UNIVERSITÀ
DEGLI STUDI
DI MILANO

Networks of Group Equivariant Non-Expansive Operators for Artificial Intelligence

Models, applications and interpretability

Candidate: Giovanni Bocchi

Supervisors: Prof. Alessandra Micheletti, Prof. Patrizio Frosini, Doct. Carmine Talarico

18th February 2025

www.unimi.it

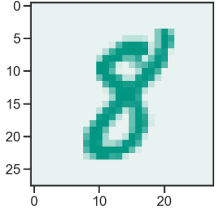
Outline

UNIVERSITÀ DEGLI STUDI DI MILANO

Introduction

Some initial motivation

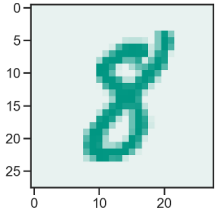
Many data that we encounter can be expressed using real valued **functions**. The similarities between such functions usually strongly depend on the observer.

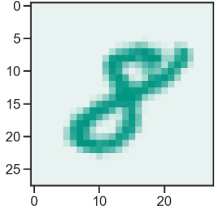


Some initial motivation

Many data that we encounter can be expressed using real valued **functions**. The similarities between such functions usually strongly depend on the observer.

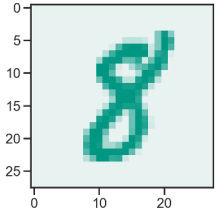
If data functions are subject to geometric transformations, we want our data processing agents to behave **consistently** to these transformations.



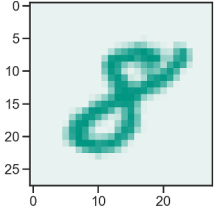


Some initial motivation

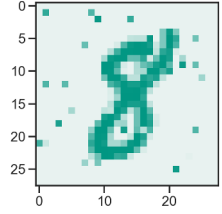
Many data that we encounter can be expressed using real valued **functions**. The similarities between such functions usually strongly depend on the observer.



If data functions are subject to geometric transformations, we want our data processing agents to behave **consistently** to these transformations.



Since, quite often, data are affected by some kind of noise, we also want to have agents that are **robust** to small perturbations in the data functions.



Ingredients

For introducing Group Equivariant Non-Expansive Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

To formally introduce Group Equivariant Non-Expansive Operators, I will consider the ingredients listed below:

- A first space of real valued functions $\Phi = \{\varphi \colon X \to \mathbb{R}\}$ with domain X, $\mathrm{dom}(\Phi) = X$.

Ingredients

For introducing Group Equivariant Non-Expansive Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

To formally introduce Group Equivariant Non-Expansive Operators, I will consider the ingredients listed below:

- A first space of real valued functions $\Phi = \{\varphi \colon X \to \mathbb{R}\}$ with domain X, $\mathrm{dom}(\Phi) = X$.
- A second space of real valued functions $\Psi = \{\psi \colon Y \to \mathbb{R}\}$ with domain Y, $\mathrm{dom}(\Psi) = Y$.

Ingredients

For introducing Group Equivariant Non-Expansive Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

To formally introduce Group Equivariant Non-Expansive Operators, I will consider the ingredients listed below:

- A first space of real valued functions $\Phi = \{\varphi \colon X \to \mathbb{R}\}$ with domain X, $\mathrm{dom}(\Phi) = X$.
- A second space of real valued functions $\Psi = \{\psi \colon Y \to \mathbb{R}\}$ with domain Y, $\mathrm{dom}(\Psi) = Y$.
- Φ and Ψ will be, respectively, the domain and codomain of the operators.

