

Unsupervised Learning

Alberto Krone-Martins

CENTRA/SIM - Faculdade de Ciências
Universidade de Lisboa

“Crude classifications and false generalizations are the curse of organized life.” G. B. Shaw

In this talk

Unsupervised Learning

Three astronomical examples

In this talk

Unsupervised Learning

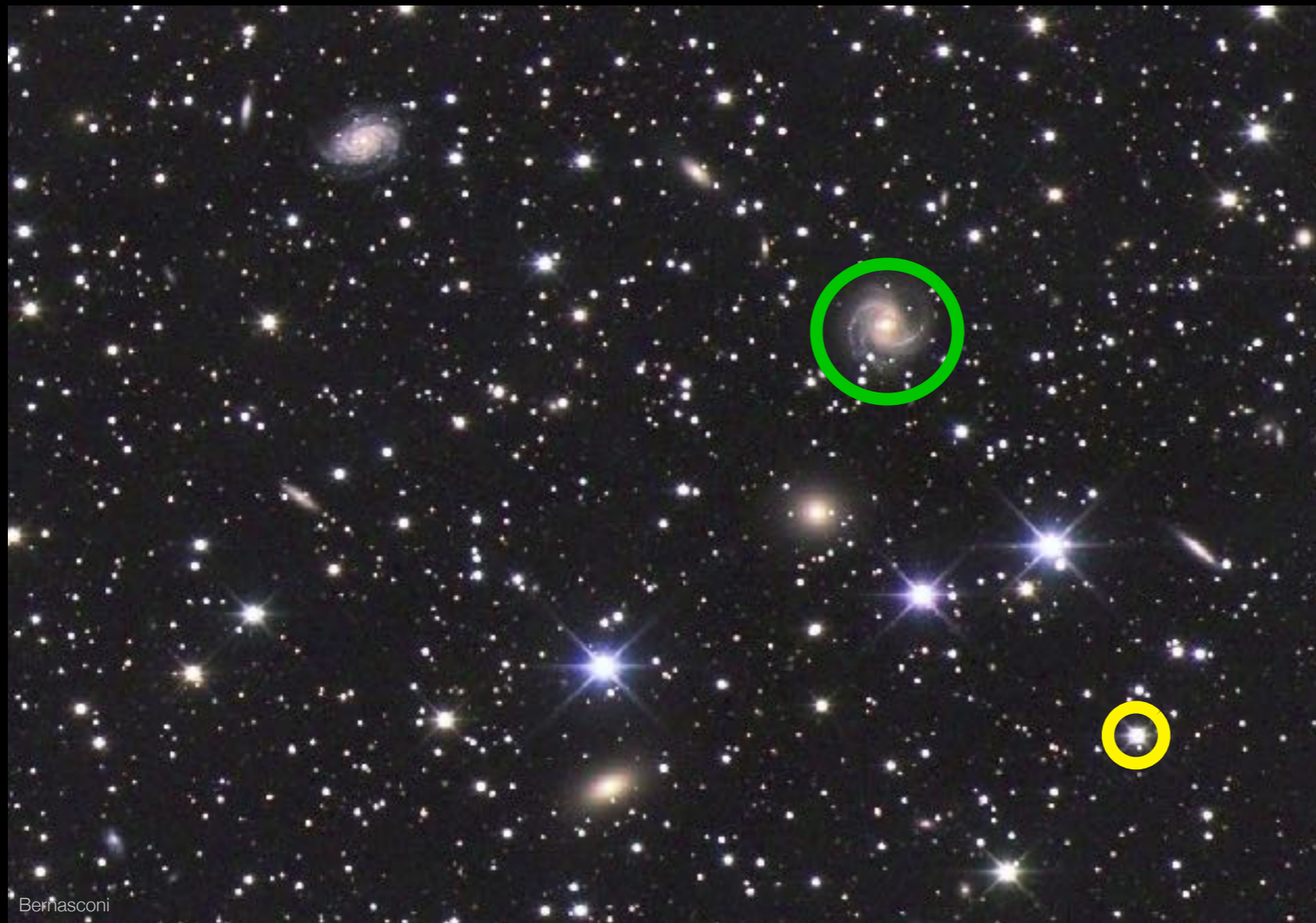
Three astronomical examples

Explicit methods : talking with machines



**What would you do
to find stars and galaxies
in this image?**

Explicit methods : talking with machines



**Human
crafted
decision
rules**

**If not compatible
with a PSF
= galaxy**

**If compatible with
a PSF
= star**

Explicit methods : limitations



**We need to write
the underlying
rules...**

**What happens
with small
galaxies?**

**Are there only
stars & galaxies in
the Universe?**

Explicit methods : limitations

Supervised machine learning...

We need to write the underlying rules...

What happens with small galaxies?

Are there only stars & galaxies in the Universe?

Supervised methods : limitations



Biased results



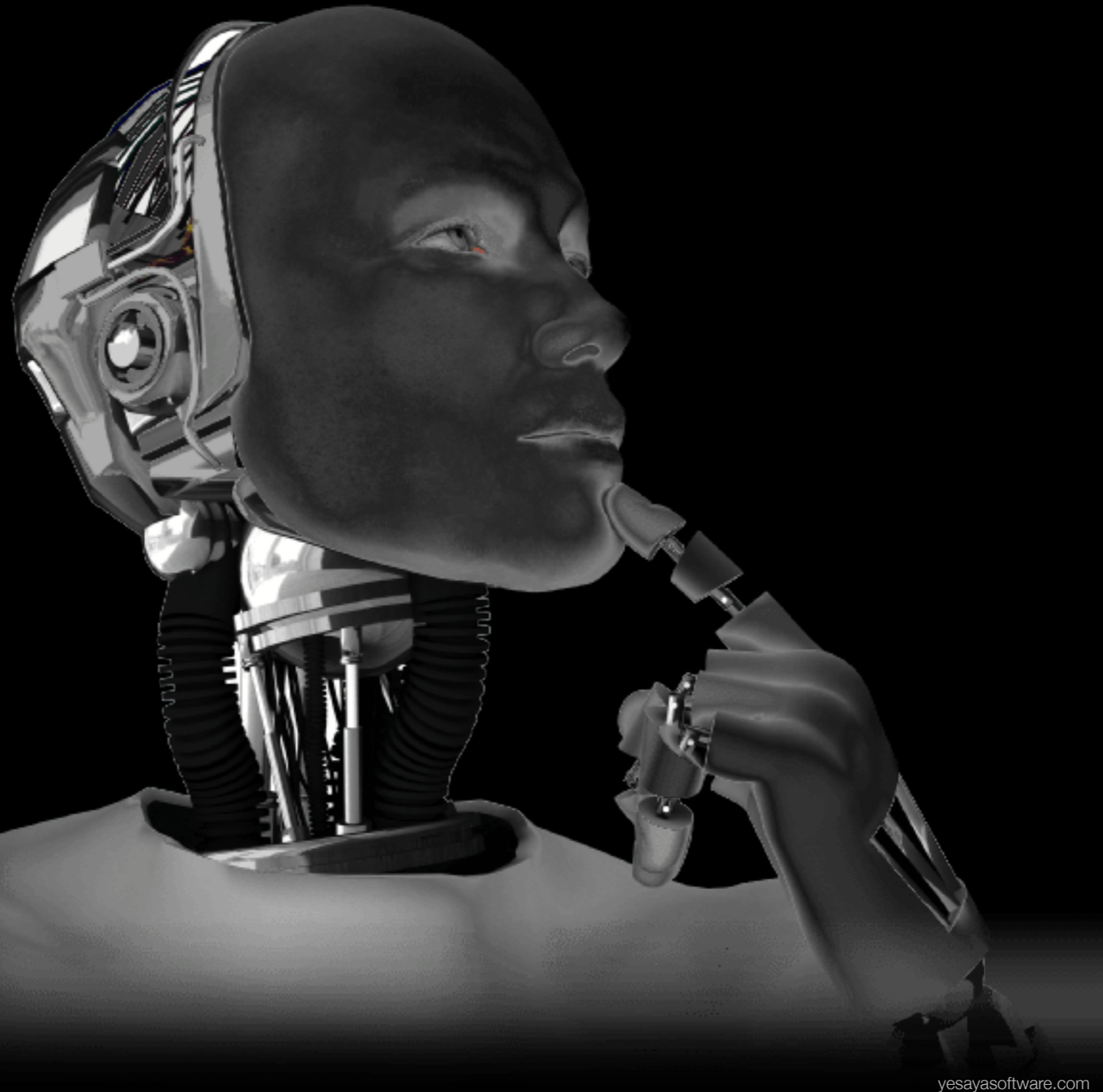
Are there only stars and galaxies in the Universe?

What if there is no training data available?



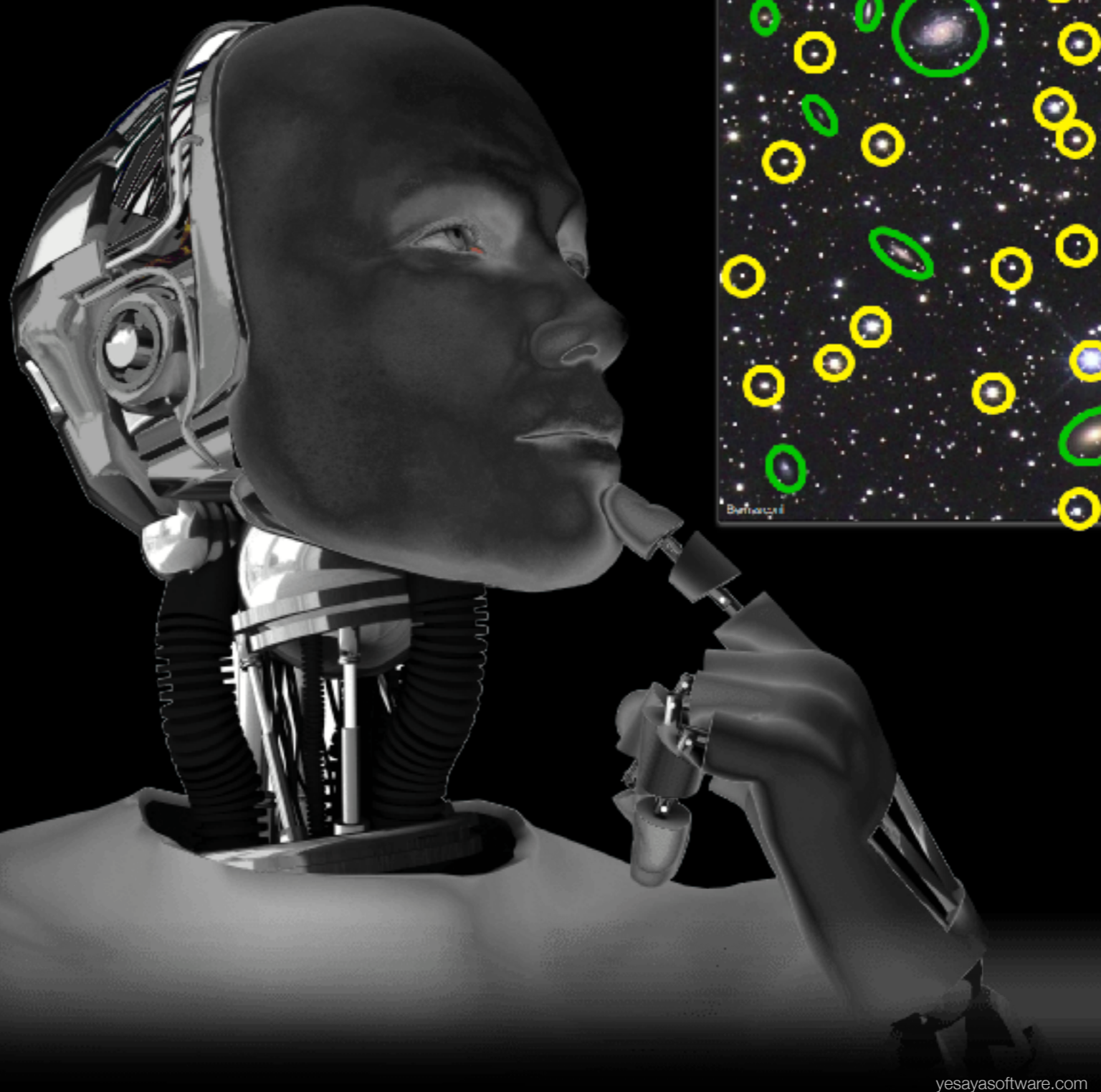
Classically, impossible

Unsupervised learning



yesayasoftware.com

Unsupervised learning

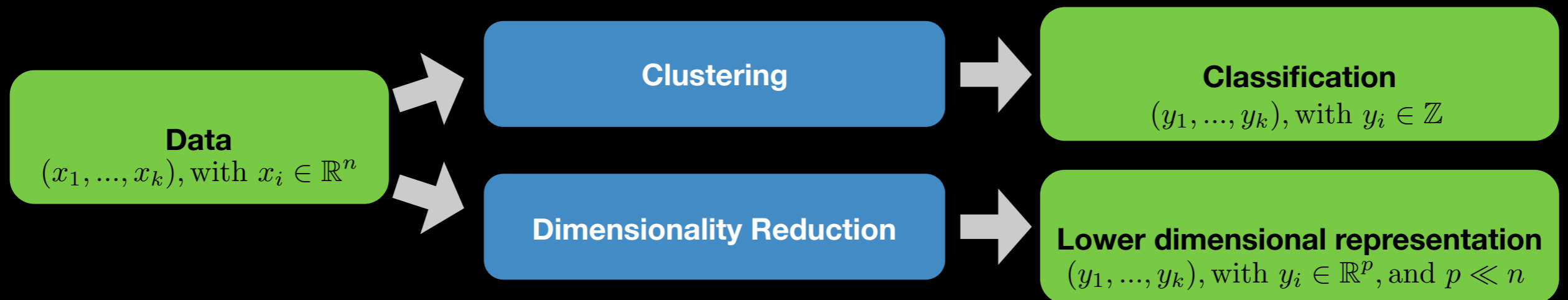


yesayasoftware.com

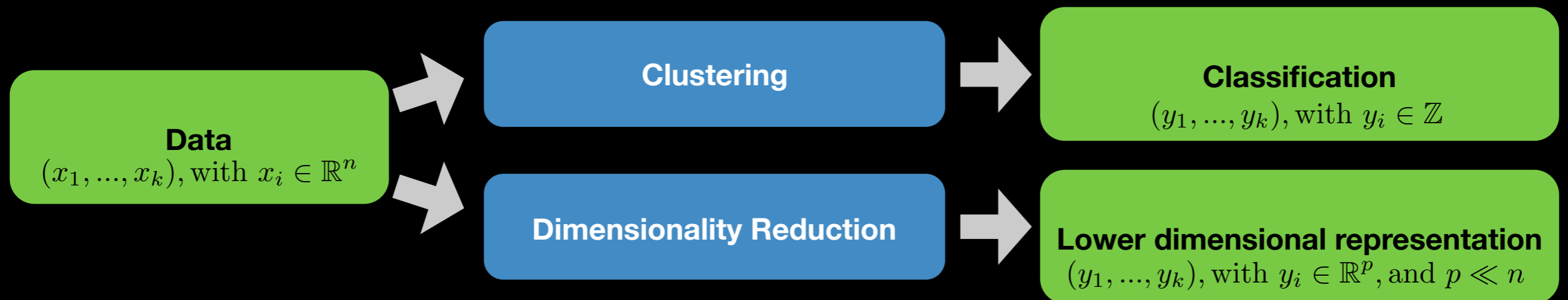
Unsupervised learning : Cornerstones



Unsupervised learning : Cornerstones



Unsupervised learning : Cornerstones



Clustering can be seen as a limit case of dimensionality reduction

Unsupervised learning : Ingredients



Unsupervised learning : Ingredients



Similarity measure?

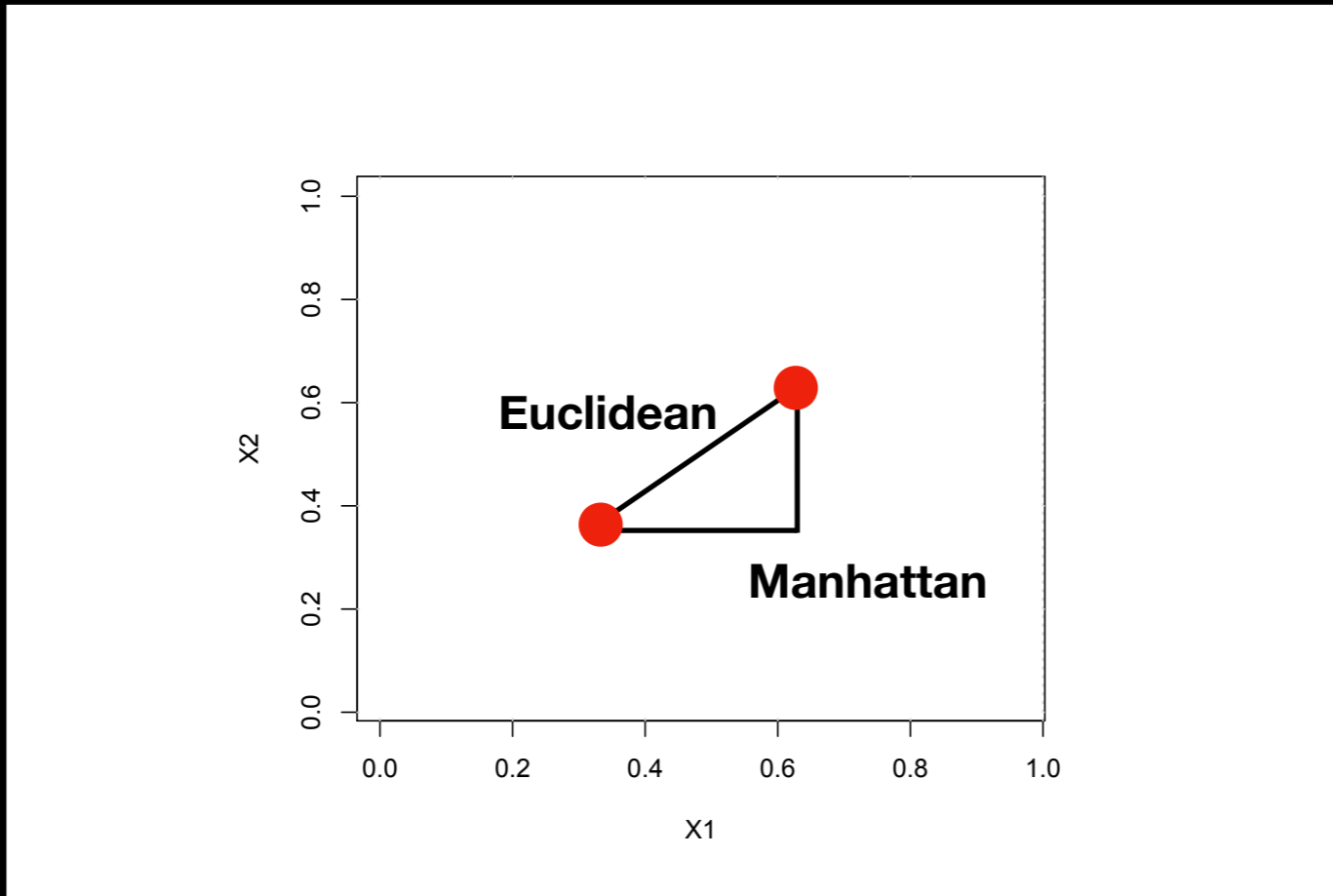
Euclidean (p=2)

Manhattan (p=1)

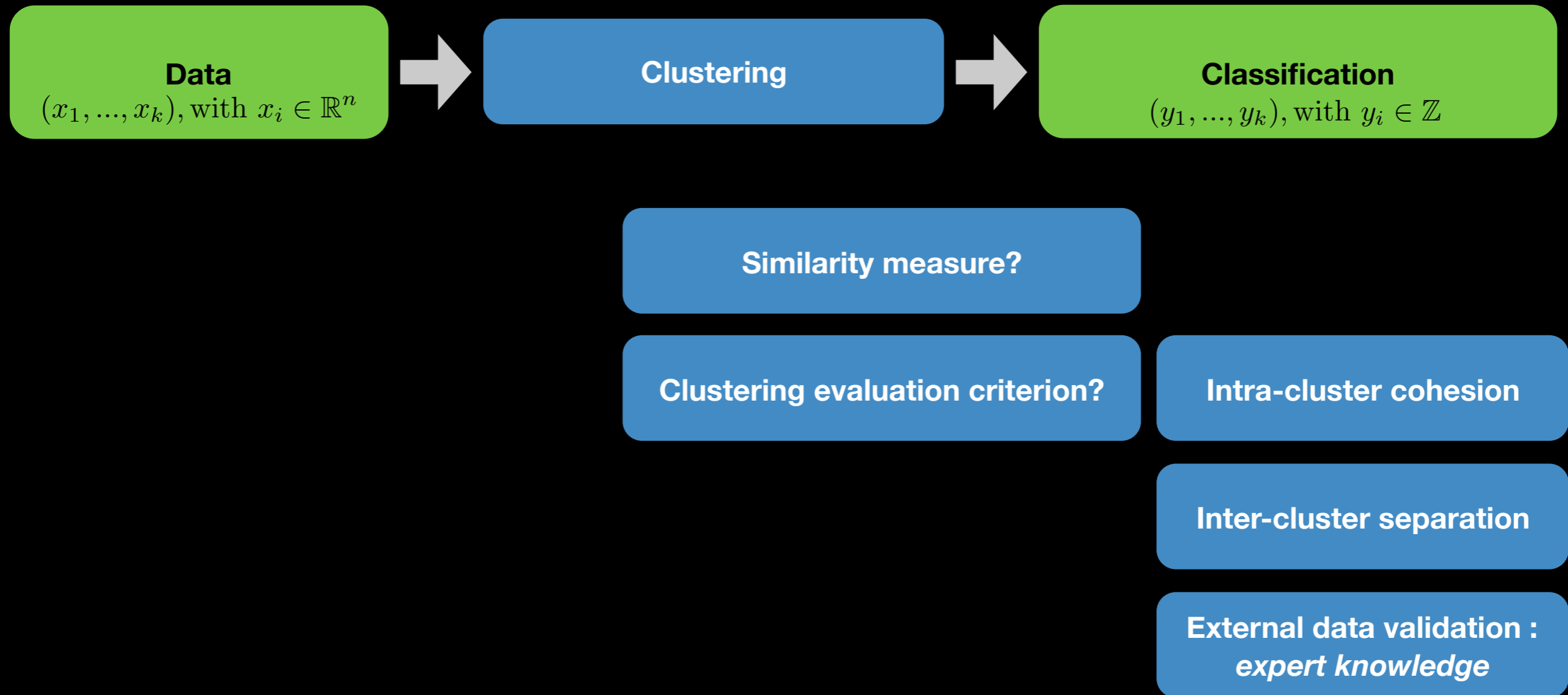
Minkowski

$$d_p(\mathbf{x}_i, \mathbf{x}_j) = \left(\sum_{k=1}^m |x_{ik} - x_{jk}|^p \right)^{1/p}$$

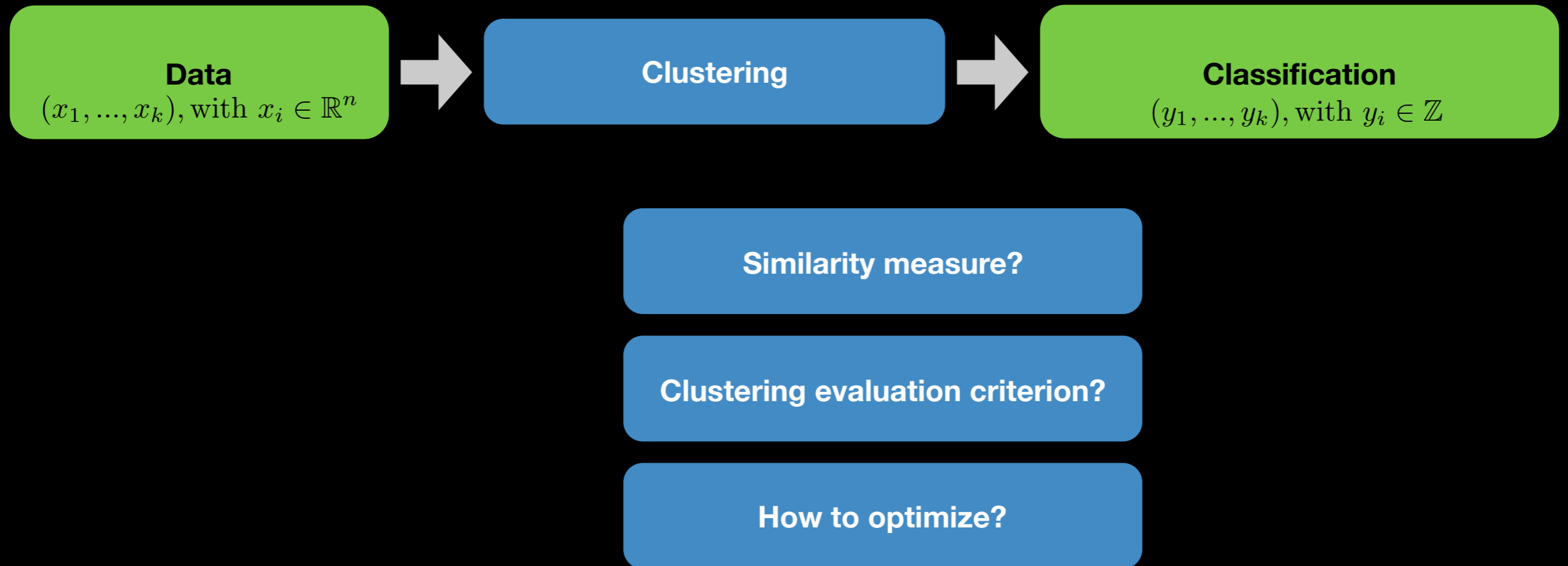
...



