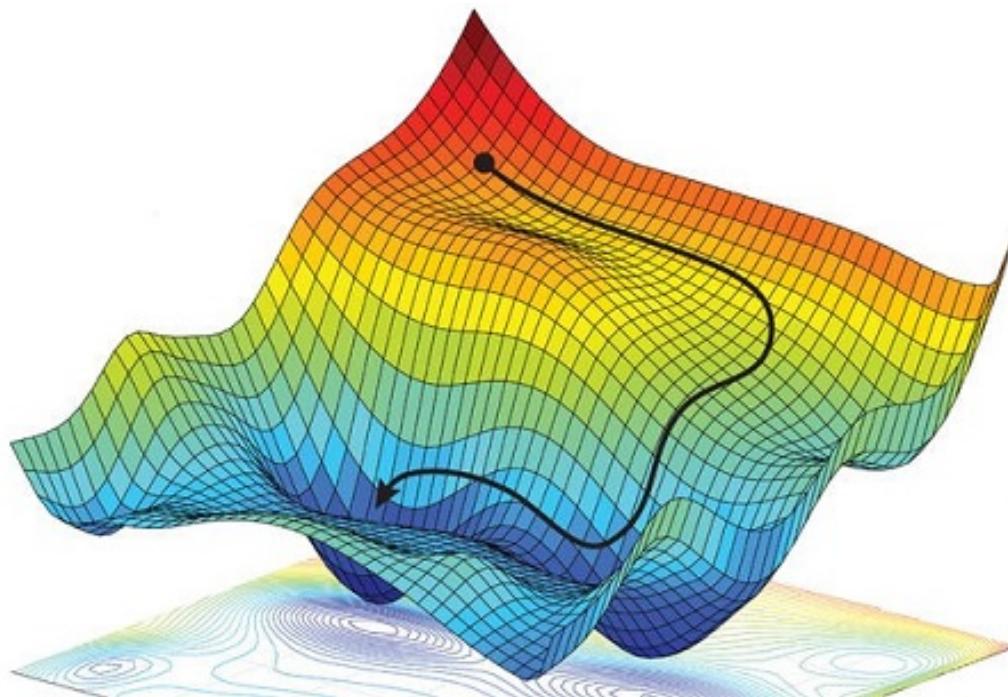
# Teoría Deep Learning

Profesores: Javier López

Daniel Cano

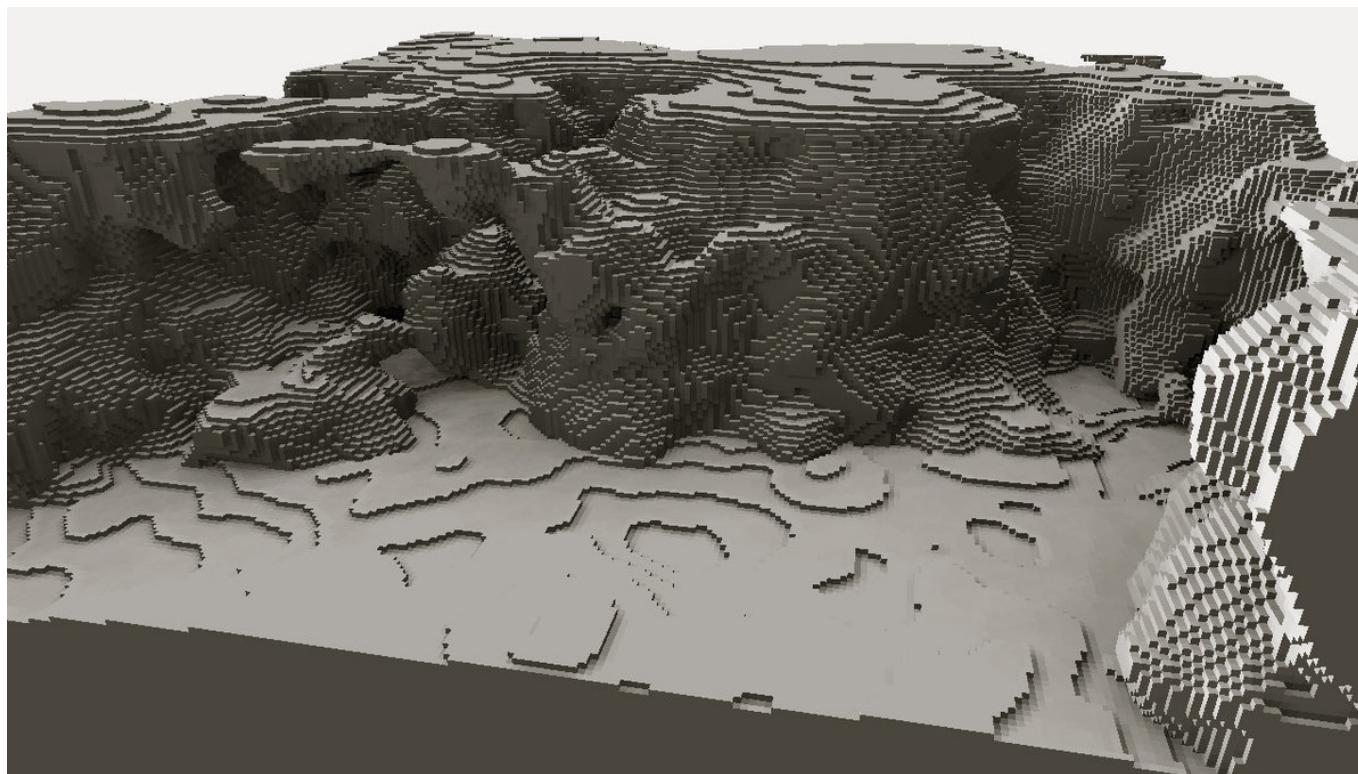