Ingredients

For introducing Group Equivariant Non-Expansive Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

To formally introduce Group Equivariant Non-Expansive Operators, I will consider the ingredients listed below:

- A first space of real valued functions $\Phi = \{\varphi \colon X \to \mathbb{R}\}$ with domain X, $\mathrm{dom}(\Phi) = X$.
- A second space of real valued functions $\Psi = \{\psi \colon Y \to \mathbb{R}\}$ with domain Y, $\mathrm{dom}(\Psi) = Y$.
- Φ and Ψ will be, respectively, the domain and codomain of the operators.
- A subgroup G of $\mathrm{Homeo}_\Phi(X) = \{g \in \mathrm{Homeo}(X) \mid \varphi \circ g \in \Phi, \varphi \circ g^{-1} \in \Phi, \forall\, \varphi \in \Phi\}$.

Ingredients

For introducing Group Equivariant Non-Expansive Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

To formally introduce Group Equivariant Non-Expansive Operators, I will consider the ingredients listed below:

- A first space of real valued functions $\Phi = \{\varphi \colon X \to \mathbb{R}\}$ with domain X, $\mathrm{dom}(\Phi) = X$.
- A second space of real valued functions $\Psi = \{\psi \colon Y \to \mathbb{R}\}$ with domain Y, $\mathrm{dom}(\Psi) = Y$.
- Φ and Ψ will be, respectively, the domain and codomain of the operators.
- A subgroup G of $\mathrm{Homeo}_\Phi(X) = \{g \in \mathrm{Homeo}(X) \mid \varphi \circ g \in \Phi, \varphi \circ g^{-1} \in \Phi, \forall\, \varphi \in \Phi\}$.
- An analogous subgroup K of $\mathrm{Homeo}_\Psi(Y)$.

Ingredients

For introducing Group Equivariant Non-Expansive Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

To formally introduce Group Equivariant Non-Expansive Operators, I will consider the ingredients listed below:

- A first space of real valued functions $\Phi = \{\varphi \colon X \to \mathbb{R}\}$ with domain X, $\mathrm{dom}(\Phi) = X$.
- A second space of real valued functions $\Psi = \{\psi \colon Y \to \mathbb{R}\}$ with domain Y, $\mathrm{dom}(\Psi) = Y$.
- Φ and Ψ will be, respectively, the domain and codomain of the operators.
- A subgroup G of $\mathrm{Homeo}_\Phi(X) = \{g \in \mathrm{Homeo}(X) \mid \varphi \circ g \in \Phi, \varphi \circ g^{-1} \in \Phi, \forall\, \varphi \in \Phi\}$.
- An analogous subgroup K of $\mathrm{Homeo}_\Psi(Y)$.
- A linking group homomorphism $T \colon G \to K$.

GENEOs definition

Given such ingredients, the formal definition of Group Equivariant Non-Expansive Operator (GENEO) was first given in (Bergomi, Frosini, et al., 2019):

Definition (GENEO)

A Group Equivariant Non-Expansive Operator F is a map between Φ and Ψ that, for a fixed homomorphism of groups T, has these two properties:

- **Equivariance:** $F(\varphi \circ g) = F(\varphi) \circ T(g)$ for all $\varphi \in \Phi$ and for all $g \in G$.
- **Non-Expansivity:** $||F(\varphi_1) - F(\varphi_2)||_\infty \leq ||\varphi_1 - \varphi_2||_\infty$ for all $\varphi_1, \varphi_2 \in \Phi$.

GENEOs definition

Given such ingredients, the formal definition of Group Equivariant Non-Expansive Operator (GENEO) was first given in (Bergomi, Frosini, et al., 2019):

Definition (GENEO)

A Group Equivariant Non-Expansive Operator F is a map between Φ and Ψ that, for a fixed homomorphism of groups T, has these two properties:

- **Equivariance:** $F(\varphi \circ g) = F(\varphi) \circ T(g)$ for all $\varphi \in \Phi$ and for all $g \in G$.
- **Non-Expansivity:** $||F(\varphi_1) - F(\varphi_2)||_\infty \leq ||\varphi_1 - \varphi_2||_\infty$ for all $\varphi_1, \varphi_2 \in \Phi$.