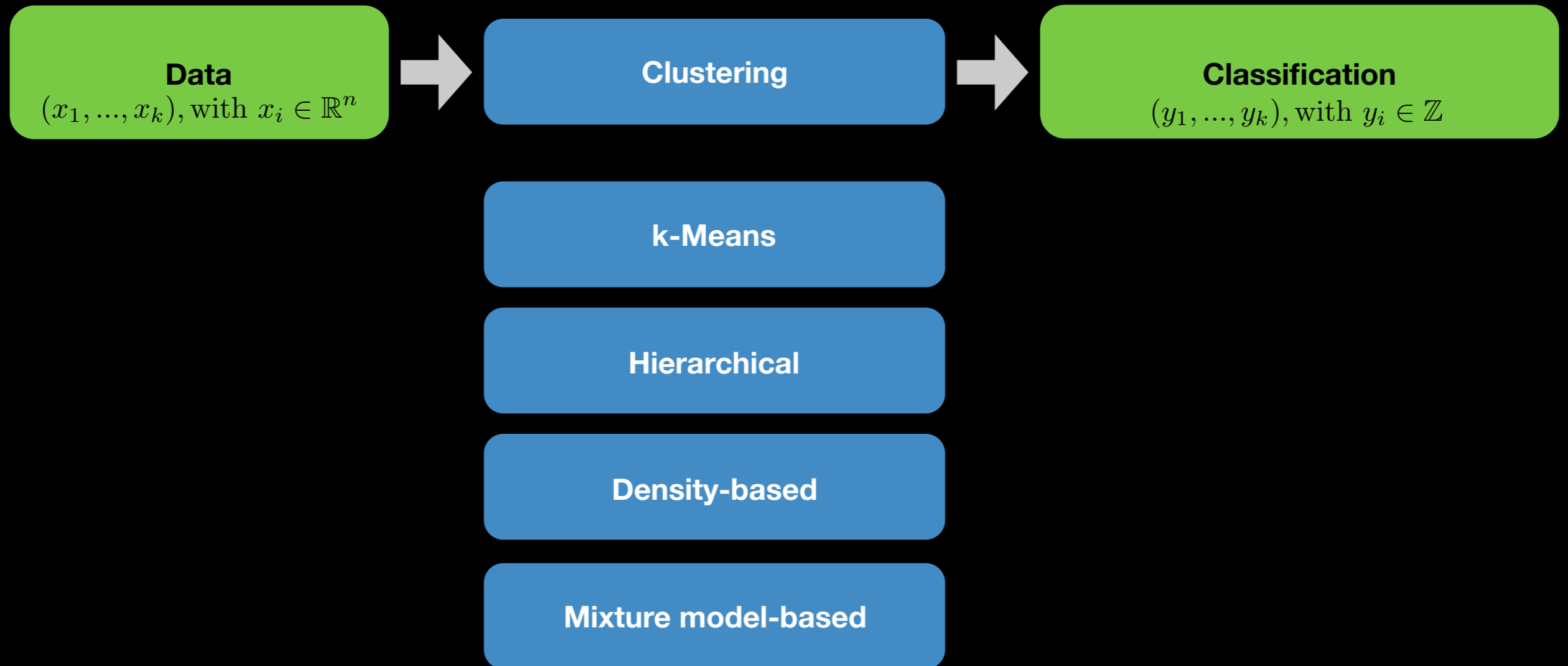
Unsupervised learning : Ingredients



Unsupervised learning : Ingredients



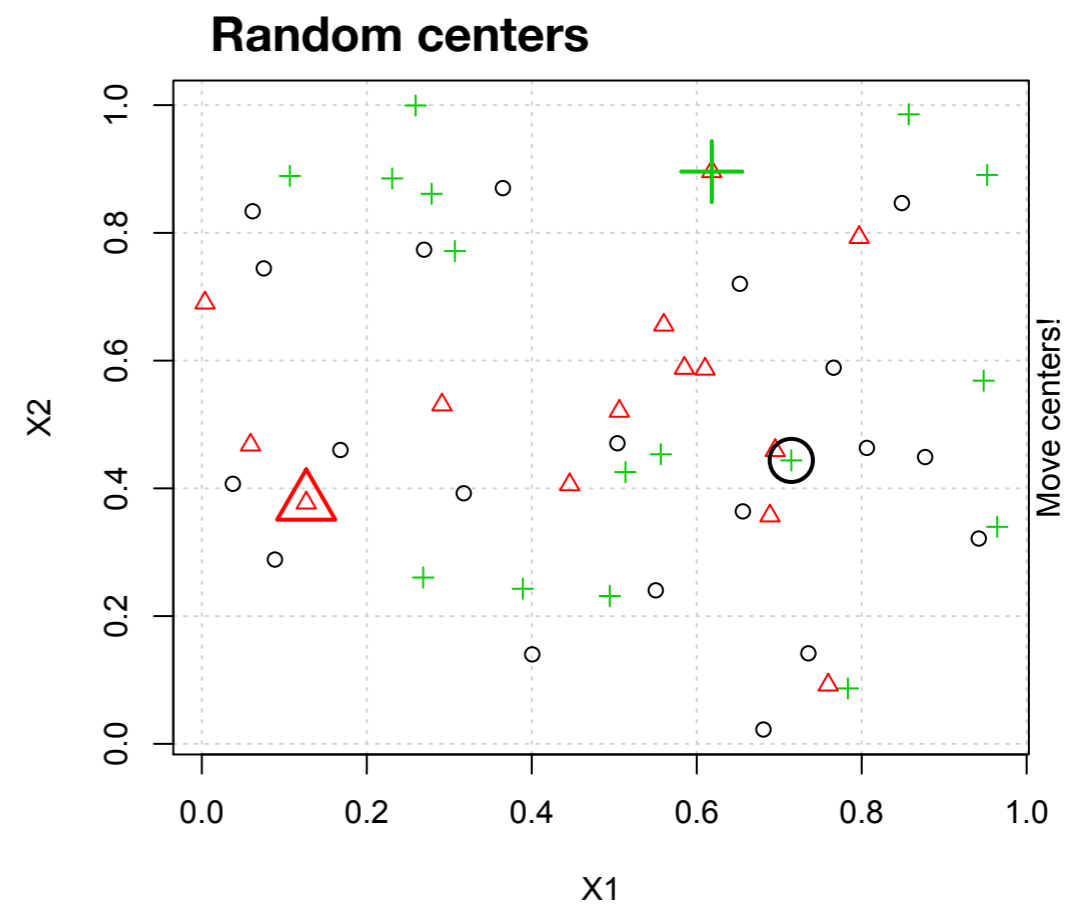
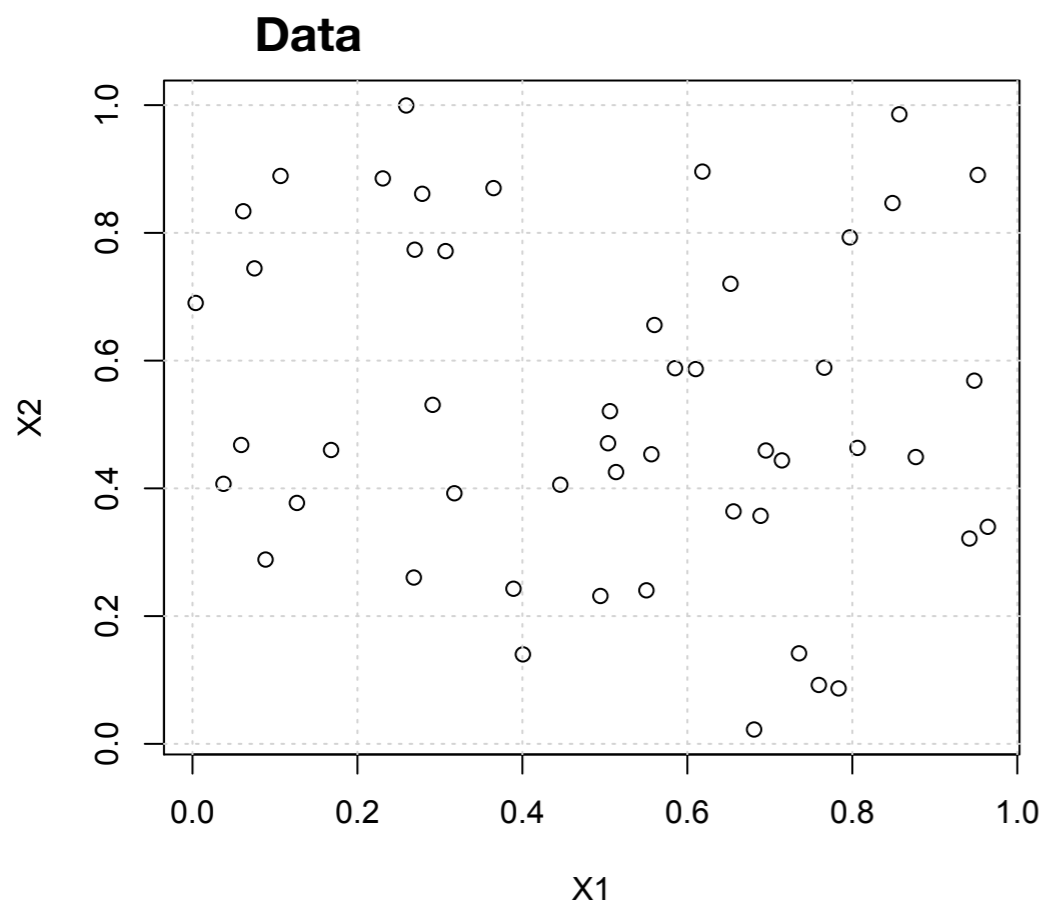
Unsupervised learning



