Fourth Mexican AstroComoStatistic School
June 28 - July 2, 2021

Machine Learnig Basics I

Bases del ML y redes neuronales

IV Mexican AstroCosmoStatistics School

Isidro Gómez Vargas
José Alberto Vázquez

Instituto de Ciencias Físicas
Universidad Nacional Autónoma de México

Contenido

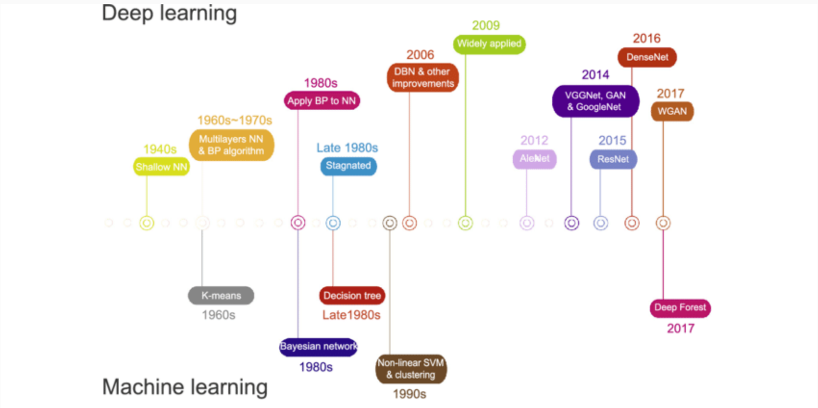
Machine Learning Basics I

CDT	Monday June 28	Tuesday June 29	Wednesday June 30	Thursday July 1	Friday July 2
9:00 - 10:0	Welcome	Plenary Activity	Plenary (Elisa Chisari)	Plenary (Agnès Ferté)	Plenary (Alma González)
10:00 - 11:30	Introduction to Bayesian inference and ML in cosmology (David Kirkby)	Introduction to Bayesian inference and ML in cosmology (David Kirkby)	Introduction to Bayesian inference and ML in cosmology (David Kirkby)	Strong Lensing (Simon Birrer) Machine Learning Basics (A. Vazquez)	Neural Networks (Miguel Aragon) Photo-z (S. Fromenteau)
11:30 - 12:00	Break: wonder.me	Break: wonder.me	Break: wonder.me	Break: wonder.me	Break: wonder.me
12:00 - 13:30	Photo-z (S. Fromenteau) Machine Learning Basics (A- Vazquez)	Strong Lensing (Simon Birrer) Machine Learning Basics (A- Vazquez)	Neural Networks (Miguel Aragon)(Lecture)	Photo-z (S. Fromenteau) Neural Networks (Miguel Aragon)	Neural Networks (Miguel Aragon) Strong Lensing (Simon Birrer)
13:30 -	Work on your own	Work on your own	Work on your own	Work on your own	Work on your own

