From the β-skeleton to the Cosmic Web elements.

John F. Suárez-Pérez

Advisor
Jaime Forero-Romero

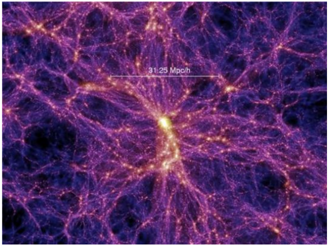
Los Andes University
Astroandes

September 25th, 2020

One of the main goals in cosmology is understand the distribution of dark matter in the local Universe.

The problem: The distribution of Dark Matter (DM) is not possible to observe directly.

A solution: Make an inference of the DM distribution using observational measurements of galaxies distributions like SDSS or DESI (Working).



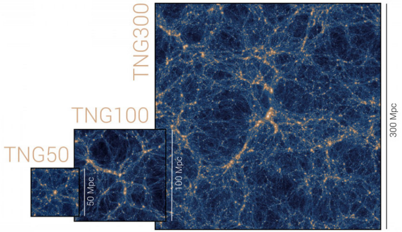
[1] Credits: V.Springel, Max-Planck Institut für Astrophysik, Garching bei München

First Step:
Simulation

What is the Illustris-TNG project?

"It is a great set of simulations magneto-hydro-dynamics of galaxies formation, completed in 2019... it uses numerical algorithms and physical models. The simulation represents a combination of high resolution and high physical fidelity"[2]..

- It includes different elements (dark matter particles, galaxies, gas cells, stars, wind stellar particles, super massive black holes, diffuse gas), in a redshift from $z = 127$ to the present $z = 0$.
- The simulation data includes 100 snapshots.
- Each simulation have a volume of $(302{,}6Mpc)^3$.



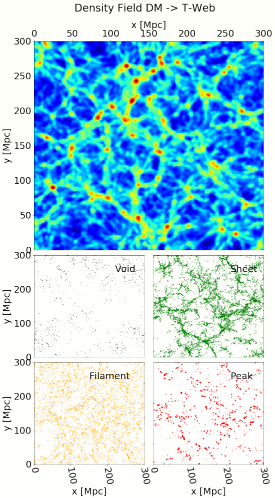
[2] Website: https://illustris-tng.org

Classification of the Cosmic Web: T-Web

"It is to possible make a classification of the cosmic web as a function of the local density, for make this classification is used the gravitational potential"

$$\boxed{T_{\alpha\beta} = \frac{\partial^2 \phi}{\partial r_\alpha \partial r_\beta}.}$$

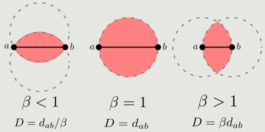
Depending on the value of the eigenvalues respect to a threshold λ_{th}, it is possible to make a classification by environments between peaks, sheets, filaments and voids. This classification is called the T-Web [3]



[3]A dynamical classification of the cosmic web. Forero–Romero J. et al. MNRAS. 2009

Characterization of the galaxies distribution using the β-skeleton.

The characterization is obtained using the β-skeleton algorithm, this algorithm allow us identify graph.



From the graph is possible to compute the number of connections by galaxy (node), the average length of connections, the eigenvalues of the inertia matrix, the pseudo-volume and the pseudo-density.

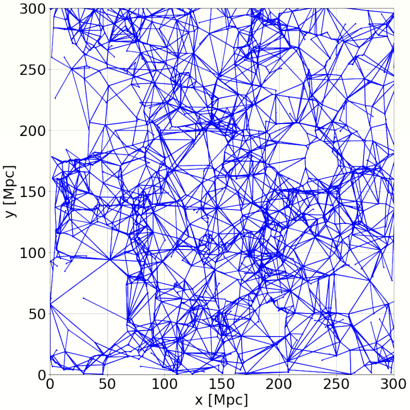


Fig. 1: Graph for the galaxies distribution of TNG for a region $z < 10Mpc$ and $\beta = 1$.

First Step: Reconstruction of the T-Web

Suárez-Pérez et. al [In prep.]

