



Inteligência Artificial em Astronomia: minha experiência com o S-PLUS

Laerte Sodré Jr.

Departamento de Astronomia

Instituto de Astronomia, Geofísica e Ciências Atmosféricas

Universidade de São Paulo

**Astronomia para a Terceira Idade
IAG/USP 06/06/2024**

resumo

- **Levantamentos de grandes áreas em Astronomia**
- **O Southern Photometric Local Universe Survey (S-PLUS)**
- **Inteligência Artificial em Astronomia**

observação astronômica hoje (no óptico)



Gemini South & SOAR

@Cerro Pachon

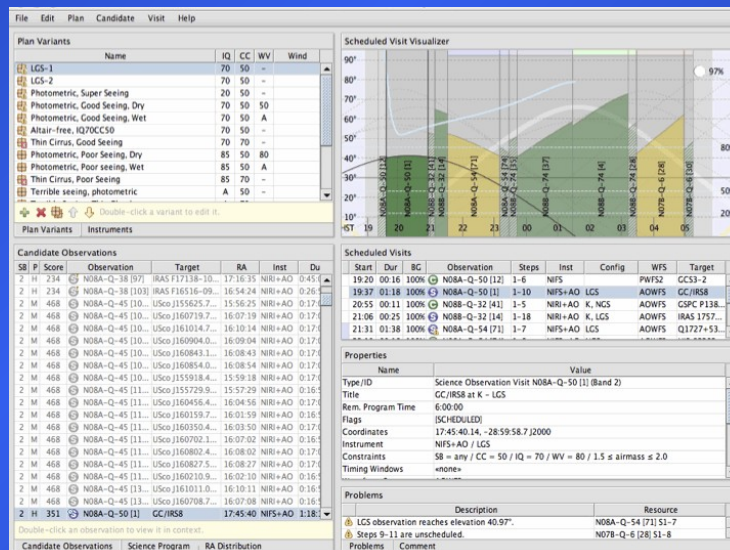
modos:

- presencial
- fila
- surveys (observação robótica)

bons tempos...



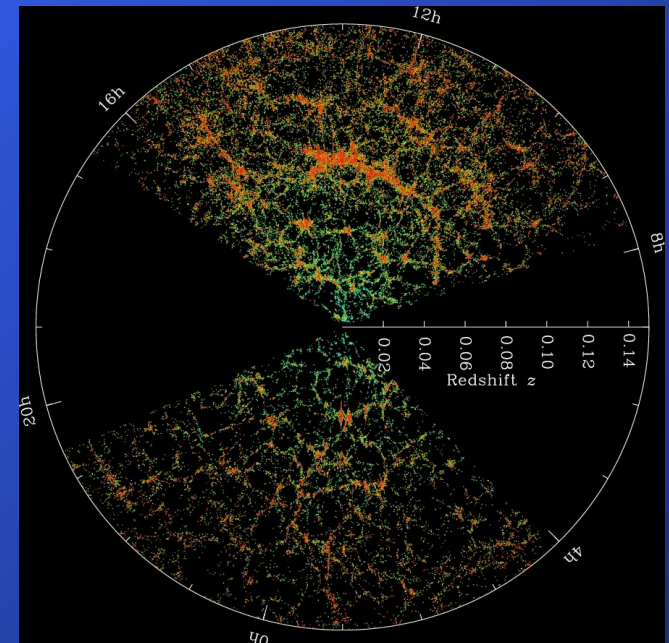
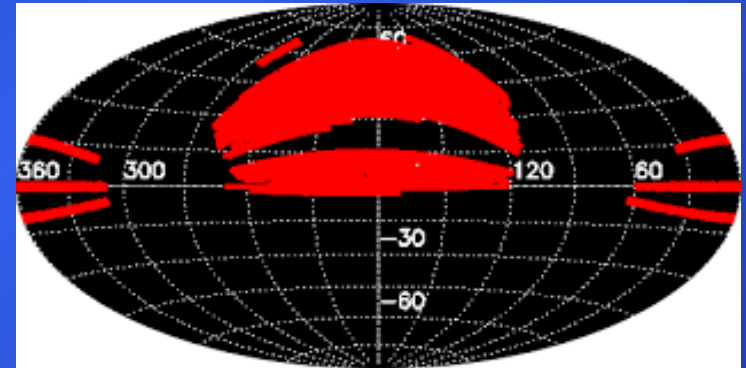
fila



observação astronômica hoje (no óptico)

- **surveys**

Sloan Digital Sky Survey
SDSS



dois tipos de observação:

- **objetos individuais**

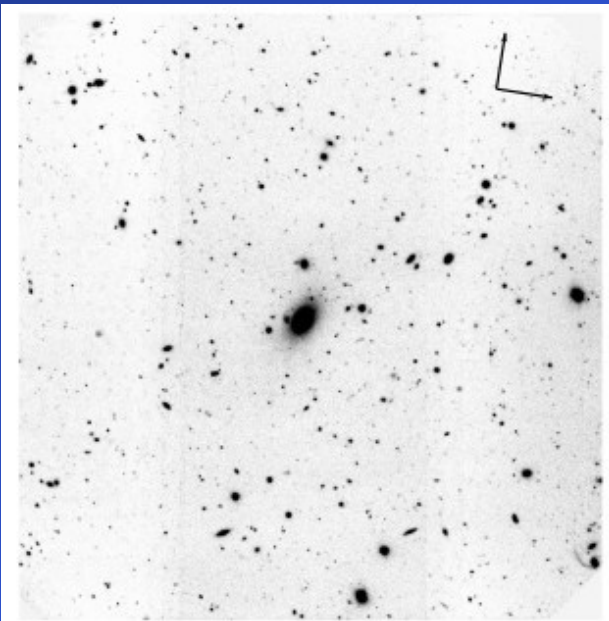


Fig. 1.— Optical i' image of RX J1340.6+4018. The field of view is 5.6 arcmin on a side, or $\sim 980 h_{70}^{-1}$ kpc at the object redshift. The field orientation is shown in the upper right corner of the figure.

Imagem obtida com o Gemini N

(Mendes de Oliveira+, 2009)

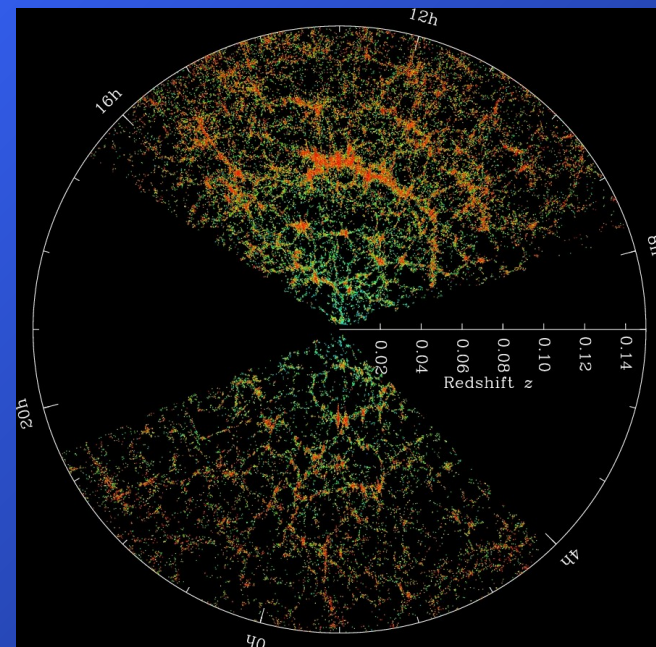
observação astronômica hoje (no óptico)

surveys

- observações de grandes áreas do céu ou
- observação profunda de uma área pequena
- estudo de populações:
distribuição espacial e propriedades físicas
- levantamentos fotométricos
- levantamentos espectroscópicos
- ambos

levantamentos fotométricos

- J-PAS, J-PLUS, S-PLUS
- LSST (grande área, profundo)

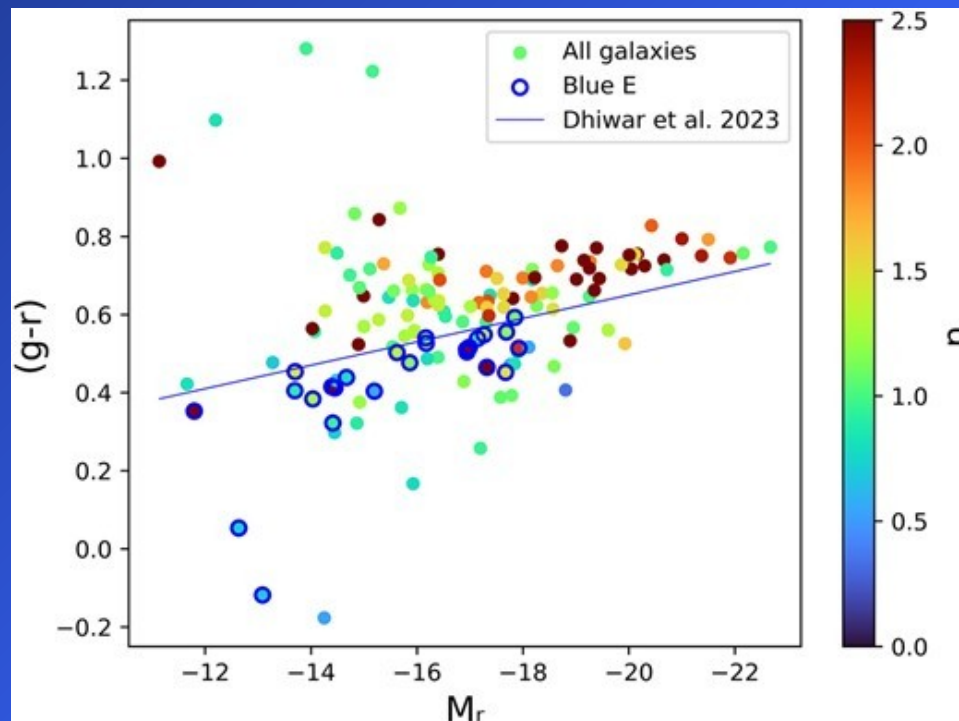


SDSS

porque surveys são importantes?