https://www.fis.unam.mx/~javazquez/MACSS2021.html

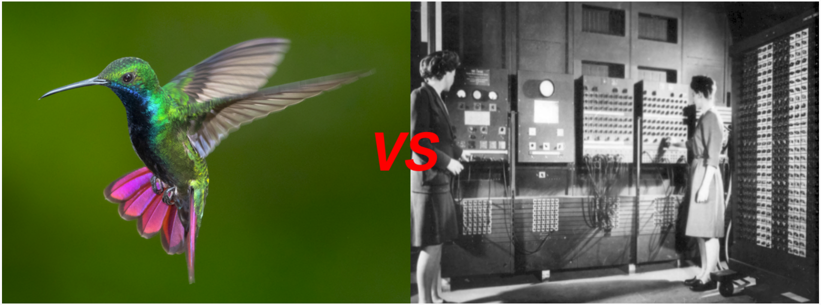
Machine Learning

Cronología ML

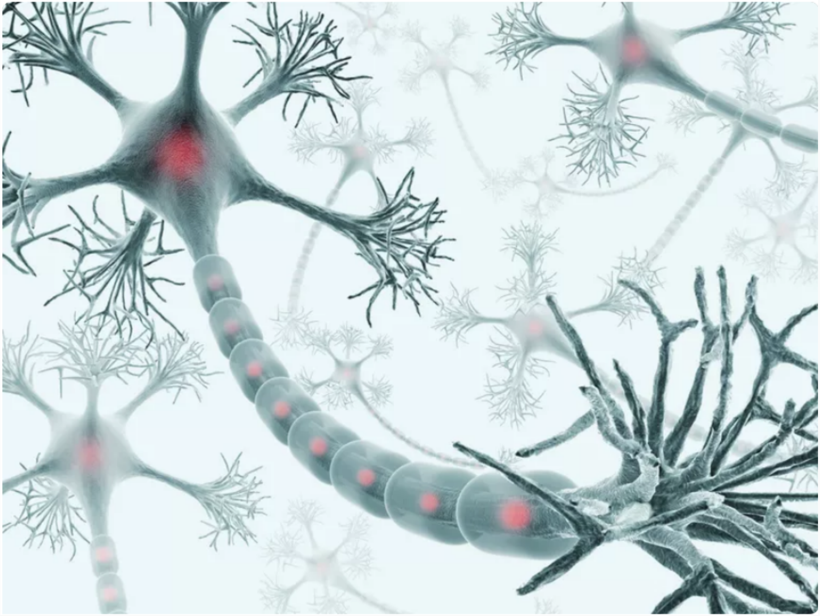


Aprendizaje máquina

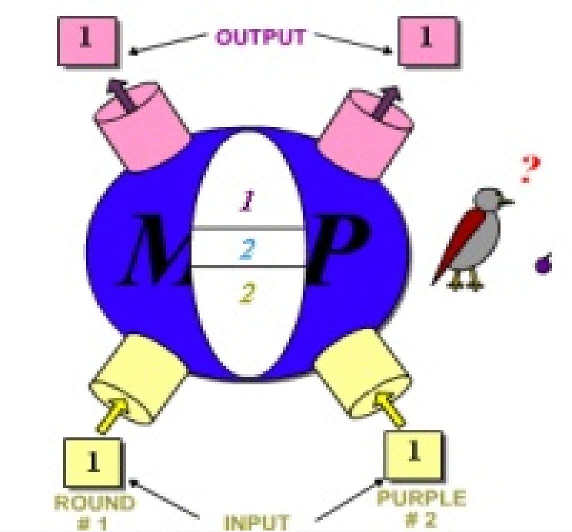
1940s



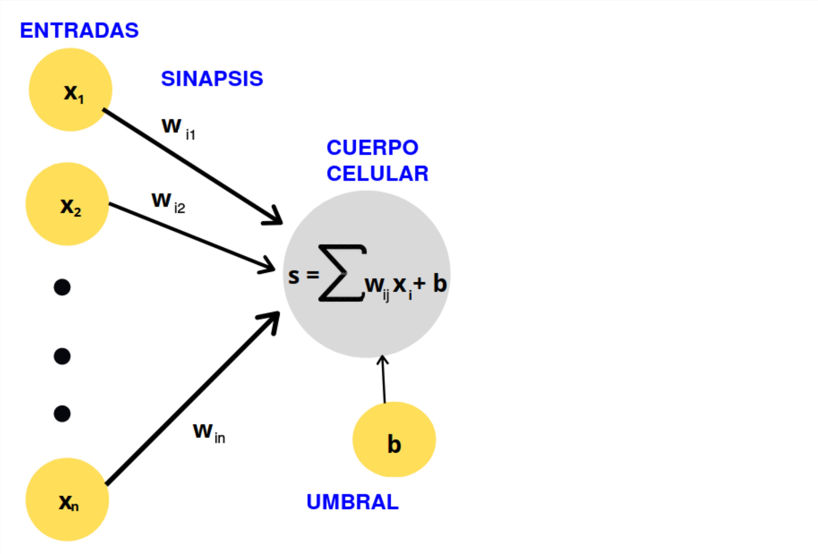
Neurona



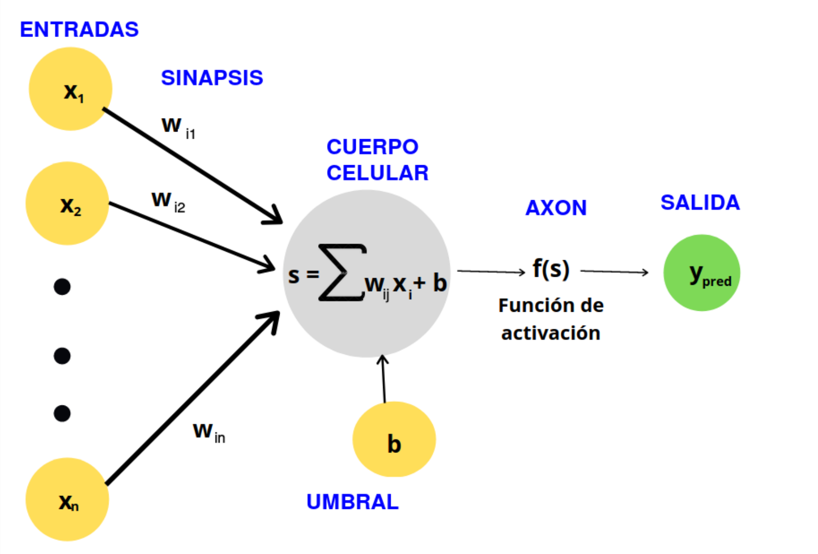
Modelo neuronal de McCullough-Pitts



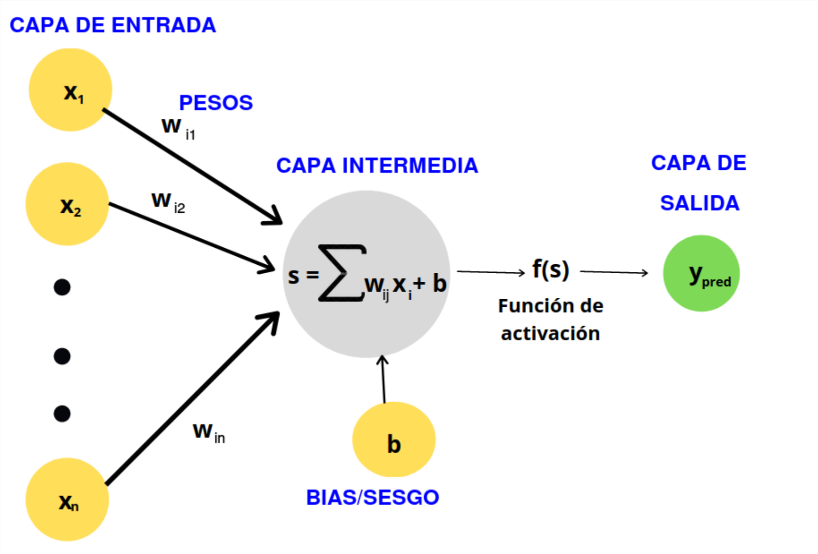
Una sola neurona artificial



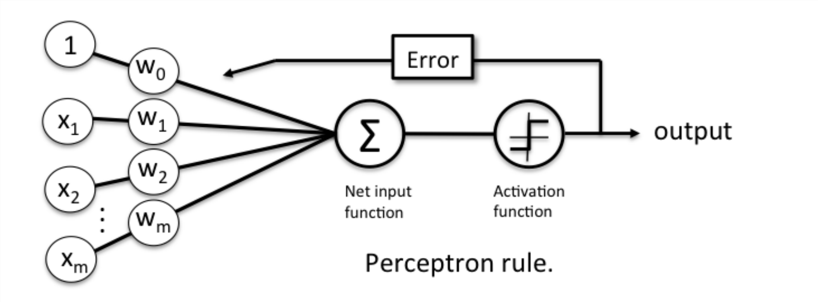
Una sola neurona artificial



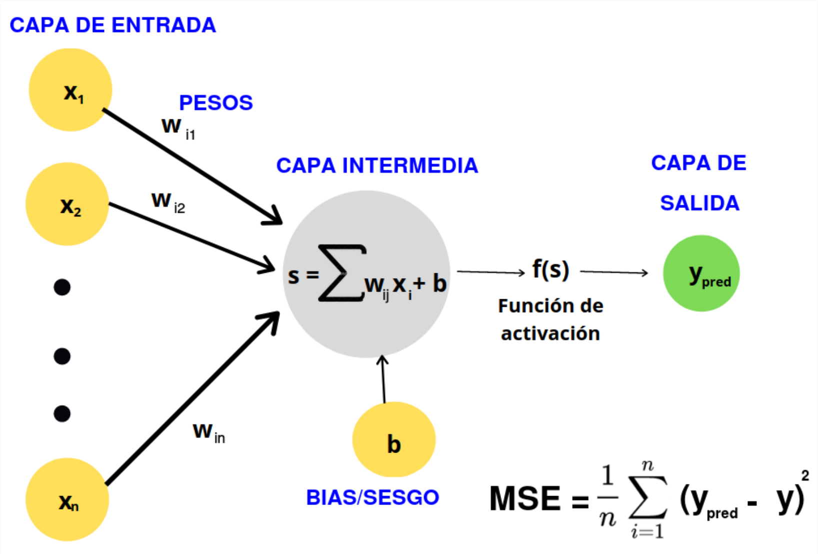
Una sola neurona artificial



Algoritmo de aprendizaje del perceptron

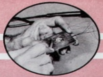


El perceptrón: una sola neurona artificial



Algoritmo de aprendizaje del perceptron

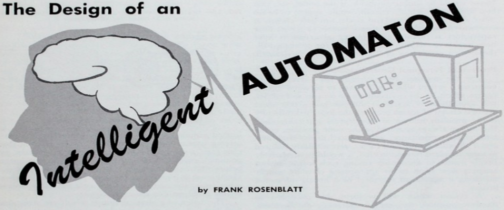
Vol. VI, No. 2, Summer 1958



research trends

CORNELL AERONAUTICAL LABORATORY, INC., BUFFALO 21, NEW YORK

The Design of an Intelligent AUTOMATON



by FRANK ROSENBLATT

