

# Probabilità e Statistica in Alta Dimensione:

## Esercitazione 4

### (Unnormalized Laplacian algorithm)

19 Dicembre 2025

Disponibile all'indirizzo:

<https://albertofachechi.com/HDP/Esercitazioni/Tracce/ES4.pdf>

Note per l'esercitazione:

[https://people.csail.mit.edu/dsontag/courses/ml14/notes/Luxburg07\\_tutorial\\_spectral\\_clustering.pdf](https://people.csail.mit.edu/dsontag/courses/ml14/notes/Luxburg07_tutorial_spectral_clustering.pdf)

Indirizzo email:

[alberto.fachechi@uniroma1.it](mailto:alberto.fachechi@uniroma1.it)

Dataset MNIST:

[https://albertofachechi.com/HDP/Esercitazioni/Materiale/MNIST\\_Dataset/mnist.mat](https://albertofachechi.com/HDP/Esercitazioni/Materiale/MNIST_Dataset/mnist.mat)

#### 1. Task 1.

Scrivere il codice per l'implementazione dell'algoritmo *unnormalized Laplacian clustering*, implementando almeno due modalità tra:  $\epsilon$ -neighborhood, mutual KNN, Gaussian similarity.

#### 2. Task 2.

Scrivere una funzione per generare un dataset in  $d = 2$  dimensioni costituito da  $n_c = 3$  nubi Gaussiane (con media e covarianza casuali) di  $n_p = 500$  punti ognuna. Determinare il clustering finale utilizzando l'algoritmo *unnormalized Laplacian* fissando  $k = n_c$ .

#### 3. Task 3.

Applicare l'algoritmo di clustering al dataset MNIST. Quale scelta nella costruzione del grafo risulta più affidabile?