Unsupervised learning



Iteration 1

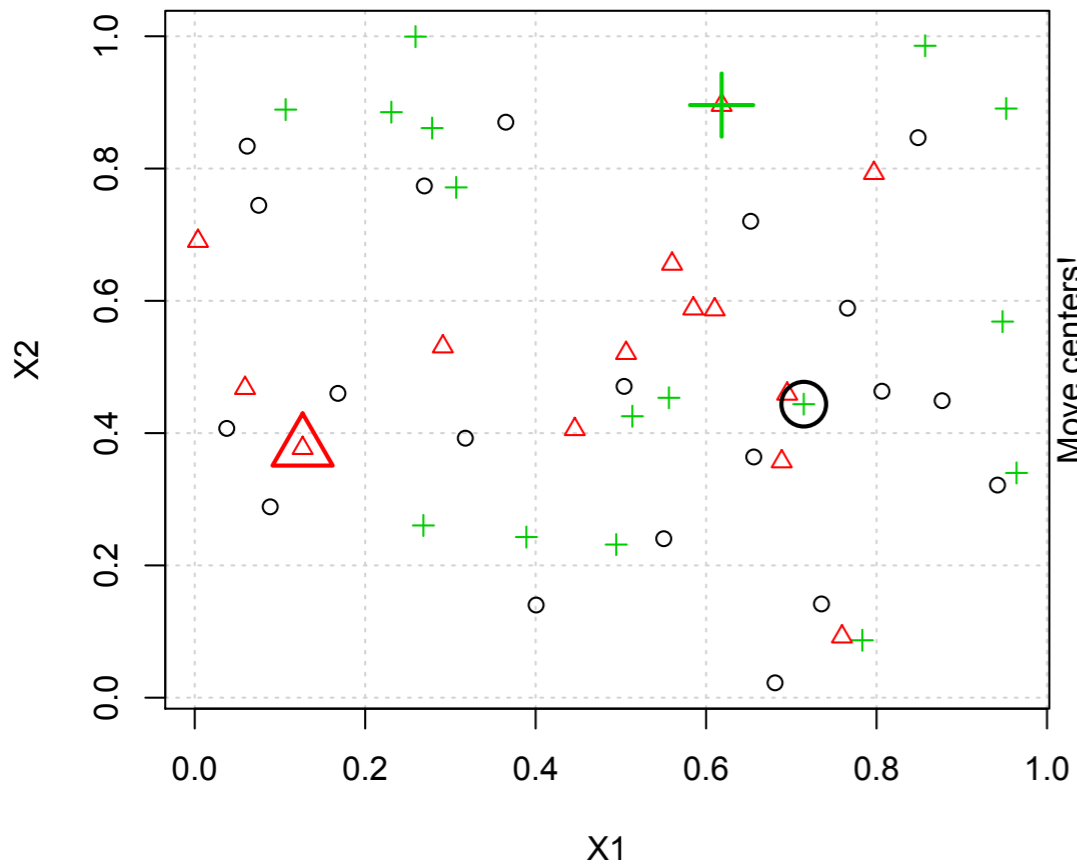


Unsupervised learning



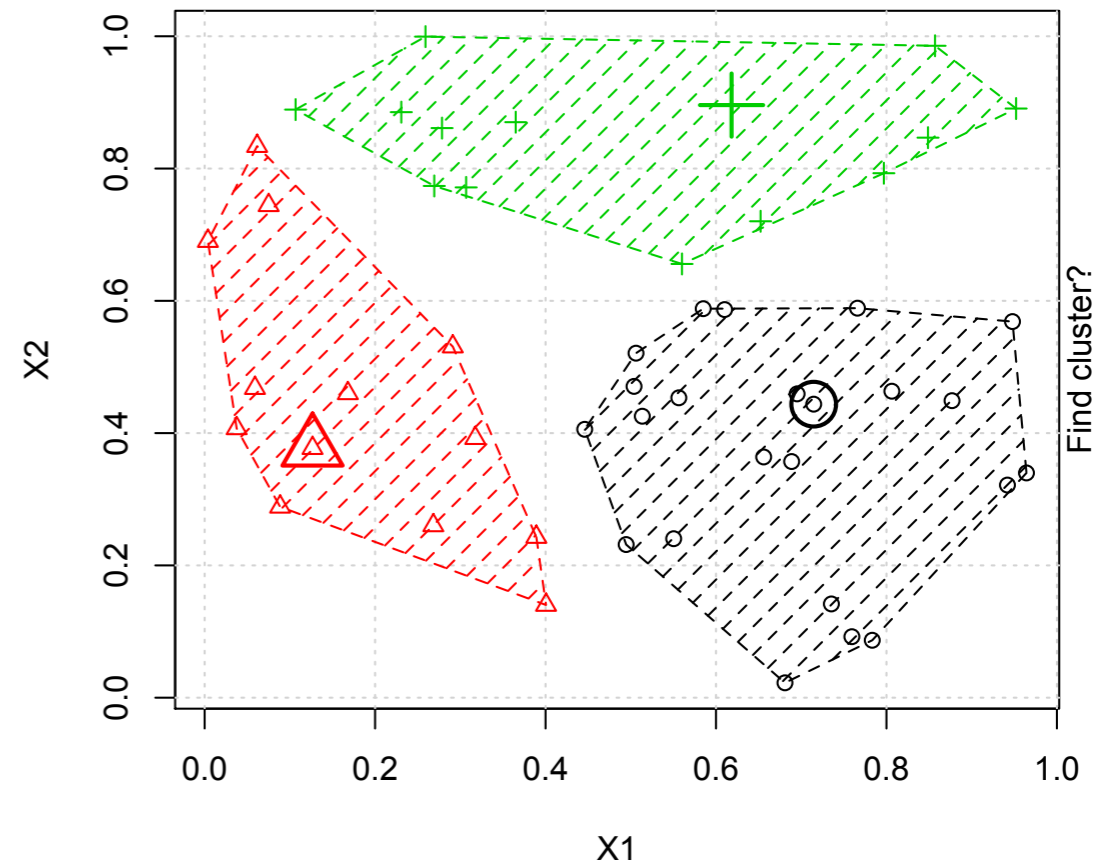
Iteration 2

Random centers



Euclidean

First class assignment

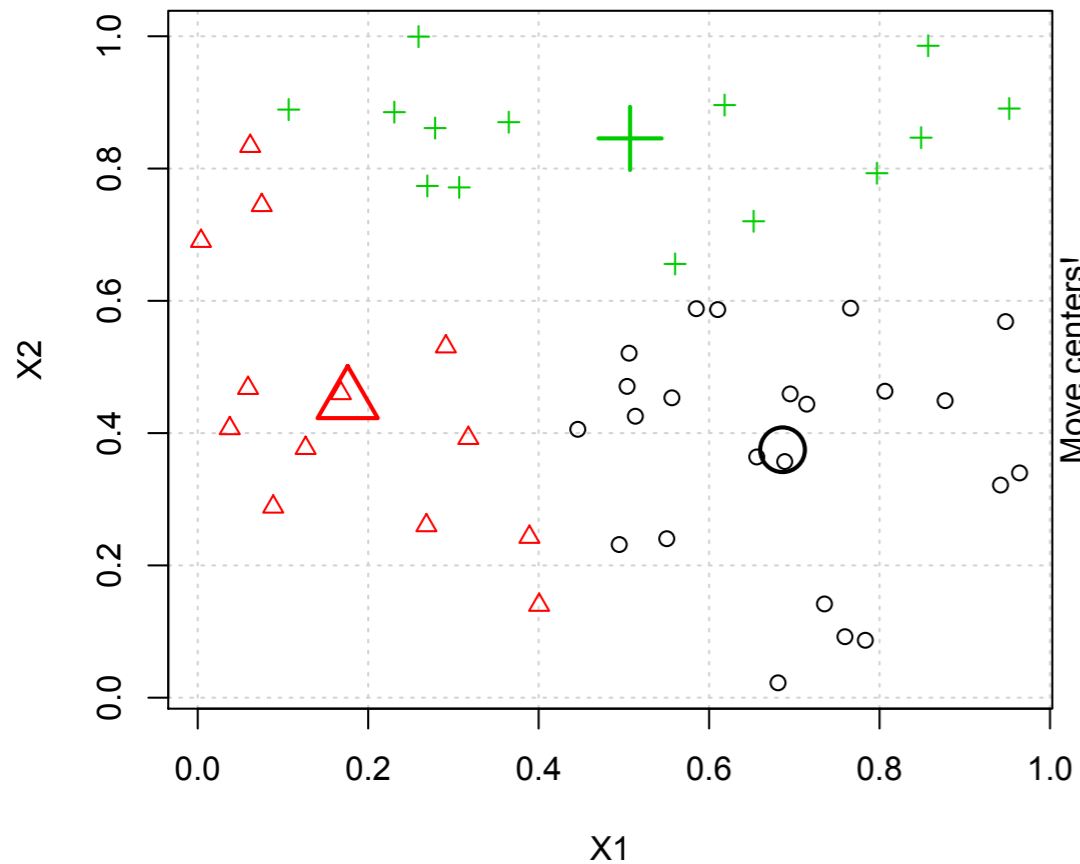


Unsupervised learning



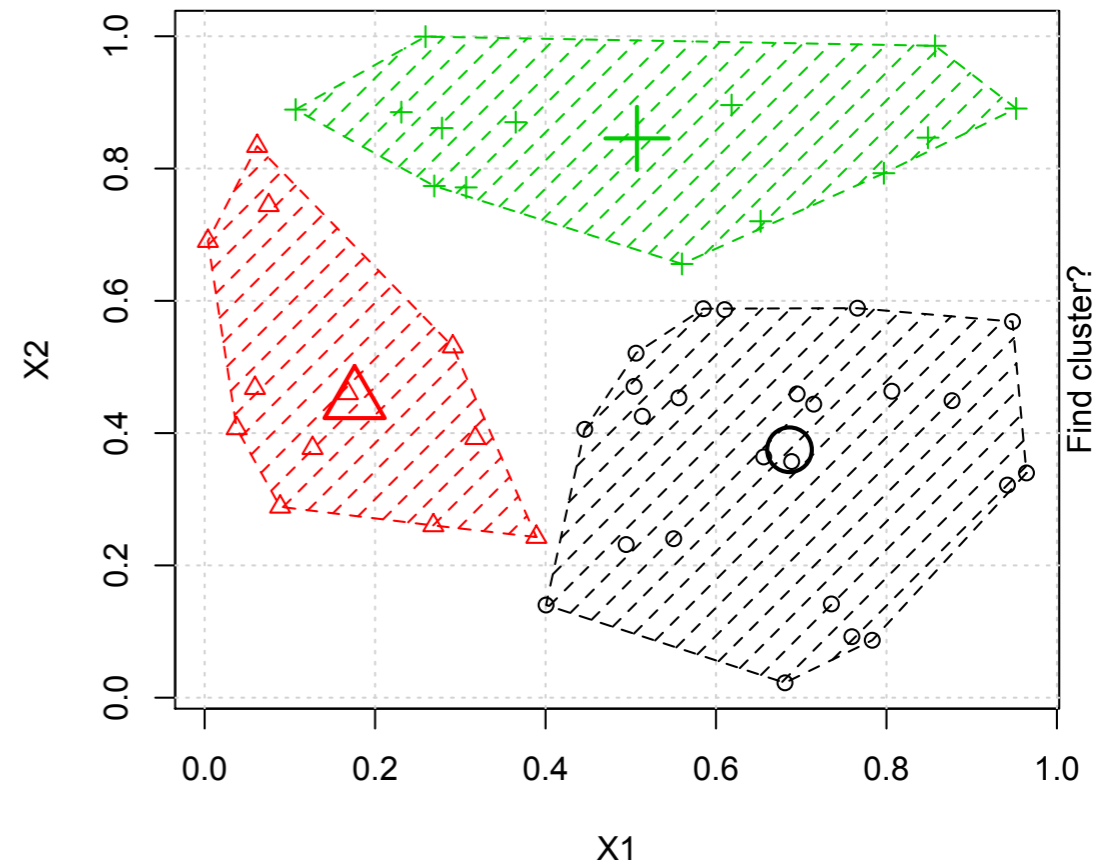
Iteration 3

New Centers



Euclidean

Class assignment

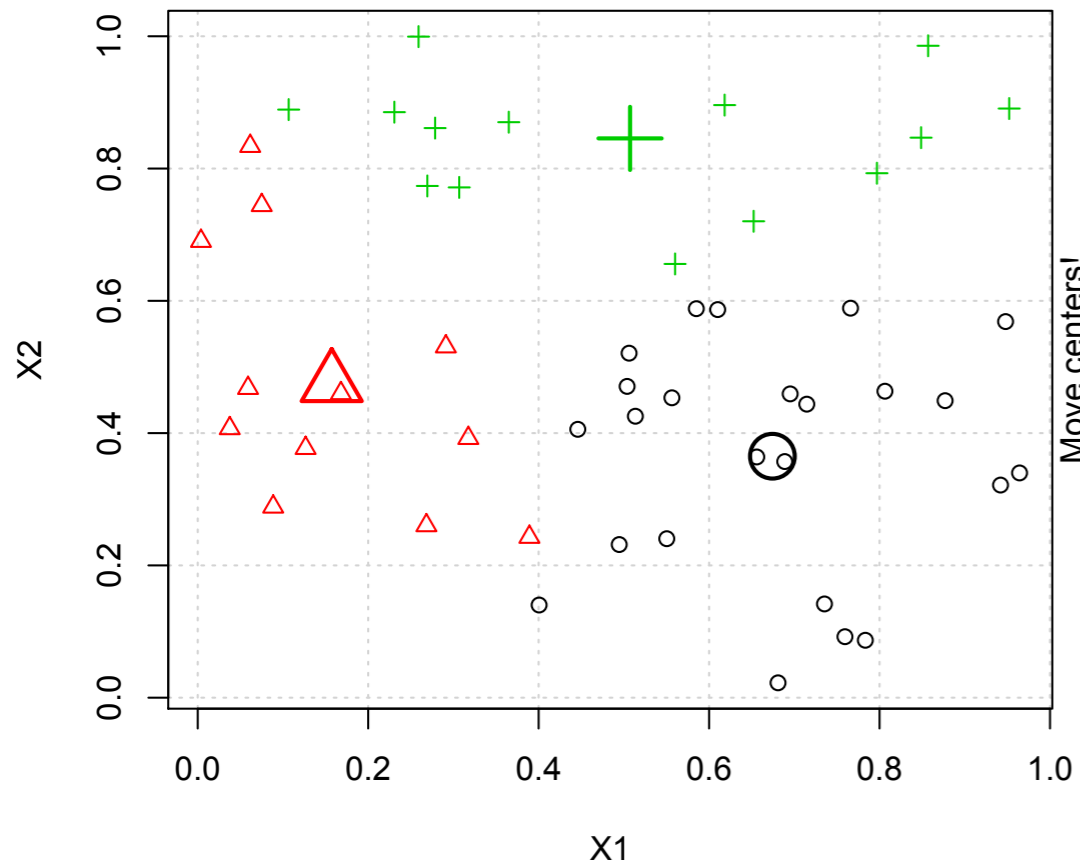


Unsupervised learning



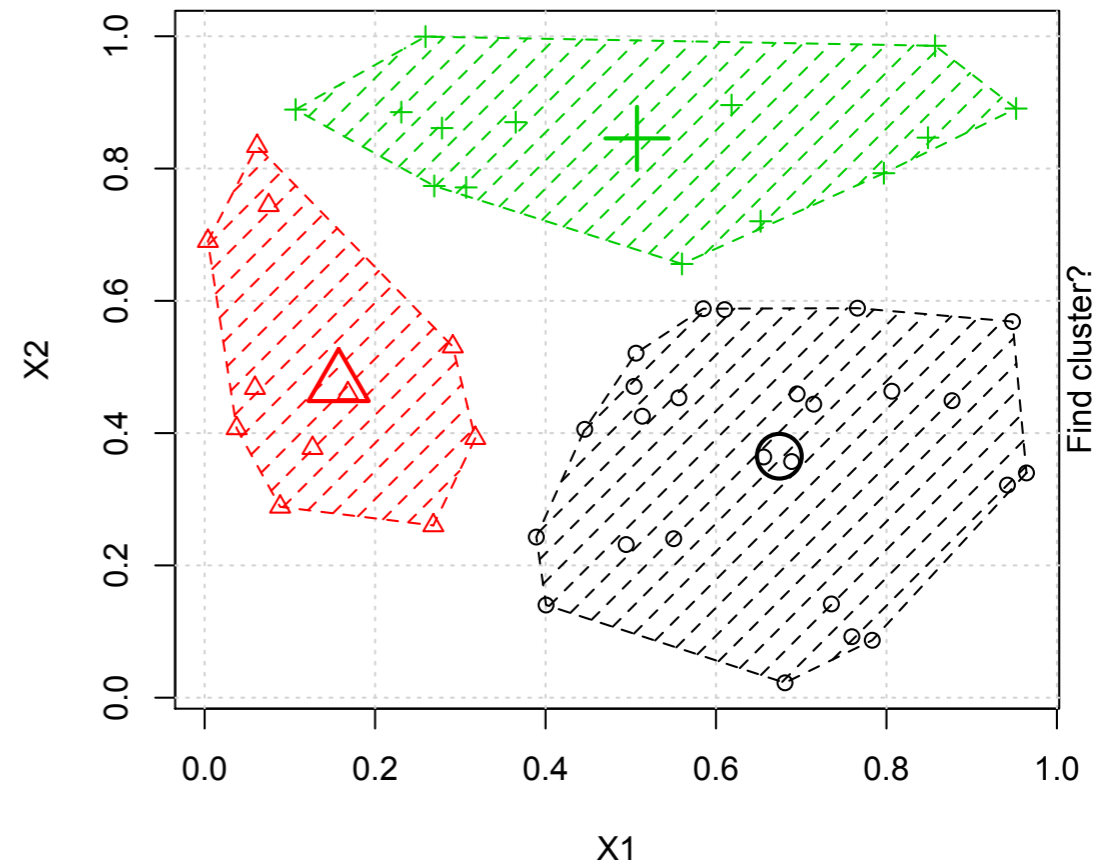
Iteration 4

New Centers



Euclidean

Class assignment

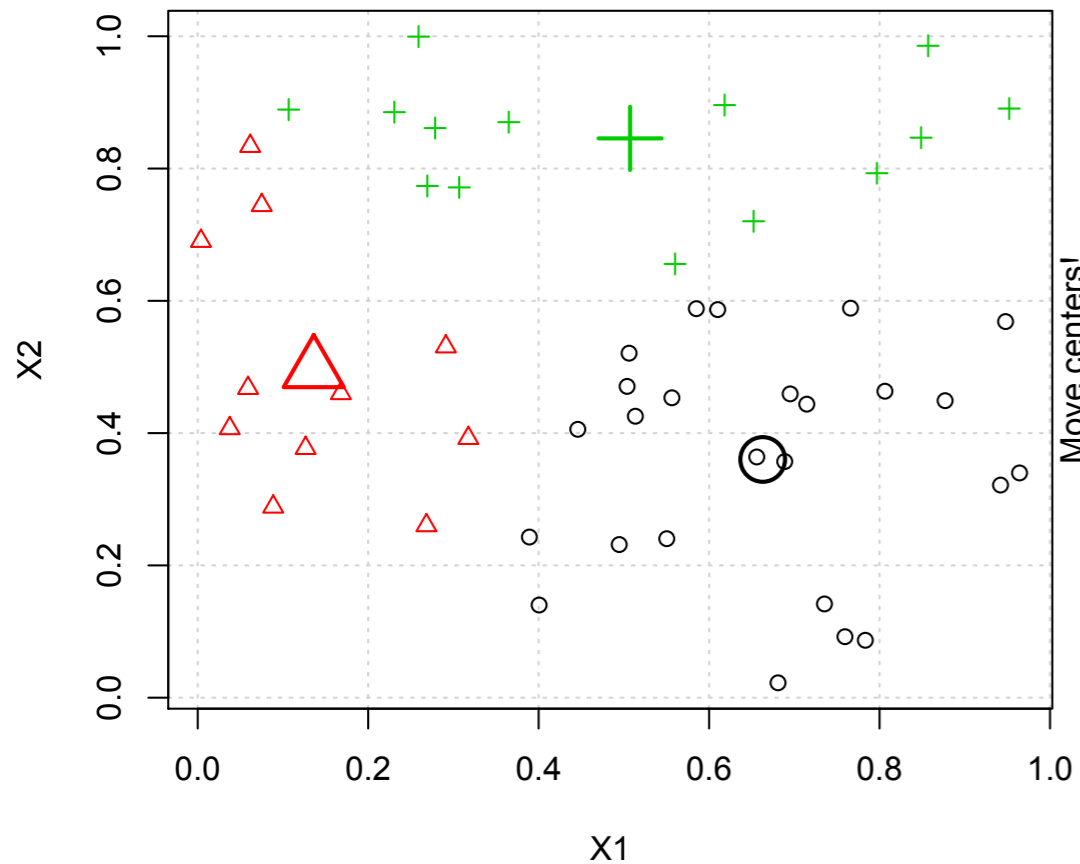


Unsupervised learning



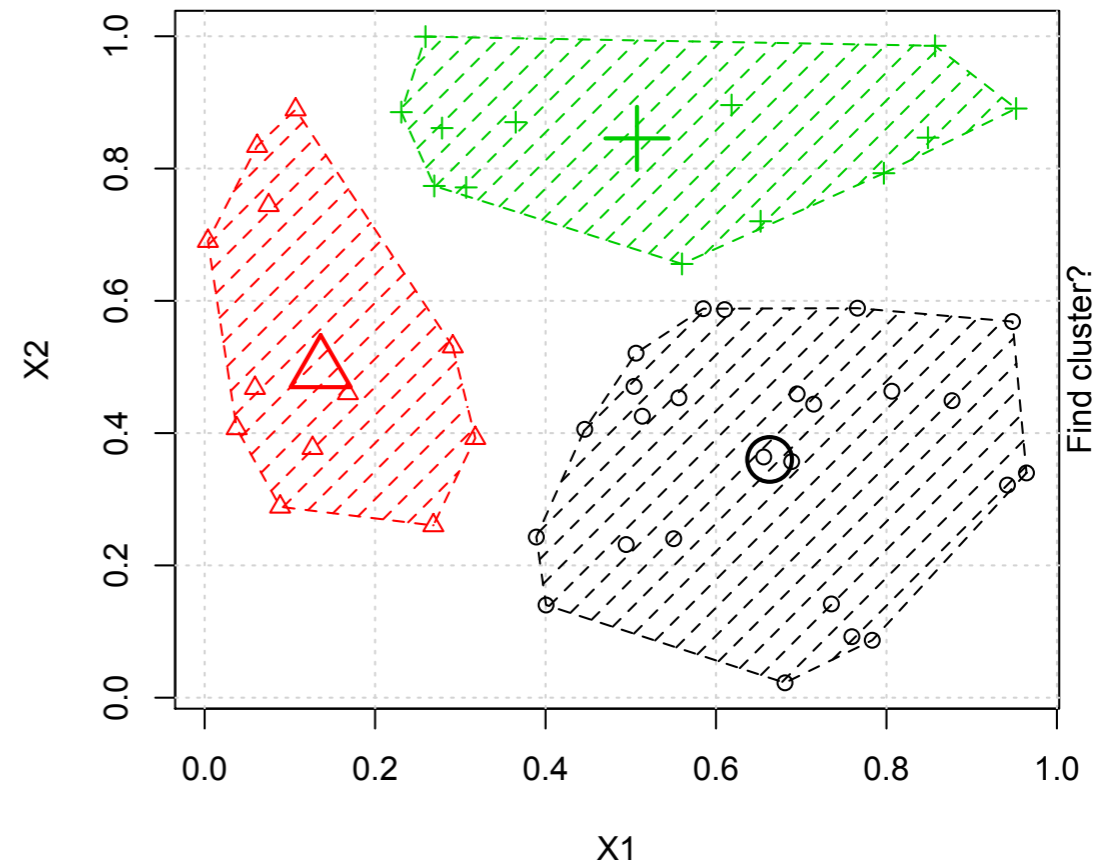
Iteration 5

New Centers



Euclidean

Class assignment

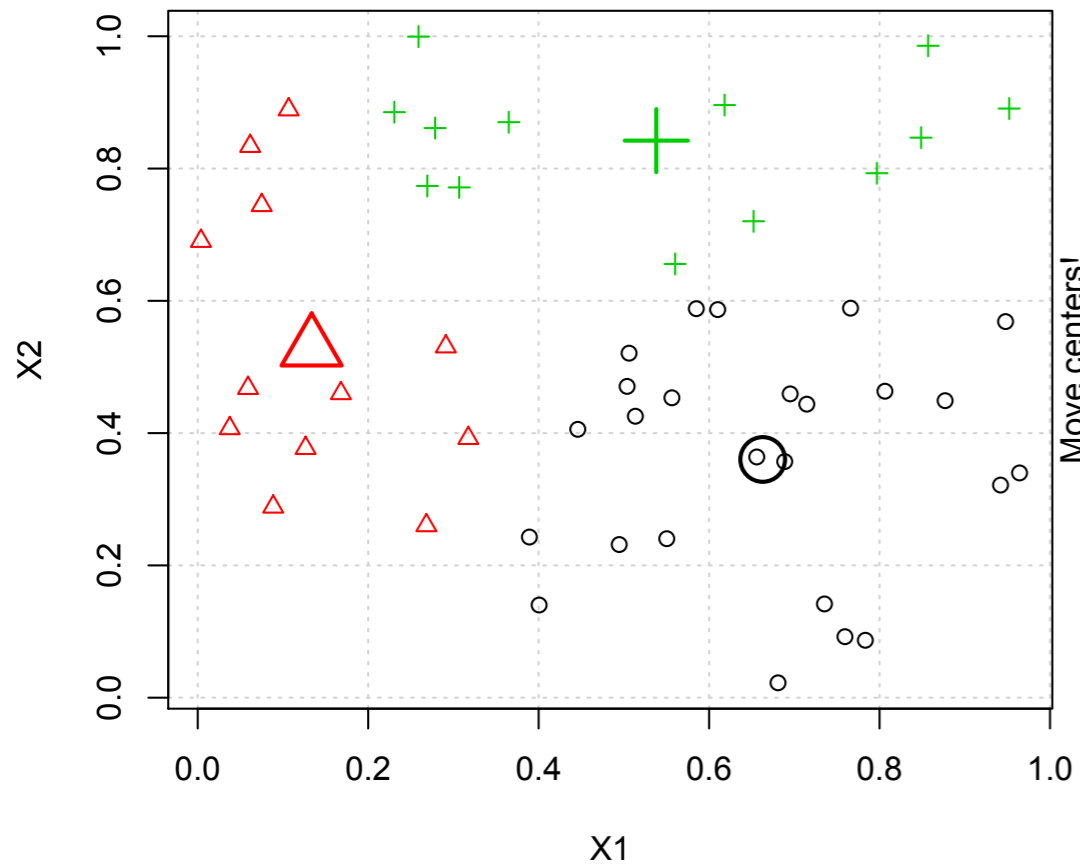


Unsupervised learning



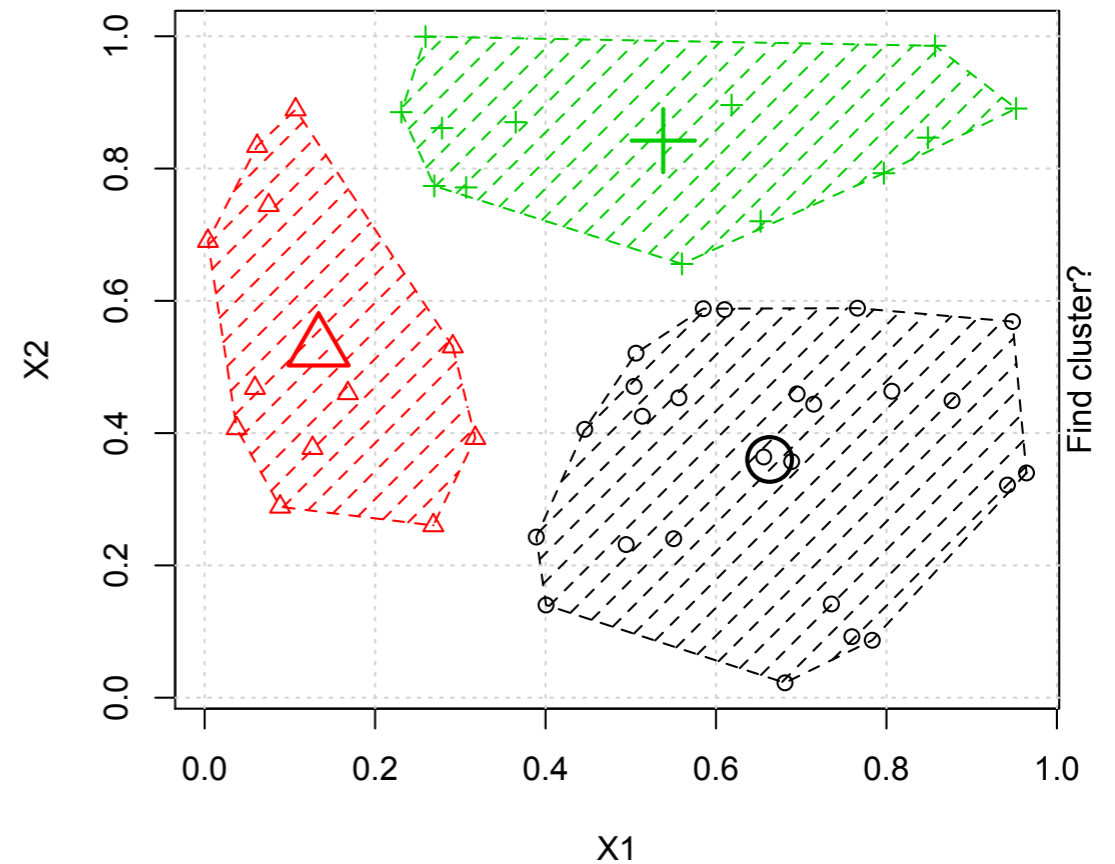
Final iteration

New Centers

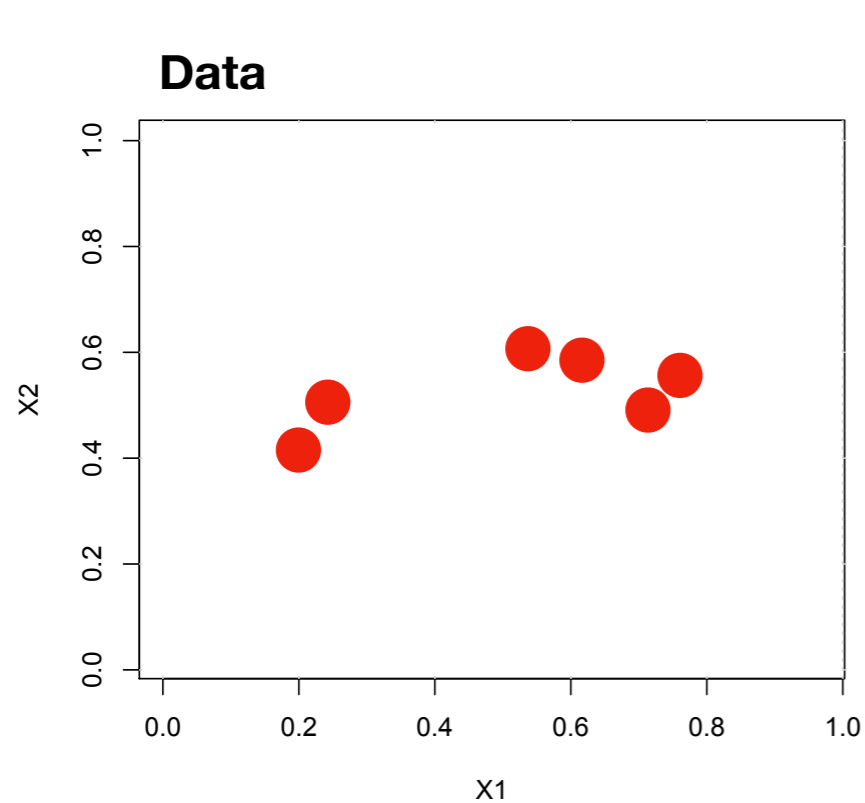
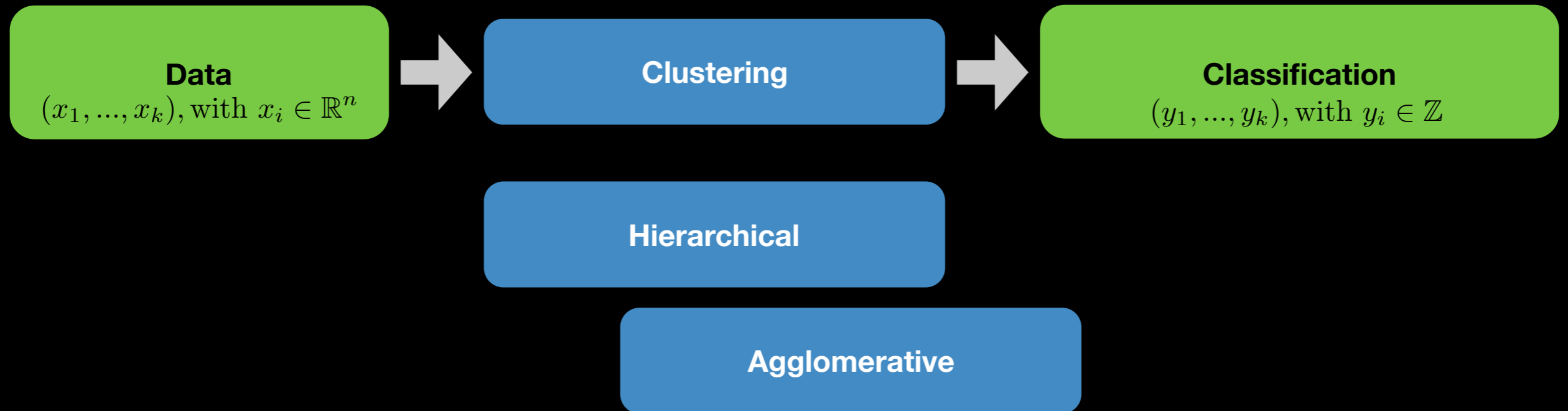


Euclidean

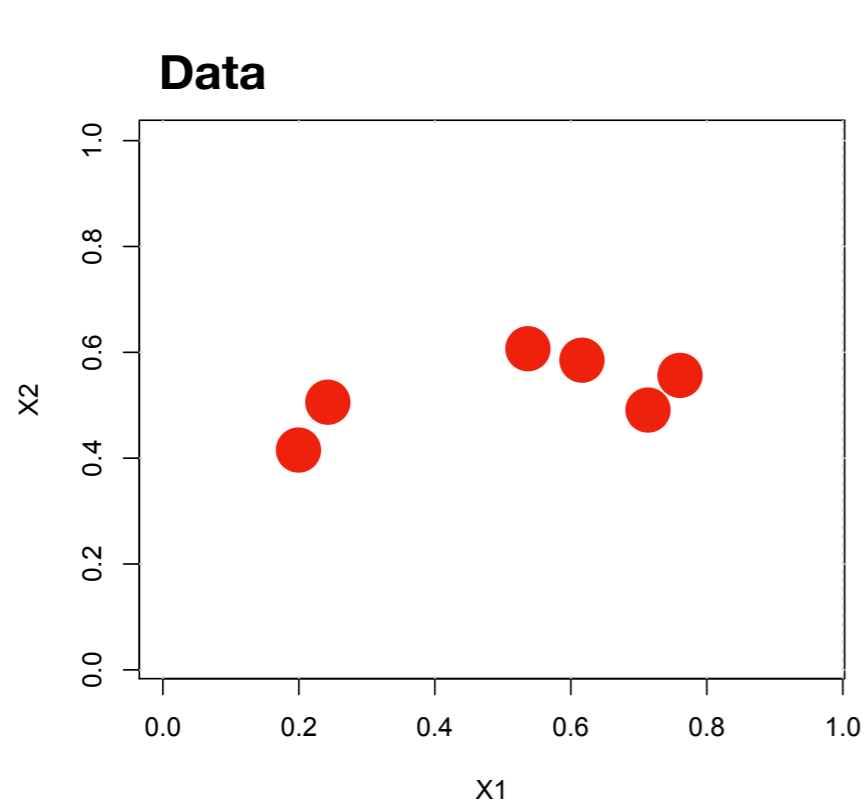
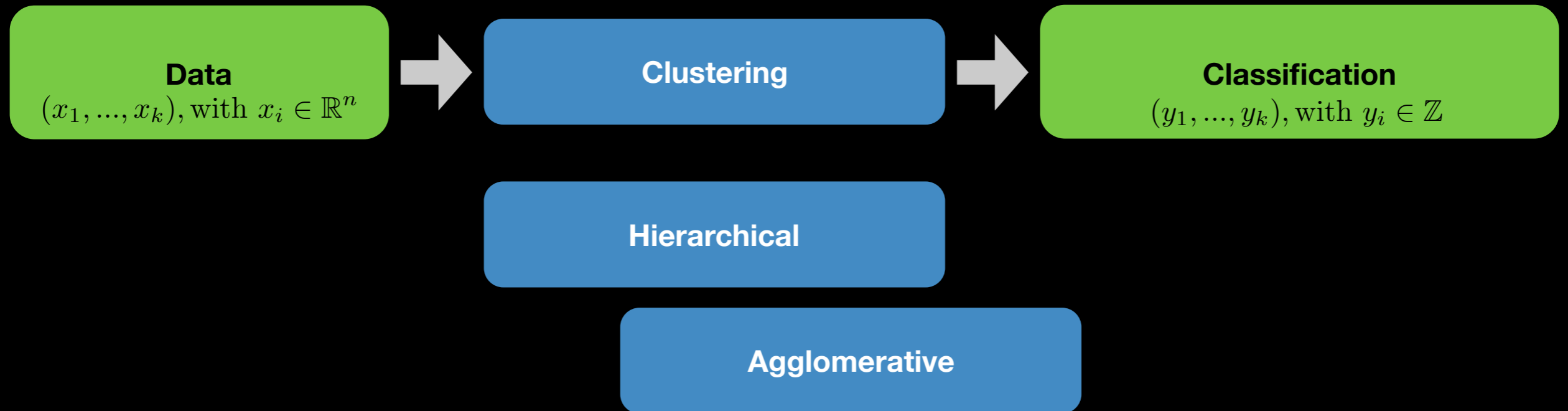
Class assignment



Unsupervised learning

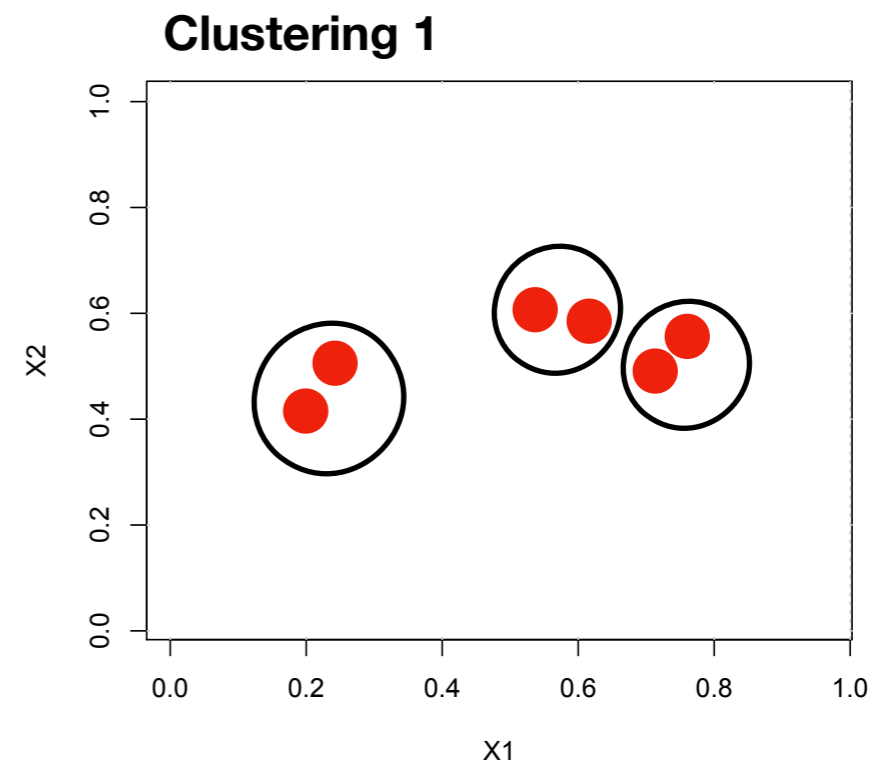


Unsupervised learning

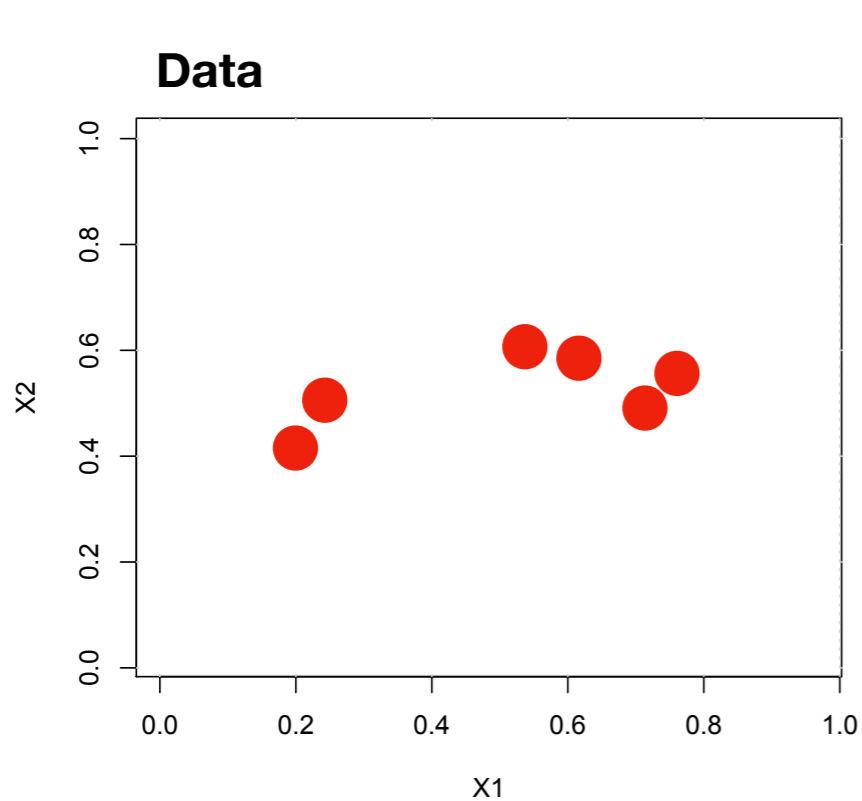
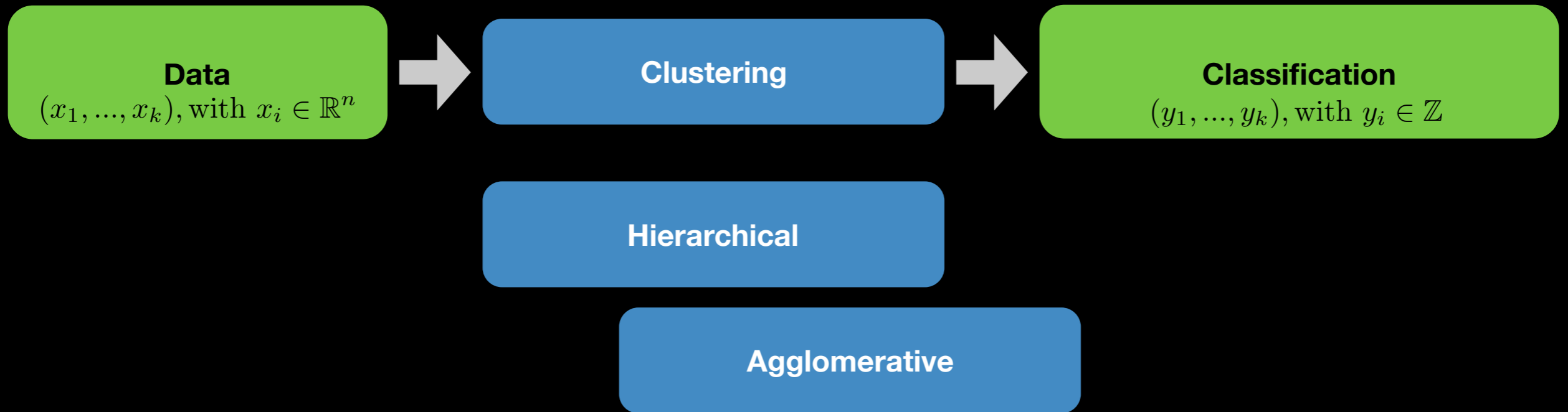


Iteration 1

Euclidean

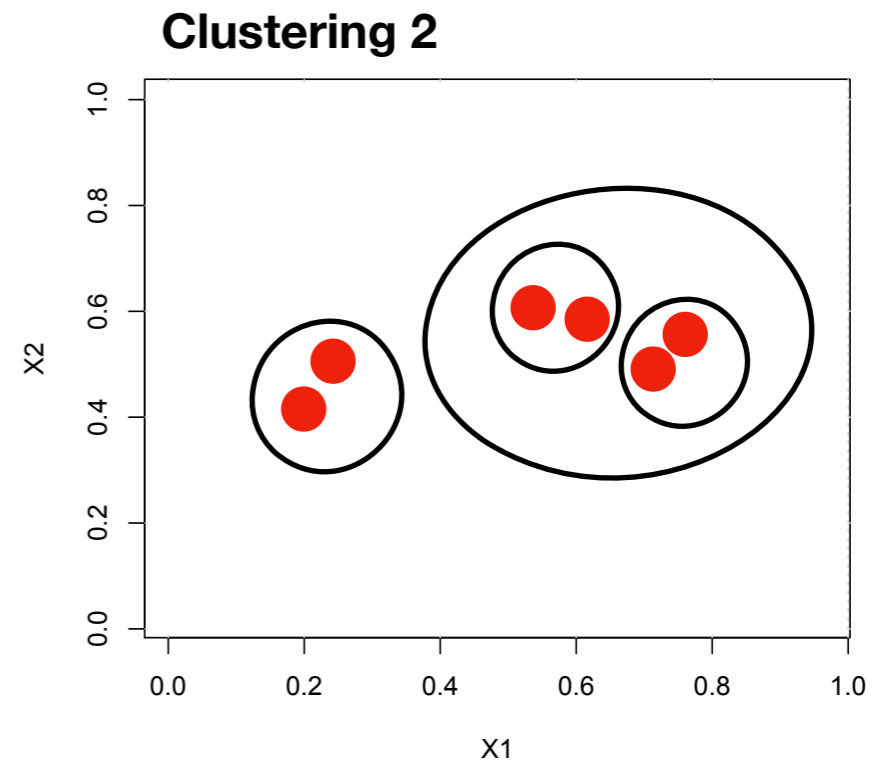


Unsupervised learning

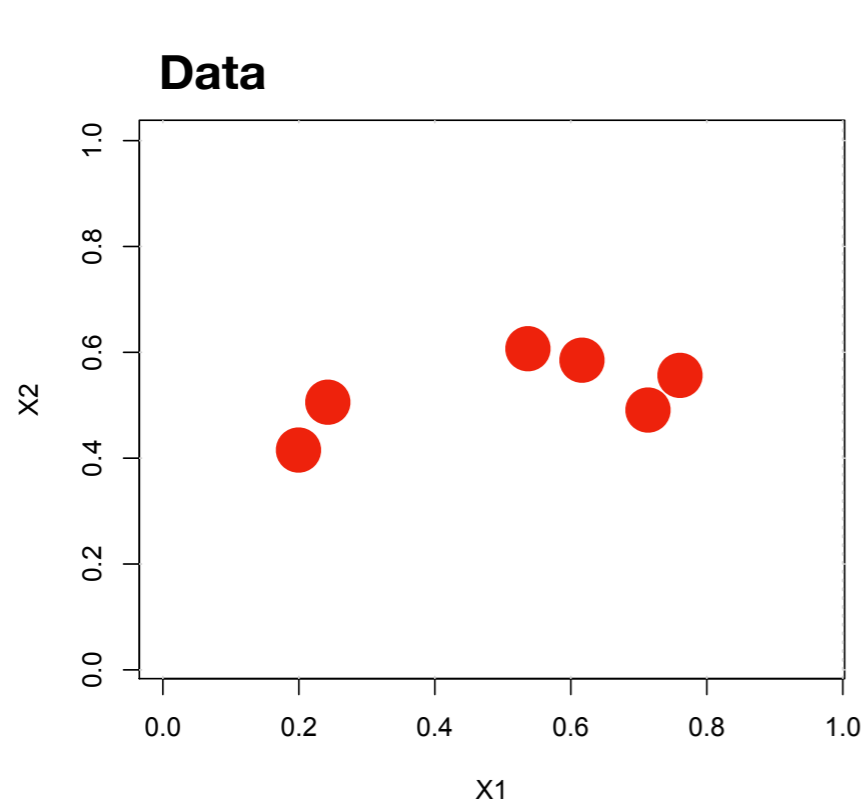
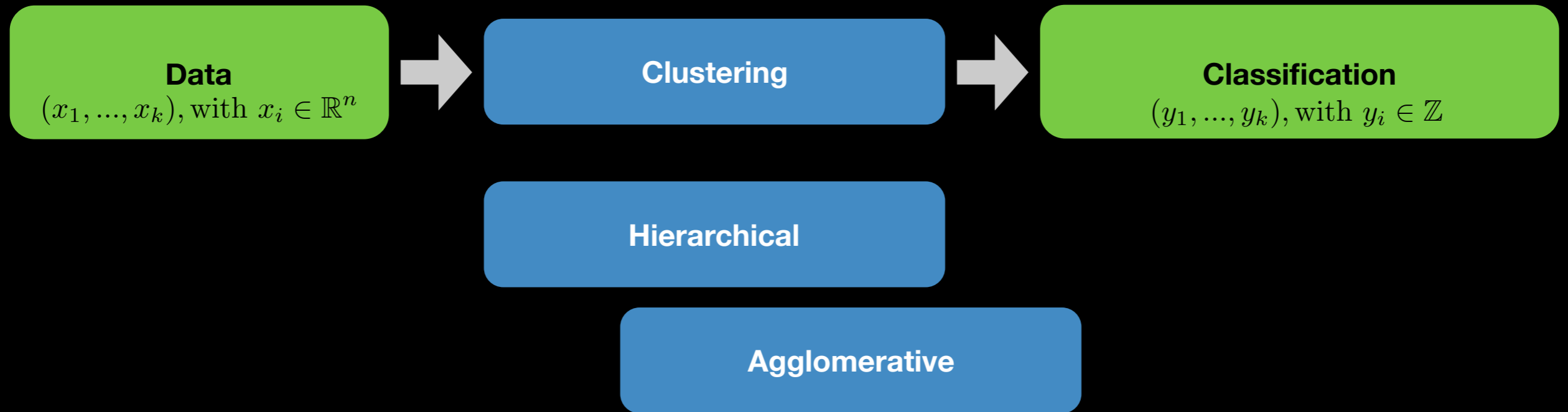