Introducing the perceptron — A machine which senses, recognizes, remembers, and responds like the human mind.

STORIES about the creation of machines having human qualities have long been a fascinating province in the realm of science fiction. Yet we are now about to witness the birth of such a machine — a machine capable of perceiving, recognizing, and identifying its surroundings without any human training or control.

Development of that machine has stemmed from a search for an understanding of the physical mechanisms which underlie human experience and intelligence. The question of the nature of these processes is at least as ancient as any other question in western science and philosophy, and, indeed, ranks as one of the greatest scientific challenges of our time.

Our understanding of this problem has gone perhaps as far as had the development of physics before Newton. We have some excellent descriptions of the phenomena to be explained, a number of interesting hypotheses, and a little detailed knowledge about events in the nervous system. But we lack agreement on any integrated set of principles by which the functioning of the nervous system can be understood.

We believe now that this ancient problem is about to yield to our theoretical investigation for three reasons:

First, in recent years our knowledge of the functioning of individual cells in the central nervous system has vastly increased.

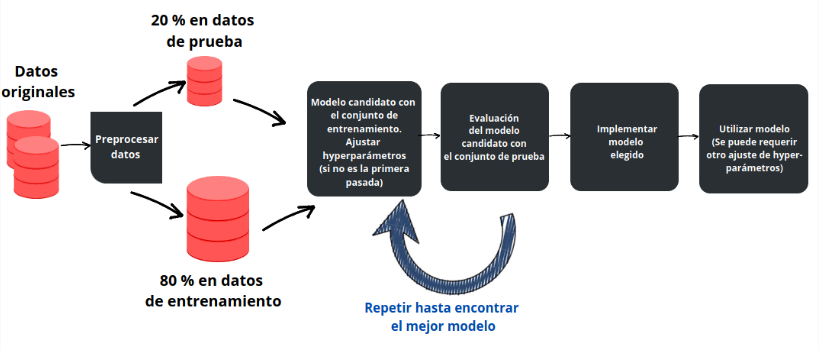
Second, large numbers of engineers and mathematicians are, for the first time, undertaking serious study of the mathematical basis for thinking, perception, and the handling of information by the central nervous system, thus providing the hope that these problems may be within our intellectual grasp.

Third, recent developments in probability theory and in the mathematics of random processes provide new tools for the study of events in the nervous system, where only the gross statistical organization is known and the precise cell-by-cell "wiring diagram" may never be obtained.