# Perceptrón

## Descenso de Gradiente



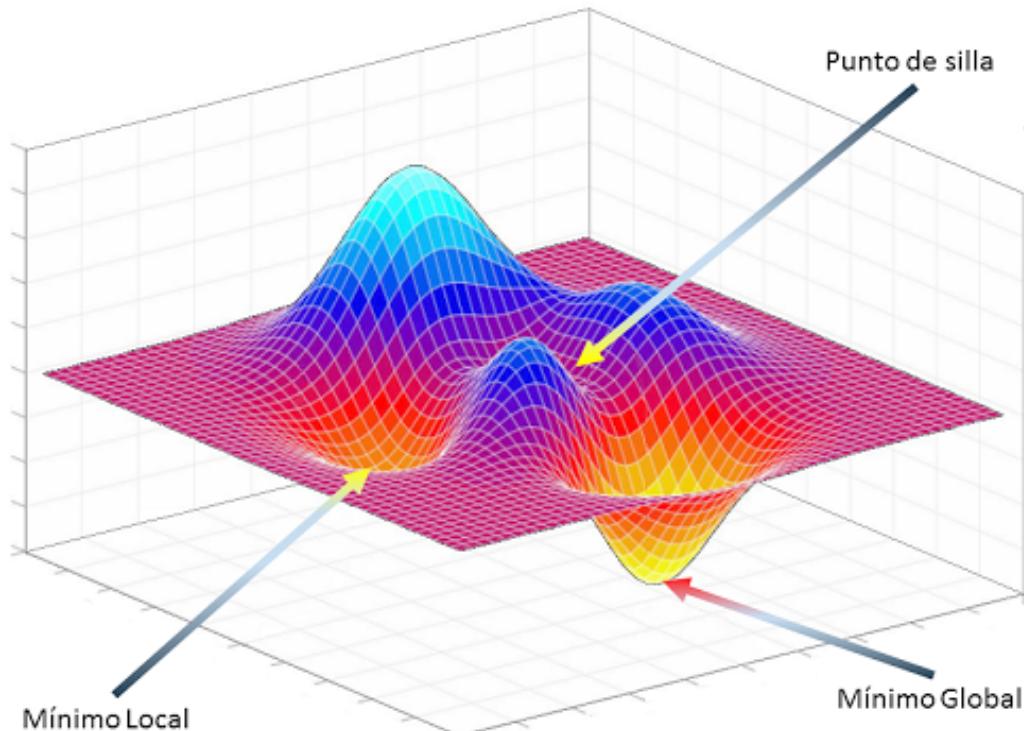
# Perceptrón

Descenso de Gradiente



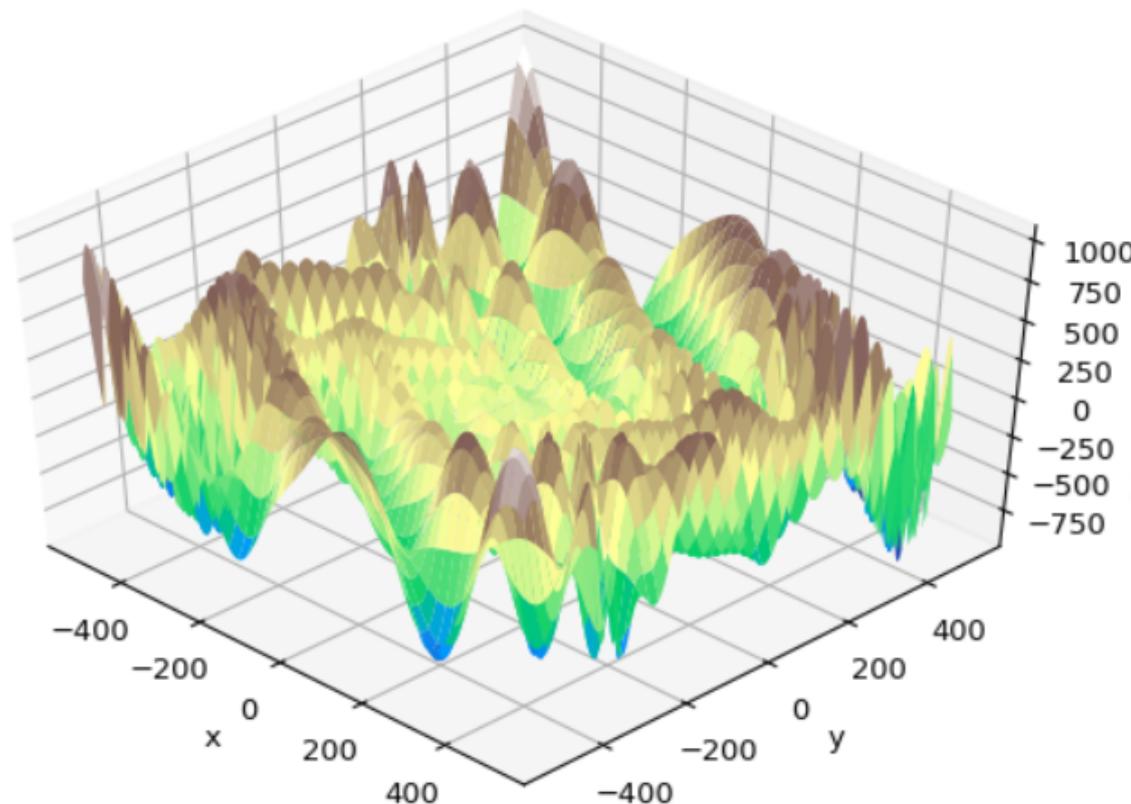