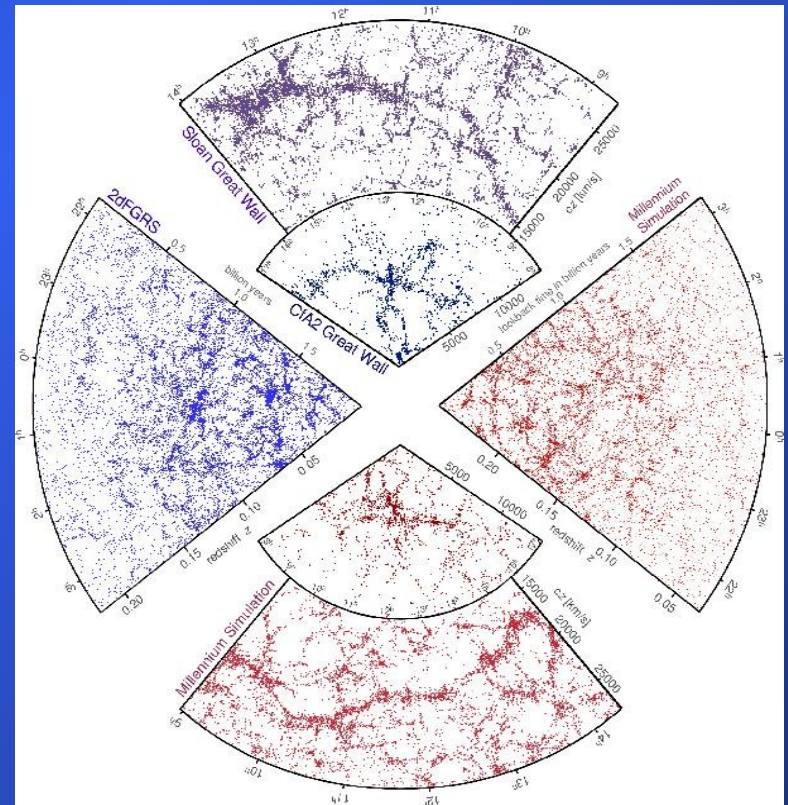
estudos de populações:
estrelas, galáxias, ...

exemplo: identificação de galáxias
elípticas azuis



Smith-Castelli et al. 2024

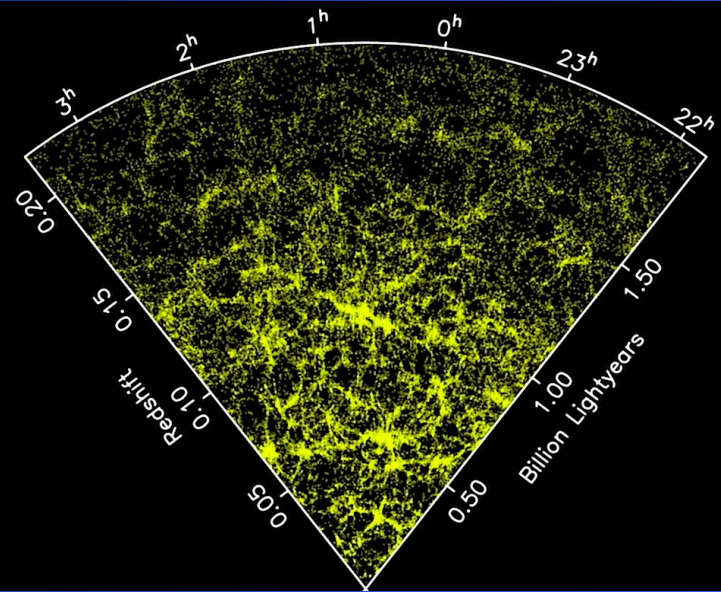
cosmologia:
investigação sobre a matéria
escura e a energia escura



Λ CDM explica muito bem a
distribuição de galáxias!

porque surveys são importantes?

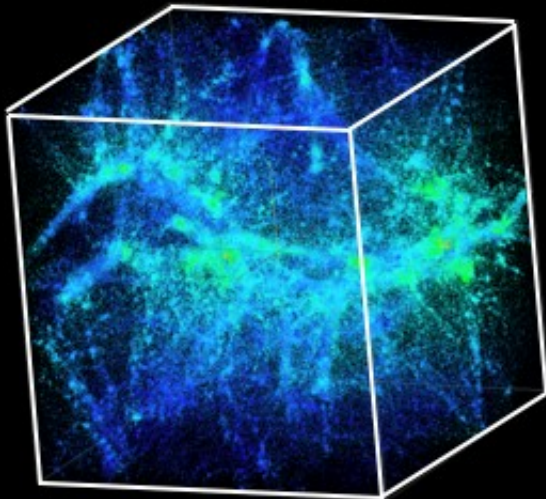
a cosmologia está impressa na distribuição de galáxias



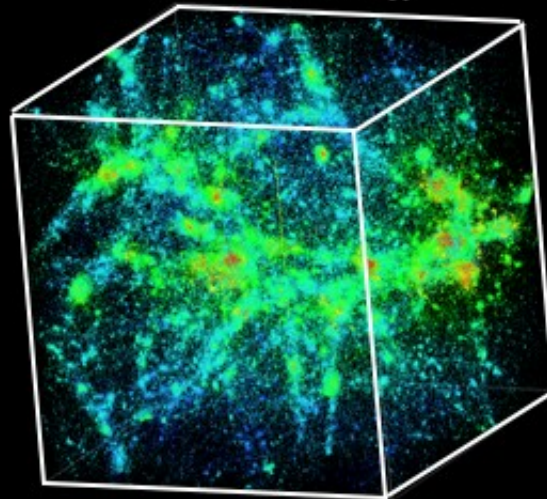
a teia cósmica:

- a distribuição 3D das galáxias informa sobre a matéria escura e a energia escura
- As propriedades das galáxias na teia cósmica informam sobre os processos que regem sua formação e evolução

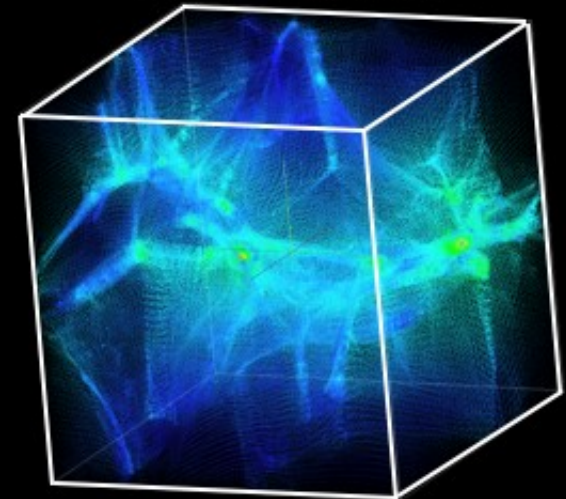
Standard Model



No dark energy



Warm dark matter



levantamentos multibandas

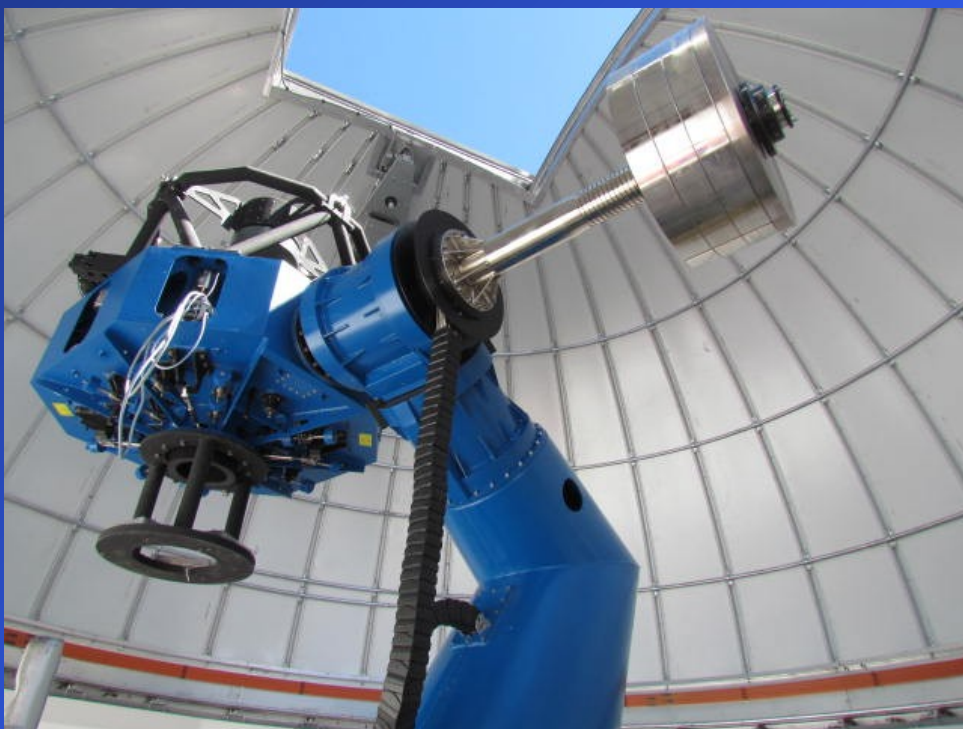
- **S-PLUS: Southern Photometric Local Universe Survey,**
9300 sq deg – 0.8m telescope @ OAJ – 12 filtros
- **J-PLUS: Javalambre Photometric Local Universe Survey**
8000 sq deg – 0.8m telescope @OAJ – 12 filtros
- **J-PAS: Javalambre Physics of the Accelerating Universe
Astrophysical Survey**
8000 sq deg – 2.5m tel @ OAJ – 59 filtros