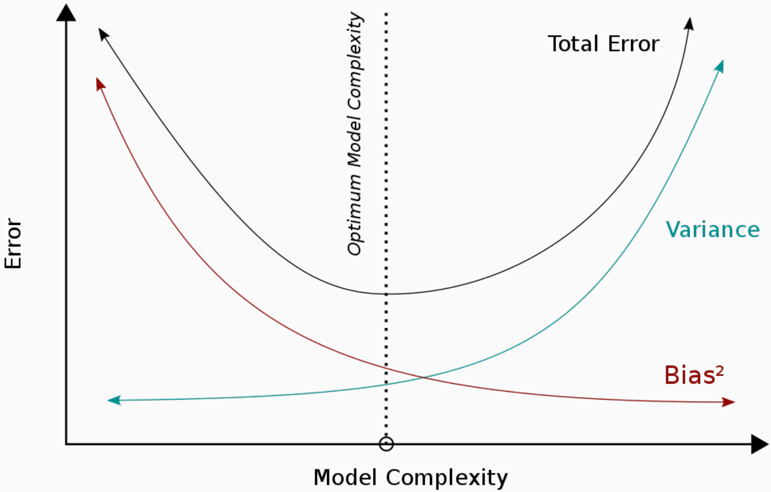
Receives Navy Support

In July, 1957, Project PARA (Perceiving and Recognizing Automaton), an internal research program which had been in progress for over a year at Cornell Aeronautical Laboratory, received the support of the Office of Naval Research. The program had been concerned primarily with the application of probability theory to

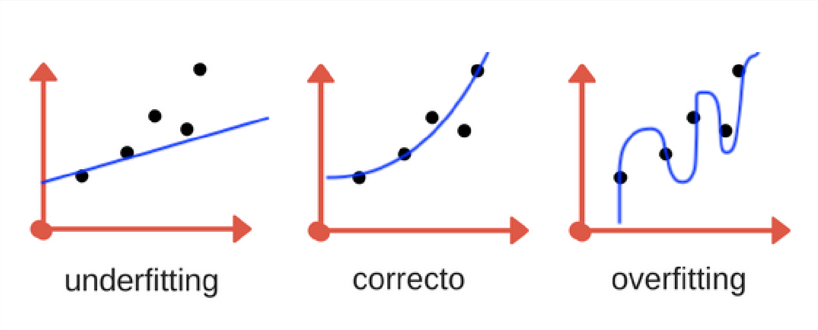
Aprendizaje máquina



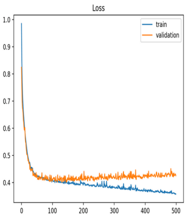
Generalización



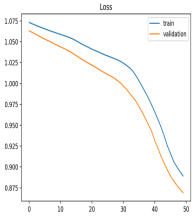
Generalización



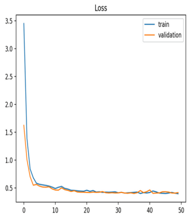
Generalización



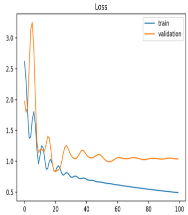
Example of Train and Validation Learning Curves Showing an Overfit Model



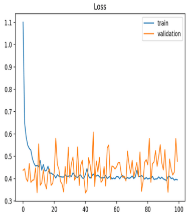
Example of Training Learning Curve Showing an Underfit Model That Requires Further Training



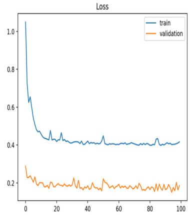
Example of Train and Validation Learning Curves Showing a Good Fit



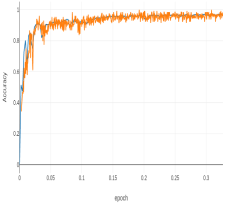
Example of Train and Validation Learning Curves Showing a Training Dataset That May Be too Small Relative to the Validation Dataset