# Perceptrón

## Descenso de Gradiente



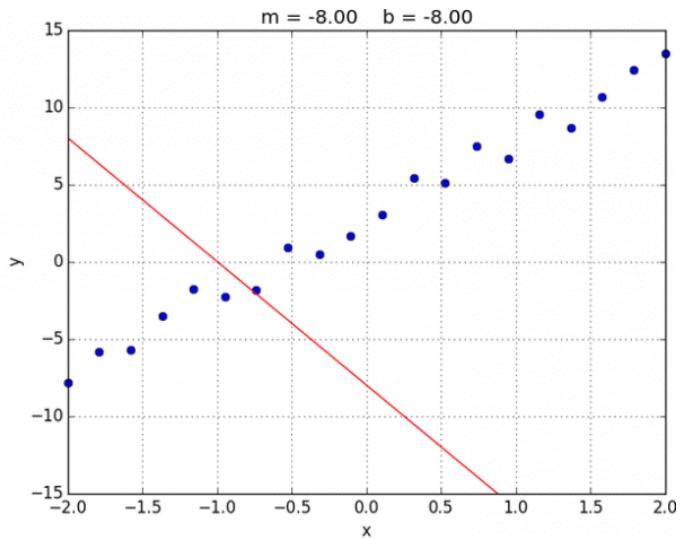
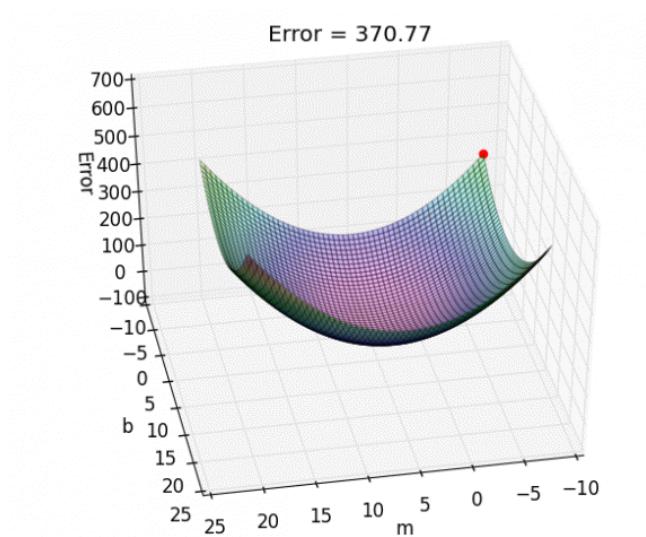
# Perceptrón