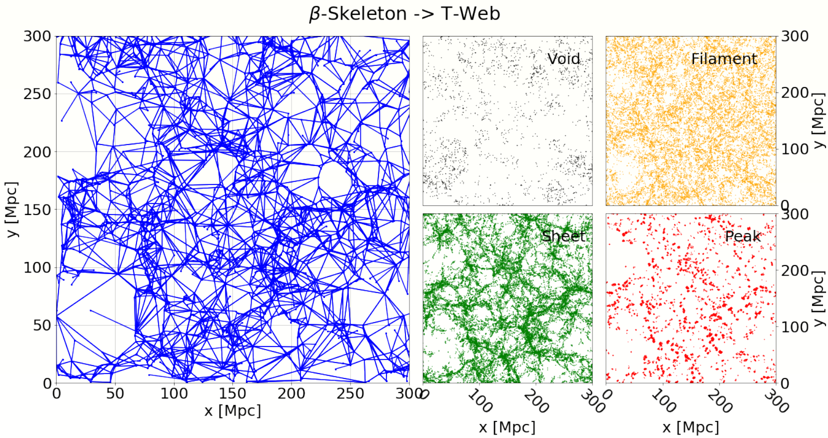


Fig. 2: Reconstruction of the T-Web from the β-skeleton

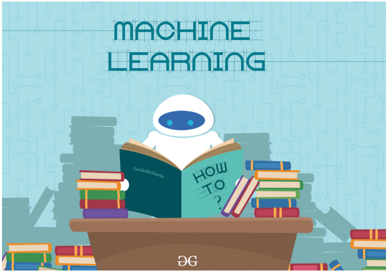
¿Machine Learning?

¿Machine Learning?

¿Why?

- It is not possible to make direct observations of the DM.
- We can to make an inference from information that can be measuring.

«Training with simulations to predict with observations».



[4]From https://www.geeksforgeeks.org/machine-learning/

Feature Space $\rightarrow$ Galaxies.

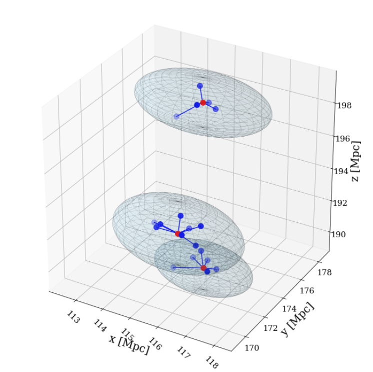


Fig. 3: Pseudo-Volumen using the parameters **a**,**b** y **c**.

- For $\beta = 1{,}0$
- By node is computed the number of connections and the average lenght.
- By structure is possible define a inertia matrix and compute its eigenvalues.(σ_1, σ_2 y σ_3).

Defined:

- $a = \sqrt{\sigma_1}$, $b = \sqrt{\sigma_2}$ y $c = \sqrt{\sigma_3}$.
- The pseudo-volume $V = abc$ and pseudo-density as $\rho = \frac{1}{abc}$.

Feature Space → Galaxies → Local Parameters.

Also was computed a set of local parameters that include the information of the first neighbors. This information is define as $\Delta f = \bar{f} - f$.

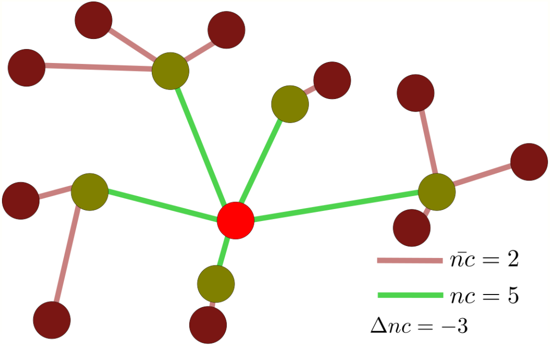


Fig. 4: Representation for the local parameters.

Feature Space $\rightarrow$ Galaxies.