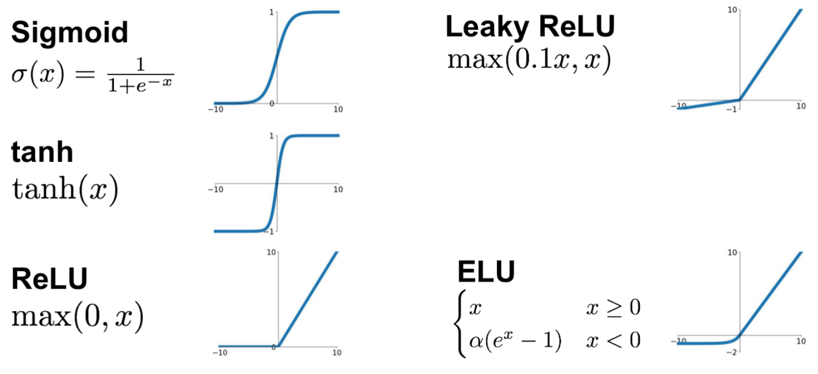
Example of Train and Validation Learning Curves Showing a Validation Dataset That May Be too Small Relative to the Training Dataset



Example of Train and Validation Learning Curves Showing a Validation Dataset That Is Easier to Predict Than the Training Dataset



Redes neuronales


Sigmoid
$\sigma(x) = \frac{1}{1+e^{-x}}$
tanh
$\tanh(x)$
ReLU
$\max(0, x)$
Leaky ReLU
$\max(0.1x, x)$
ELU
$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$

Funciones de activación

Multilayer Feedforward Networks are Universal Approximators

KURT HORNIK

Technische Universität Wien

MAXWELL STINCHCOMBE AND HALBERT WHITE

University of California, San Diego

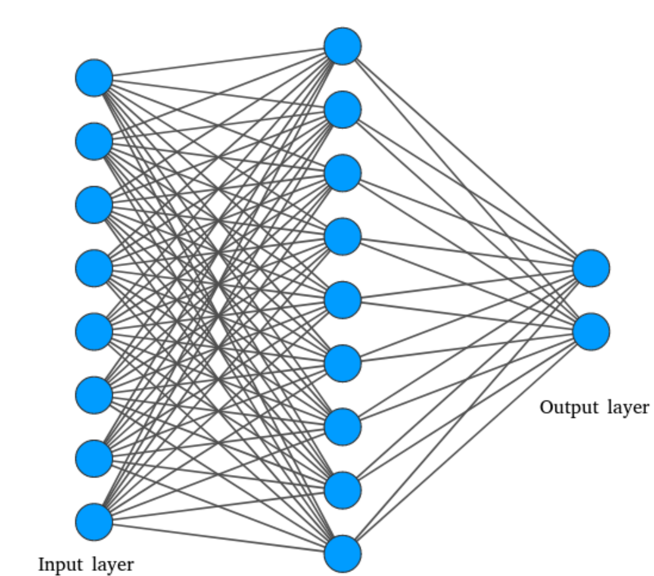
(*Received* 16 *September* 1988; *revised and accepted* 9 *March* 1989)

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.*

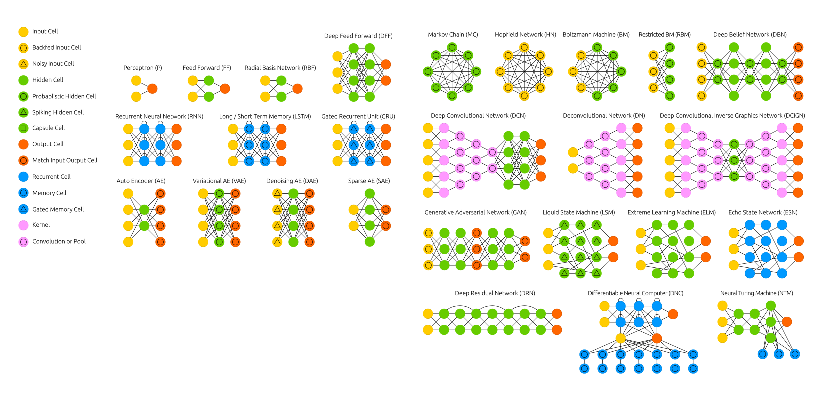
Keywords—Feedforward networks, Universal approximation, Mapping networks, Network representation capability, Stone-Weierstrass Theorem, Squashing functions, Sigma-Pi networks, Back-propagation networks.