## Descenso de Gradiente



# Perceptrón

## Descenso de Gradiente



# Perceptrón

## Descenso de Gradiente

---

nobody:

algorithm: ah yes the solution

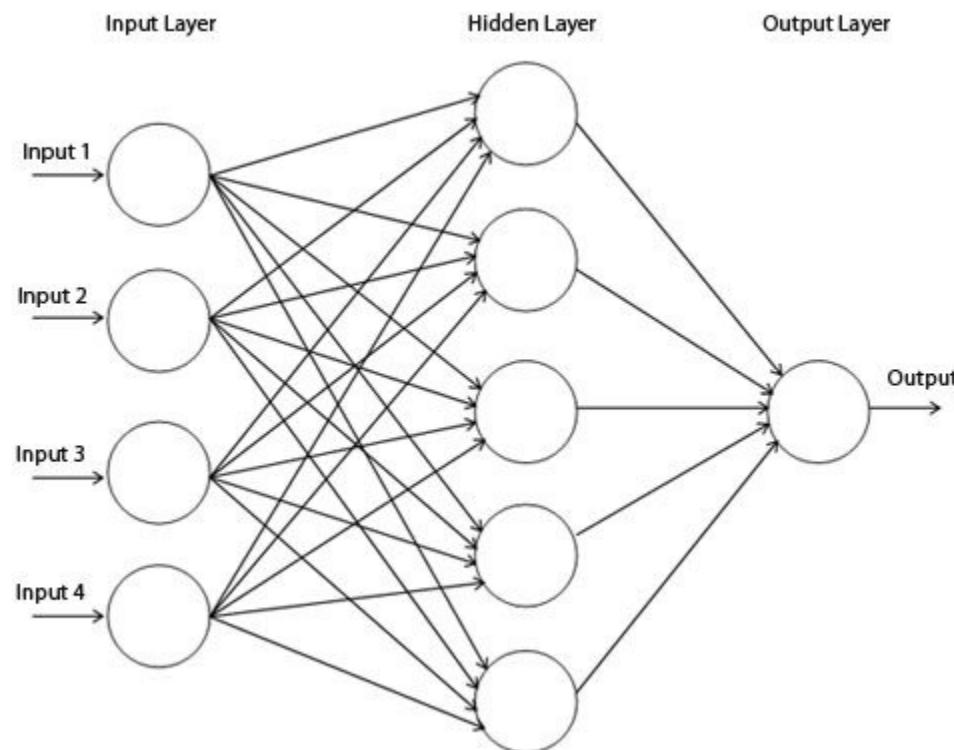


# Perceptrón

## Descenso del gradiente

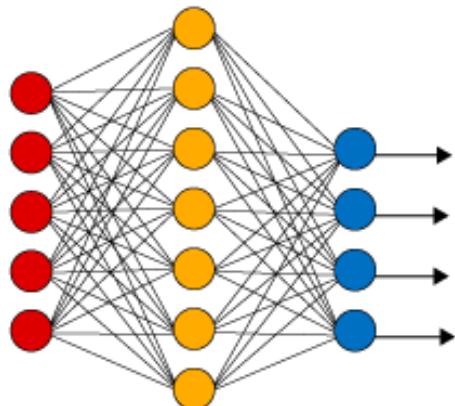
[https://www.youtube.com/embed/A6FiCDoz8\\_4?enablejsapi=1](https://www.youtube.com/embed/A6FiCDoz8_4?enablejsapi=1)

