Lab Presentation Session

Deep Learning applied to Astrophysics

Lara Lloret Iglesias & Diego Tuccillo
(IFCA-CSIC)

MACHINE LEARNING/DEEP LEARNING

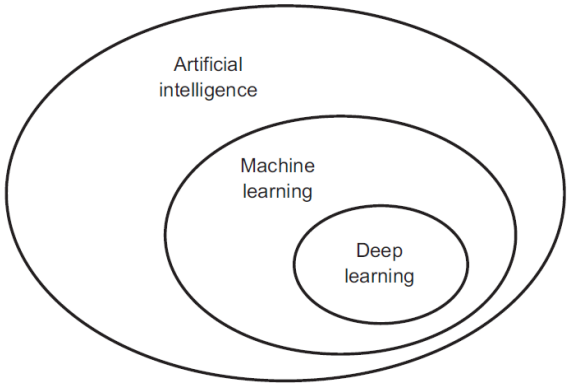
Artificial Intelligence **was already born in 1950** from a group of experts aiming to make computers “think”

This includes what was later called Machine Learning and Deep Learning, but **also system that do not actually learn**.

First chess computer programs: symbolic AI

The “learning approach” arises when trying to answer a series of questions:

- Can a computer go beyond what *we know how to order (code)*?
- Can it really learn a given task *its own way*?
- Can it even surprise us?

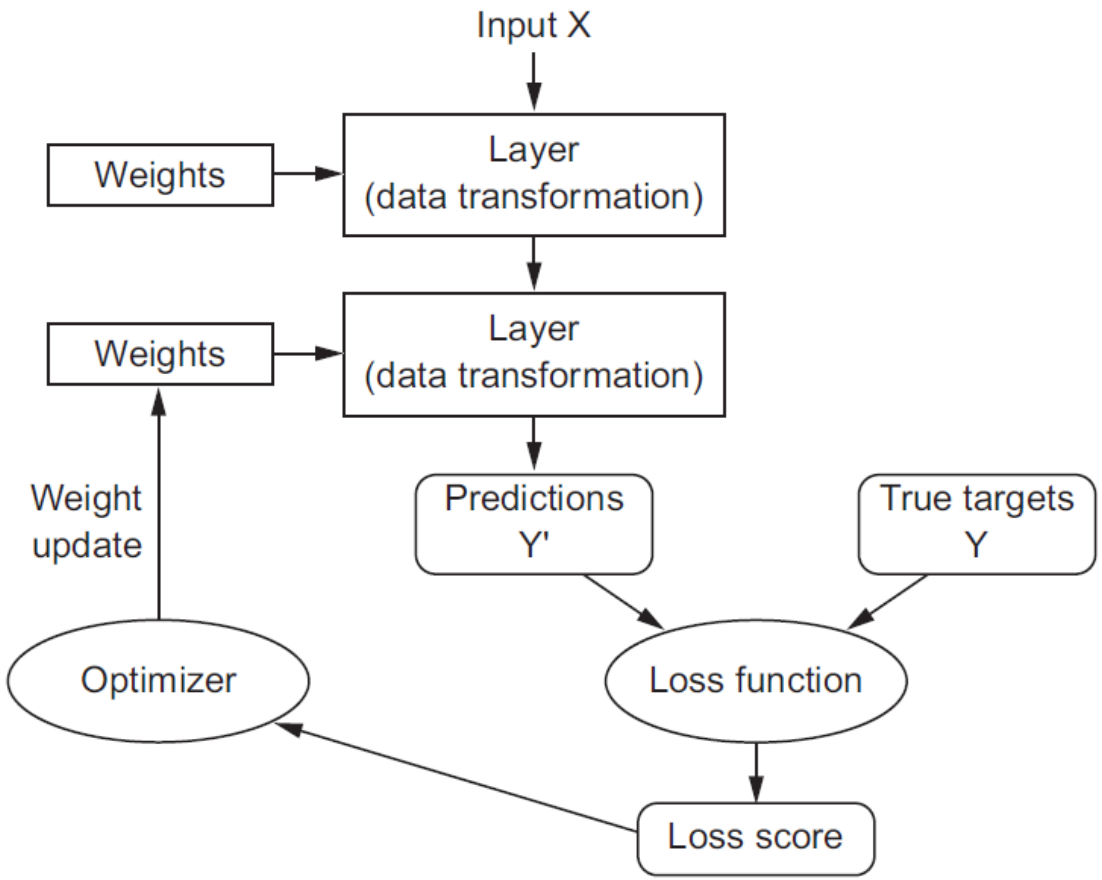


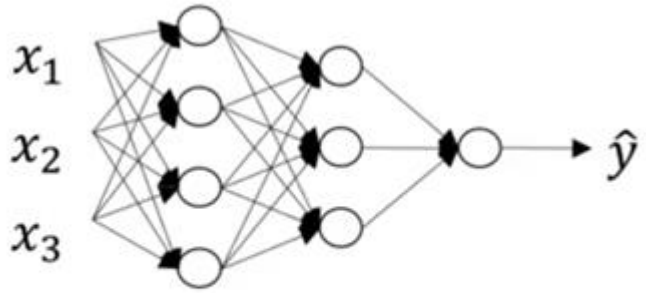
Instead of having people codifying the rules…

- Can it learn by looking at data?