Perceptrón multicapa

También llamado *red profunda de propagación hacia adelante.*



Múltiples arquitecturas-> Deep learning

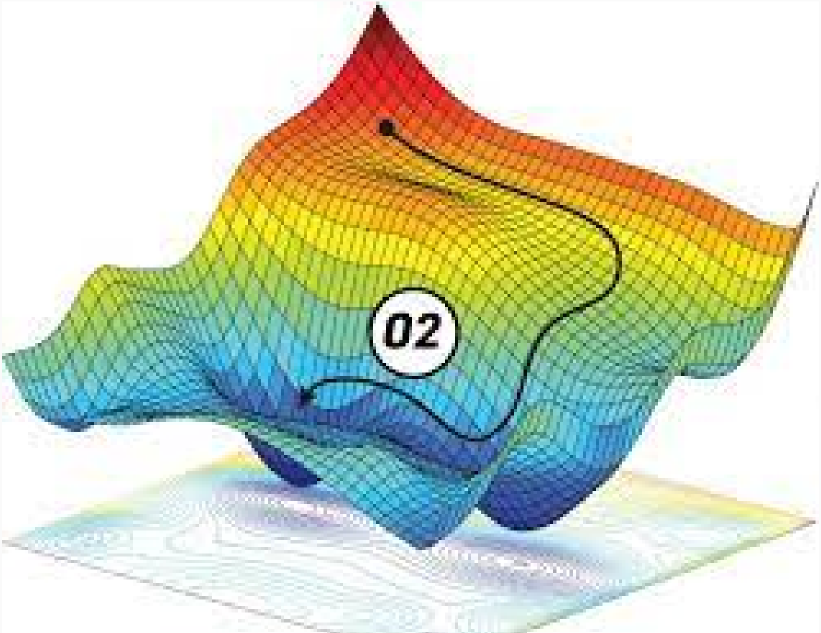


Múltiples arquitecturas-> Deep learning

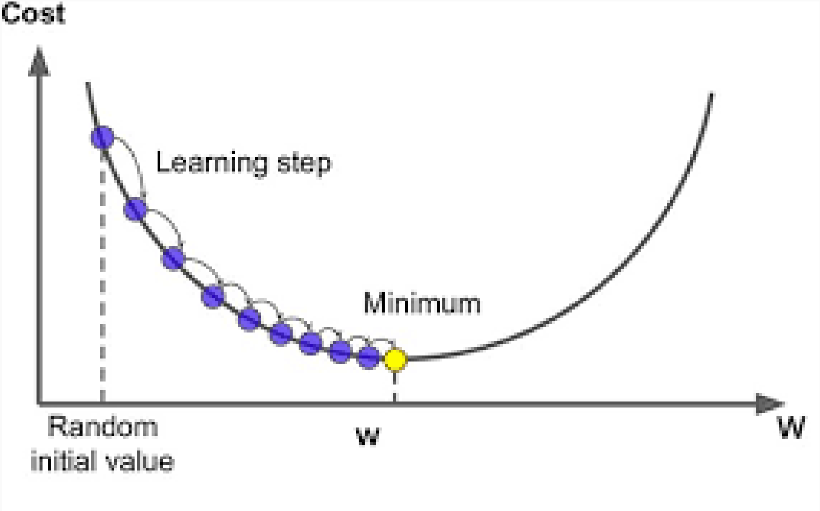
Terminología (hiper-parámetros)

- Conjunto de prueba, entrenamiento y *split*
- Épocas
- Batch size
- Número de nodos/neuronas
- Número de paredes/capas
- Learning rate (tasa de aprendizaje)

Learning rate (tasa de aprendizaje)



Learning rate (tasa de aprendizaje)