# 1 Hidden Layer

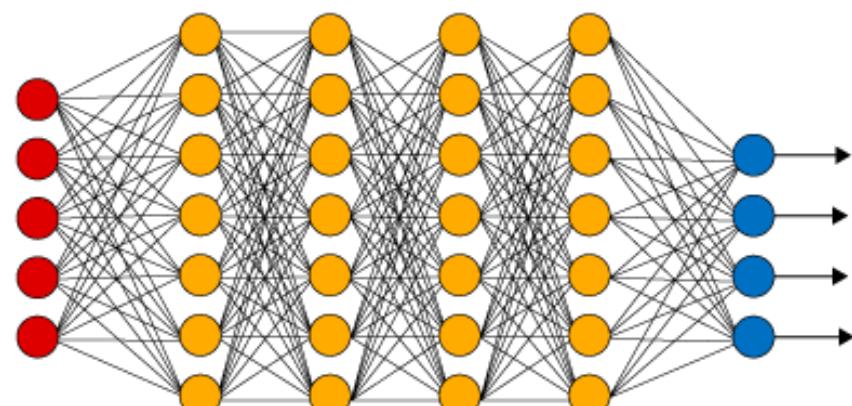


# Multilayer Perceptrón

**Simple Neural Network**



**Deep Learning Neural Network**



● Input Layer

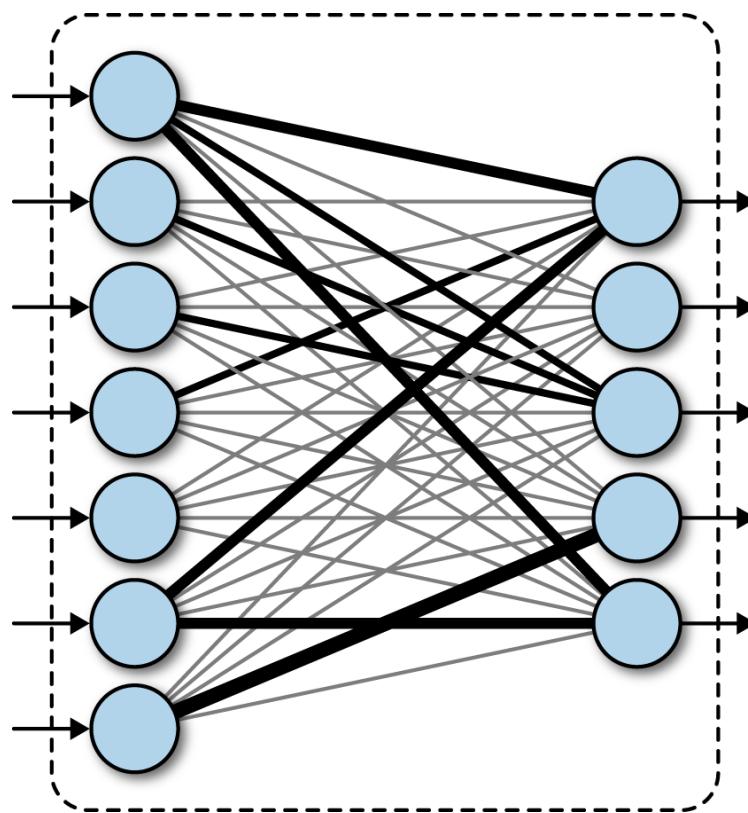
○ Hidden Layer

● Output Layer

<https://playground.tensorflow.org/>

# Densa - Full connect

<https://www.cs.ryerson.ca/~aharley/vis/>



A mostly complete chart of

# Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

Backfed Input Cell

Input Cell

Noisy Input Cell

Hidden Cell

Probabilistic Hidden Cell

Spiking Hidden Cell

Output Cell

Match Input Output Cell

Recurrent Cell

Memory Cell

Different Memory Cell

Kernel

Convolution or Pool

Perceptron (P)

Feed Forward (FF)

Radial Basis Network (RBF)

Deep Feed Forward (DFF)

Recurrent Neural Network (RNN)

Long / Short Term Memory (LSTM)

Gated Recurrent Unit (GRU)

Auto Encoder (AE)

Variational AE (VAE)

Denoising AE (DAE)

Sparse AE (SAE)

Markov Chain (MC)

Hopfield Network (HN)