Observatório Astrofísico de Javalambre, Espanha



Southern Photometric Local Universe Survey



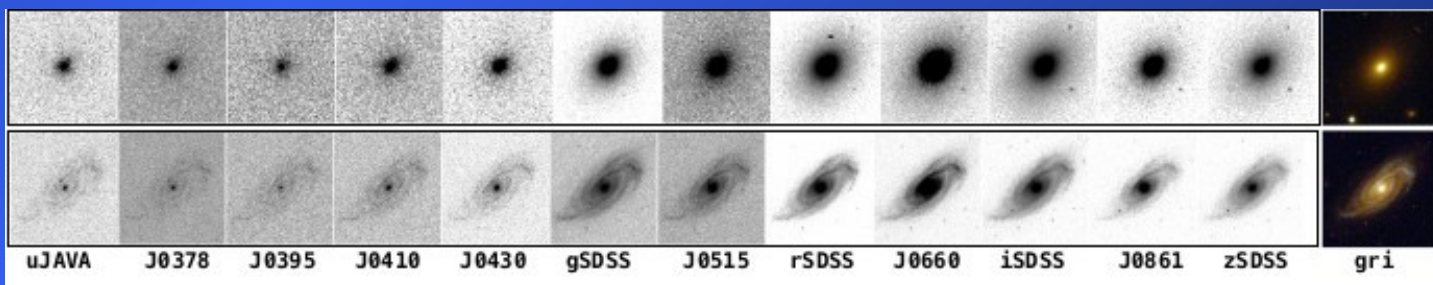
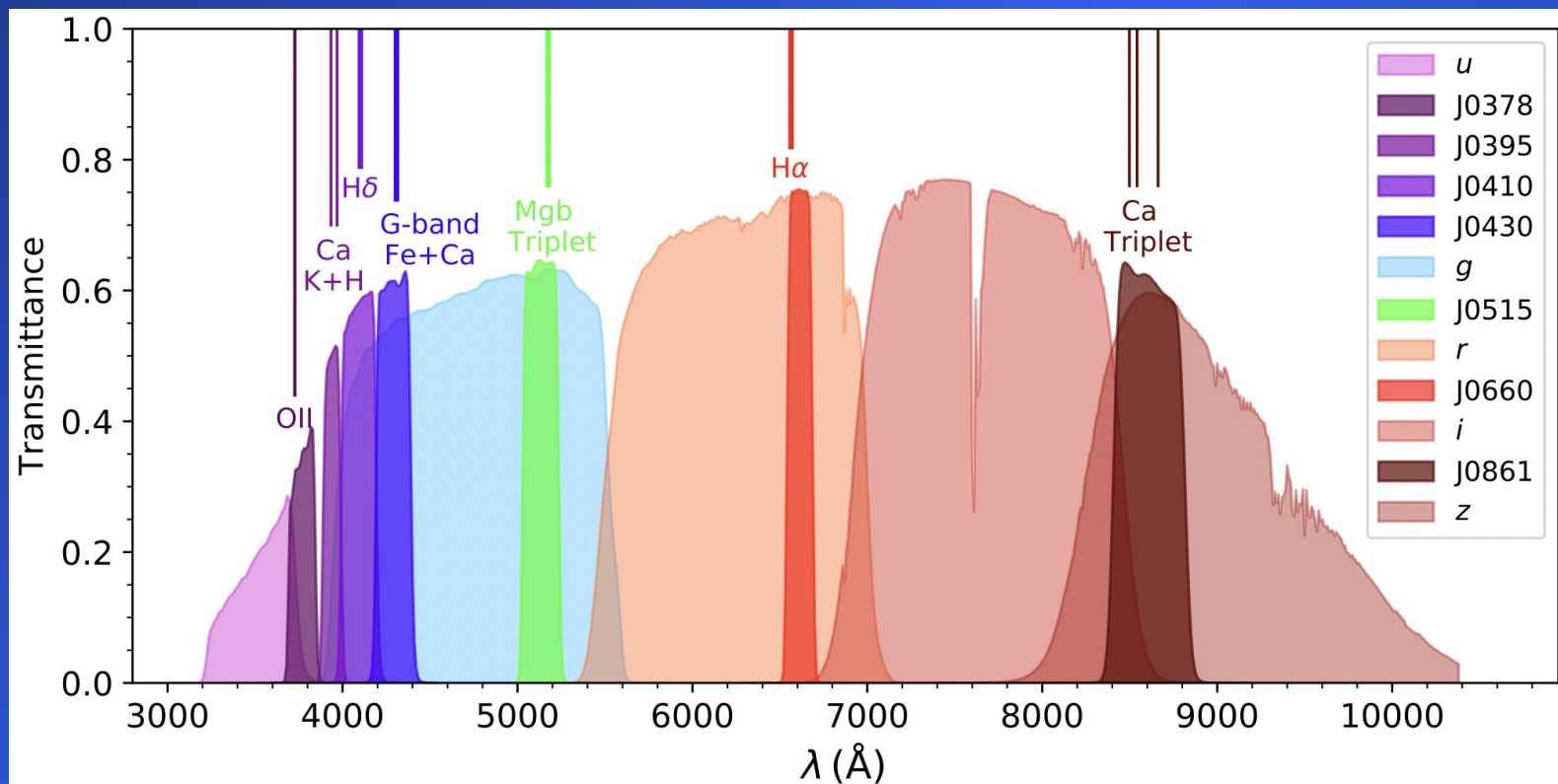
T80-South (83cm)



Cerro Tololo

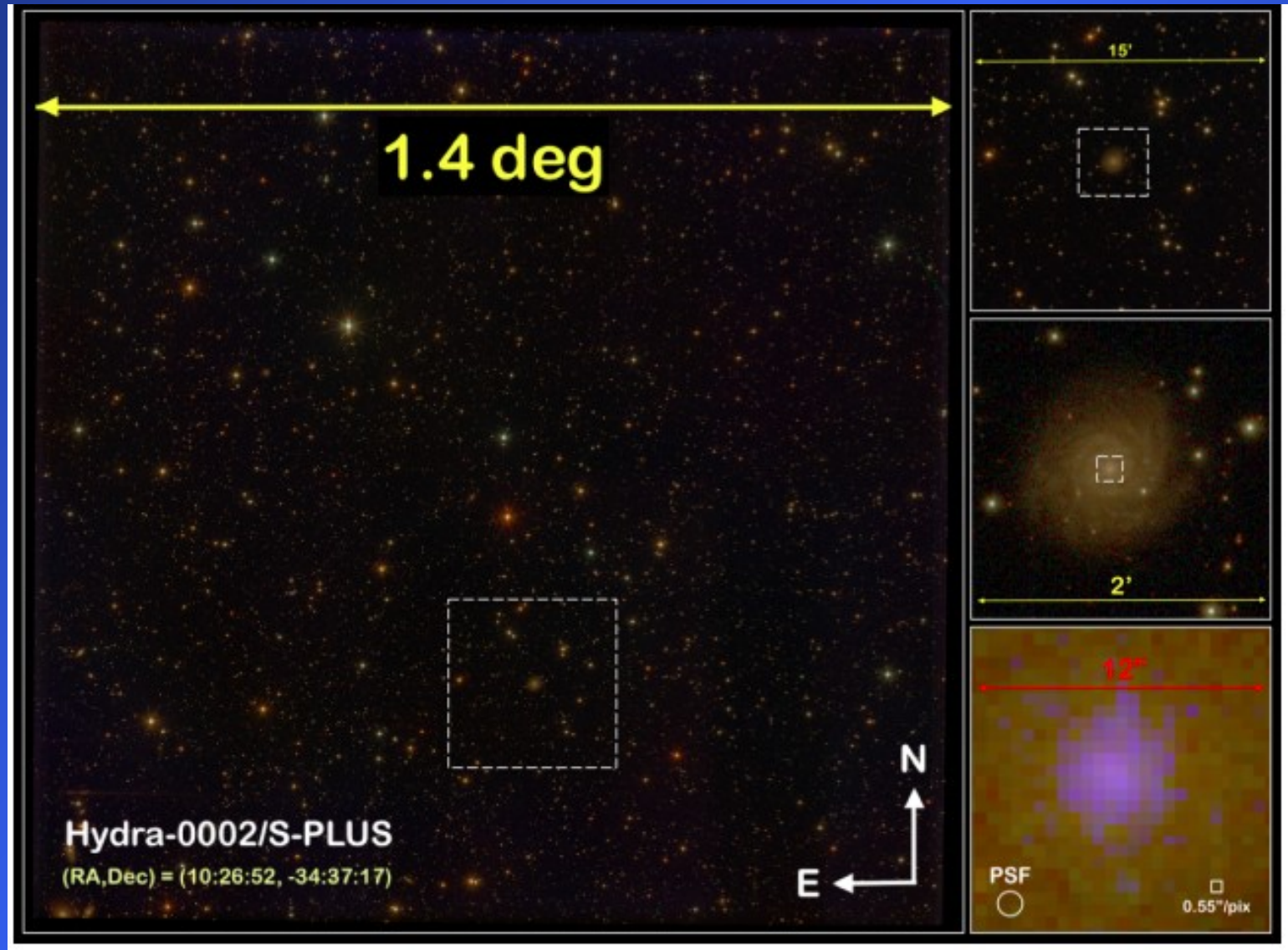
S-PLUS

**mapeamento do céu em 12 bandas:
a fotometria é equivalente a um espectro de baixa resolução**



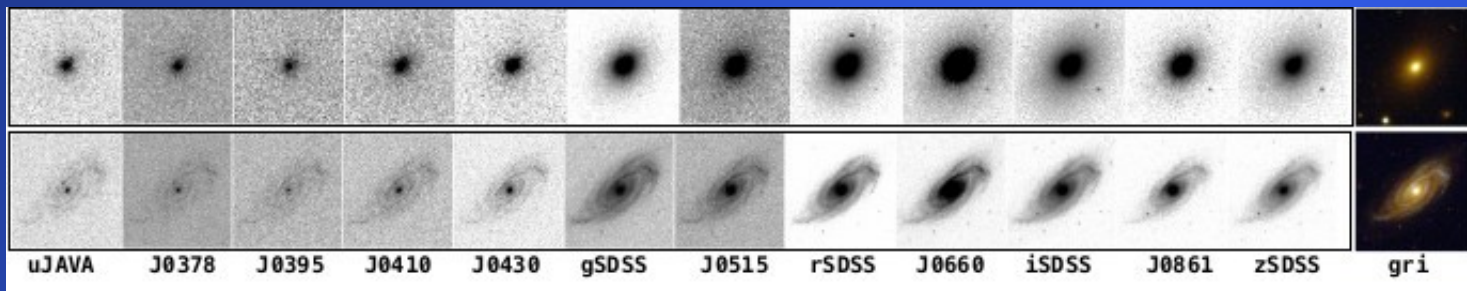
S-PLUS

campo de visão grande

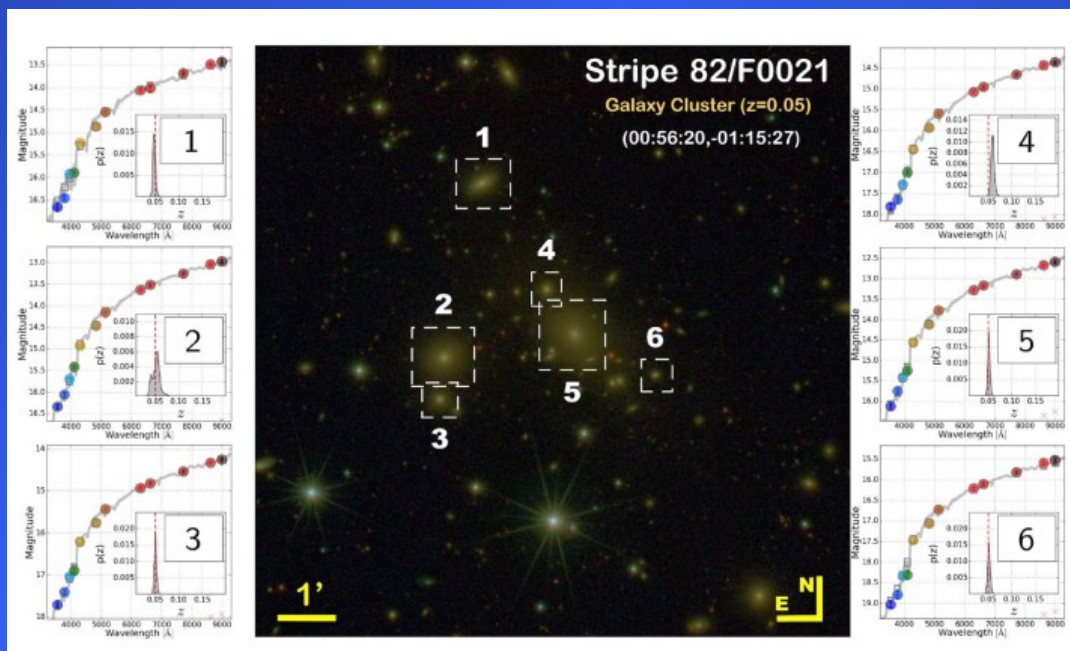


o tsunami de dados!