Usually if $\mathrm{dom}(\Phi) = \mathrm{dom}(\Psi)$ and $G = K$ the natural choice is to consider $T = \mathrm{id}_G$.

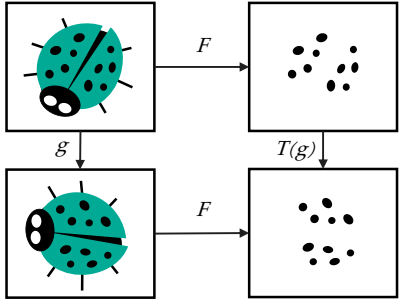
GENEOs properties
Interpretation

UNIVERSITÀ DEGLI STUDI DI MILANO

GENEOs properties

Interpretation

Equivariance: GENEOs must commute with a specific group of transformations of the data domain called equivariance group. We may say that GENEOs are able to filter out those transformations.

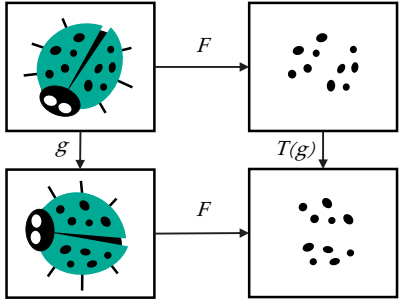


GENEOs properties

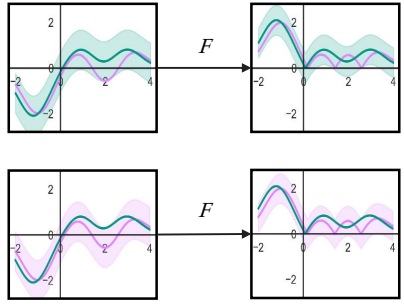
Interpretation

UNIVERSITÀ DEGLI STUDI DI MILANO

Equivariance: GENEOs must commute with a specific group of transformations of the data domain called equivariance group. We may say that GENEOs are able to filter out those transformations.



Non-expansivity: GENEOs simplify the metric structure of data, also providing stability to small perturbations. Non-expansivity is also crucial to derive topological properties of GENEO's space.



GENEO space and its properties

Once fixed two pairs (Φ, G) and (Ψ, K) and a homomorphism T we can consider $\mathscr{F}$ the space of all GENEOs between Φ and Ψ with respect to T. Moreover it is possible to introduce a metric for such space:

$$d_{\text{GENEO}}(F_1, F_2) = \sup_{\varphi \in \Phi} ||F_1(\varphi) - F_2(\varphi)||_\infty$$

It can be shown that the metric space $(\mathscr{F}, D_{\text{GENEO}})$ possesses interesting topological properties:

GENEO space and its properties

UNIVERSITÀ DEGLI STUDI DI MILANO

Once fixed two pairs (Φ, G) and (Ψ, K) and a homomorphism T we can consider $\mathscr{F}$ the space of all GENEOs between Φ and Ψ with respect to T. Moreover it is possible to introduce a metric for such space:

$$d_{\text{GENEO}}(F_1, F_2) = \sup_{\varphi \in \Phi} ||F_1(\varphi) - F_2(\varphi)||_\infty$$

It can be shown that the metric space $(\mathscr{F}, D_{\text{GENEO}})$ possesses interesting topological properties:

Theorem (Compactness of $\mathscr{F}$)

If both Φ and Ψ are compact in the topology induced by the sup norm distance then also $\mathscr{F}$ is compact in the topology induced by the metric d_{GENEO}.