Boltzmann Machine (BM)

Restricted BM (RBM)

Deep Belief Network (DBN)

Deep Convolutional Network (DCN)

Deconvolutional Network (DN)

Deep Convolutional Inverse Graphics Network (DCIGN)

Generative Adversarial Network (GAN)

Liquid State Machine (LSM)

Extreme Learning Machine (ELM)

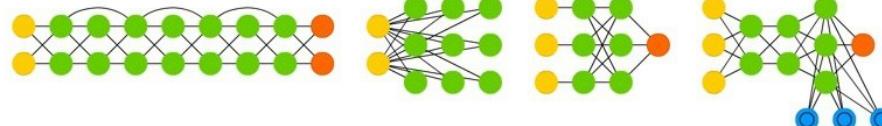
Echo State Network (ESN)

Deep Residual Network (DRN)

Kohonen Network (KN)

Support Vector Machine (SVM)

Neural Turing Machine (NTM)



# Entrenamiento

## Backpropagation

[https://www.youtube.com/embed/eNlqz\\_noix8?  
enablejsapi=1](https://www.youtube.com/embed/eNlqz_noix8?enablejsapi=1)

# Entrenamiento

## Descenso de Gradiente

- SGD Stochastic gradient descent
- BGD Batch Gradient Descent
- Mini-Batch Gradient Descent
- RMSProp
- Adam

# Entrenamiento

## Descenso de Gradiente

Gradient Descent:

$$\theta = \theta - \alpha \cdot \nabla_{\theta} J(\theta)$$

Stochastic Gradient Descent

$$\theta = \theta - \alpha \cdot \nabla_{\theta} J(\theta; sample)$$

Mini-Batch Gradient Descent

$$\theta = \theta - \alpha \cdot \nabla_{\theta} J(\theta; N samples)$$

SGD + Momentum

$$v = \gamma \cdot v + \eta \cdot \nabla_{\theta} J(\theta)$$

$$\theta = \theta - \alpha \cdot v$$

SGD + Momentum + Acceleration

$$v = \gamma \cdot v + \eta \cdot \nabla_{\theta} J(\theta - \gamma \cdot v)$$

$$\theta = \theta - \alpha \cdot v$$

Adagrad

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{G_{t,ii}} + \epsilon} \nabla_{\theta_{t,i}} J(\theta_{t,i})$$

Adadelta

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{E[G_{t,ii}]} + \epsilon} \nabla_{\theta_{t,i}} J(\theta_{t,i})$$

Adam

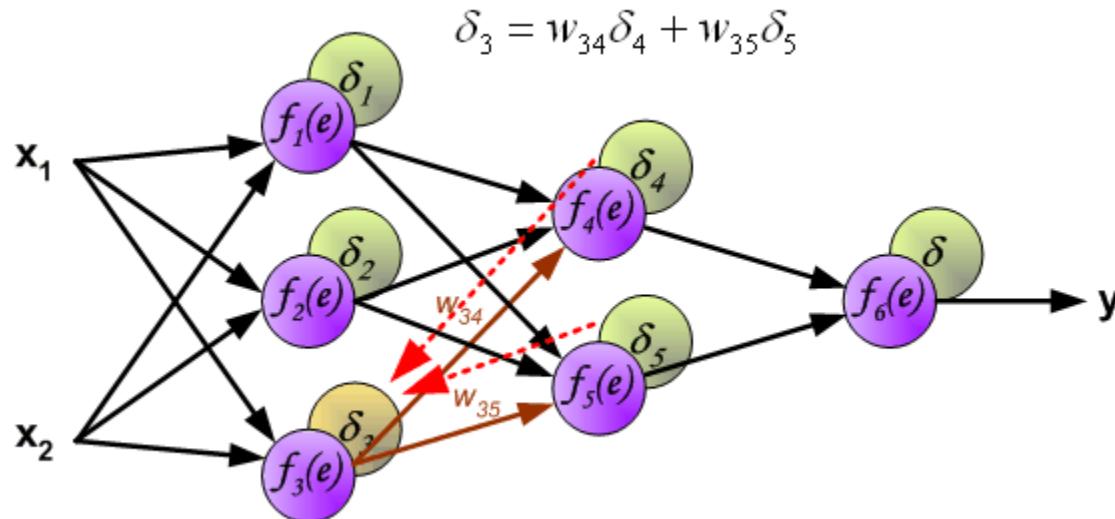
$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{E[G_{t,ii}]} + \epsilon} \times E[g_{t,i}]$$