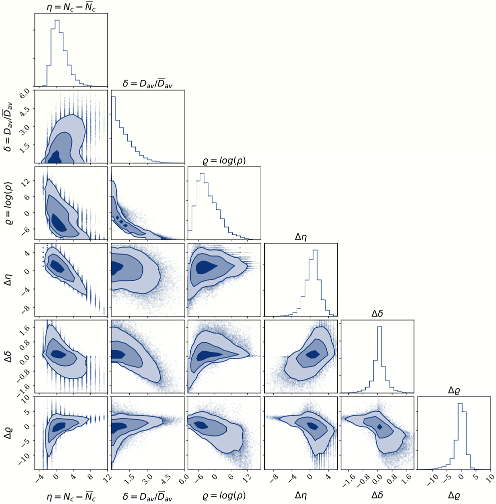


Fig. 5: Final correlations between the features extracted from the galaxies.

Feature Space → Density field of DM.

- The smoothing (sm) is a tuning parameter over the density field.

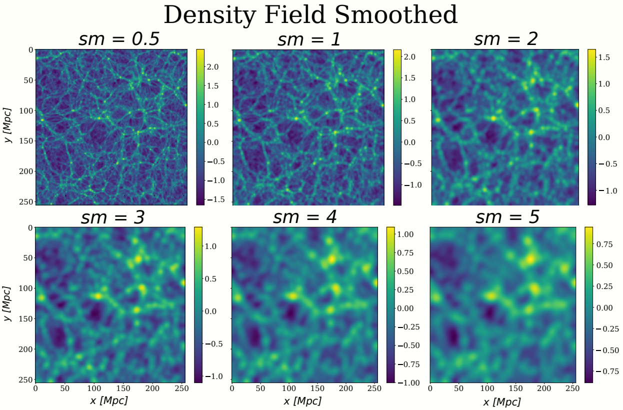


Fig. 6: Density field for different smoothing ***sm***.

Feature Space → Deformation Tensor.

It is possible tuning a cut value λ_{th} that allow us make a classification by environments between peaks, sheets, filaments and voids with the eigenvales computed from the deformation tensor $T_{\alpha\beta}$.

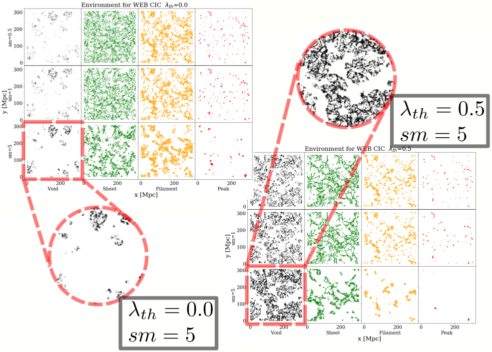


Fig. 7: Classification by environments for different cuts in λ_{th} y sm.

Feature Space → Galaxies.

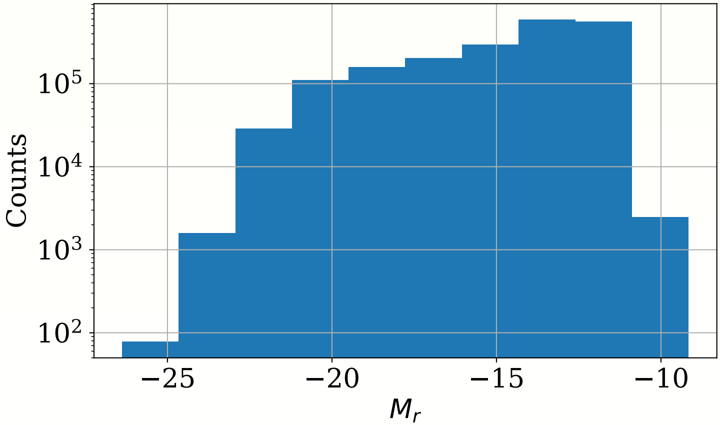
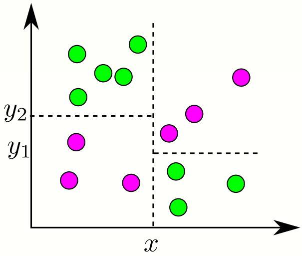
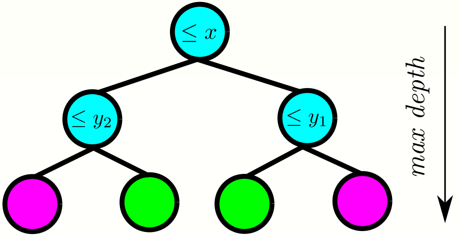


Fig. 8: Making cuts in the R band Luminosity.