Dinámica de trabajo

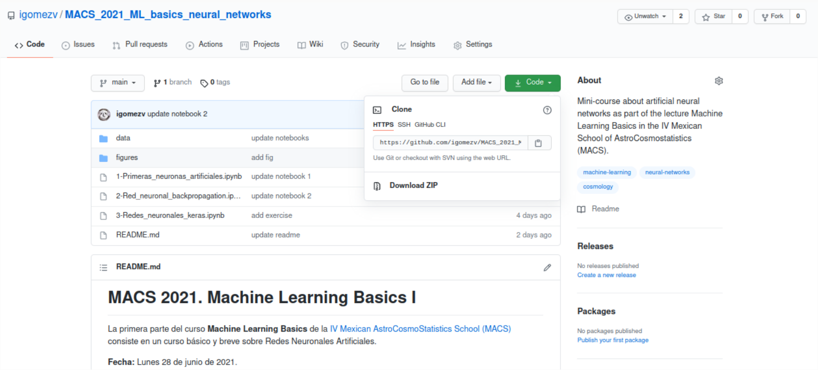
Plan del mini-curso

- 12:00-12:10. Introducción.
- 12:10-12:30. Primeras neuronas artificiales (Notebook 1).
- 12:30-12:50. Red neuronal con diferentes funciones de activacion y propagación de errores hacia atrás (backpropagation). (Notebook 2).
- 12:50-13:10. Redes neuronales con múltiples capas (Notebook 3).
- 13:10-13:40. Sesión de preguntas y/o trabajar algún ejercicio en grupo.

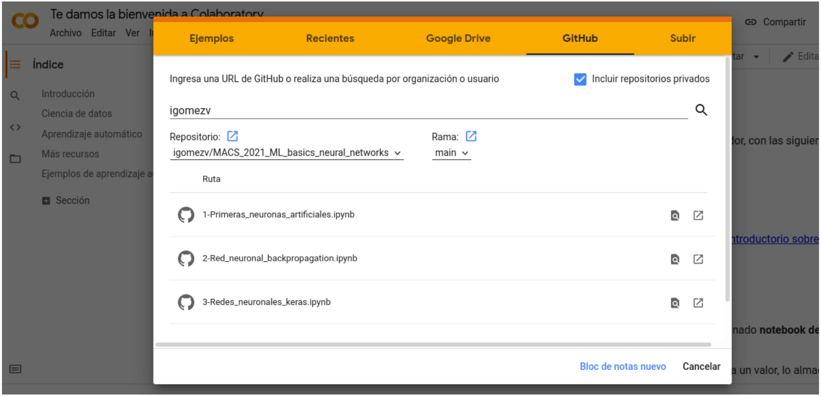
Preparando las notebooks

- Se requiere una cuenta de Google para usar Google Colab.
- Hay 5 opciones para usar el repositorio con las notebooks:
    1. Clonar repositorio (requiere tener instalado git).
    2. Descargar repositorio.
    3. **Fork del repositorio (requiere tener cuenta de GitHub)**.
    4. **Copiar repositorio desde Google Colab en Drive**.
    5. Tan solo abrir y ejecutar en Google Colab (no se guardan cambios).

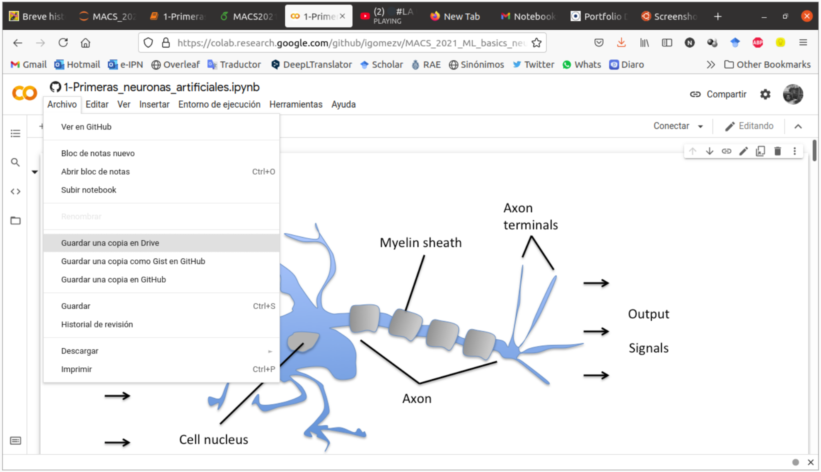
Preparando las notebooks



Google Colab



Google Colab (copiar en Drive) para poder guardar cambios



Ejercicios opcionales

Por votación se elegirá uno para trabajarlo la última media hora.

1. Compuertas lógicas (Notebook 1).
2. Clasificación lineal (Notebook 1).
3. Aproximar función con red neuronal de una capa (Notebook 2).
4. Aprender funciones con redes de múltiples capas y neuronas (Notebook 3).
5. Reducir tiempos (Notebook 3).
6. Modelar datos de Supernovas del tipo IA de la compilación JLA (Notebook 3).
7. Clasificar estrellas, galaxias y cuásares del SDSS-DR14 (Notebook 1 y notebook 3).

Acción