GENEO space and its properties

UNIVERSITÀ DEGLI STUDI DI MILANO

Once fixed two pairs (Φ, G) and (Ψ, K) and a homomorphism T we can consider $\mathscr{F}$ the space of all GENEOs between Φ and Ψ with respect to T. Moreover it is possible to introduce a metric for such space:

$$d_{\text{GENEO}}(F_1, F_2) = \sup_{\varphi \in \Phi} ||F_1(\varphi) - F_2(\varphi)||_\infty$$

It can be shown that the metric space $(\mathscr{F}, D_{\text{GENEO}})$ possesses interesting topological properties:

Theorem (Compactness of $\mathscr{F}$)

If both Φ and Ψ are compact in the topology induced by the sup norm distance then also $\mathscr{F}$ is compact in the topology induced by the metric d_{GENEO}.

Theorem (Convexity of $\mathscr{F}$)

If Ψ is convex than the space $\mathscr{F}$ is also convex.

GENEO space and its properties
Interpretation

GENEO space and its properties

Interpretation

UNIVERSITÀ DEGLI STUDI DI MILANO

Compactness: must be intended as finite approximability of the space $\mathscr{F}$. In the sense that, for every tolerance $\varepsilon > 0$ it exists a finite number n of operators $F_1, \ldots, F_n$ such that:

$$\mathscr{F} \subseteq \bigcup_{i=1}^{n} B(F_i, \varepsilon)$$

This suggests that we can consider only a finite number of GENEOs to have a satisfactory approximation of the whole space.

GENEO space and its properties

Interpretation

UNIVERSITÀ DEGLI STUDI DI MILANO

Compactness: must be intended as finite approximability of the space $\mathscr{F}$. In the sense that, for every tolerance $\varepsilon > 0$ it exists a finite number n of operators $F_1, \ldots, F_n$ such that:

$$\mathscr{F} \subseteq \bigcup_{i=1}^{n} B(F_i, \varepsilon)$$

This suggests that we can consider only a finite number of GENEOs to have a satisfactory approximation of the whole space.

Convexity: Given non-negative weights $\alpha_1, \ldots, \alpha_n$ such that $\sum_{i=1}^{n} \alpha_i = 1$ and the same number of GENEOs, the operator

$$F_\alpha(\cdot) = \sum_{i=1}^{n} \alpha_i F_i(\cdot)$$

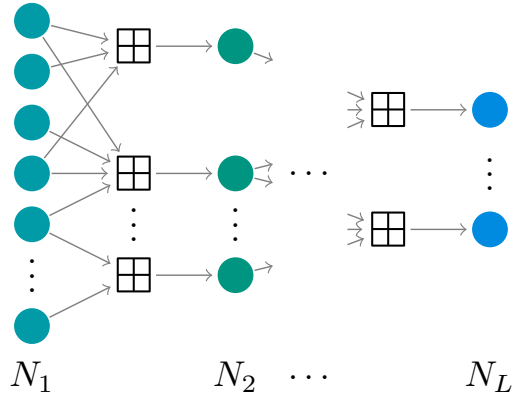
is still a GENEO. This property can be seen as the precursor of other compositional rules that in the thesis I called **aggregation functions**:

- min, max, etc.

GENEO networks

Aggregation functions allow to aggregate the information coming from different operators, in particular they enable to link different operators in order to obtain networks where:

1. **nodes** are families of GENEOs (possibly parameter dependent).
2. **arrows** represents aggregation functions between elements of different families.



Permutants

[Generalized] permutants

UNIVERSITÀ DEGLI STUDI DI MILANO

In the applications it is crucial to find the adequate GENEOs, a first option to explore the space is to use aggregation functions. Another noteworthy algebraic tool for obtaining GENEOs are **permutants**. Permutants and their generalizations have been introduced and studied in (Camporesi, Frosini, and Quercioli, 2018; Conti, Frosini, and Quercioli, 2022; Ahmad, Ferri, and Frosini, 2023).

Definition (Generalized permutant)

A finite subset $H \subseteq X^Y$ is a generalized permutant for $T\colon G \to K$ if $g\,H\,T(g^{-1}) \subseteq H$ for every $g \in G$ or if H is the empty set.