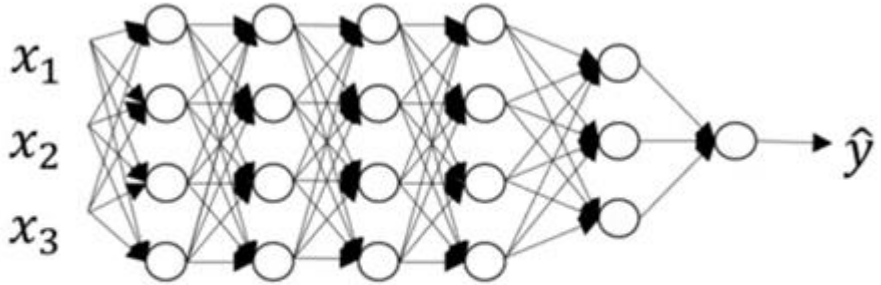
In DEEP Learning the “learning” is **almost always done using Neural Networks**

(DENSE) NEURAL NETWORKS





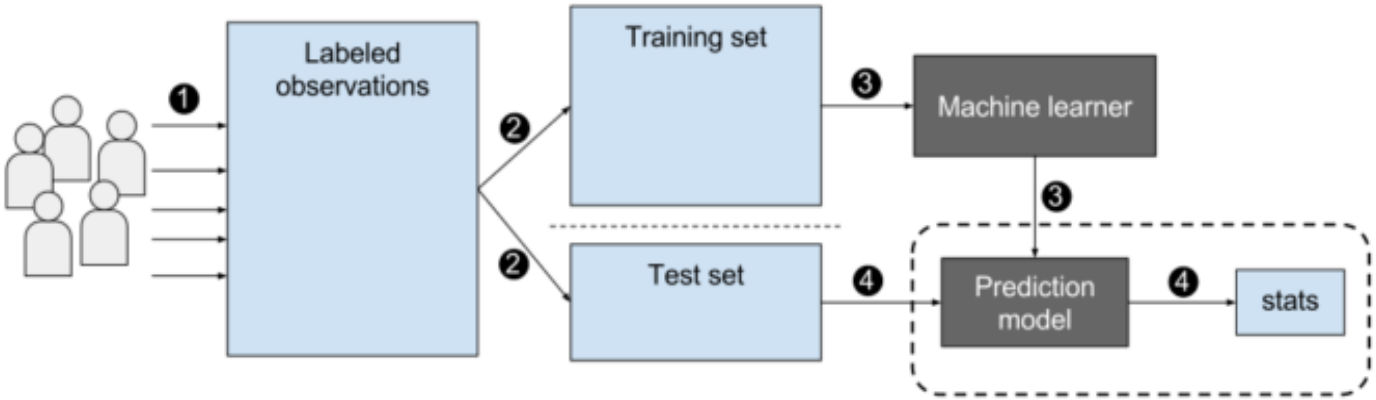
2 hidden layers



5 hidden layers

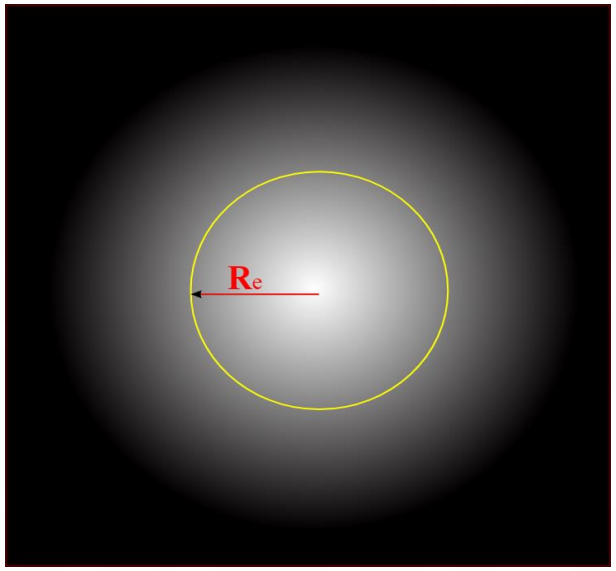
TYPES OF LEARNING: SUPERVISED LEARNING

- In supervised learning, one has a **full set of labelled data:**
  - Each example in the training dataset is **tagged with the answer the algorithm should come up with on its own.**
  - A labelled dataset of flower images would tell the model which photos were of roses or daisies.
- There are two main areas where supervised learning is useful:
  - **Classification problems** → predict an integer (class)
  - **Regression problems** → predict a real number



BRIGHTNESS PROFILE FITTING

- The half light radius (R_e) of a galaxy is the **radius at which half of the total light of the system is emitted.**
- This assumes the galaxy has either **intrinsic spherical symmetry** or is at least circularly symmetric as viewed in the plane of the sky.