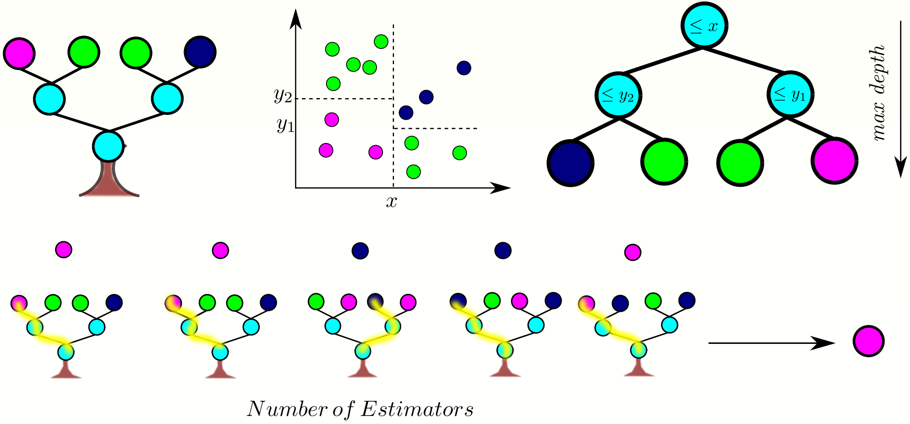
Classification Trees





Train: 50 % **Valid:** 30 %

Random Forest



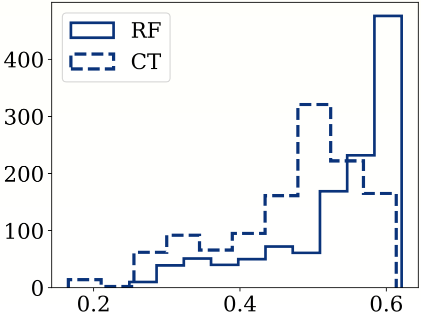
Train: 50 % **Valid:** 30 %

Universidad de los Andes

Results

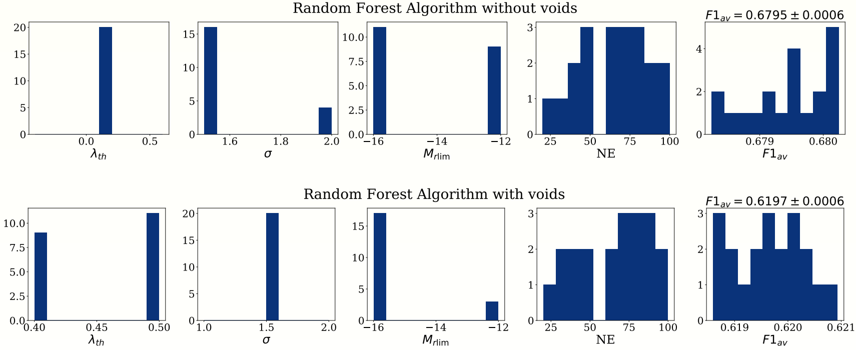
Classification Trees

	F1-score RF	F1-score CT
Peaks	0.53 ± 0.26	0.47 ± 0.25
Filaments	0.71 ± 0.05	0.65 ± 0.07
Sheets	0.58 ± 0.06	0.51 ± 0.11
Voids	0.31 ± 0.19	0.26 ± 0.17
Average including Voids	0.53 ± 0.09	0.47 ± 0.09
Average excluding Voids	0.61 ± 0.09	0.54 ± 0.09