[Generalized] permutants

UNIVERSITÀ DEGLI STUDI DI MILANO

In the applications it is crucial to find the adequate GENEOs, a first option to explore the space is to use aggregation functions. Another noteworthy algebraic tool for obtaining GENEOs are **permutants**. Permutants and their generalizations have been introduced and studied in (Camporesi, Frosini, and Quercioli, 2018; Conti, Frosini, and Quercioli, 2022; Ahmad, Ferri, and Frosini, 2023).

Definition (Generalized permutant)

A finite subset $H \subseteq X^Y$ is a generalized permutant for $T\colon G \to K$ if $g\, H\, T(g^{-1}) \subseteq H$ for every $g \in G$ or if H is the empty set.

If particular if $\mathrm{dom}(\Phi) = \mathrm{dom}(\Psi) = X$ and $T = \mathrm{id}_G$ we instead require that $H \subseteq \mathrm{Homeo}(X)$ and we call them simply permutants.

Building GENEOs with permutants

UNIVERSITÀ DEGLI STUDI DI MILANO

Given a permutant, we have a straightforward way to obtain a GENEO for the corresponding homomorphism T.

Proposition (A permutant defines a GENEO)

If H is a generalized permutant for $T\colon G \to K$ then the operator $F_H\colon \Phi \to \Psi$

$$F_H(\varphi) = \begin{cases} \frac{1}{|H|} \sum_{h \in H} \varphi \circ h & \text{if } |H| > 0 \\ \mathbf{0} & \text{otherwise} \end{cases}$$

is a GENEO provided that $F_H(\Phi) \subseteq \Psi$.

Building GENEOs with permutants

Given a permutant, we have a straightforward way to obtain a GENEO for the corresponding homomorphism T.

Proposition (A permutant defines a GENEO)

If H is a generalized permutant for $T\colon G \to K$ then the operator $F_H\colon \Phi \to \Psi$

$$F_H(\varphi) = \begin{cases} \frac{1}{|H|} \sum_{h \in H} \varphi \circ h & \text{if } |H| > 0 \\ \mathbf{0} & \text{otherwise} \end{cases}$$

is a GENEO provided that $F_H(\Phi) \subseteq \Psi$.

Since H must be finite, the previous proposition provides a practical and possibly also efficient way of **computing** GENEOs: when G is infinite, but also when G is finite and very large provided that $|H| \ll |G|$.

Representation theorem

UNIVERSITÀ DEGLI STUDI DI MILANO

Given a permutant, we can always define a GENEO, is it also true that a GENEOs are always obtained from permutants ? In the context of a finite X with $\Phi = \Psi = \mathbb{R}^X$, we could establish the following **representation** result appeared in (Bocchi, Botteghi, et al., 2023):

Representation theorem

Given a permutant, we can always define a GENEO, is it also true that a GENEOs are always obtained from permutants ? In the context of a finite X with $\Phi = \Psi = \mathbb{R}^X$, we could establish the following **representation** result appeared in (Bocchi, Botteghi, et al., 2023):

Theorem (G.B., S. Botteghi, M. Brasini, P. Frosini, N. Quercioli)

If G acts transitively on the finite set X then an operator F from Φ to Φ is a linear GENEO if and only if it exists a permutant measure μ such that $\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| \leq 1$ that allows writing F as:

$$F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} (\varphi \circ h^{-1}) \mu(h)$$

Representation theorem

Given a permutant, we can always define a GENEO, is it also true that a GENEOs are always obtained from permutants ? In the context of a finite X with $\Phi = \Psi = \mathbb{R}^X$, we could establish the following **representation** result appeared in (Bocchi, Botteghi, et al., 2023):

Theorem (G.B., S. Botteghi, M. Brasini, P. Frosini, N. Quercioli)

If G acts transitively on the finite set X then an operator F from Φ to Φ is a linear GENEO if and only if it exists a permutant measure μ such that $\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| \leq 1$ that allows writing F as:

$$F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} (\varphi \circ h^{-1}) \mu(h)$$

Definition (G.B., S. Botteghi, M. Brasini, P. Frosini, N. Quercioli)

A finite signed measure μ on $\mathrm{Aut}(X)$ is called a permutant measure with respect to G if each subset H of $\mathrm{Aut}(X)$ is measurable and $\mu(H) = \mu(gHg^{-1})$ for every $g \in G$.