Iteration 2

Euclidean

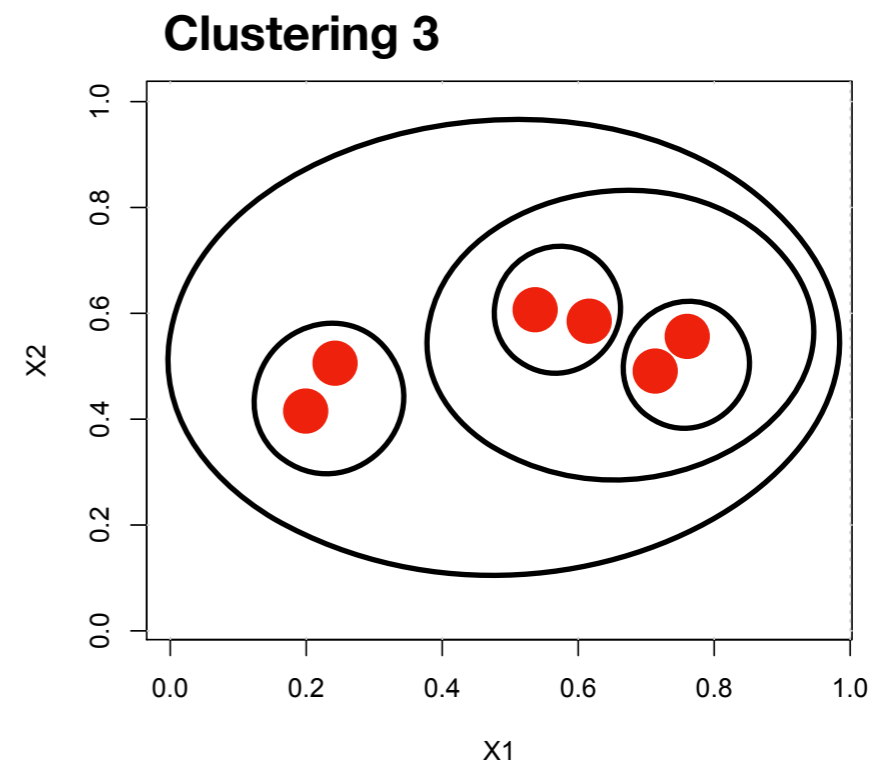


Unsupervised learning

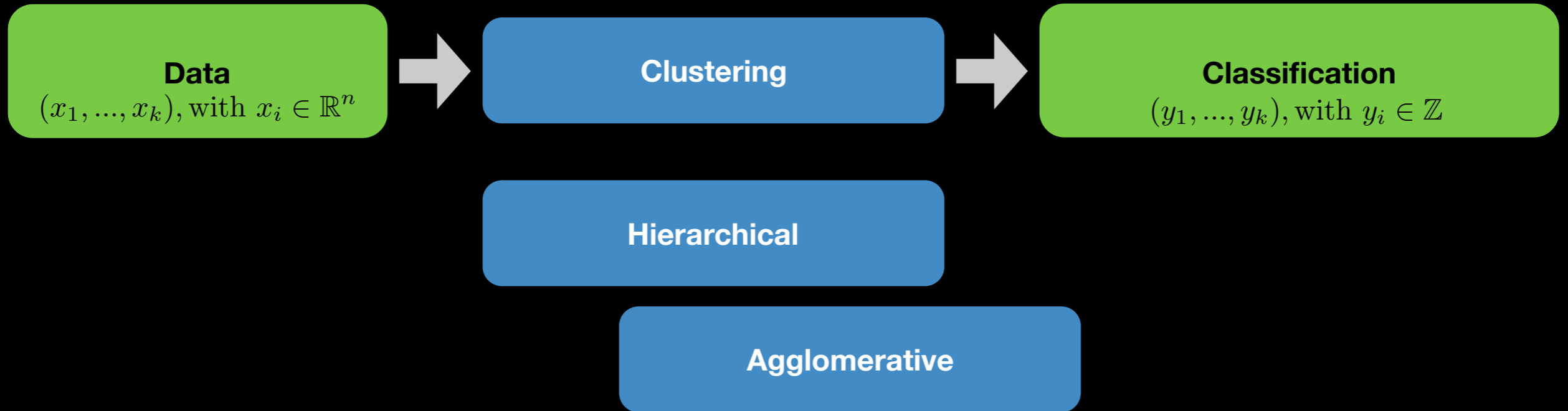


Iteration 3

Euclidean

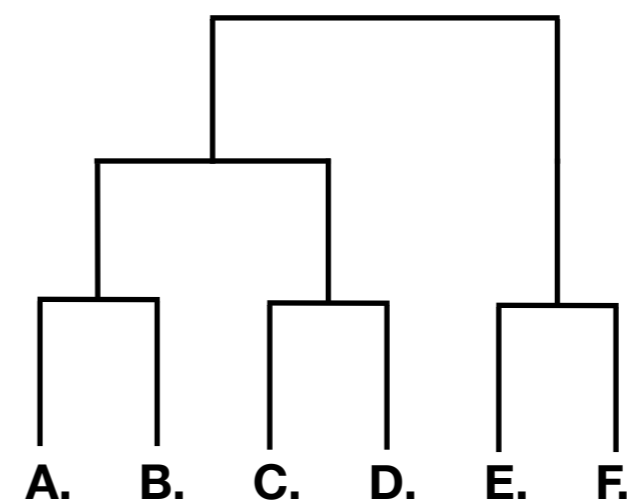
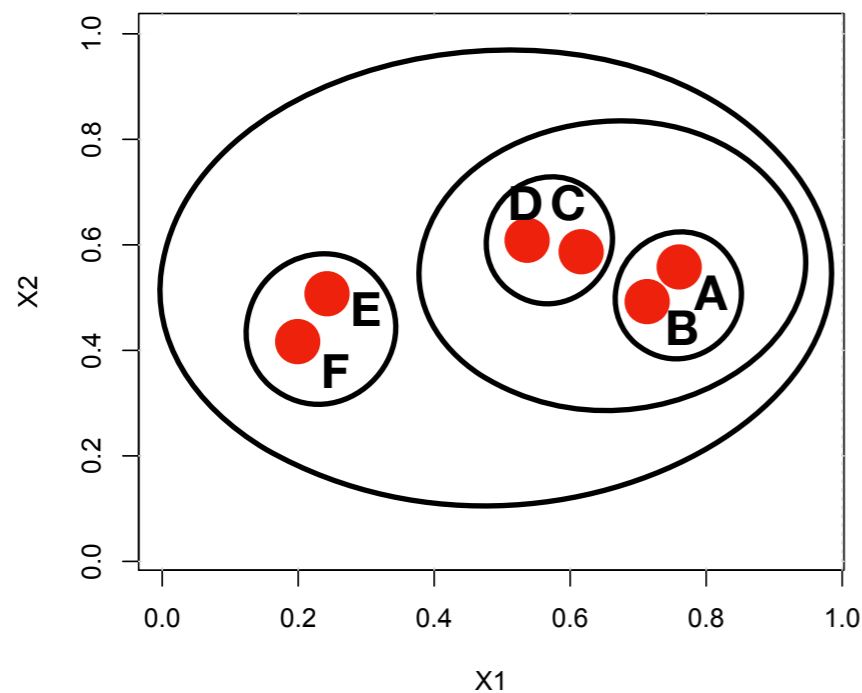


Unsupervised learning

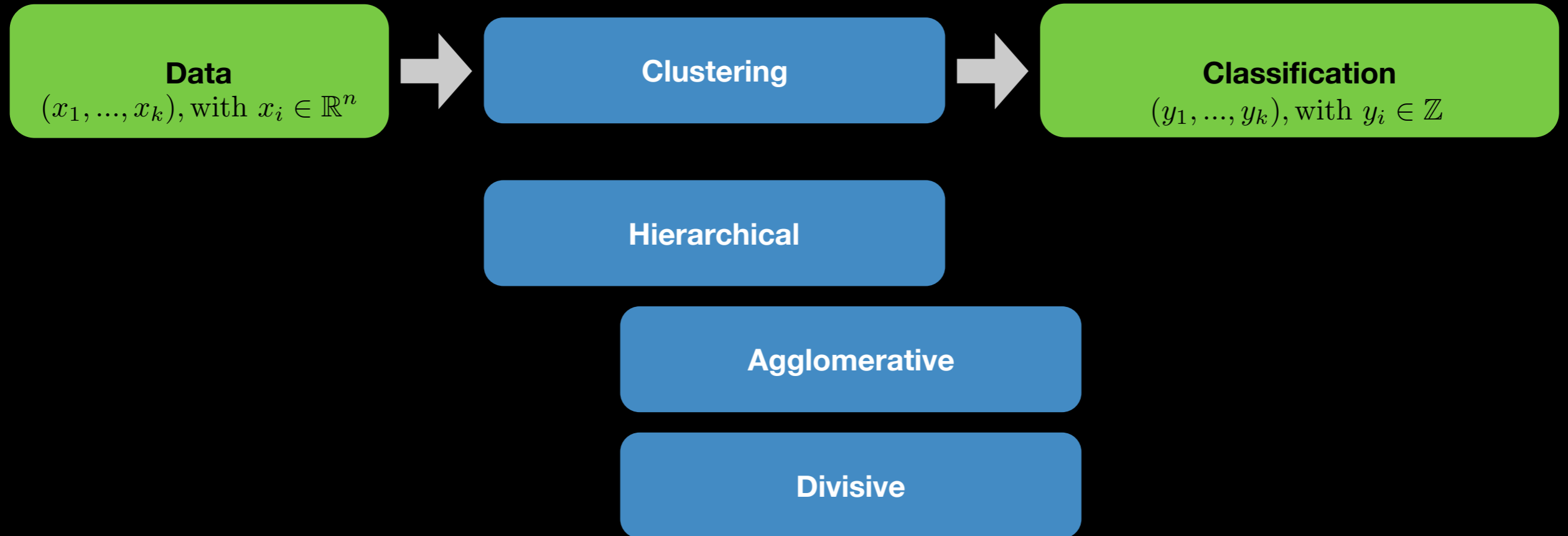


Representation : Dendrograms

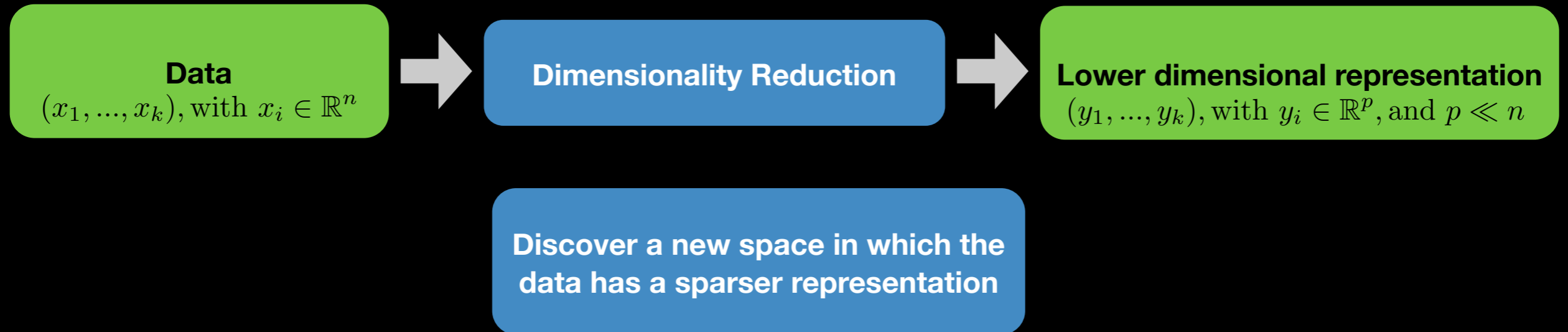
Clustering 3



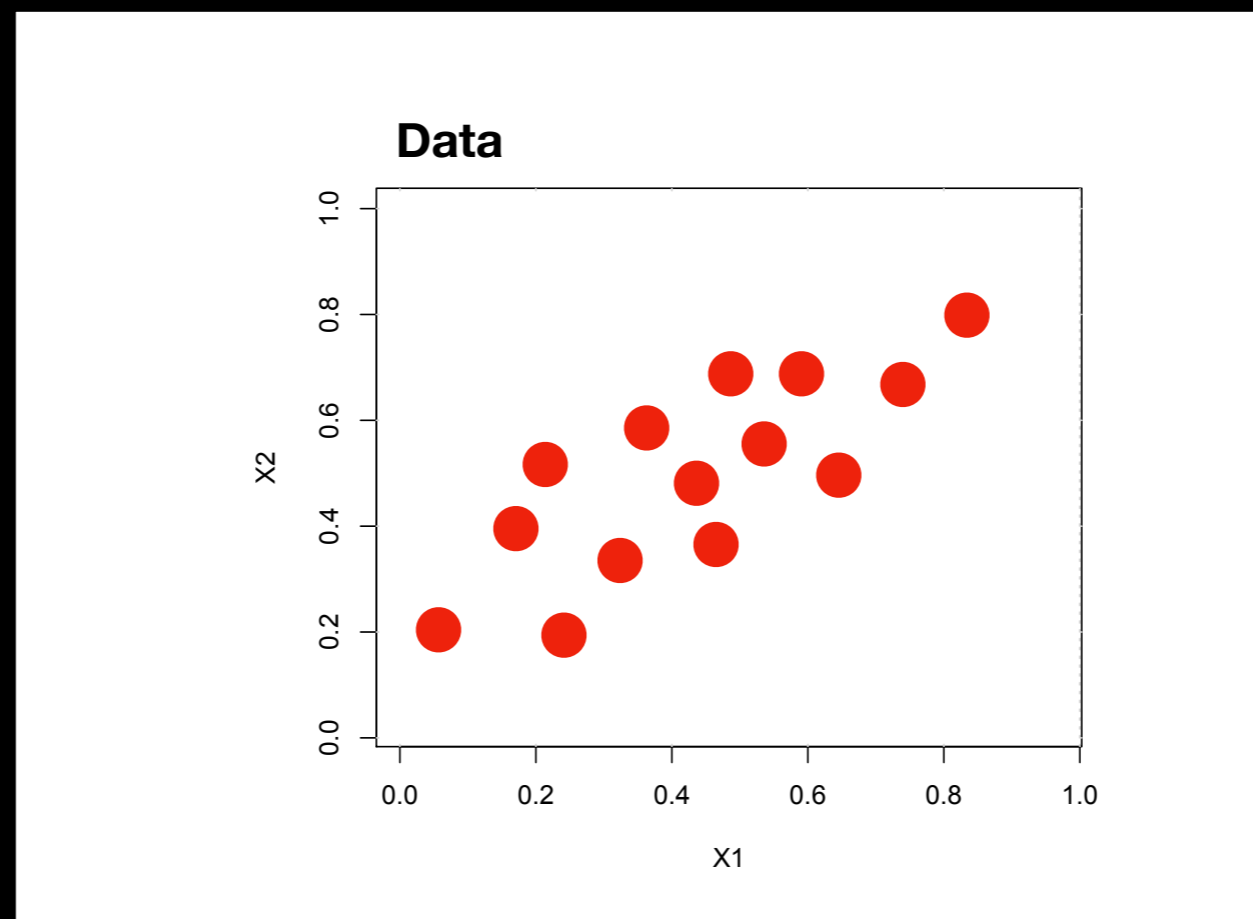
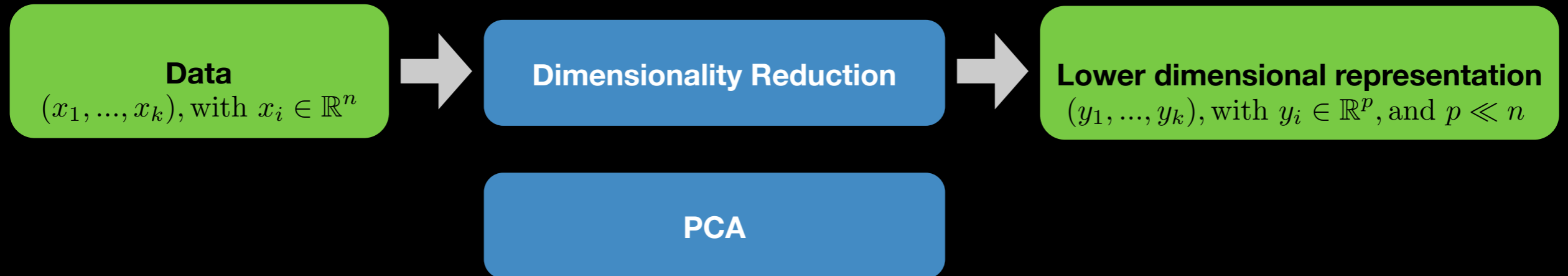
Unsupervised learning



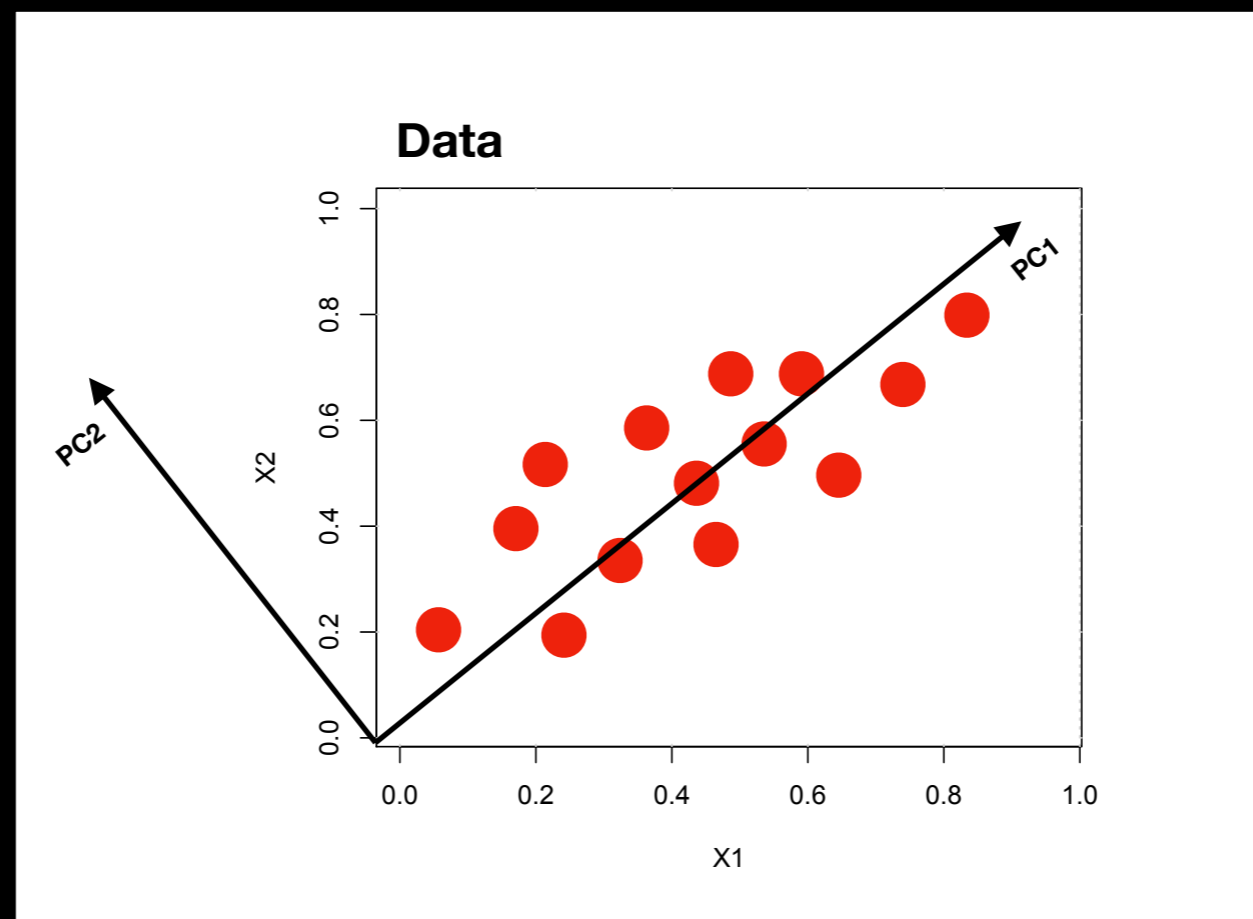
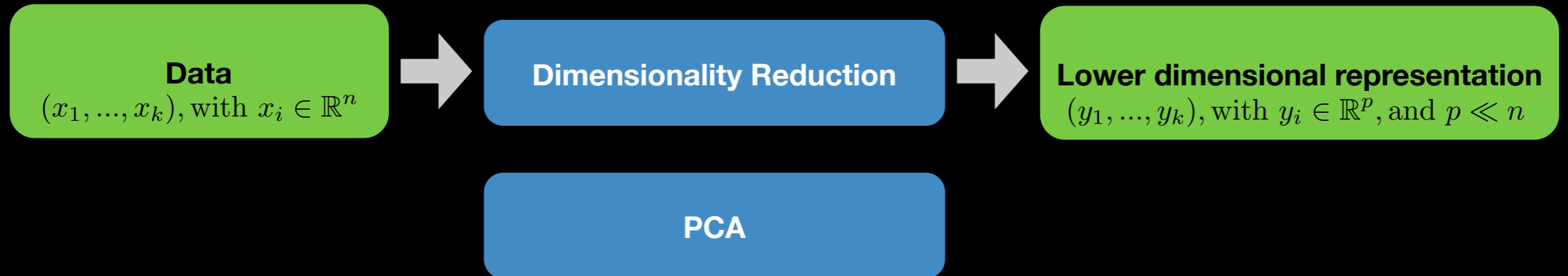
Unsupervised learning