- entrada: cores, posições, imagens ... para milhões de objetos



- saída: propriedades de estrelas e galáxias



A IA vem
para
ajudar!

o tsunami de dados!

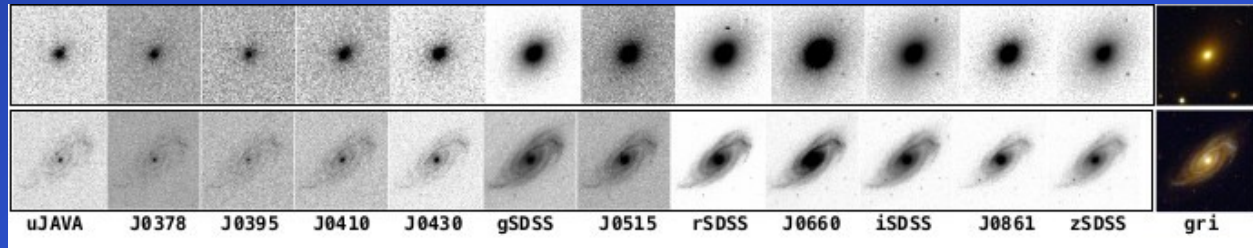
S-PLUS:

~9300 graus quadrados

~500 Tb de dados

> 1 milhão de imagens

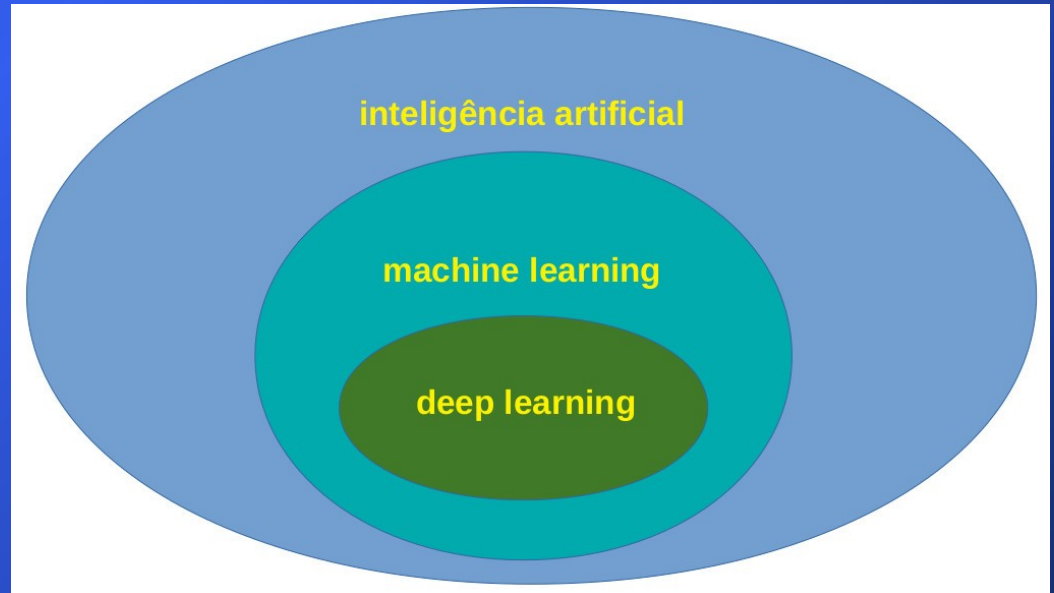
~200 milhões de objetos



- **nossa abordagem: aprendizado de máquina para estimar parâmetros interessantes**
- **input: magnitudes/cores, posições, imagens para milhares/milhões de objetos**
- **output: photo-z, parâmetros de populações estelares de galáxias ex.: massa em estrelas, idades e metalicidades médias, morfologia, objetos interessantes, ...**

inteligência artificial em Astronomia

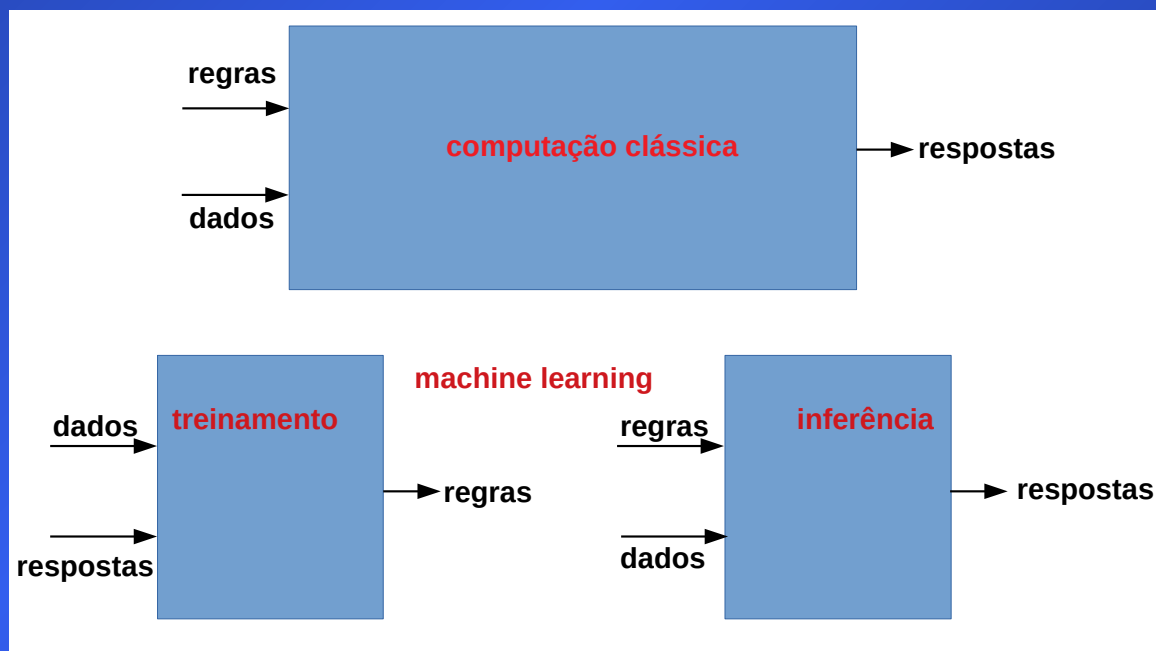
- *aprendizado de máquina/machine learning:*
permite lidar com o tsunami de dados
- são métodos flexíveis: permitem tratar um grande número de problemas
- ML é (relativamente) fácil de usar: muitos recursos disponíveis
- o mais importante:
ML é poderoso!
Permite resolver problemas intratáveis por métodos convencionais



computação convencional x aprendizado de máquina

- *programação clássica: sequência de regras específicas ou algoritmos que um computador deve seguir para resolver um problema*

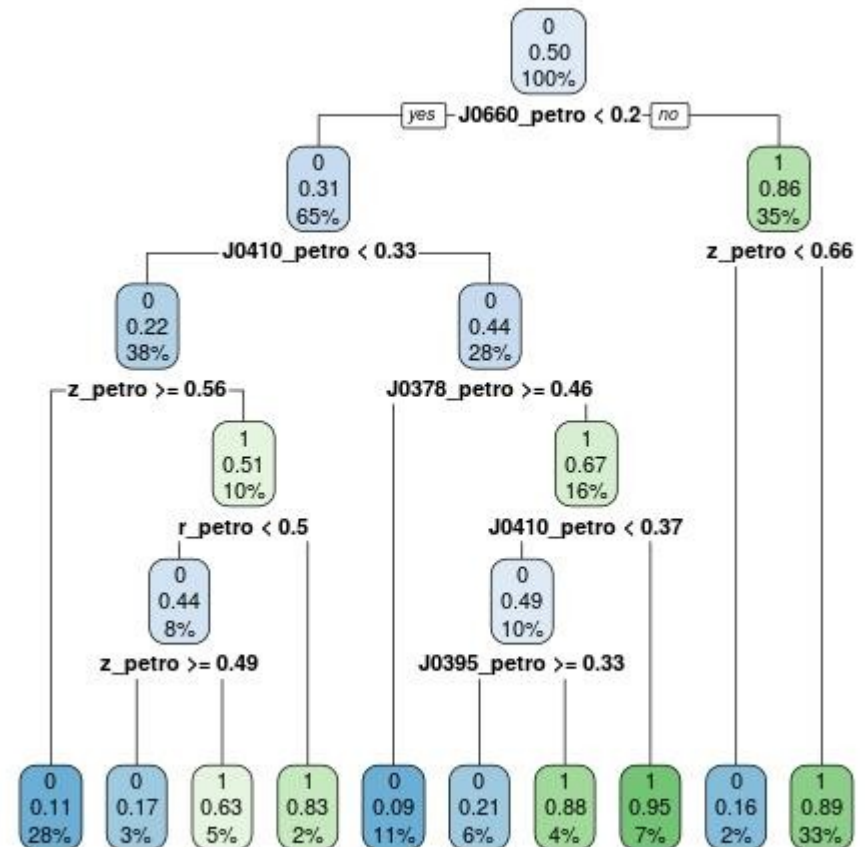
- *aprendizado de máquina (machine learning): sistemas que aprendem diretamente dos dados, sem serem explicitamente programados para uma tarefa específica*