# Entrenamiento

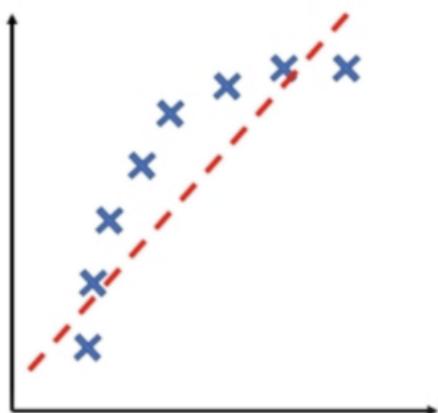
## Backpropagation

·[https://web.archive.org/web/20190603114227/http://galaxy.agh.edu.pl/~vlsi/AI/backp\\_t\\_en/backprop.html](https://web.archive.org/web/20190603114227/http://galaxy.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html)

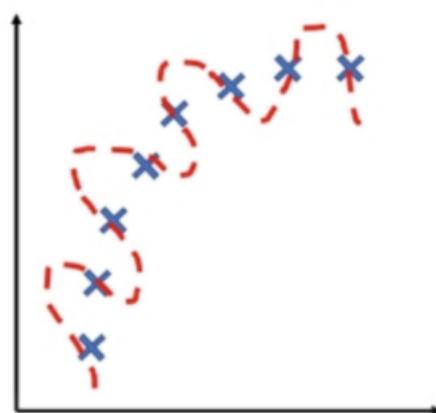


# Overfitting

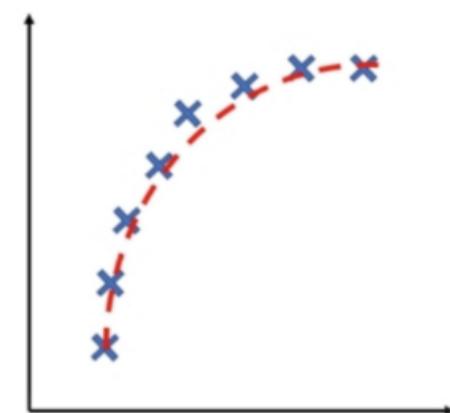
Underfitting



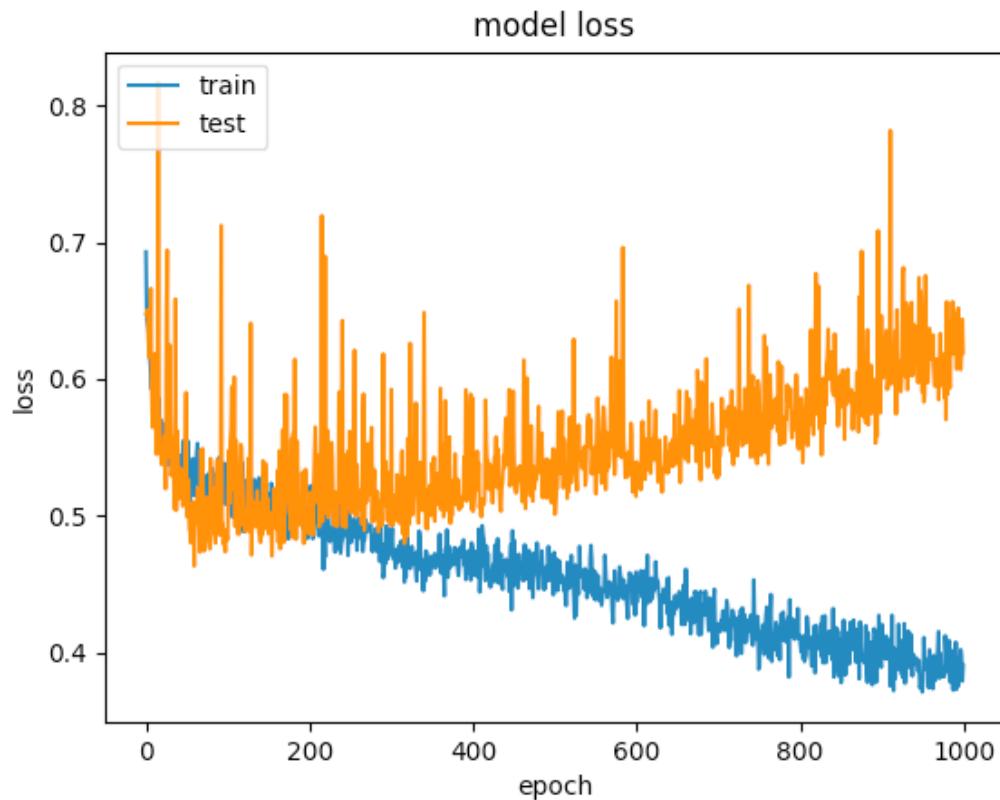
Overfitting



Ideal Balance



# Overfitting

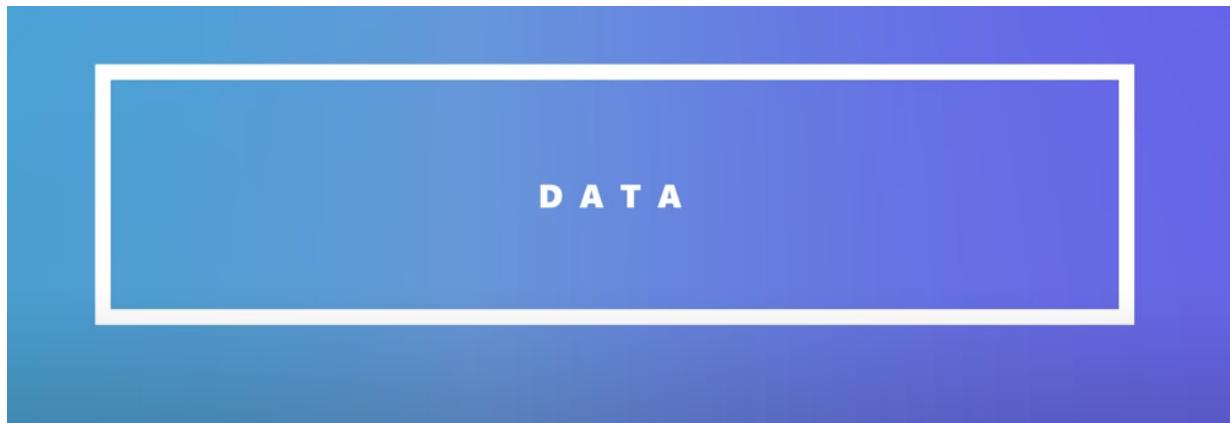


# Overfitting

[https://www.youtube.com/embed/7-6X3DTt3R8?  
enablejsapi=1](https://www.youtube.com/embed/7-6X3DTt3R8?enablejsapi=1)

# Overfitting

División de datos



# Overfitting

División de datos