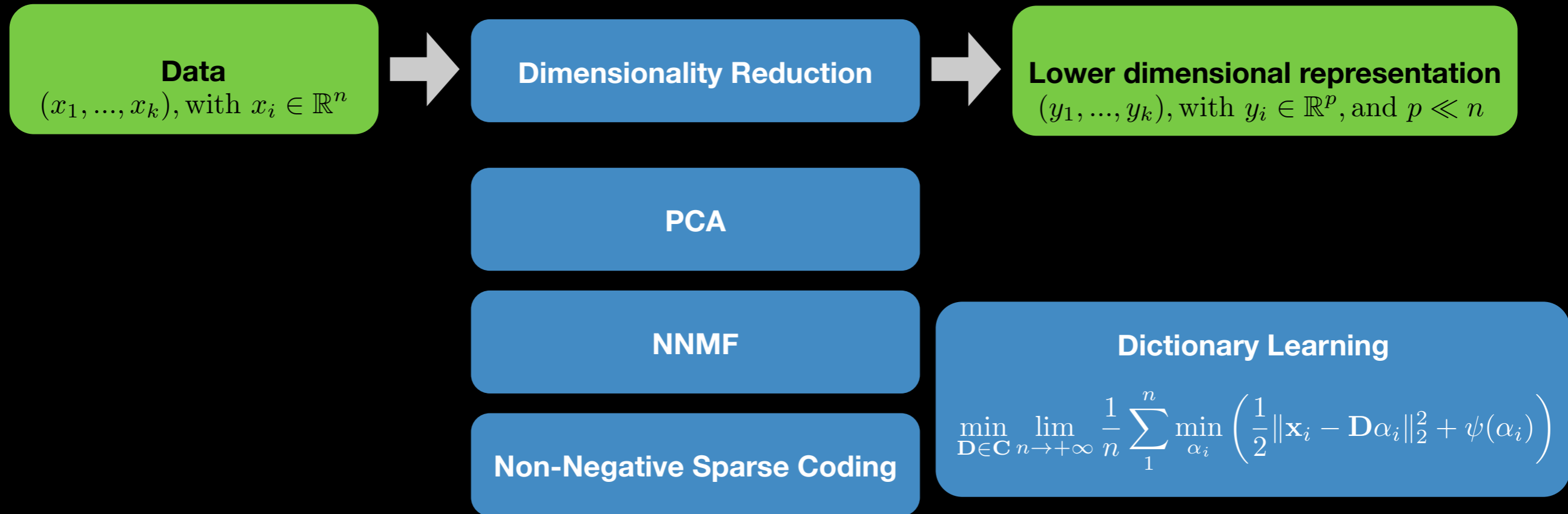
Unsupervised learning



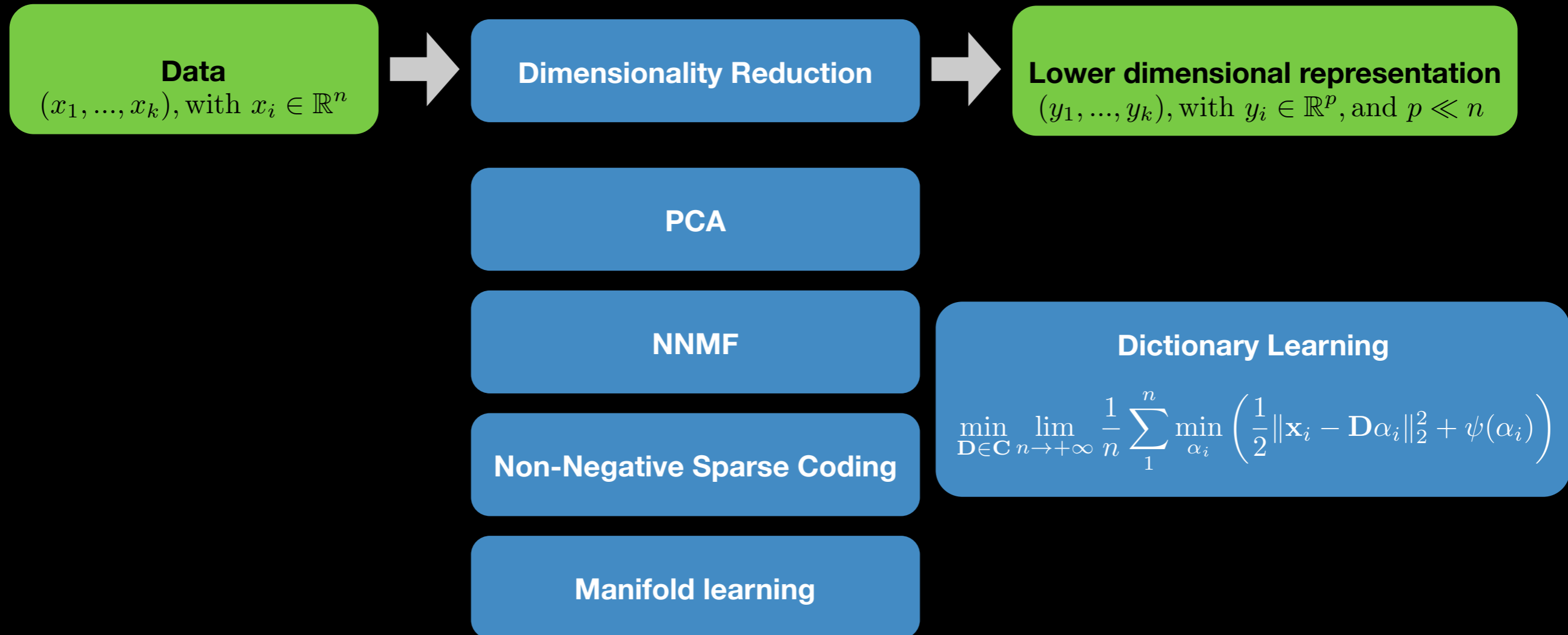
Unsupervised learning



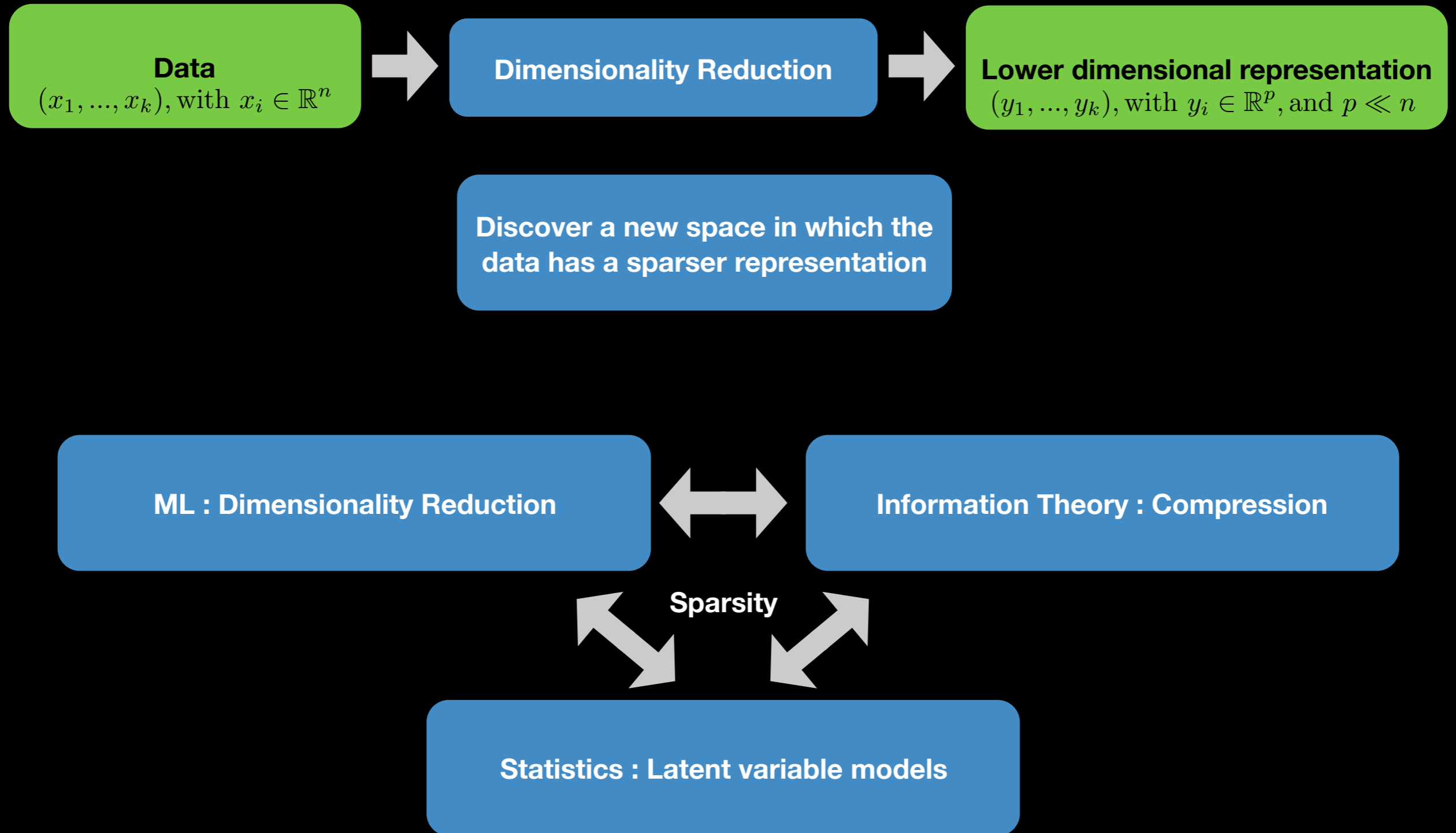
Unsupervised learning



Unsupervised learning



Unsupervised learning

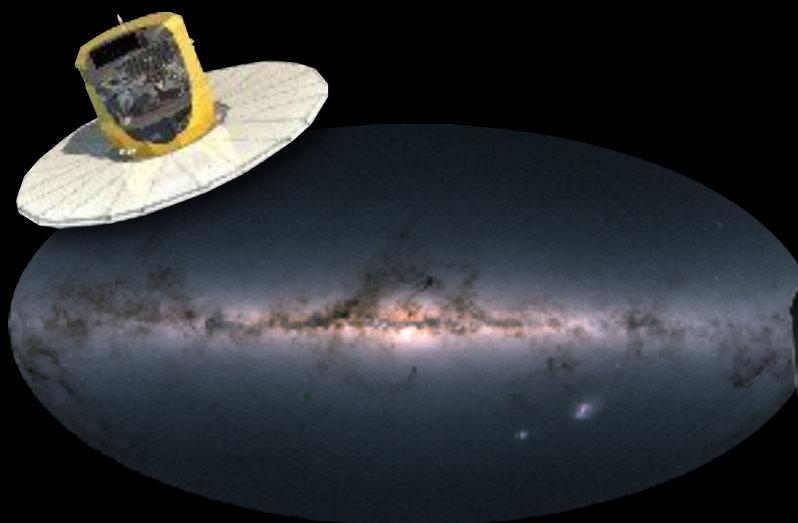


In this talk

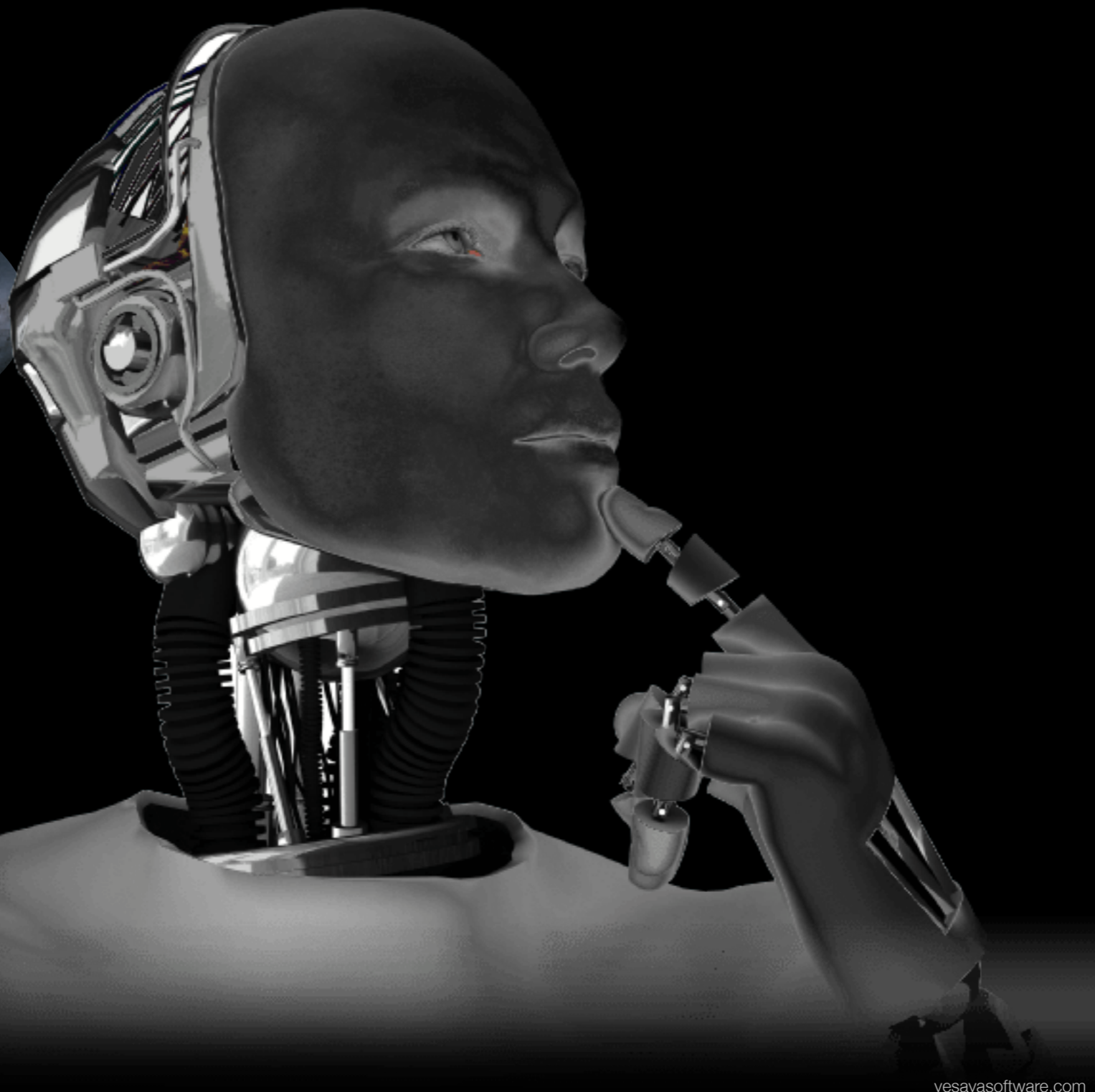
Unsupervised Learning

Three astronomical examples

Unsupervised learning : Stellar Clusters

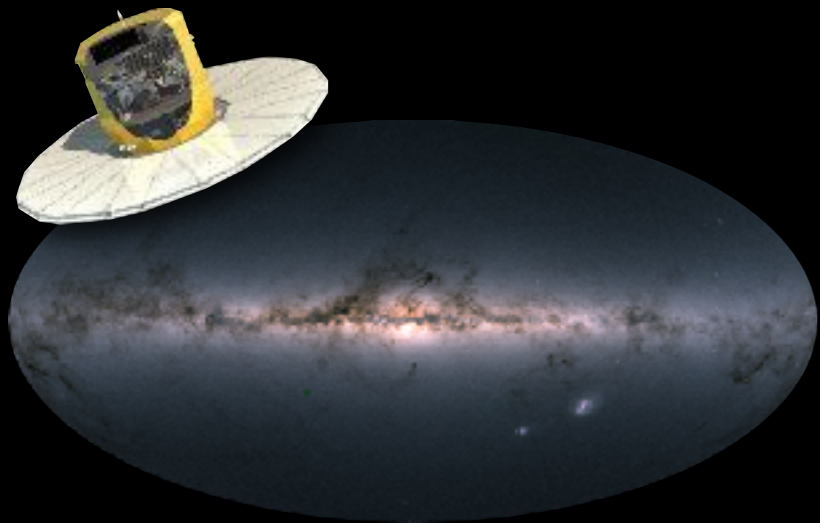


ESA/DPAC/U.Lisboa/CENTRA

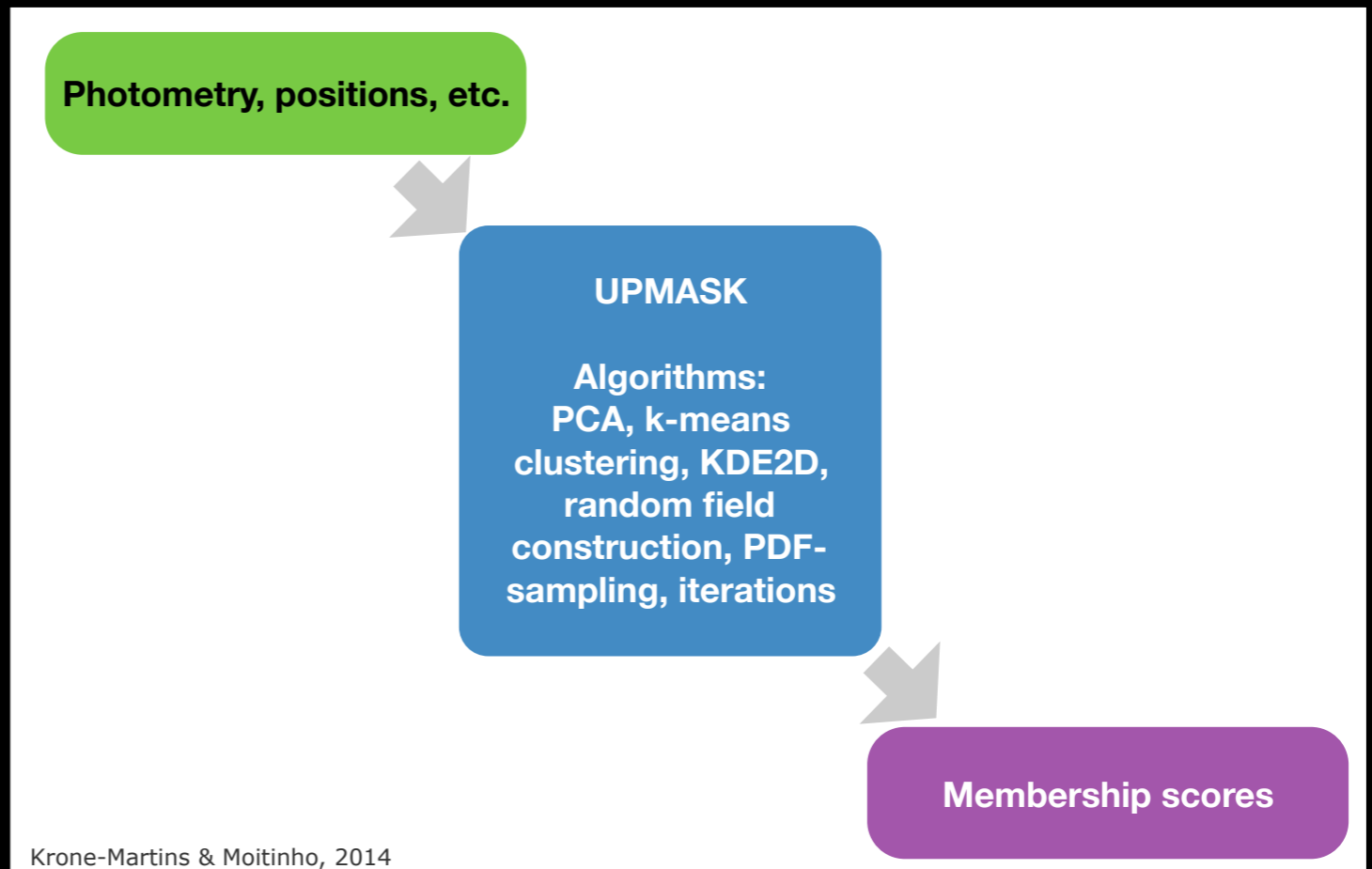


yesayasoftware.com

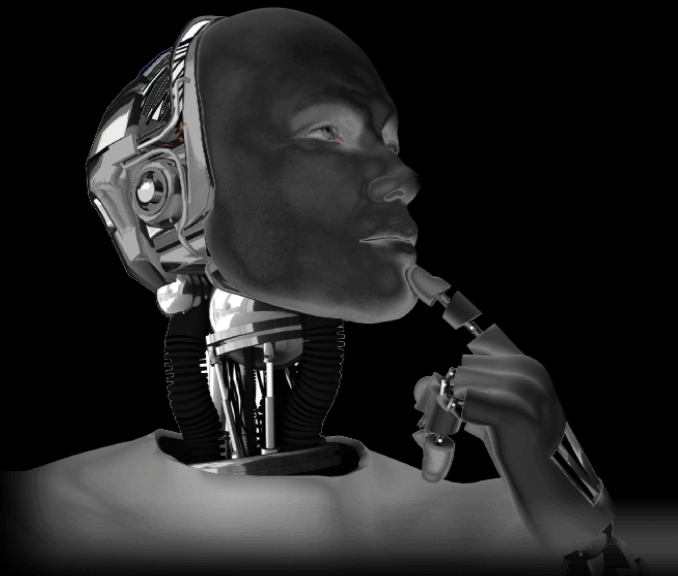
Unsupervised learning : Stellar Clusters



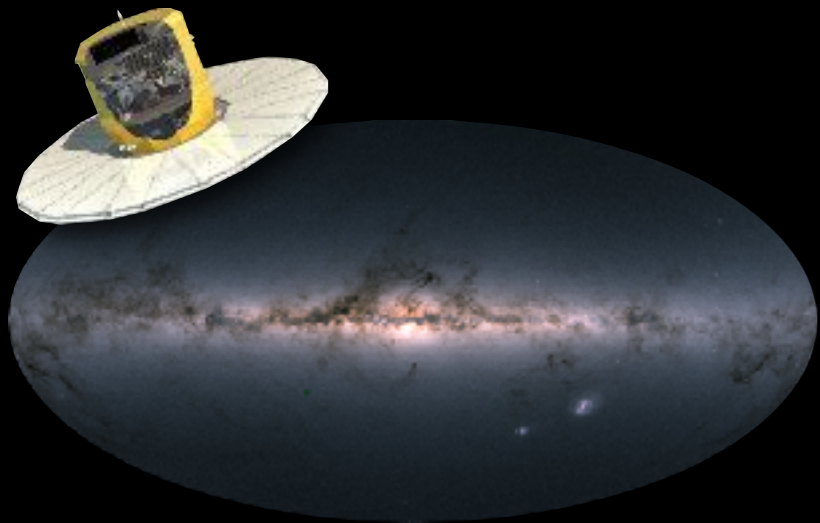
ESA/DPAC/U.Lisboa/CENTRA



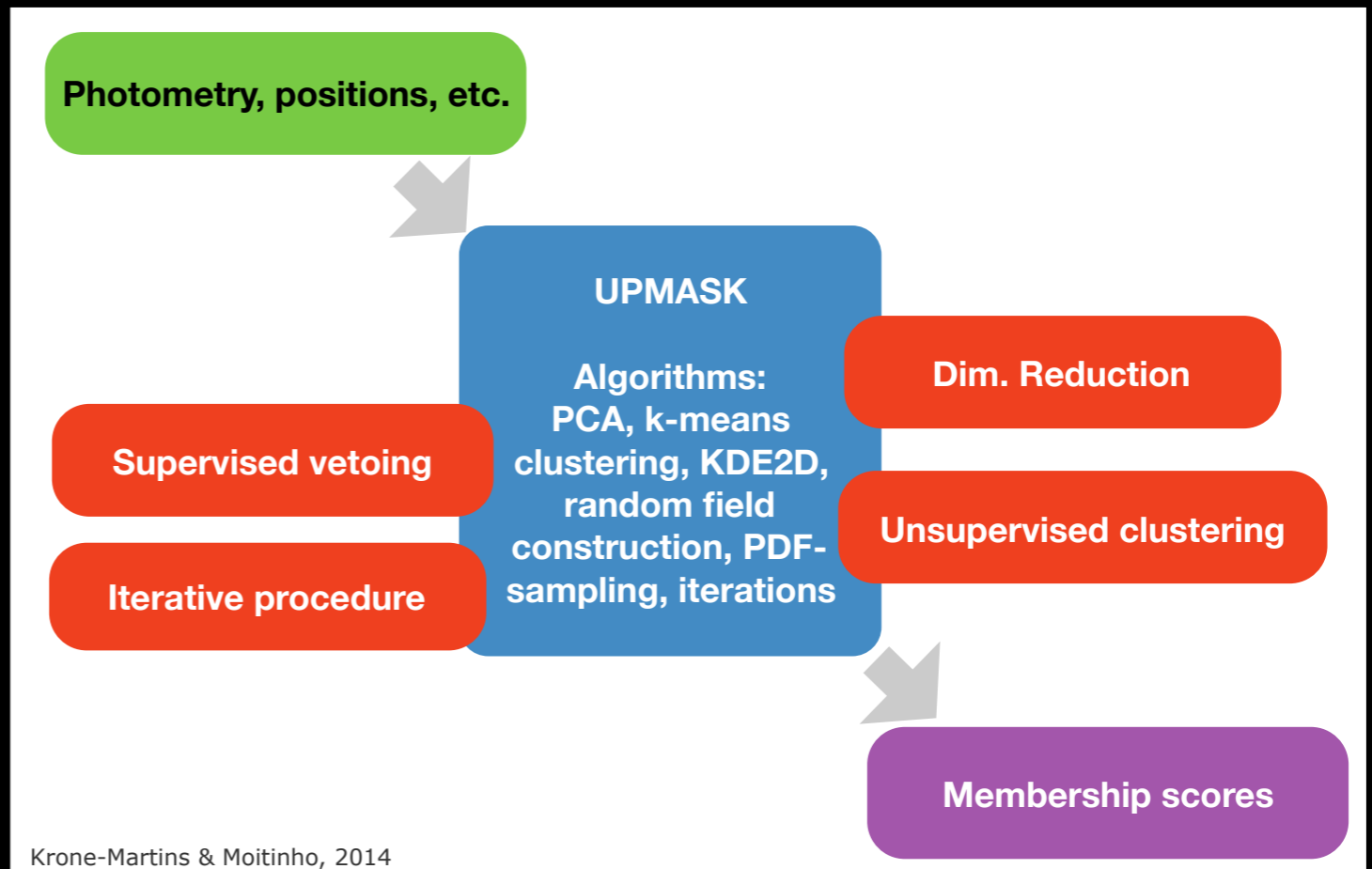
Krone-Martins & Moitinho, 2014



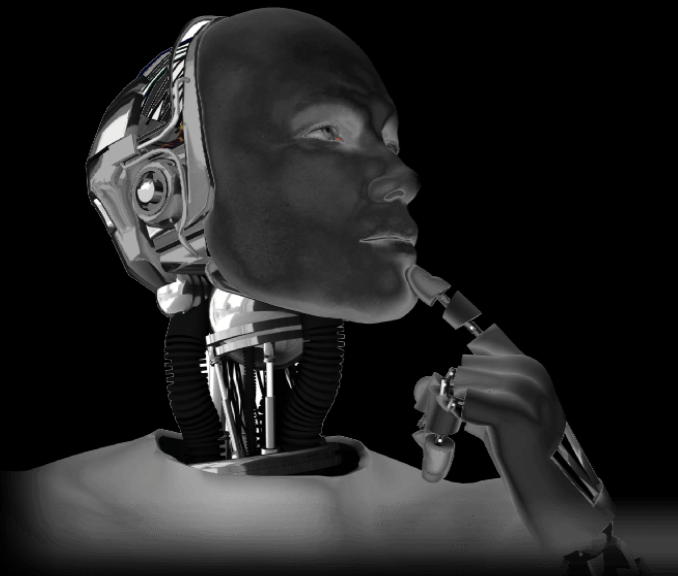
Unsupervised learning : Stellar Clusters



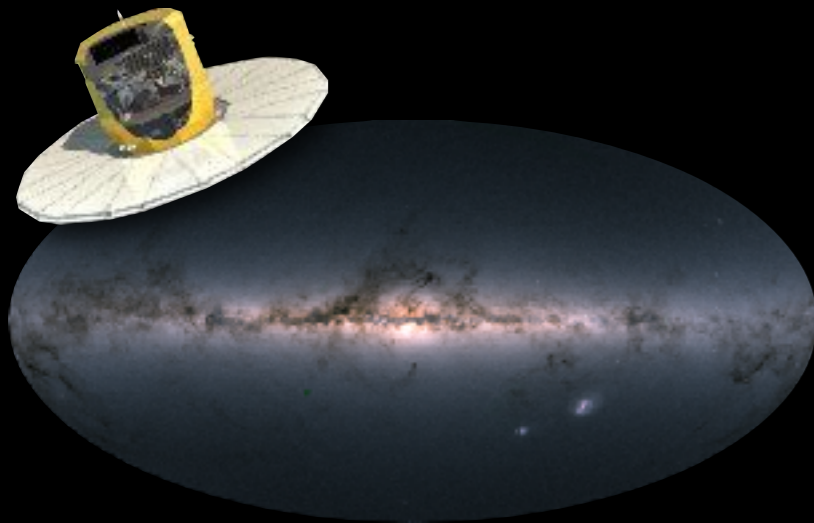
ESA/DPAC/U.Lisboa/CENTRA



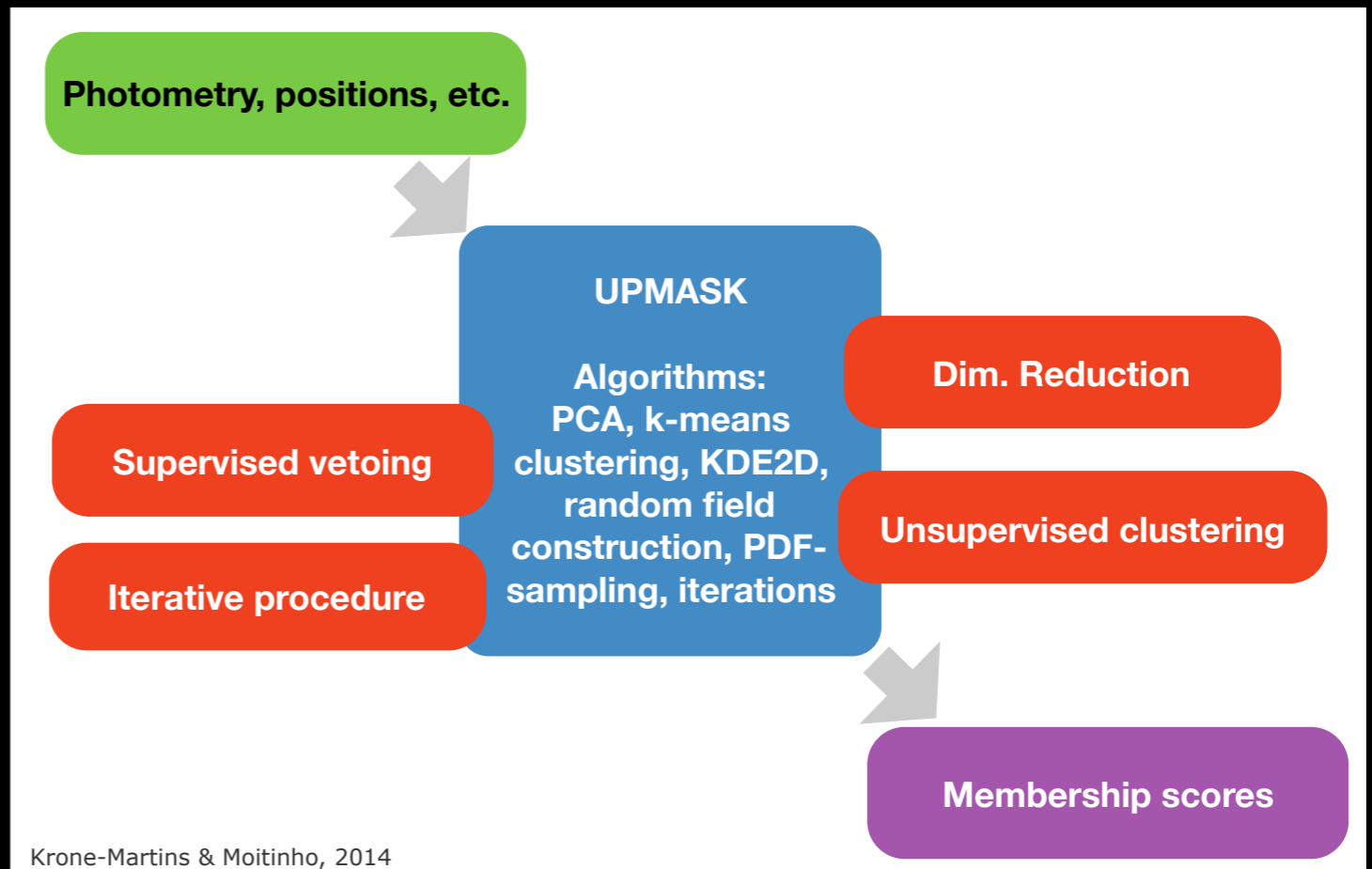
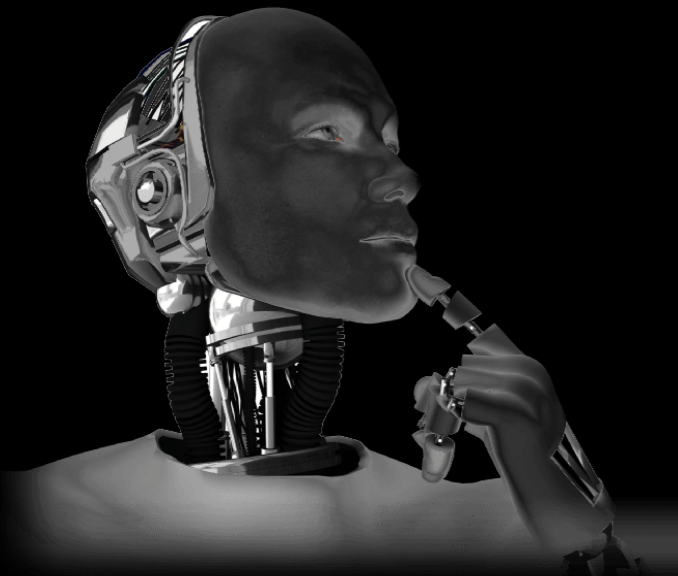
Krone-Martins & Moitinho, 2014



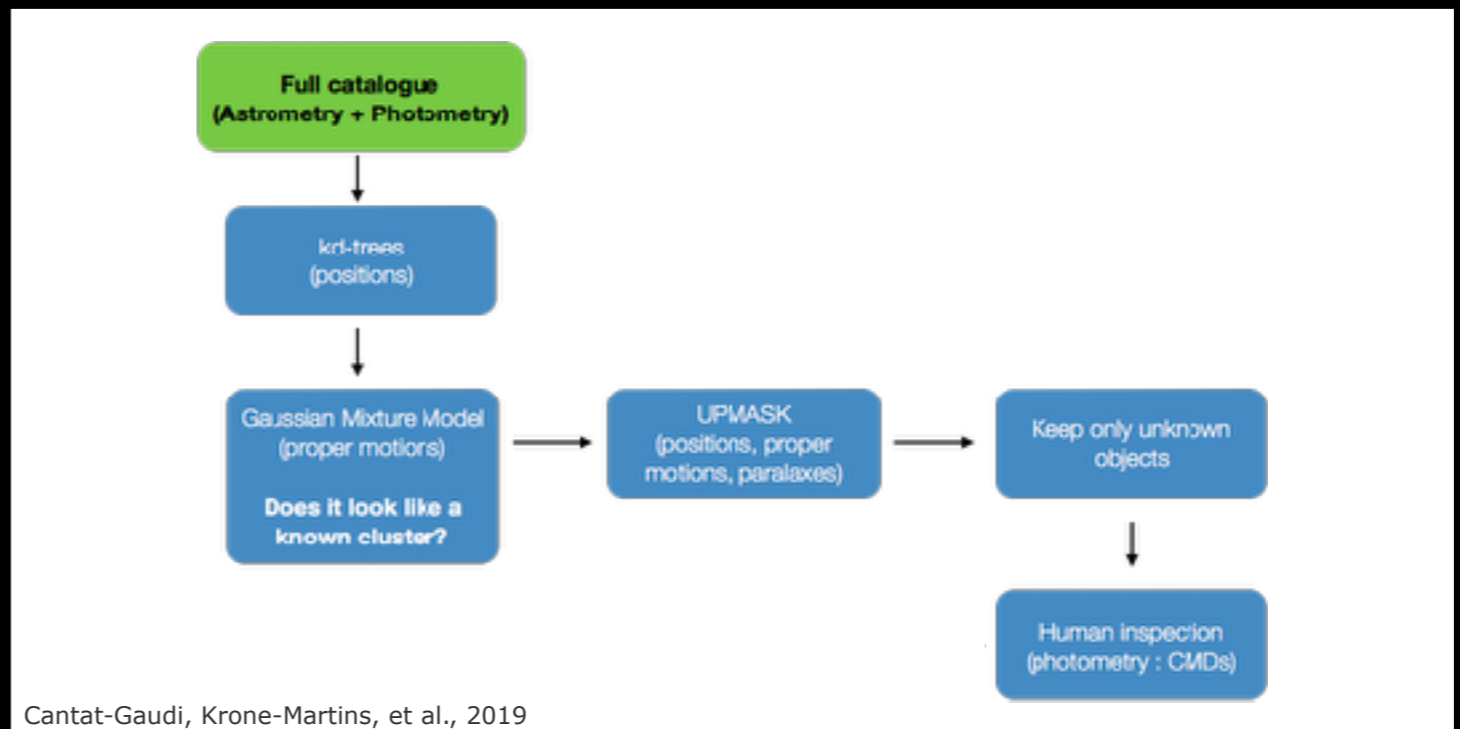
Unsupervised learning : Stellar Clusters