Exemplo: classificação estrela/galáxia com árvores de decisão



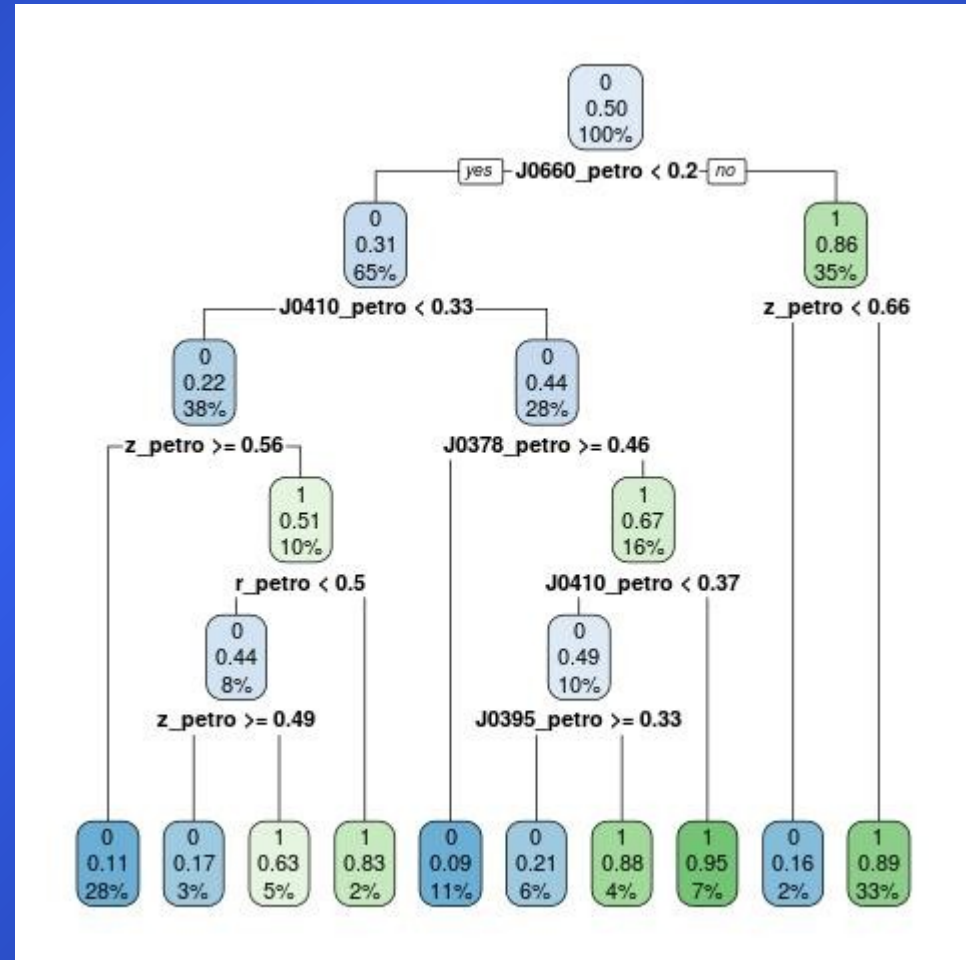
0 – estrela
1 – galáxia



Exemplo: classificação estrela/galáxia com árvores de decisão

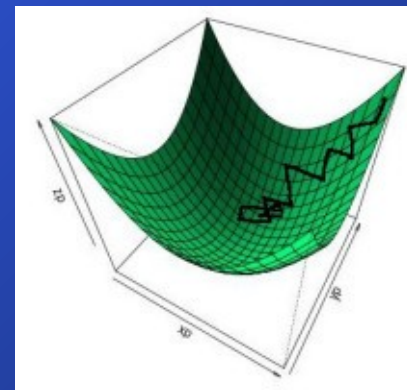
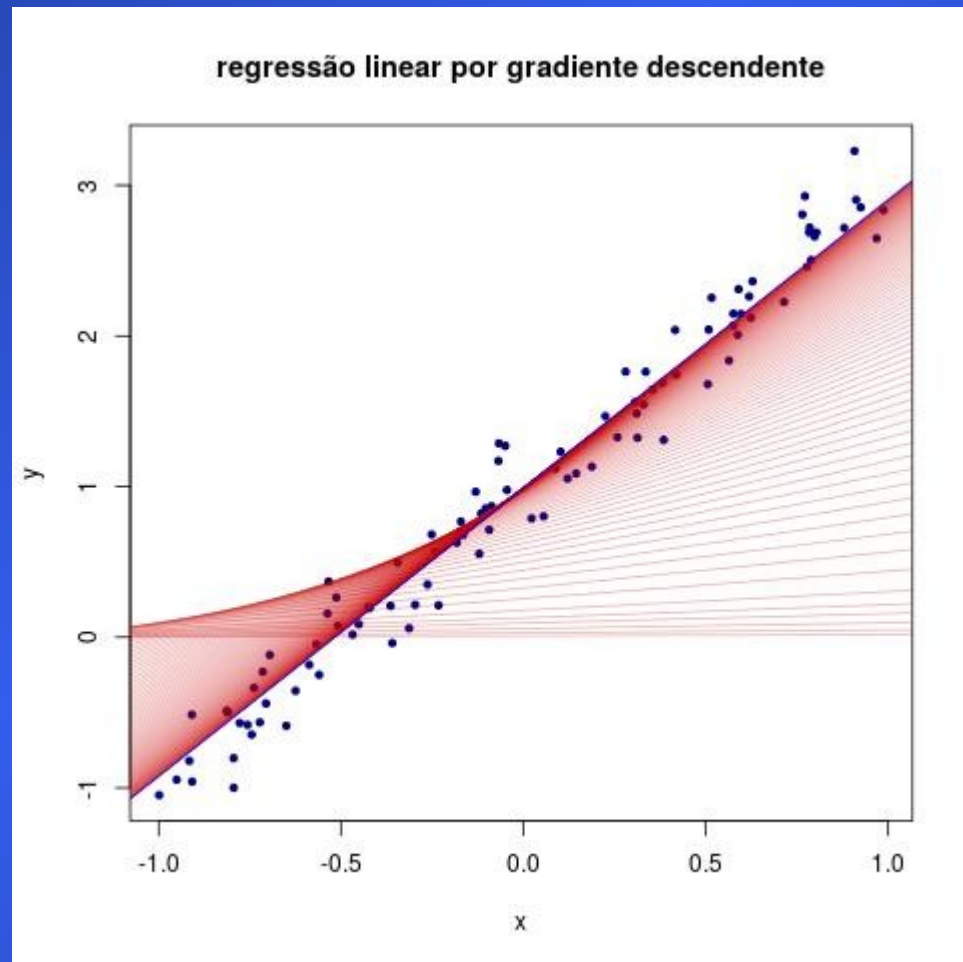
como o modelo “aprende”?

- *usa-se um conjunto de treinamento com objetos com classificação conhecida*
- *usa-se um processo iterativo para encontrar os parâmetros do modelo*
- *os parâmetros são modificados iterativamente de modo a melhorar o desempenho do modelo*

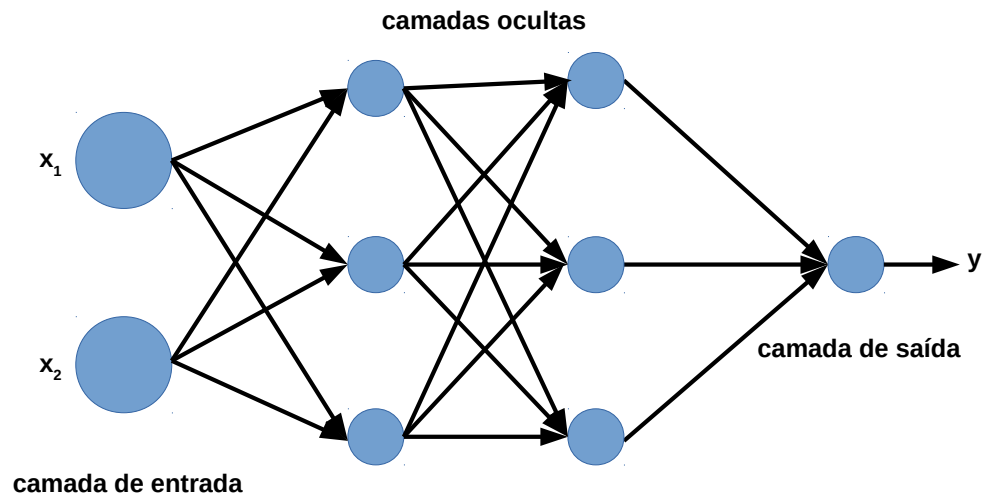


como o modelo “aprende”?

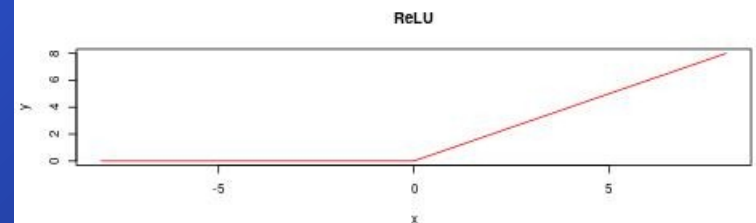
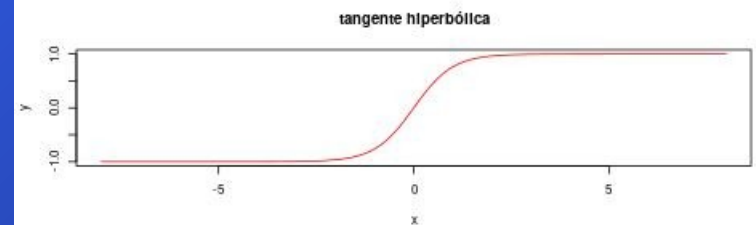
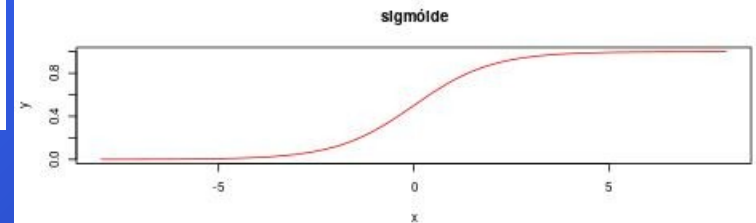
- *usa-se um conjunto de treinamento com objetos com classificação conhecida*
- *usa-se um processo iterativo para encontrar os parâmetros do modelo*
- *os parâmetros são modificados de modo a melhorar o desempenho do modelo*
- *minimiza-se uma função custo com gradiente descendente*