- In this first part of the Lab we will use **simulated images (GALSIM) for training** a **convolutional neural network** and perform **transfer learning using real data** from Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDLES) that is part of the Hubble Space Telescope.

http://adsabs.harvard.edu/abs/2018MNRAS.475..894T

STAR-GALAXY SEPARATION

- Most existing star-galaxy classifiers require **careful feature extraction and selection.**
- The latest advances in deep learning that use convolutional neural networks **allow a machine to automatically learn the features directly from data**, minimizing the need for input from human experts.
- In this part of the lab we present a **star-galaxy classification framework** that uses deep **convolutional neural networks** to solve this problem
- Using **real data** from the Sloan Digital Sky Survey (**SDSS**)
- The SDSS is a major **multi-spectral imaging and spectroscopic redshift survey** using a dedicated 2.5-m wide-angle optical telescope at Apache Point Observatory in New Mexico, United States.

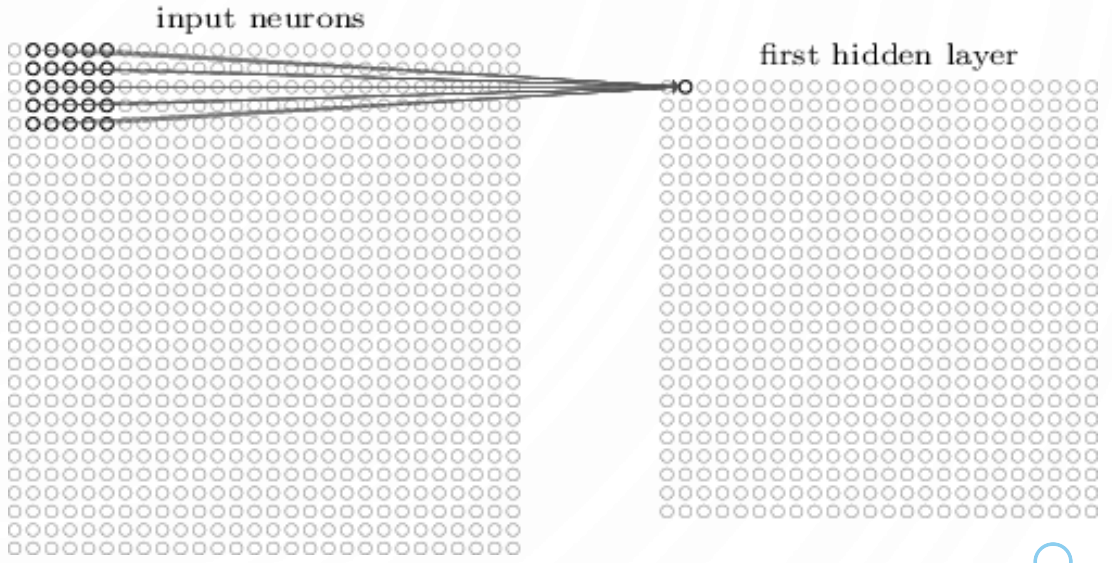
https://arxiv.org/pdf/1608.04369.pdf

CONVOLUTIONAL NEURAL NETWORKS

The main differences between a dense neural network and a convolutional neural network:

- The dense network **learn global patterns** in the input features space. In the case of images they use all the pixels in the image.

- The convolutional neural networks **learn local patterns**. In the case of the images the patterns are found using small 2D windows (filters)



State of the art for working with images!

SUMMARY

- In this datalab we will be using Convolutional Neural Networks for solving two supervised learning problems in Astrophysics using real data
- We will have an introductory session explaining the basics on Convolutional Neural networks and how to optimize them
- The task of the students will be:
  - Take a “ready to run” neural network for Half Ligh Radius determination and optimize it as much as possible with the techniques explained in the introduction
  - With the knowledge from the first part of the lab, create a model (almost) from scratch in order to classify Galaxies and Stars
- We will be using Keras
- Some knowledge of Python (matplotlib, numpy…) is suitable

Thank you!

And hope to see you in the Lab!