ESA/DPAC/U.Lisboa/CENTRA

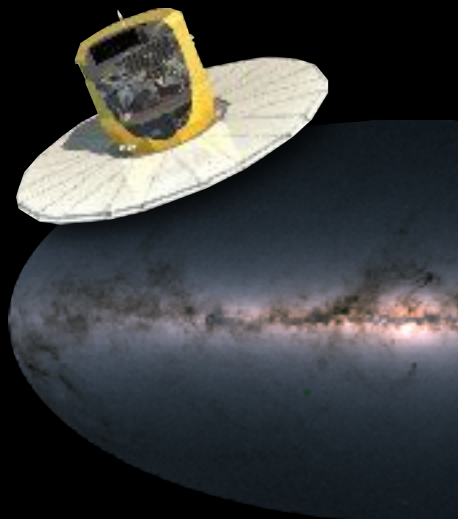


Krone-Martins & Moitinho, 2014

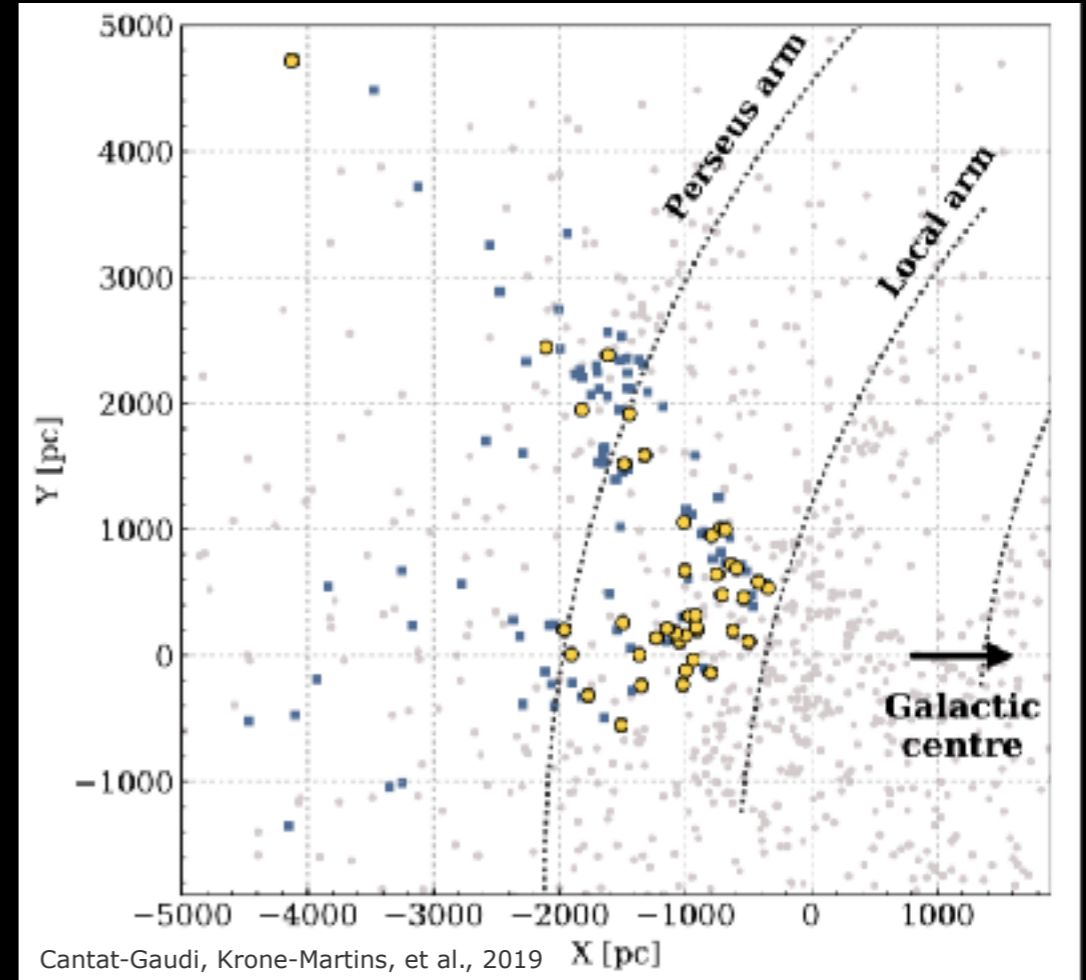
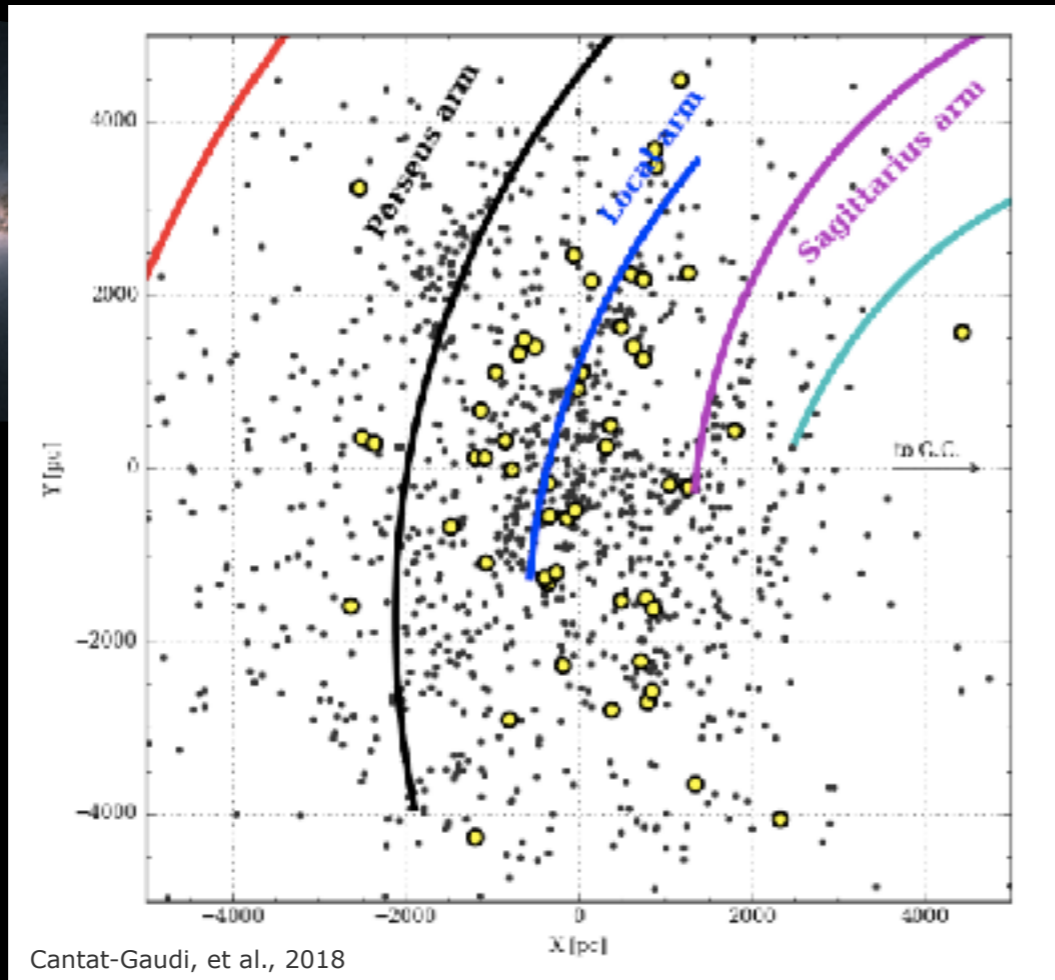


Cantat-Gaudi, Krone-Martins, et al., 2019

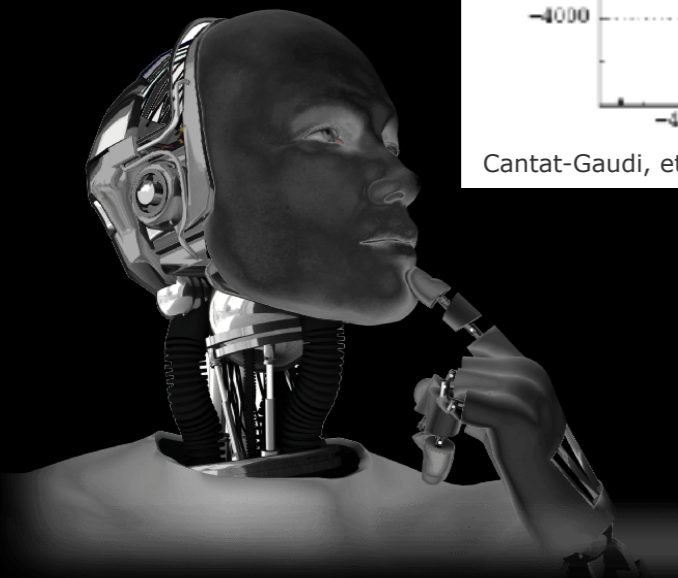
Unsupervised learning : Stellar Clusters