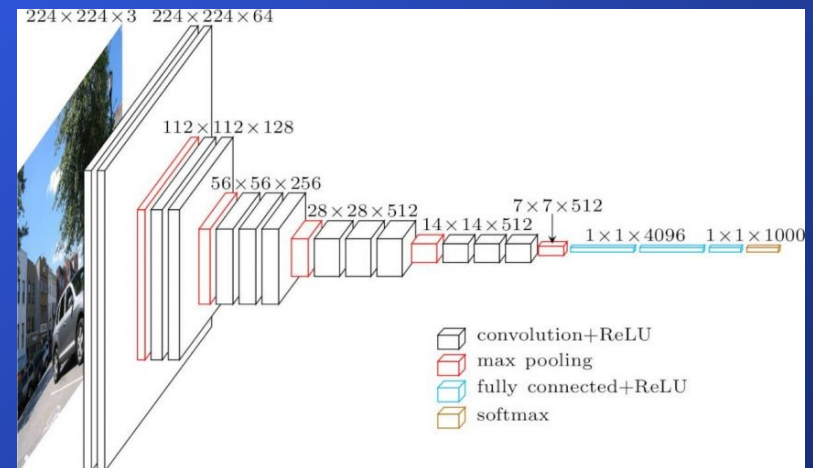
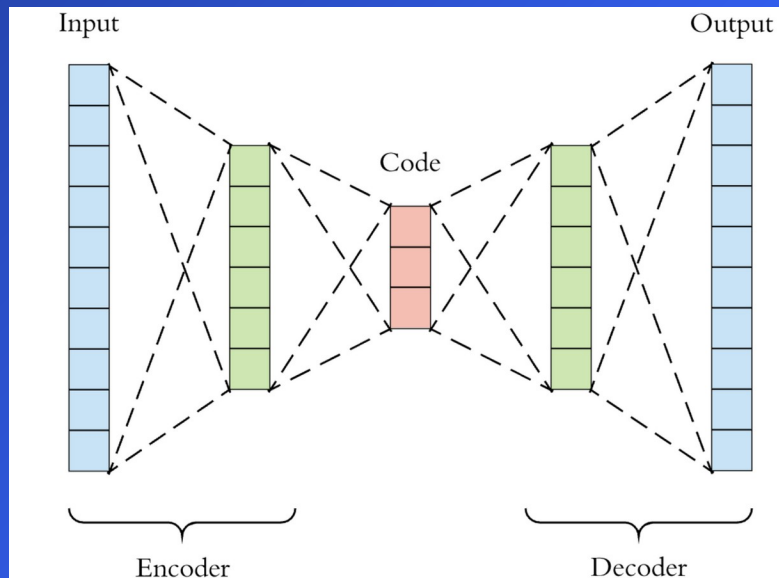
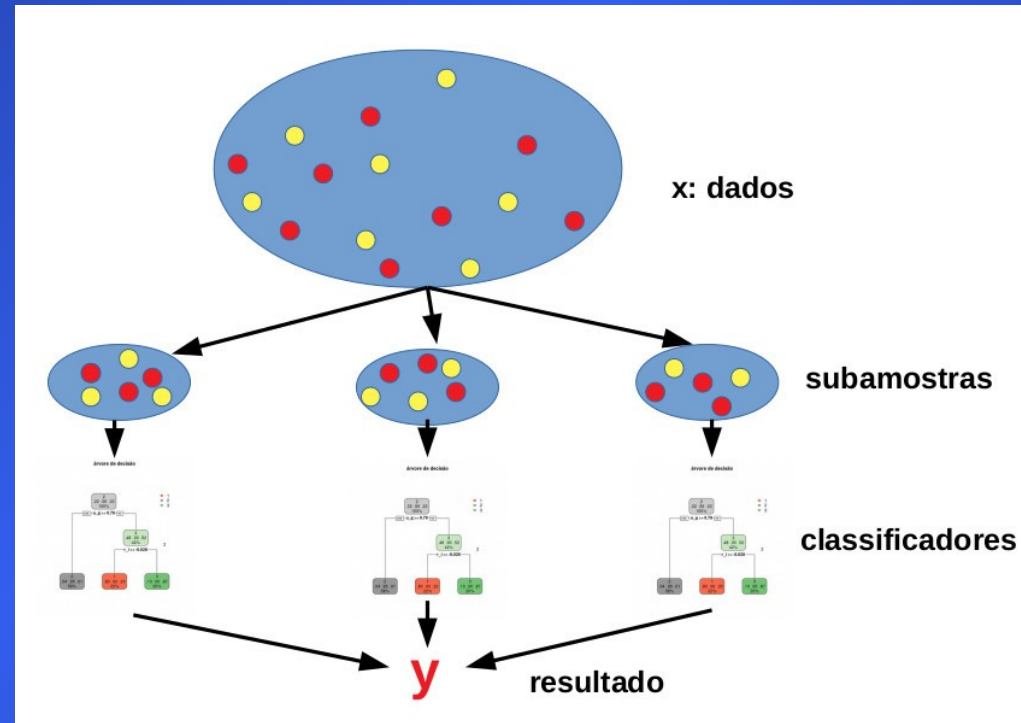
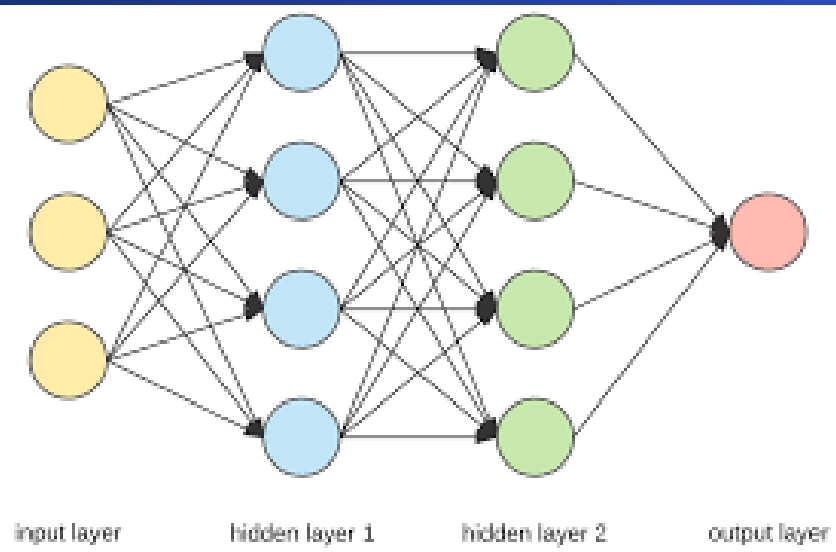
redes de neurônios artificiais



o que é uma rede profunda?
muitas camadas ocultas, com
funções de ativação não-lineares