- sm = 0.5, 1.0, 1.5, 2.0, 2.5
- λ_{th} = 0.1, 0.2, 0.3, 0.4, 0.5
- M_r < -20 , -18, -16, -14, -12
- NE = 20-100

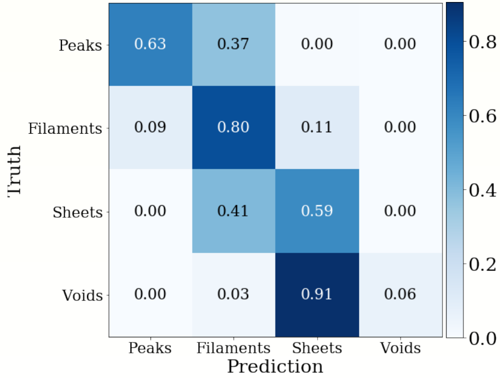
Evaluation in the Feature Space



Best Parameters

- sm = 1.5
- λ_{th} = 0.1
- M_r < -16
- NE = 80

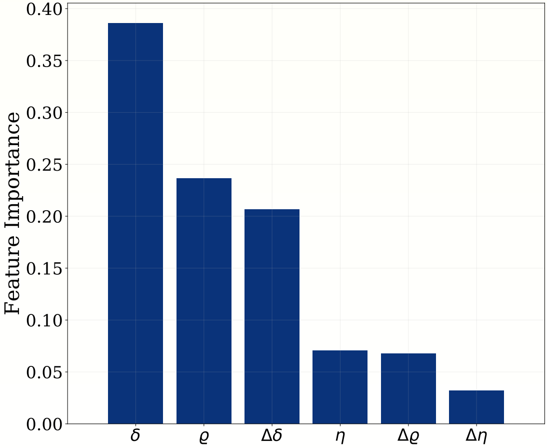
Confusion Matrix



- 63/100 peaks was correctly predicted.
- 6/100 voids was correctly predicted.

Features Importance

- δ: Average Distance
- ϱ: Pseudo-density
- $\Delta\delta$: Local average distance
- η: Number of connections



Qualitative comparison

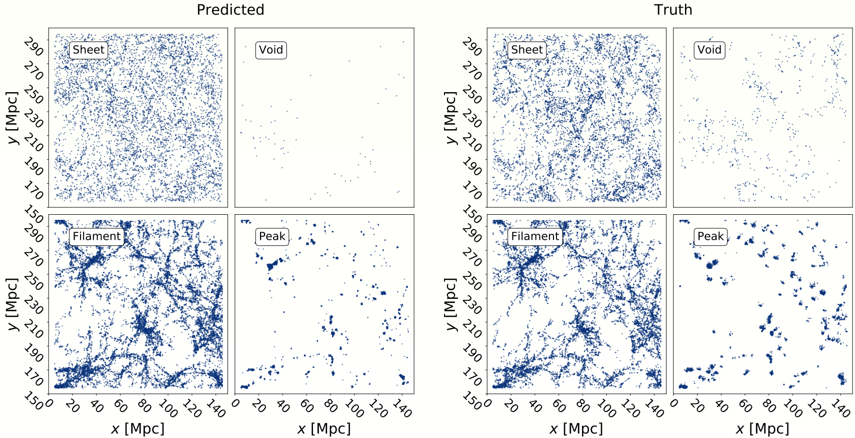
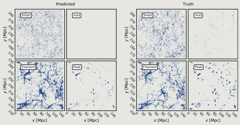


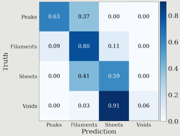
Fig. 9: Qualitative comparison between the predicted (right) and the truth (left) environments.

Conclusions

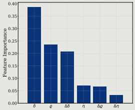
- It is possible to make a characterization of the T-Web through the implementation of the β-skeleton algorithm.



- The classification according to the confusion matrix is efficient to predict filaments, however, it is not good when trying to predict voids.



- The more important features to predict the cosmic web is the average distance δ and the pseudo-density ϱ.



Thanks!

jf.suarez@uniandes.edu.co

https://jsuarez314.gitlab.io