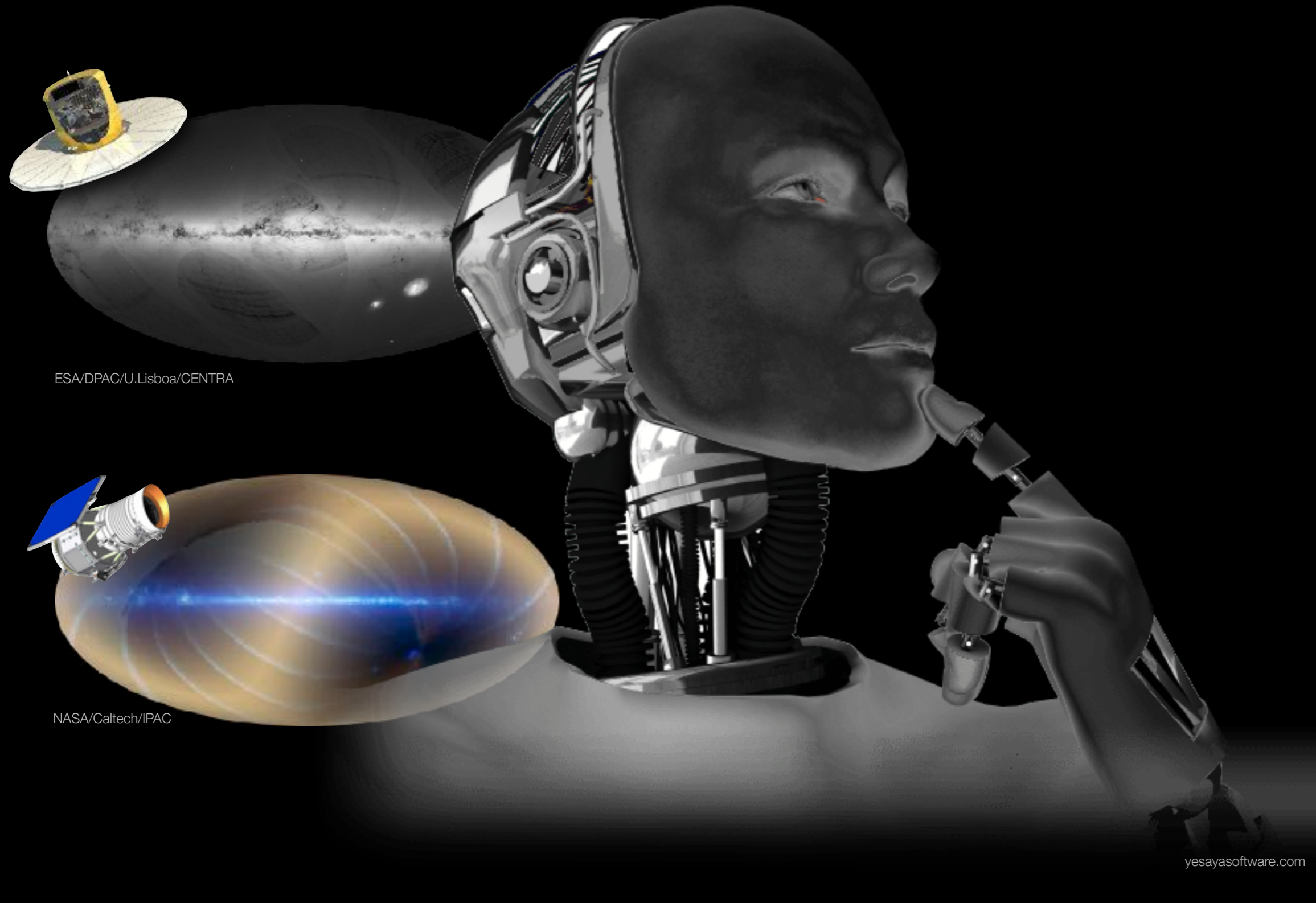
ESA/DAPAC/U.Lisboa/CENTRA



101 new stellar clusters
from Gaia DR2



Unsupervised learning : Galaxies

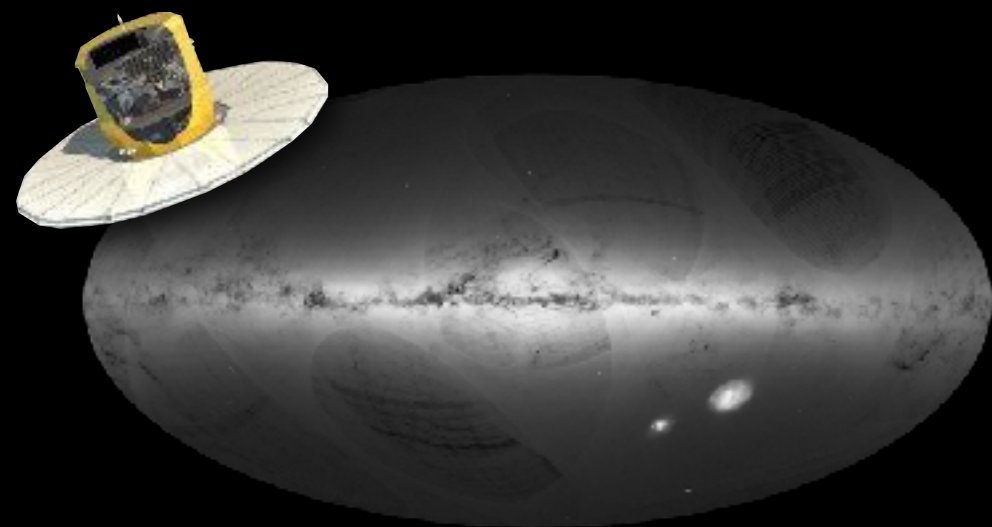


ESA/DPAC/U.Lisboa/CENTRA

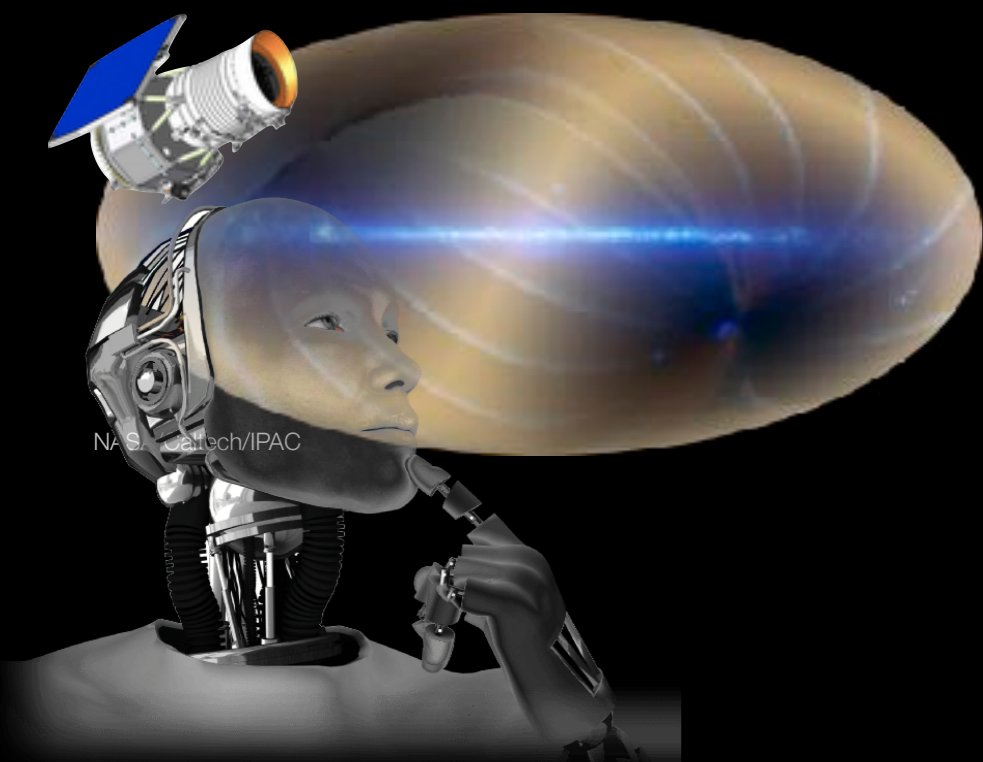
NASA/Caltech/IPAC

yesayasoftware.com

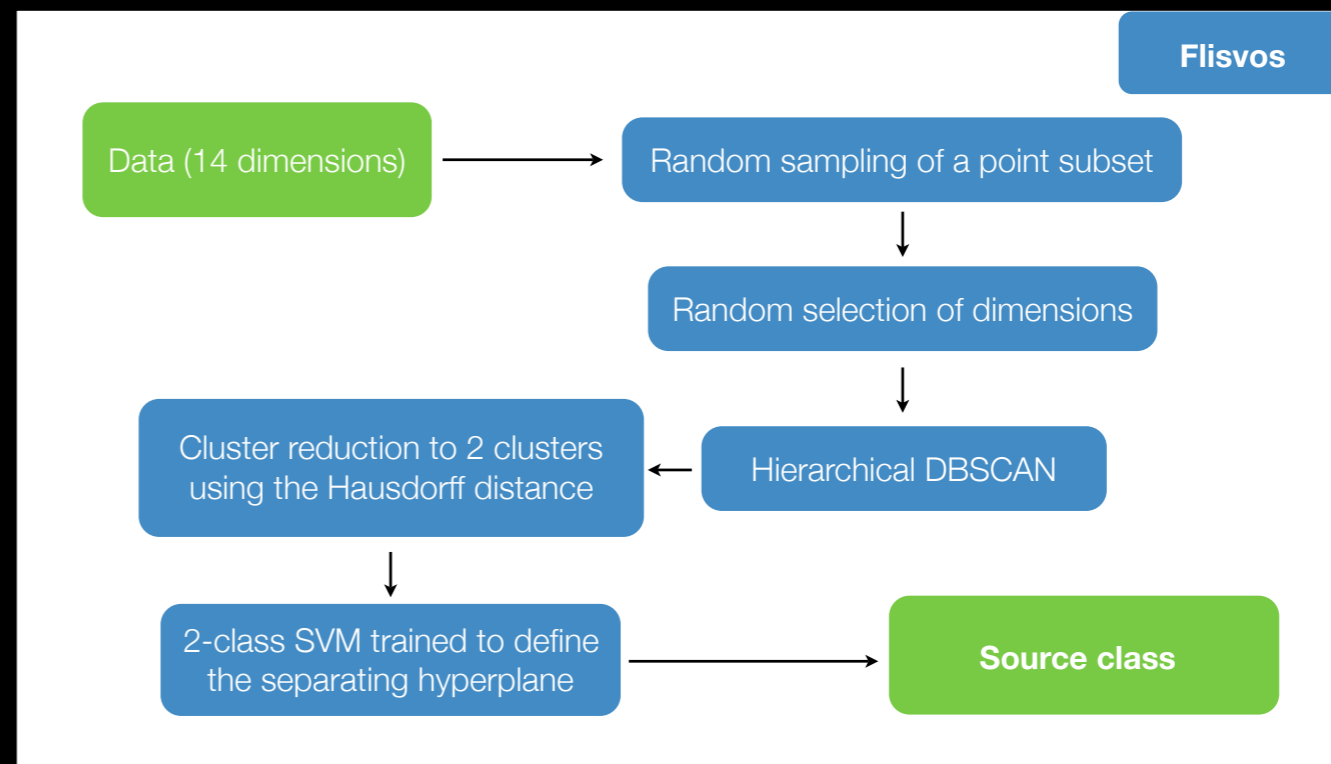
Unsupervised learning : Galaxies



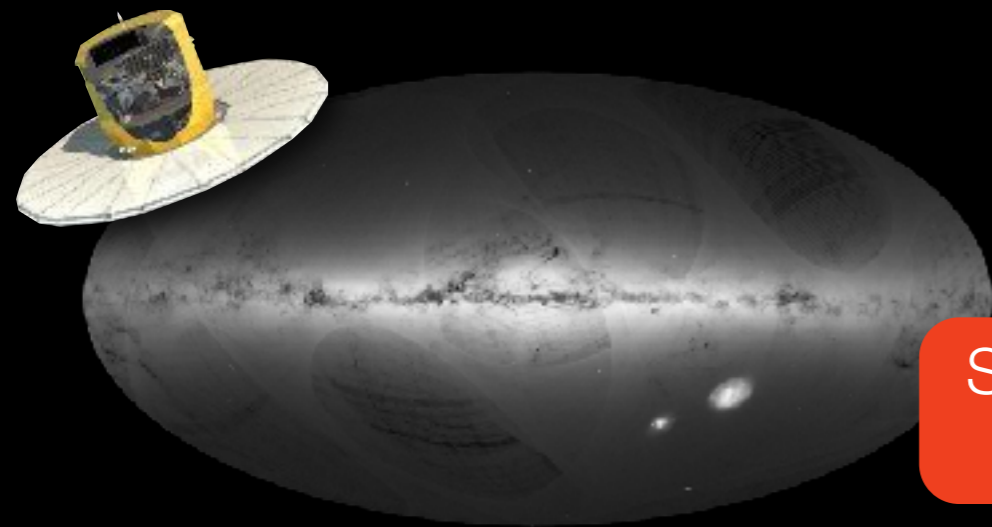
ESA/DPAC/U.Lisboa/CENTRA



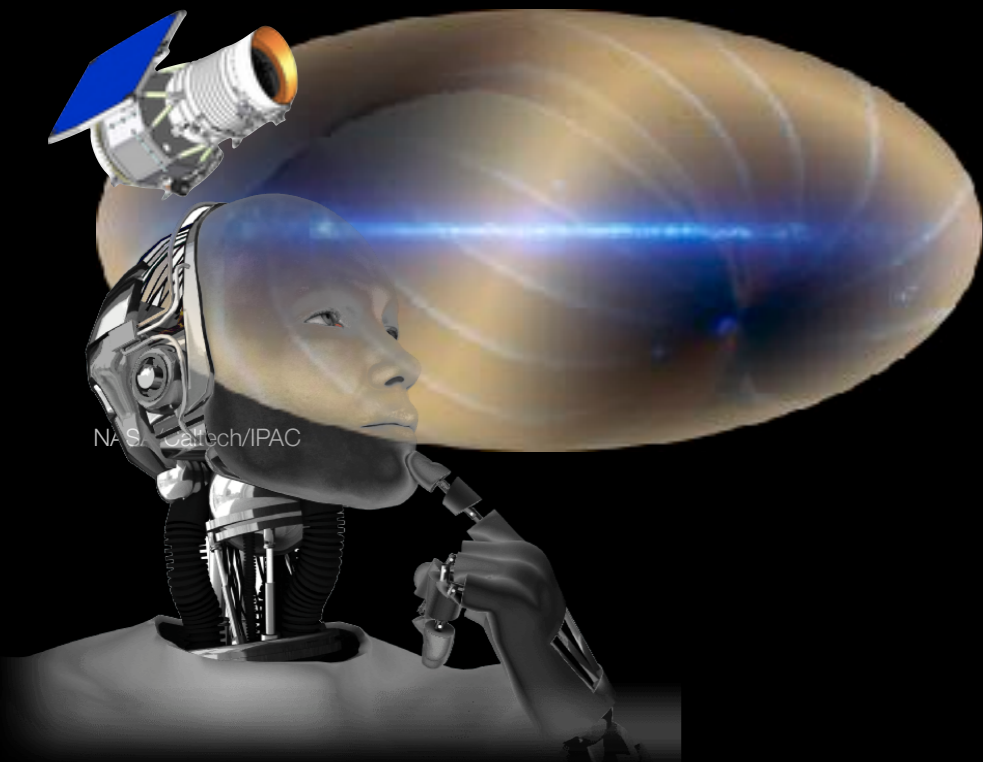
NASA/Cattech/IPAC



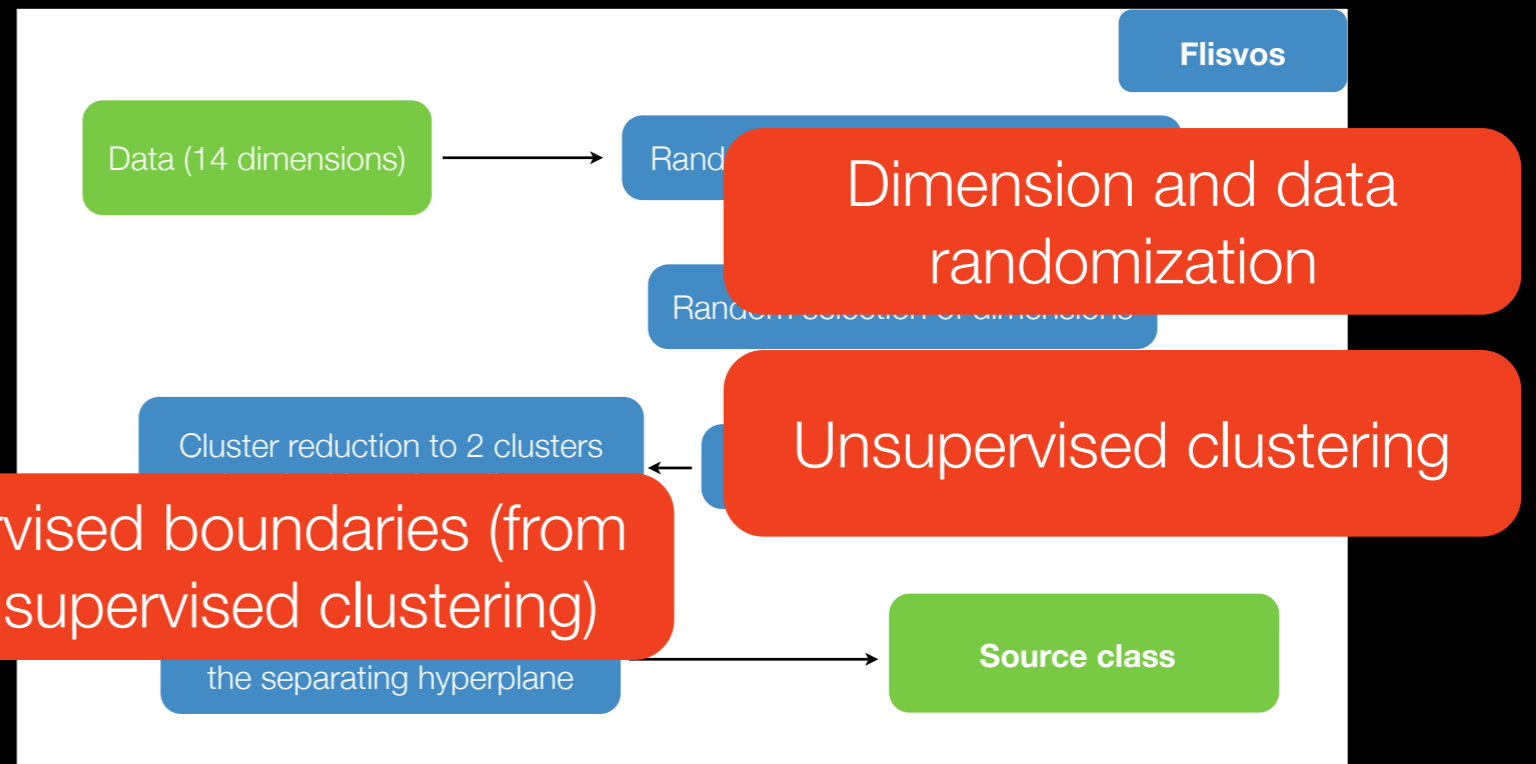
Unsupervised learning : Galaxies



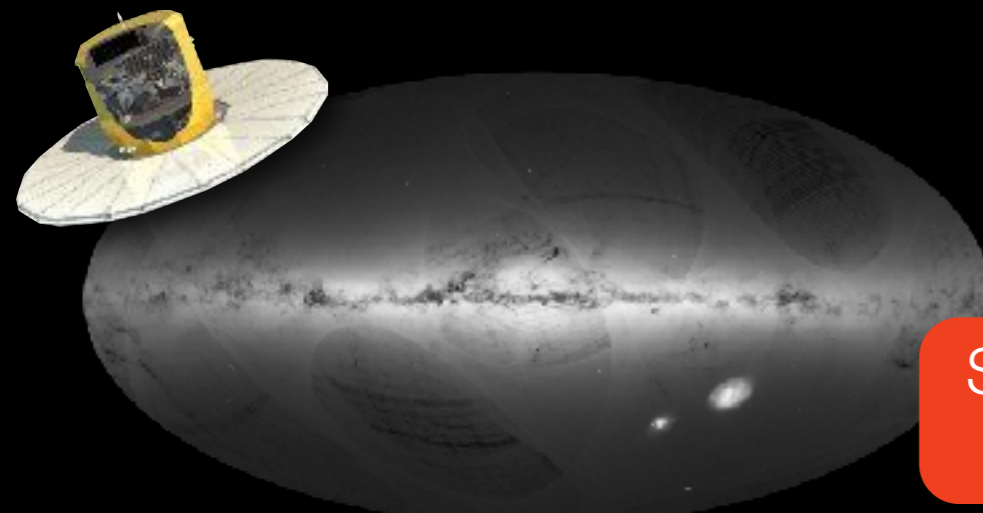
ESA/DPAC/U.Lisboa/CENTRA



NASA/JPL/Caltech/IPAC



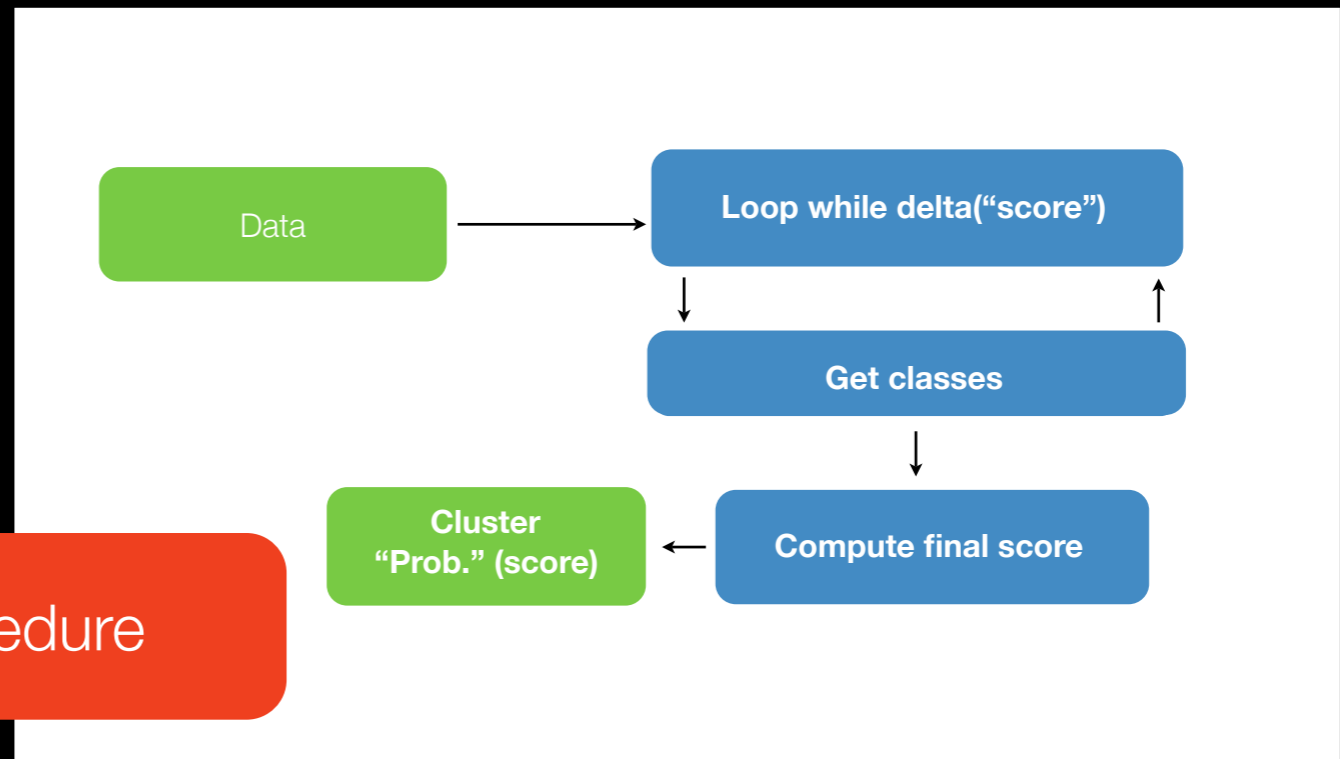
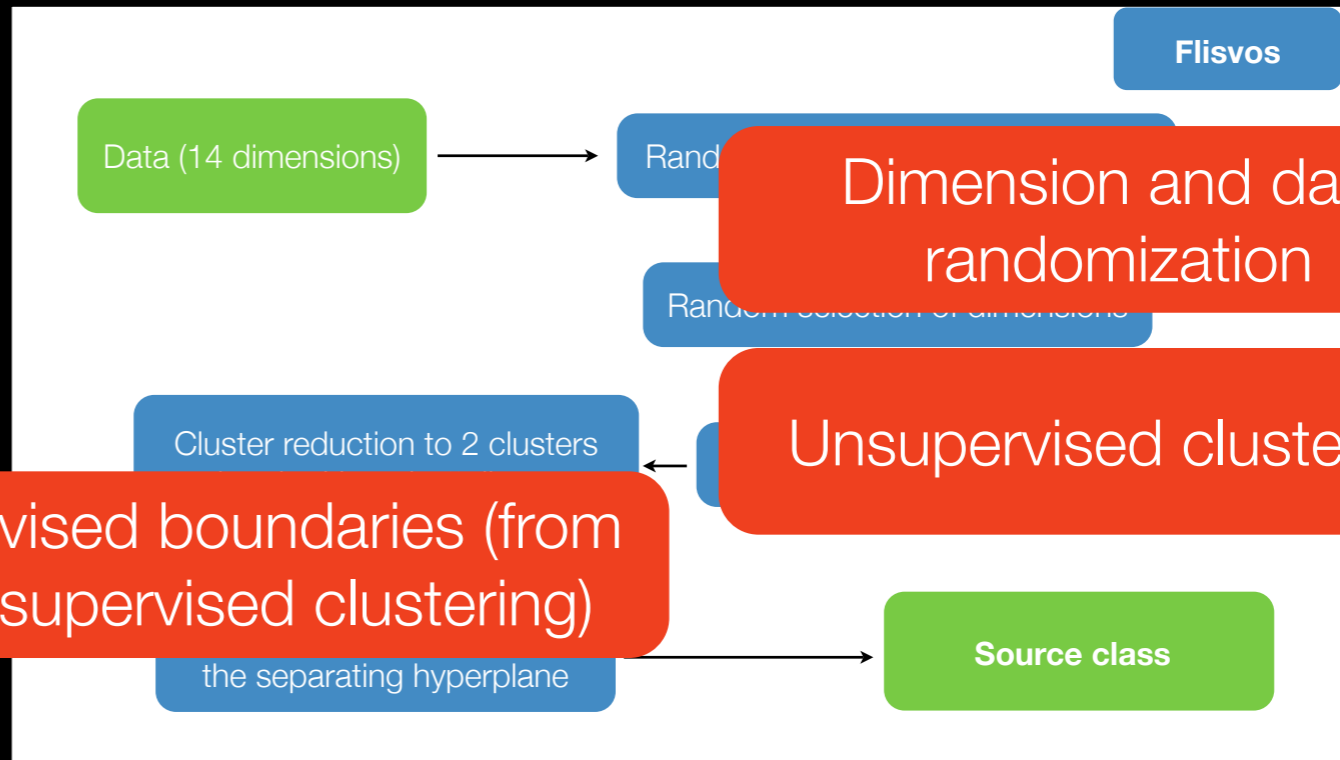
Unsupervised learning : Galaxies



ESA/DPAC/U.Lisboa/CENTRA

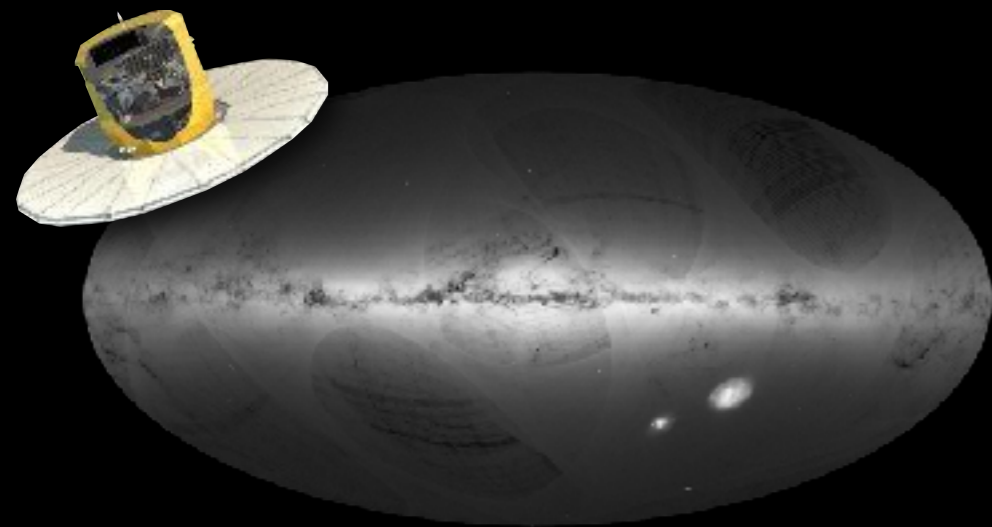


NASA/JPL/Caltech/IPAC

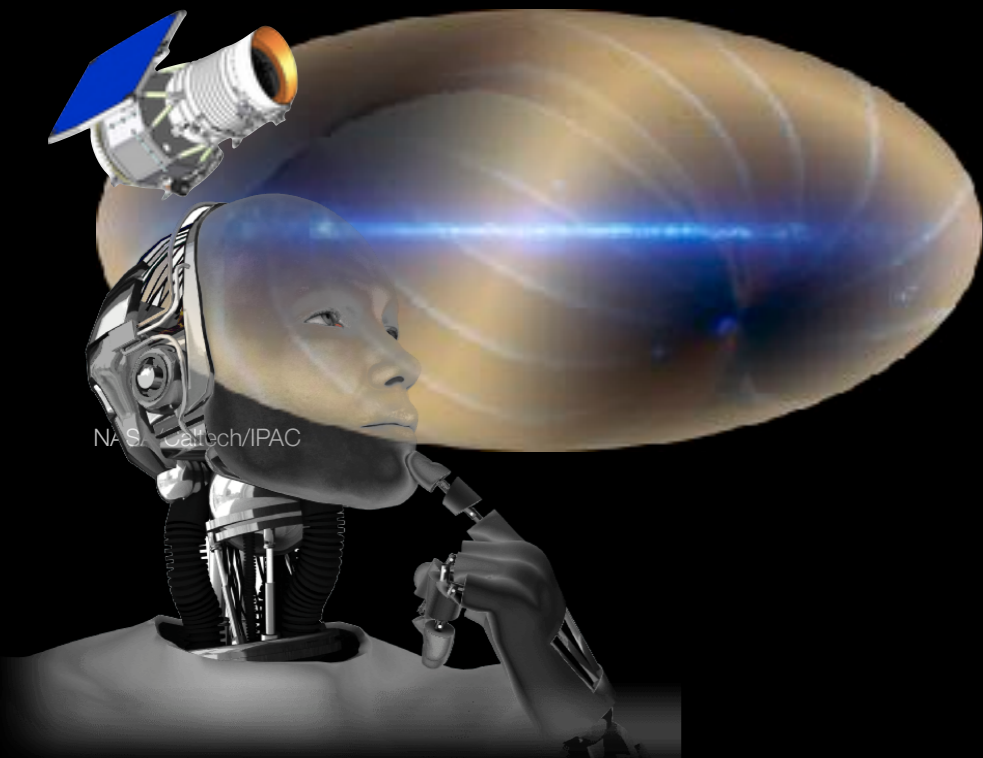


Iterative procedure

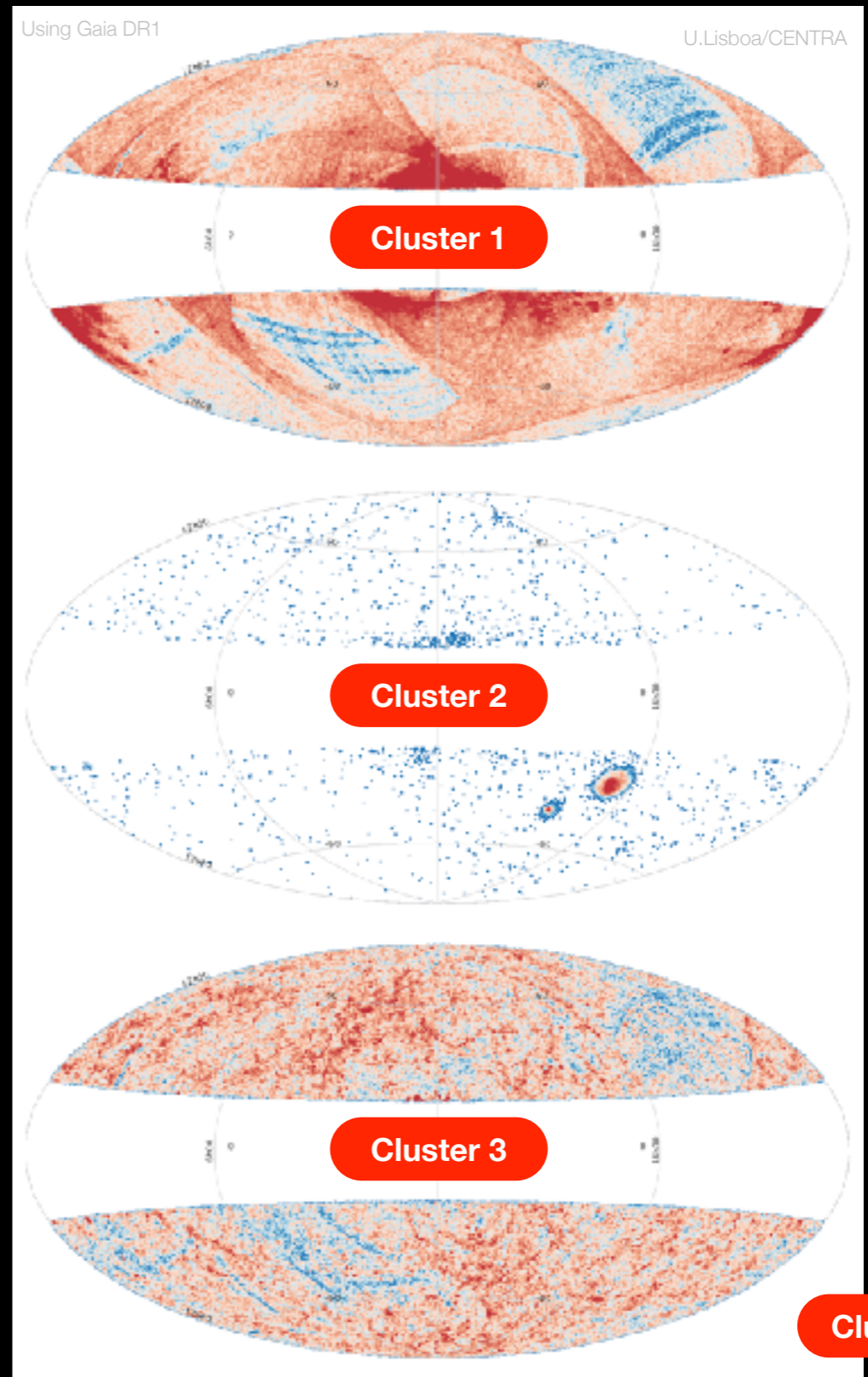
Unsupervised learning : Galaxies



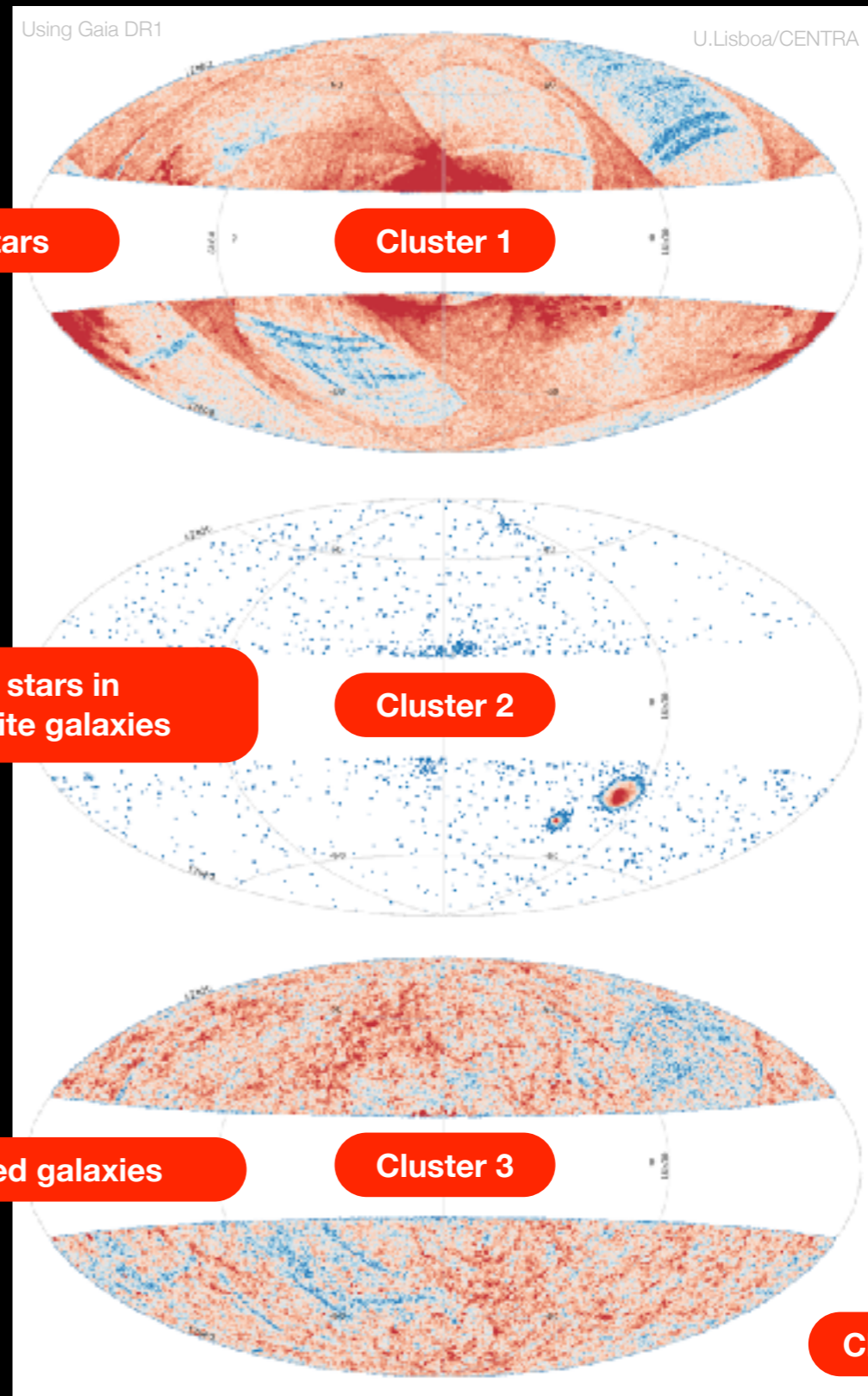
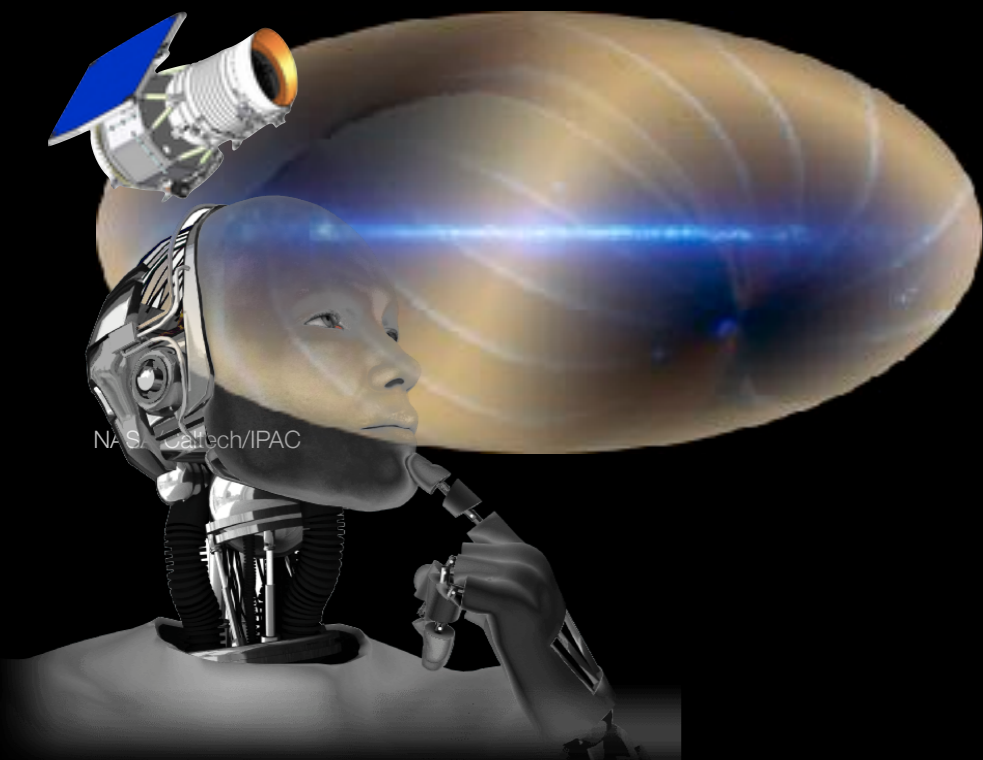
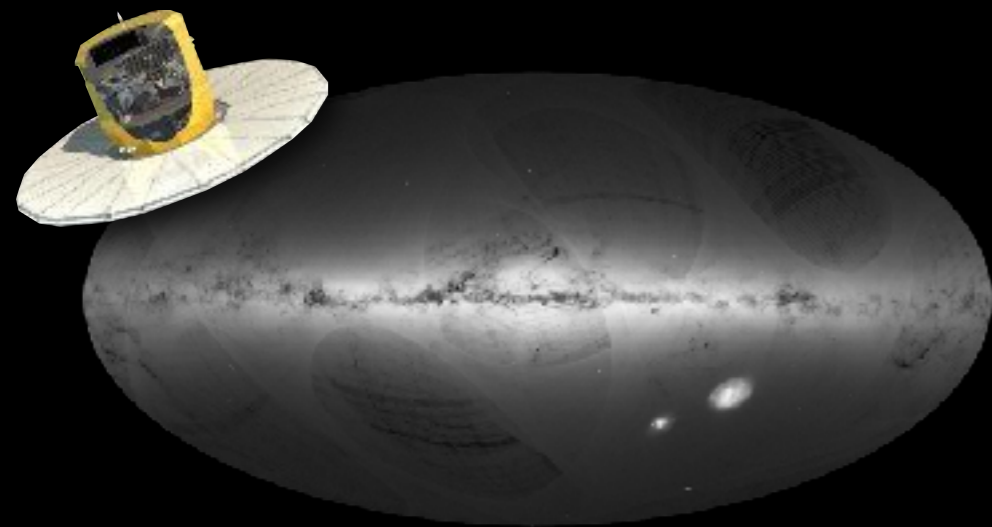
ESA/DPAC/U.Lisboa/CENTRA



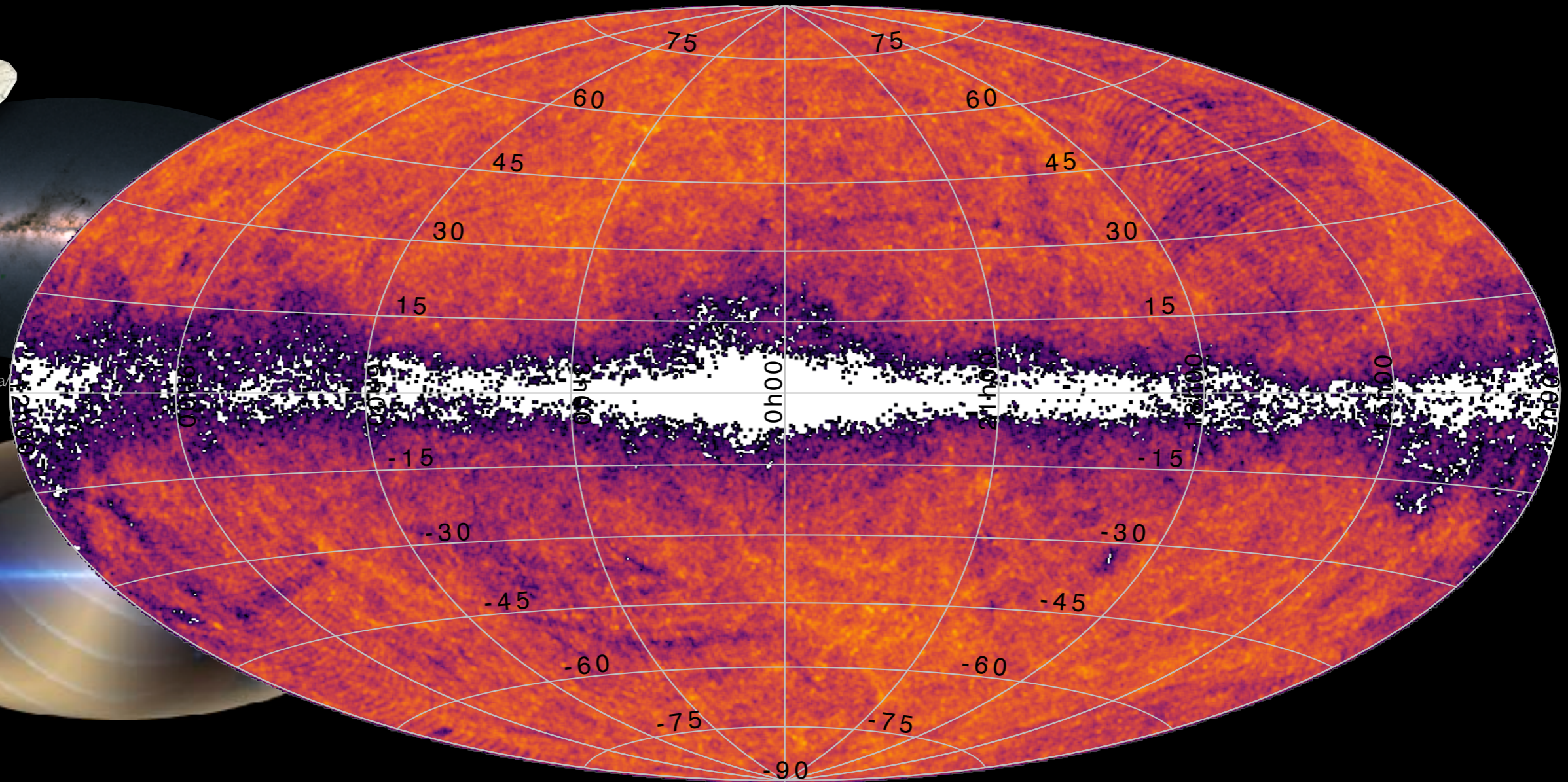
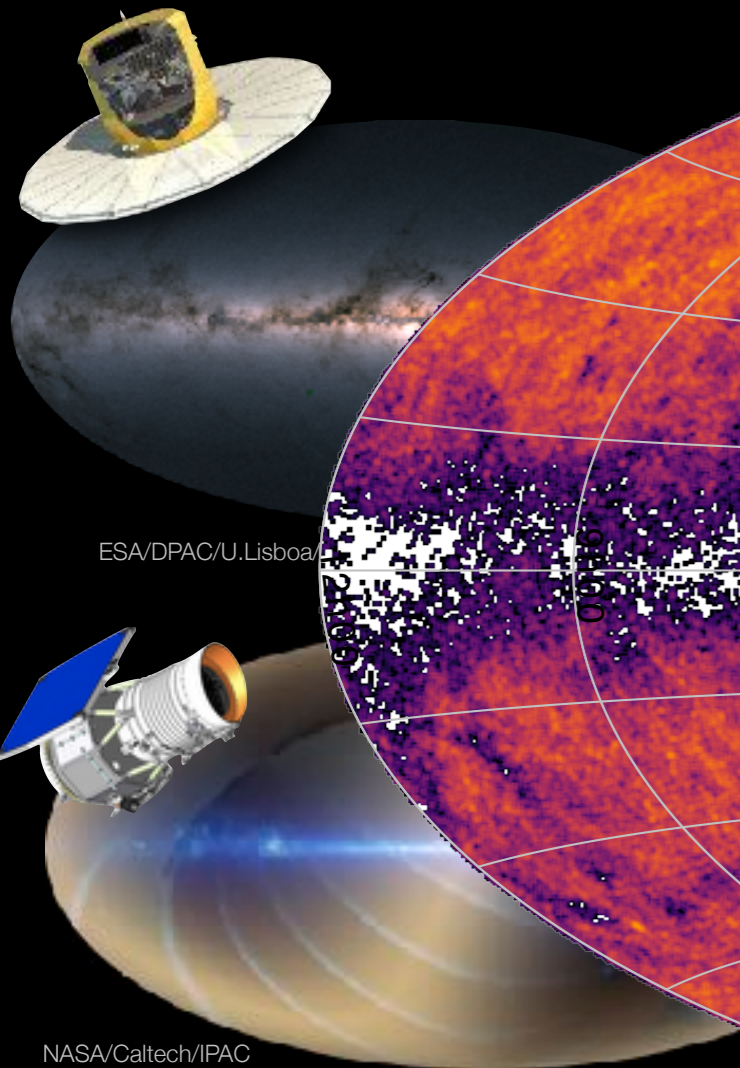
NASA/Cattech/IPAC