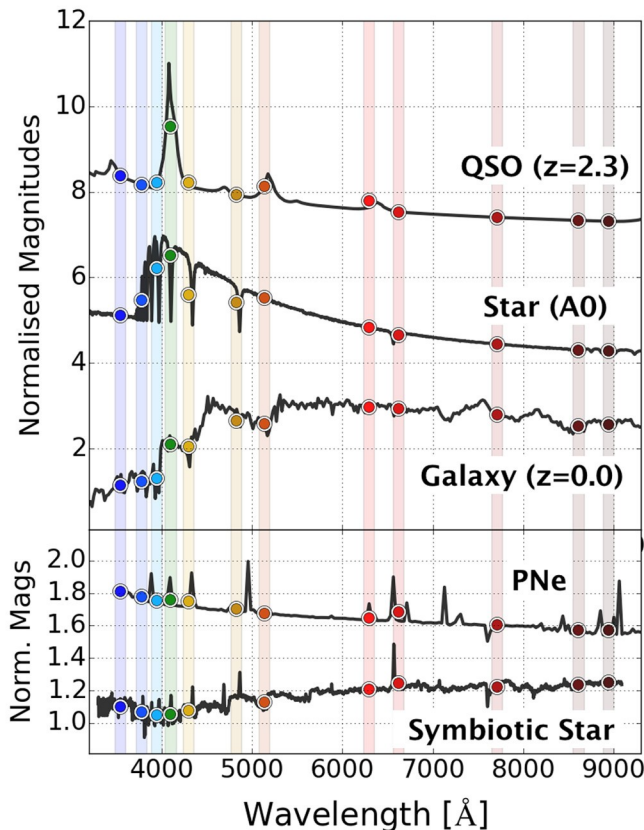


a fauna de algoritmos

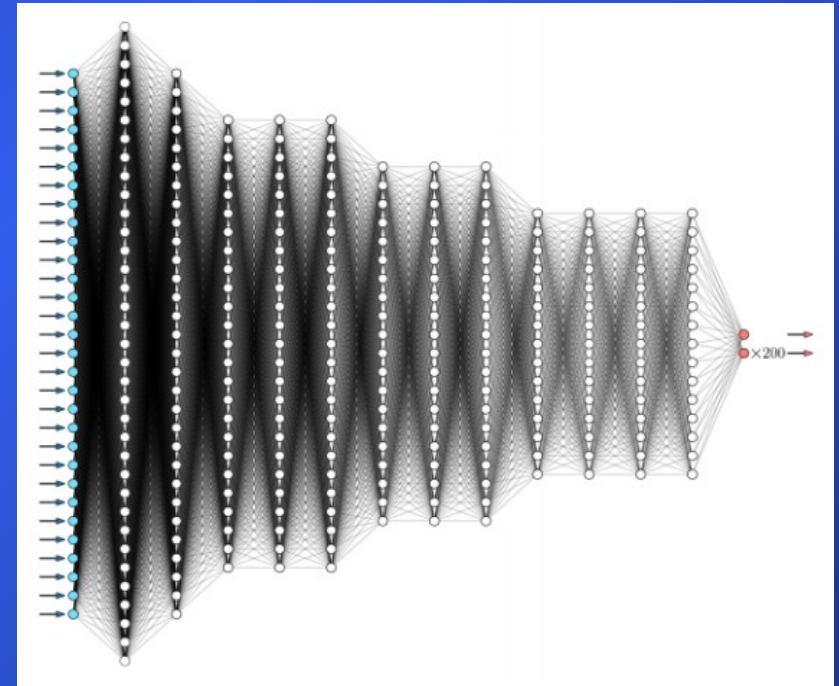


redshifts fotométricos para o S-PLUS

redshifts são normalmente medidos usando espectroscopia



Input: parâmetros extraídos das imagens



Output: $\text{Prob}(z)$

redshifts fotométricos:
estimados a partir de fotometria

A rede neural é treinada com galáxias que têm redshifts espectroscópicos

populações estelares de galáxias

Photometric SED fitting with Deep Learning

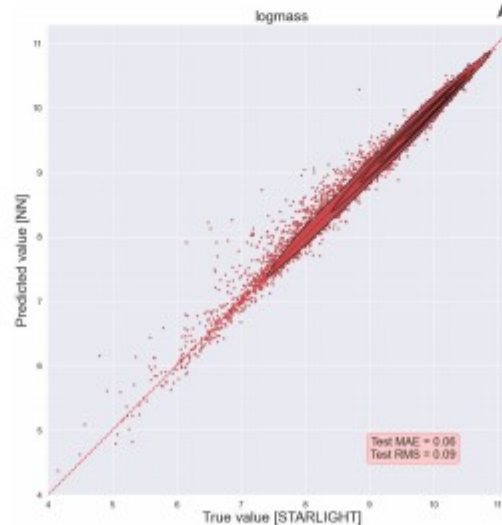
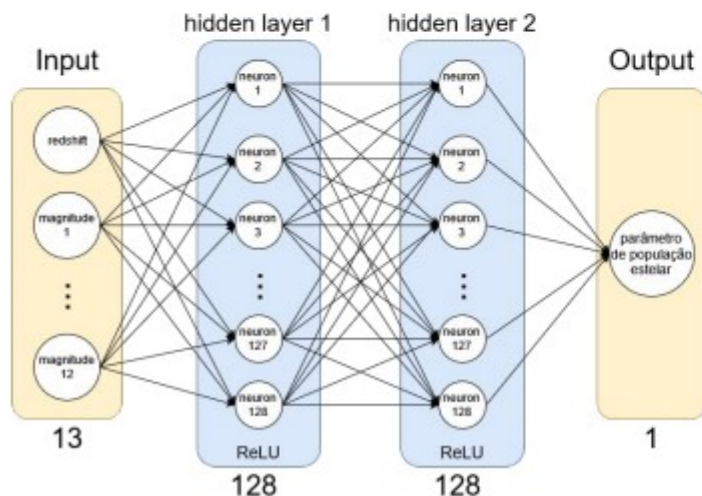
Vitor Cernic (vitorcernic@usp.br)

S-PLUS data

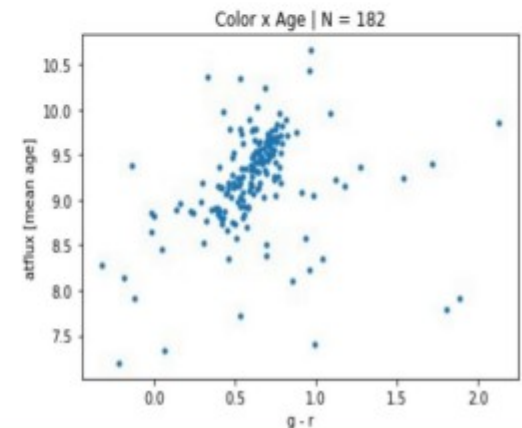
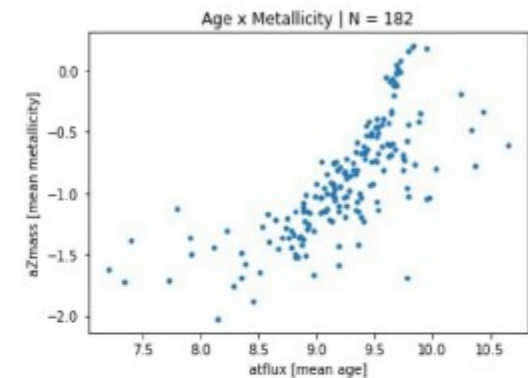
12 magnitudes
+ 1 redshift

1 STARLIGHT
parameter

- Age
- Metallicity
- Veloc. Disp.
- Reddening
- EWH α
- SFR
- Stellar mass



Application to Fornax Cluster:

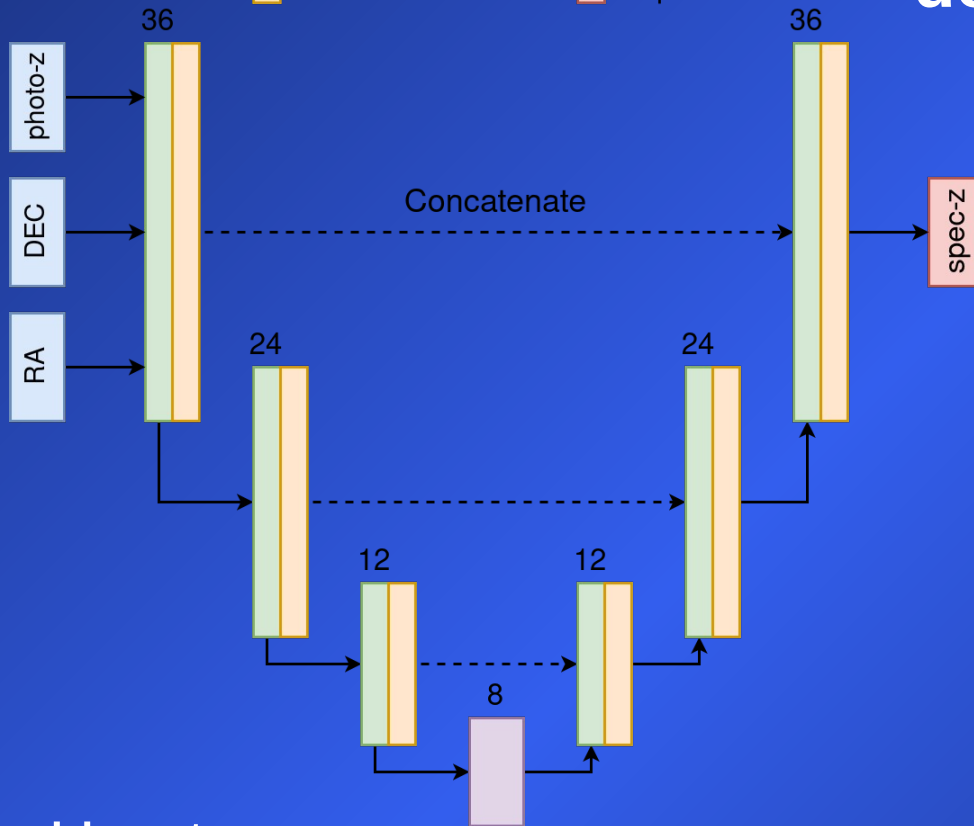


deconvolução de redshifts fotométricos

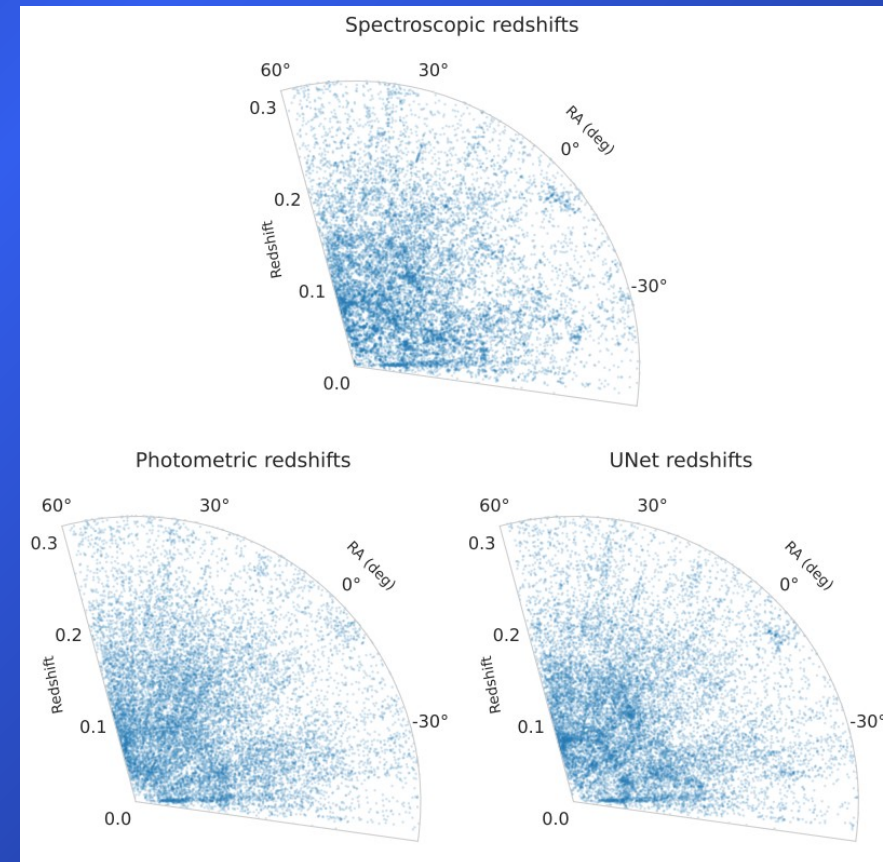
Recuperação da distribuição de galáxias em 3D