# Overfitting

## División de datos



A clipboard with a brown cover holds a white sheet of paper. The paper has a header 'LISTADO TRANSPLANTES 2018' and a table with columns: P. Id, Date, Urg., C/A/Z, and Éxito. The data consists of 15 rows of transplant records.

P. Id	Date	Urg.	C/A/Z	Éxito
#FW32	23.01.42	0.98	1/1/2	- 0.78
#RK48	13.11.54	0.95	0/1/0	- 0.64
#HH12	02.10.45	0.94	1/1/1	- 0.68
#JJ56	05.12.38	0.92	1/5/2	- 0.71
#MT27	13.06.52	0.87	0/5/1	- 0.79
#DW85	12.11.43	0.86	0/5/2	- 0.78
#DR06	30.08.54	0.85	0/1/1	- 0.81
#FF43	25.07.62	0.78	0/1/1	- 0.80
#AE55	23.12.53	0.75	1/5/2	- 0.82
#II01	29.01.51	0.71	1/3/2	- 0.85
#BT92	25.11.40	0.69	1/2/1	- 0.72
#RM12	13.10.48	0.68	1/2/1	- 0.79
#HW63	13.04.48	0.63	2/3/2	- 0.86
#WW59	05.12.57	0.63	2/6/2	- 0.88
#CF02	06.05.52	0.61	3/1/0	- 0.91
#PB91	18.11.61	0.59	2/6/0	- 0.93

# Tipos de redes

¿Que tipos de redes existen?

- Densas
  - Convolucionales
  - Recurrentes
  - Reinforcement DeepQ Learning
- \* Muchas otras