Unsupervised learning : Galaxies



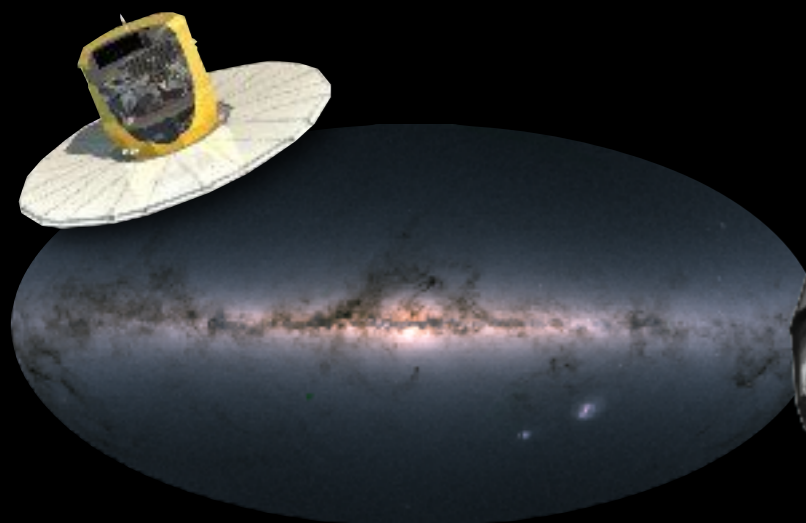
Unsupervised learning : Galaxies



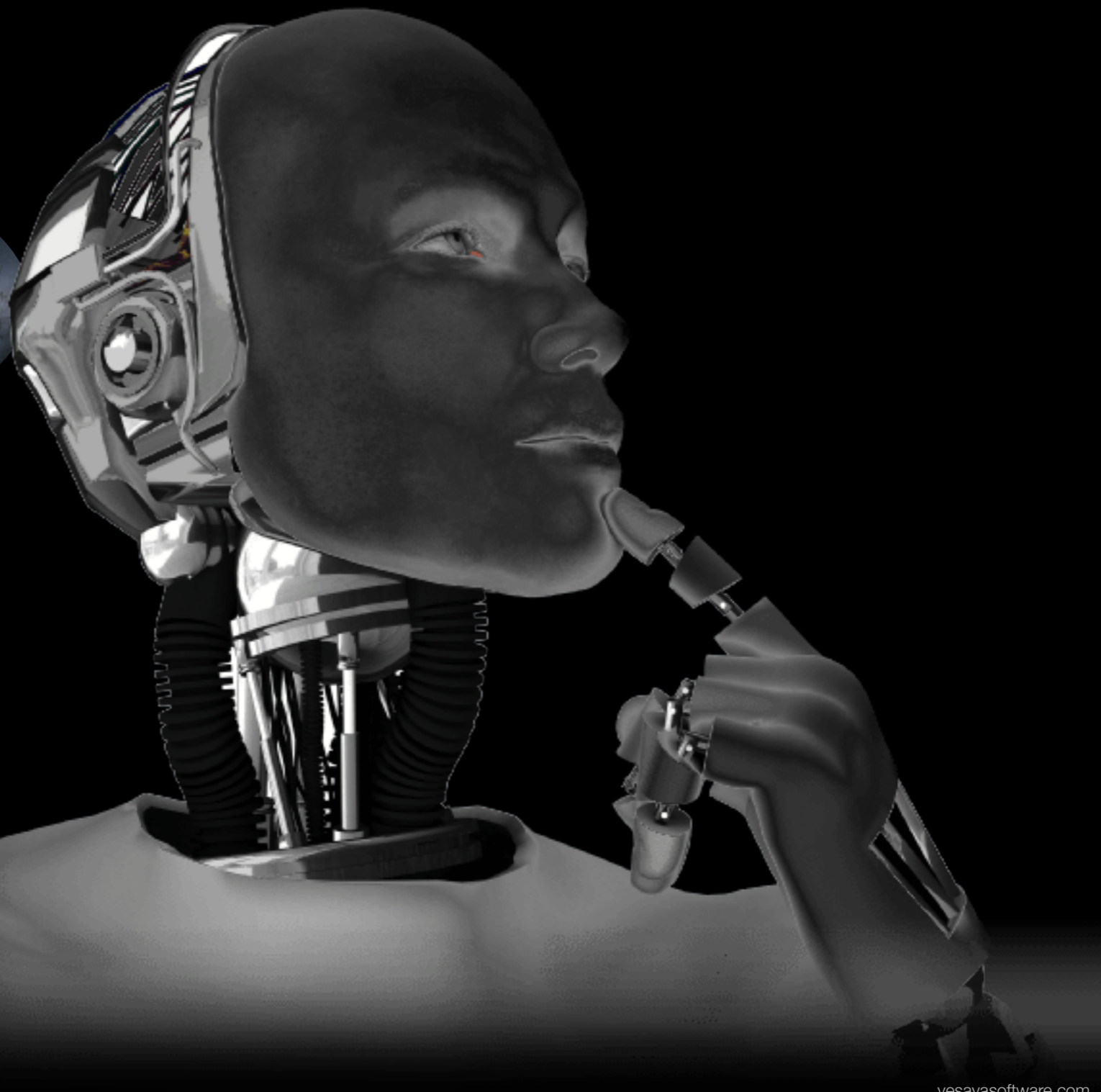
Using Gaia DR2

**Millions of galaxies
(hundreds of thousands are new)**

Unsupervised learning : Lensed Quasars



ESA/DPAC/U.Lisboa/CENTRA



yesayasoftware.com

Unsupervised learning : Lensed Quasars

Among the most interesting (and useful) extragalactic phenomena...

modelling: astrometry + photometry + spectroscopy

measurement:
photometry
time-series

$$\tau = \frac{(1 + z_L) d_L d_S}{H_0 d_{LS}} \left[\frac{1}{2} |\vec{\theta} - \vec{\beta}| - \psi_{2D}(\vec{\theta}) \right]$$



Unsupervised learning : Lensed Quasars

Among the most interesting (and useful) extragalactic phenomena...

modelling: astrometry + photometry

measurement:
photometry
time-series

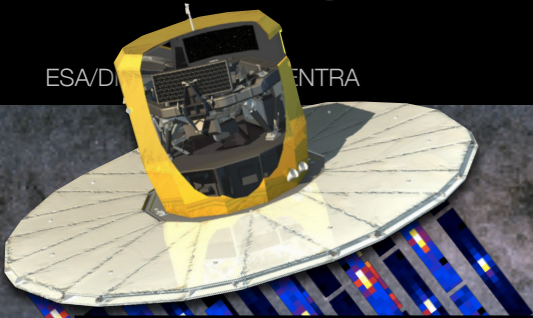
$$\tau = \frac{(1 + z_L) d_L d_S}{H_0} \left[1 - \frac{d_L}{d_S} \right]$$

Problem
Very few multiply-imaged QSOs are known!



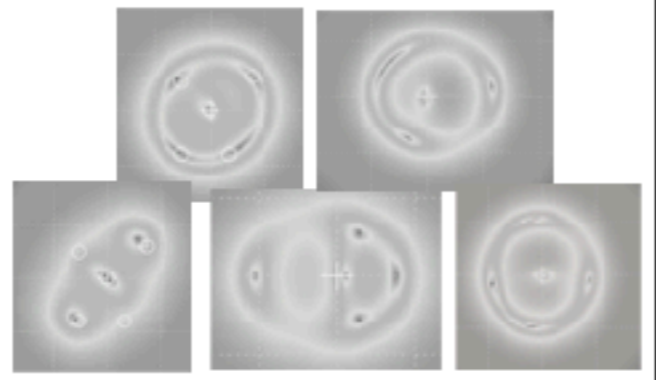
Unsupervised learning : Lensed Quasars

ESA/D... ENTRA



Lens candidate selection

Astrometric + photometric patterns + ERTs



Astrometry + lower lightcurve entropy + SVM

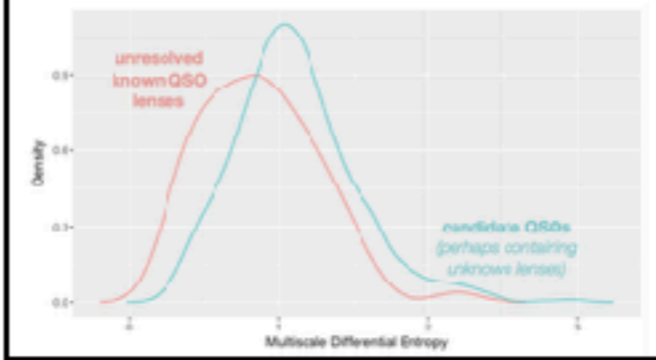
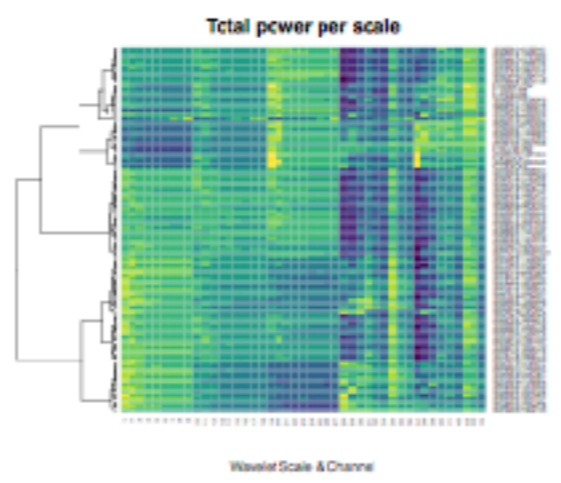
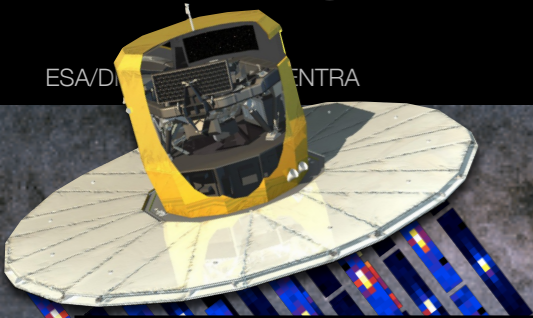


Image wavelet power spectrum signatures + hierarchical clustering + Astrometry



Unsupervised learning : Lensed Quasars

ESA/D... ENTRA



Lens candidate selection

Astrometric + photometric patterns + ERTs



Astrometry + lower lightcurve entropy + SVM

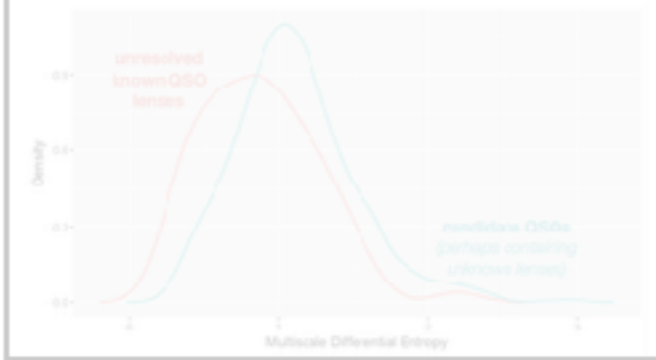
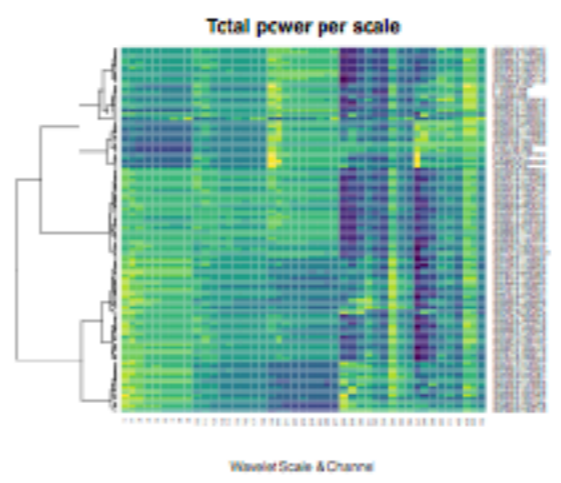
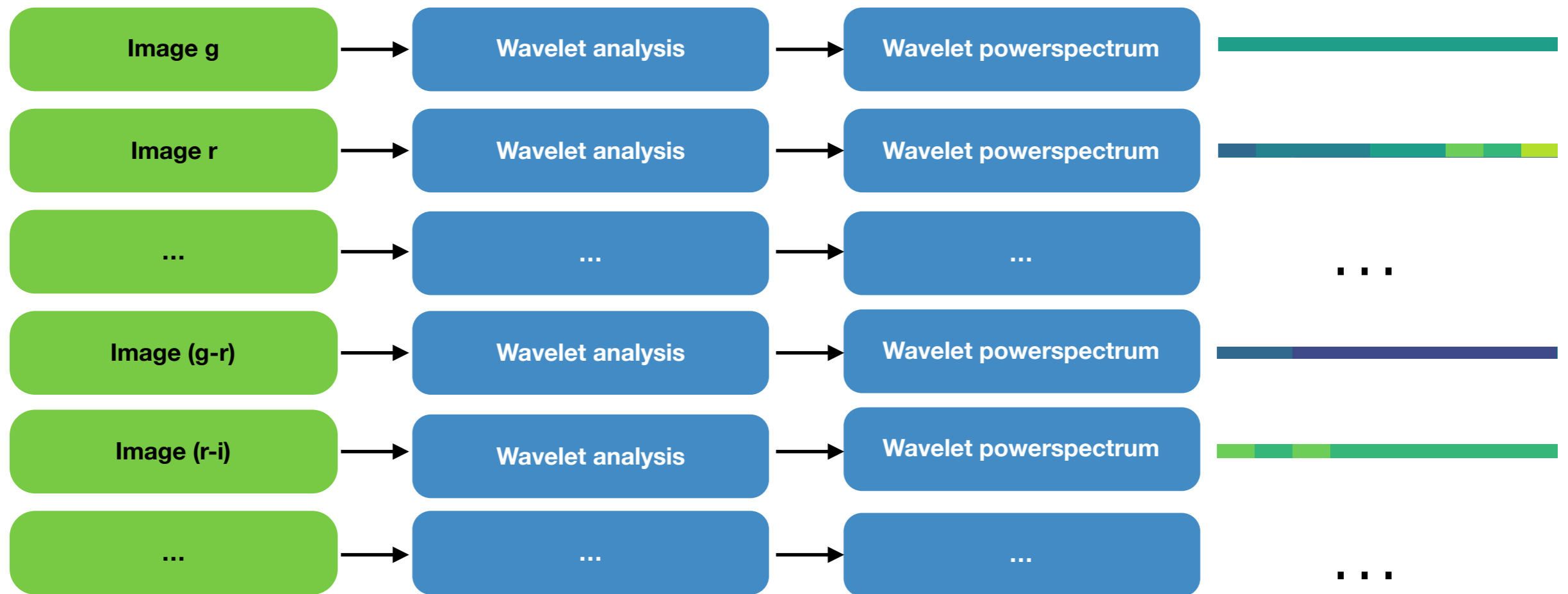
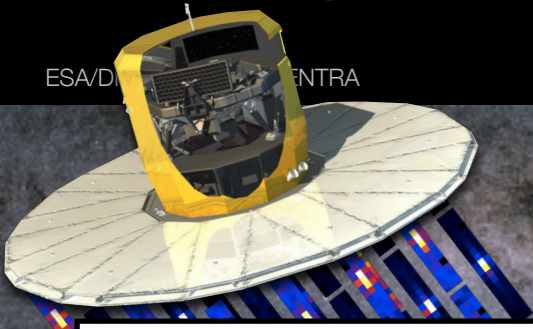


Image wavelet power spectrum signatures + hierarchical clustering
+Astrometry



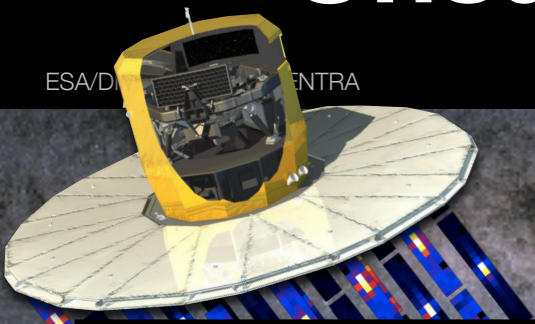
Unsupervised learning : Lensed Quasars

ESA/D... ENTRA

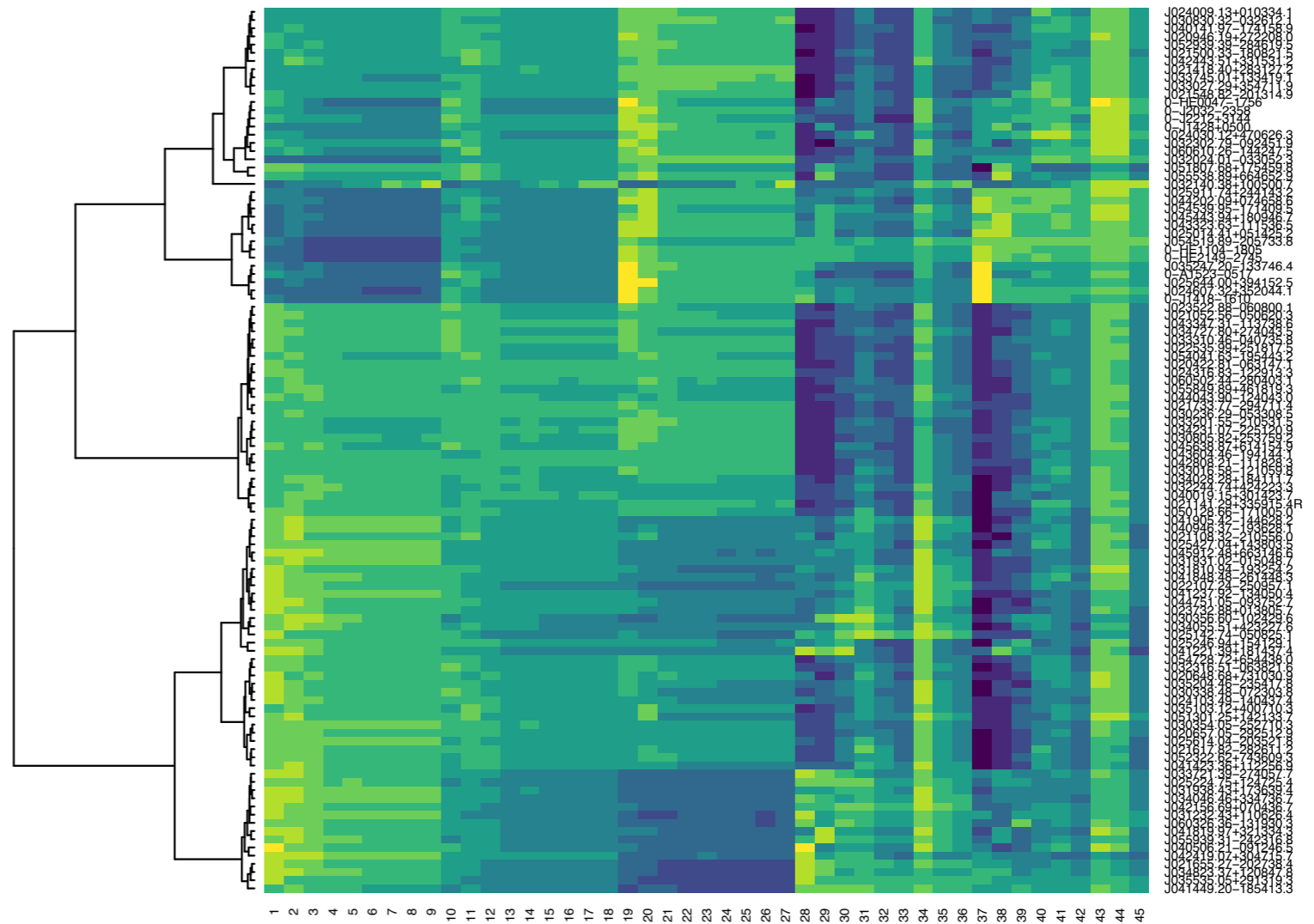
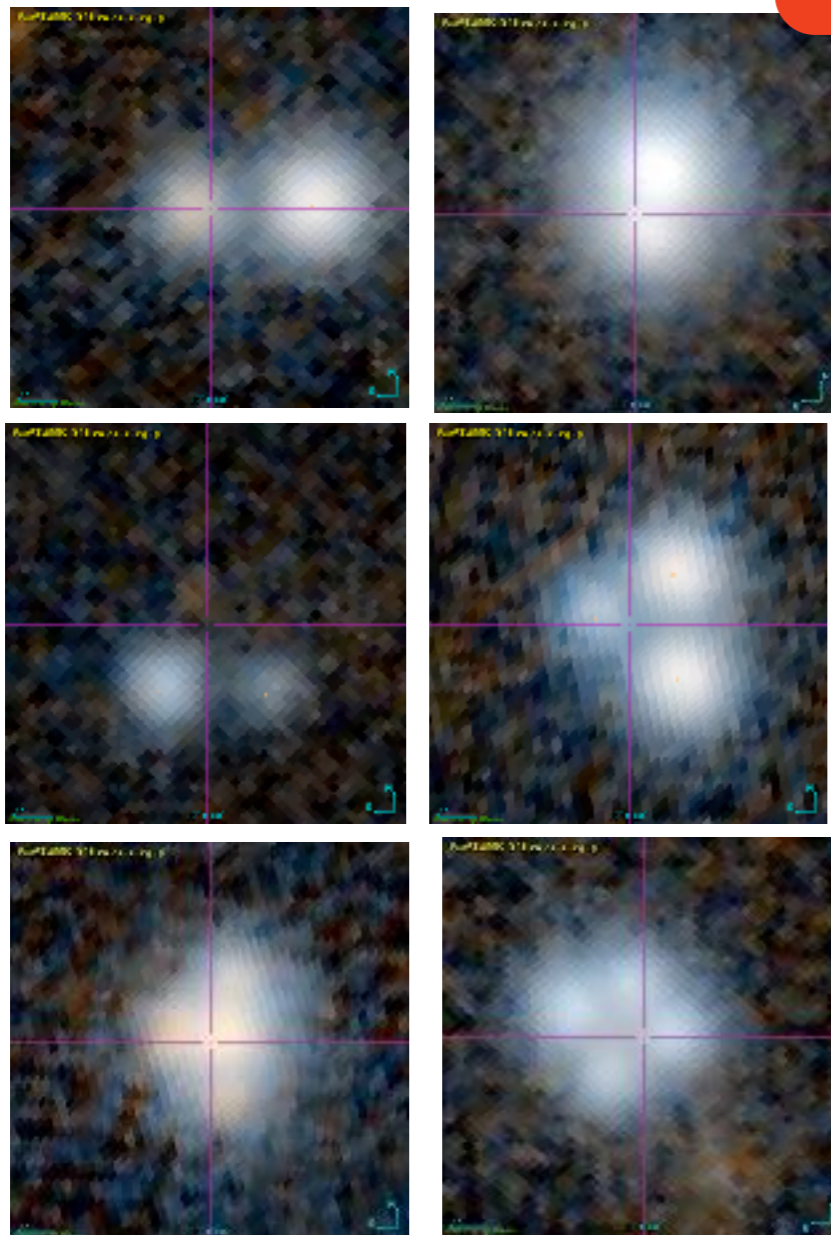


Unsupervised learning : Lensed Quasars

ESA/DLR SCIENCE CENTER



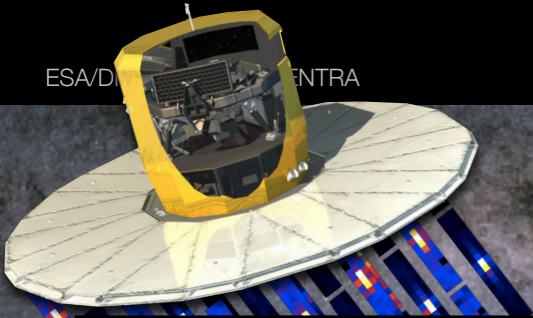
Keck run : 5 candidates -> 2 new lenses, 1 unresolved



Wavelet Scale & Channel

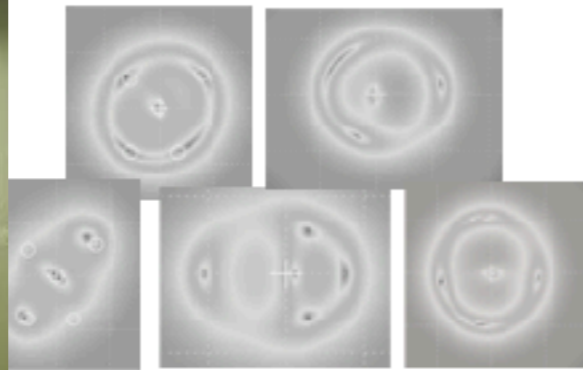
Inside the machinery...

ESA/D... ENTRA

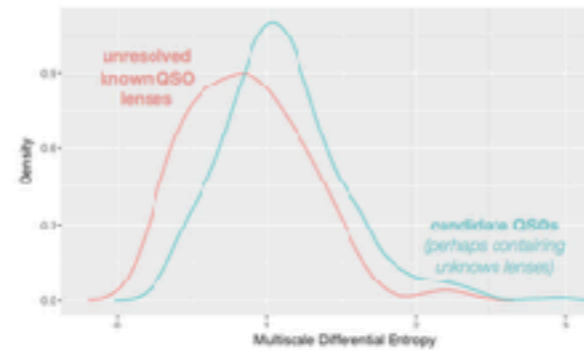


Lens candidate selection

Astrometric + photometric patterns + ERTs



Astrometry + lower lightcurve entropy + SVM

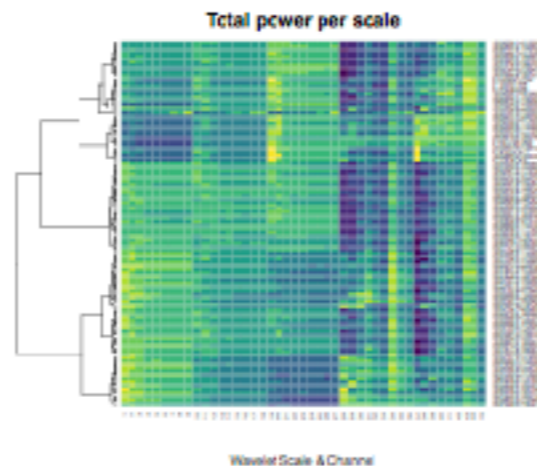


Candidates

Lens spectroscopic confirmation



Image wavelet power spectrum signatures + hierarchical clustering + Astrometry

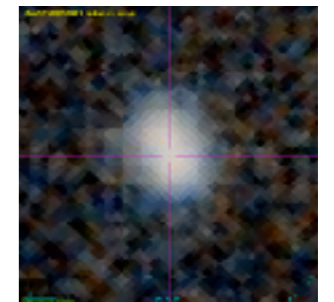
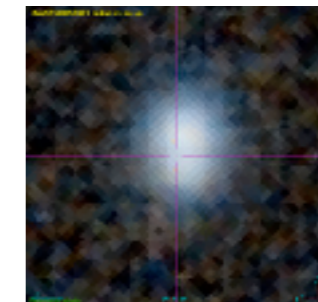
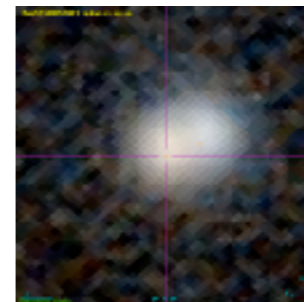
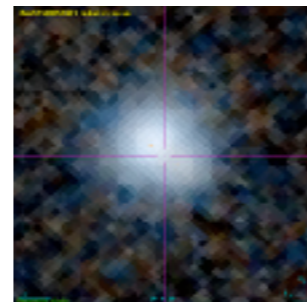
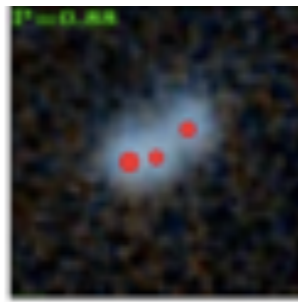
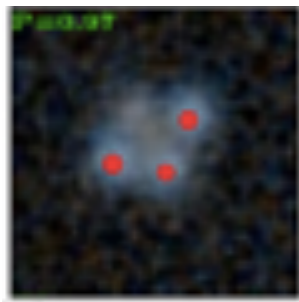
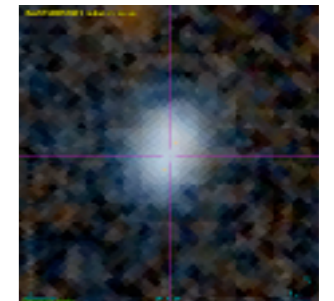
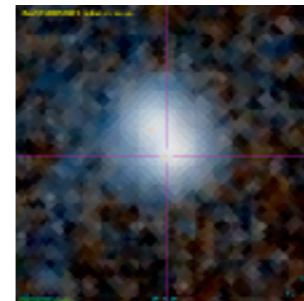
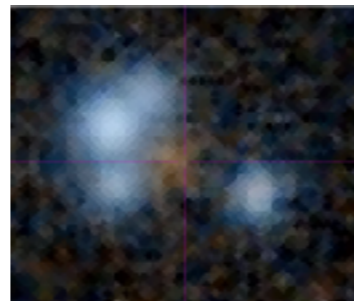
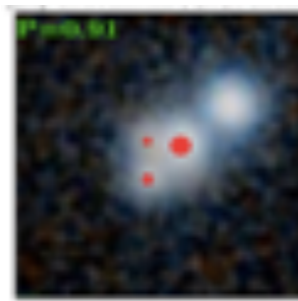
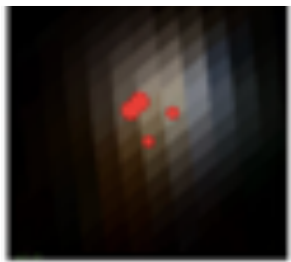
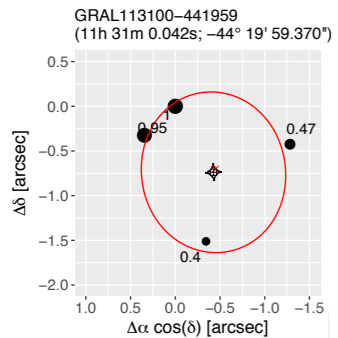
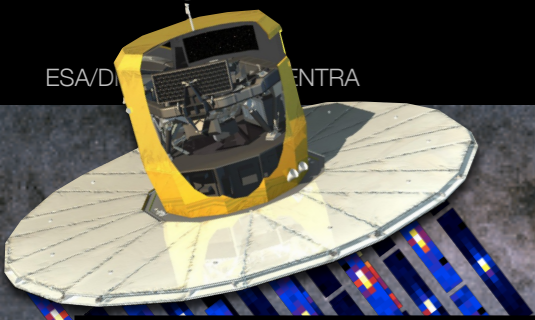


Results

An Agile Software Engineering methodology approach while doing Science...

GraL confirmed lenses: May/2018 – April/2019

ESA/DLR SCIENCE CENTRE



References

- Krone-Martins, A. et al., in prep
- Stern, D. et al., in prep
- Wertz, O.; Stern, D.; Krone-Martins, A. et al., A&A accepted.
- Delchambre, L.; Krone-Martins, A.; Wertz, O., et al., A&A, 622, A165, 2019.
- Ducourant, C.; Wertz, O.; Krone-Martins, A., et al., A&A, 618, A56, 2018.
- Krone-Martins, A.; Delchambre, L.; Wertz, O. et al., A&A, 616, L11, 2018.

And many more...

Thank you!

Unsupervised Learning is powerful,

but it requires some taming

(i.e. metric, clustering evaluation criterion, dictionary constraints).

**SL/ML methods are “just” engines,
we can profit much more by building “ships”**