Input Dense layer Bottleneck

BatchNormalization Output



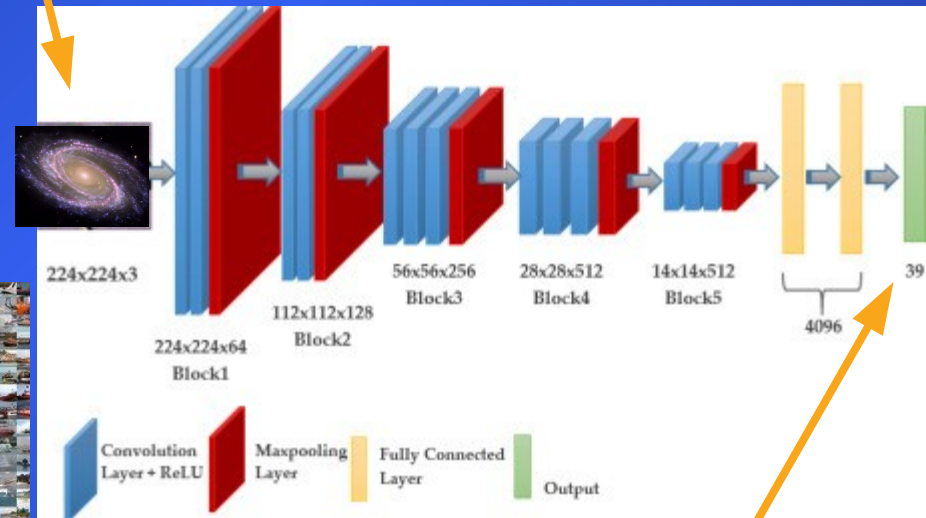
U-nets
diffusion networks



classificação morfológica de galáxias com redes pré-treinadas

Input: imagens de galáxias

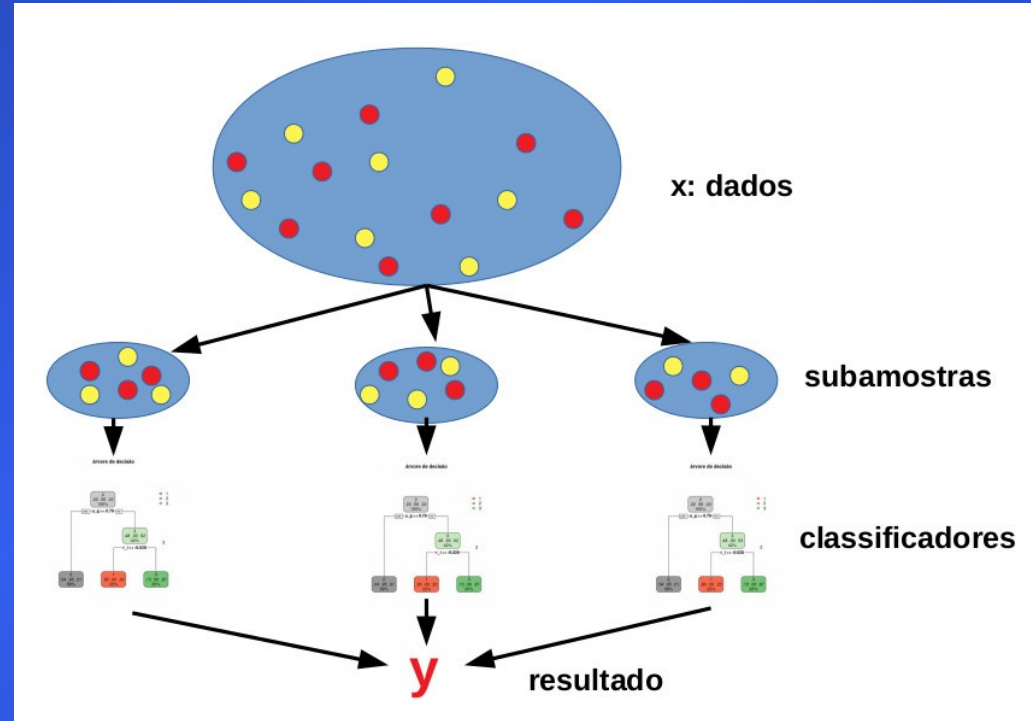
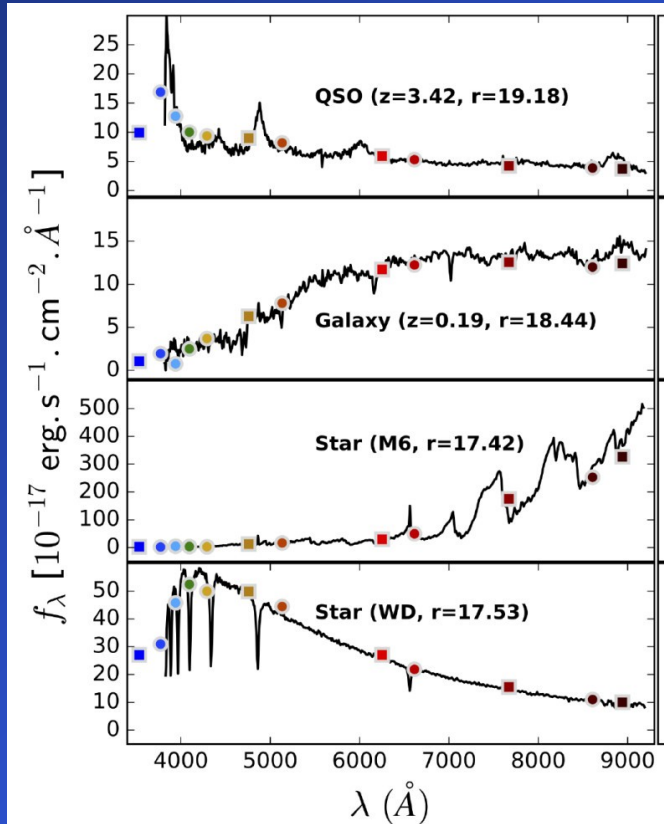
- EfficientNet B2 architecture
- (transfer learning)



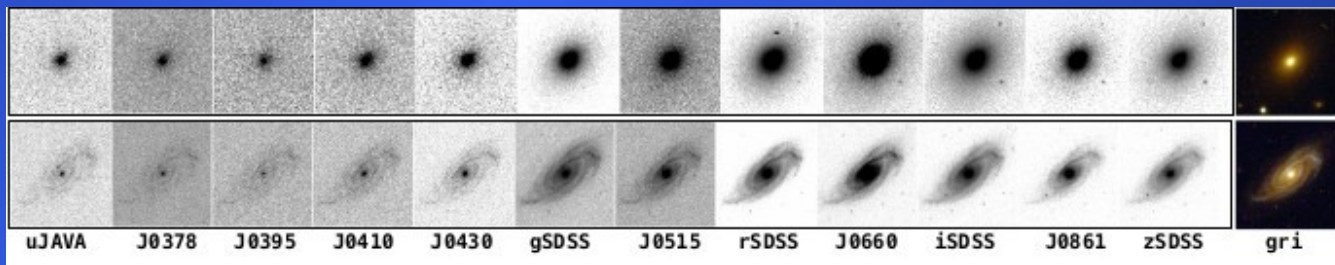
Output: E, S, Ambíguo

conjunto de treinamento: Galaxy Zoo 1

classificação estrela/galáxia/quasar com *random forest*

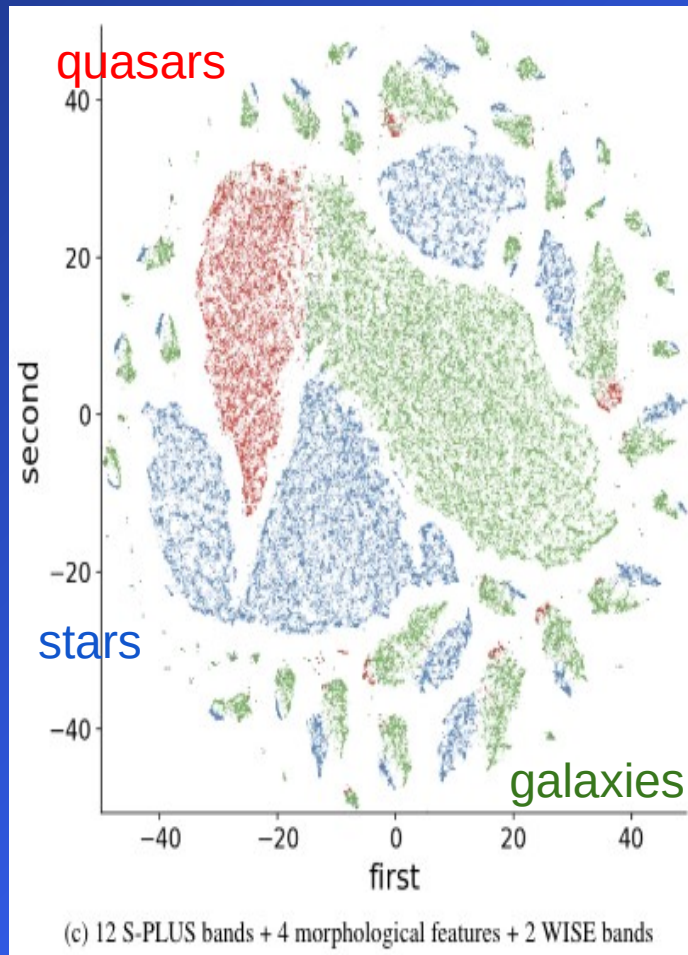


O algoritmo dá a probabilidade de um objeto ser estrela, galáxia ou quasar

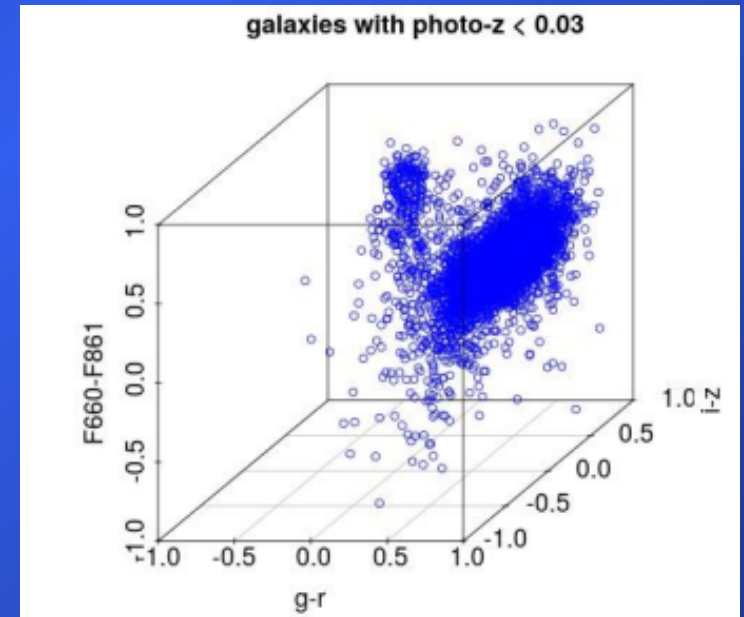


Compressão de informação

o espaço de dados é multi-dimensional!
como visualizá-lo?



exemplo de um espaço de dados
tri-dimensional



t-Distributed Stochastic Neighbor
Embedding (t-SNE)

Inteligência artificial em Astronomia

- métodos de *machine learning* são eficientes para lidar com o tsunami de dados
- flexíveis: podem ser aplicados a um grande número de problemas
- muitos recursos disponíveis

mais importante:
podemos resolver
problemas intratáveis ou
computacionalmente muito
intensivos com métodos
convencionais