GENEOnet

Protein pocket detection

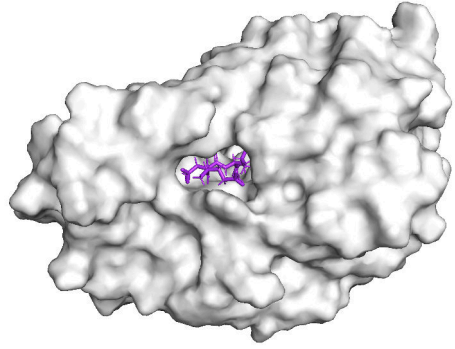
The second and main instance of a GENEO network considered in the thesis, is linked to an industrial problem that was studied in collaboration with the pharmaceutical company **Dompé Farmaceutici S.p.A.**

https://geneonet.exscalate.eu

Protein pocket detection

The second and main instance of a GENEO network considered in the thesis, is linked to an industrial problem that was studied in collaboration with the pharmaceutical company **Dompé Farmaceutici S.p.A.**

https://geneonet.exscalate.eu



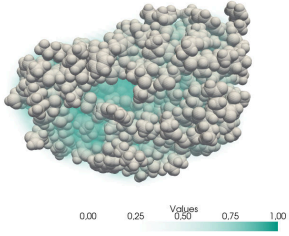
Model construction
Data and GENEOs

UNIVERSITÀ DEGLI STUDI DI MILANO

Model construction
Data and GENEOs

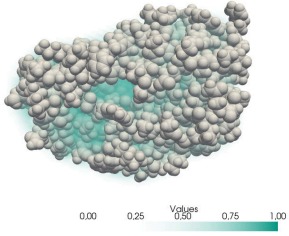
UNIVERSITÀ DEGLI STUDI DI MILANO

The input functions $\varphi_i \colon B \subseteq \mathbb{R}^3 \to \mathbb{R}$ are modeled as discretizations of 8 potential functions on a grid of voxels.

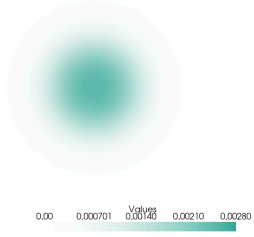


Model construction
Data and GENEOs

The input functions $\varphi_i \colon B \subseteq \mathbb{R}^3 \to \mathbb{R}$ are modeled as discretizations of 8 potential functions on a grid of voxels.

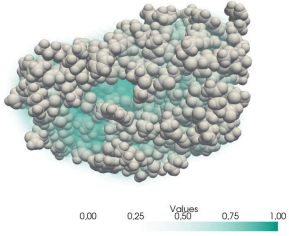


We chose convolutional operators $F_i(\cdot) = \cdot * K_i$ with radial kernels $K_i \colon \mathbb{R}^3 \to \mathbb{R}$. Each kernel depends on a shape parameter σ_i.

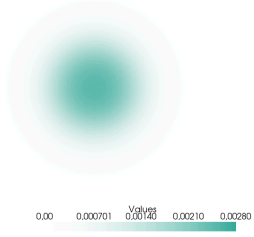


Model construction
Data and GENEOs

The input functions $\varphi_i \colon B \subseteq \mathbb{R}^3 \to \mathbb{R}$ are modeled as discretizations of 8 potential functions on a grid of voxels.



We chose convolutional operators $F_i(\cdot) = \cdot * K_i$ with radial kernels $K_i \colon \mathbb{R}^3 \to \mathbb{R}$. Each kernel depends on a shape parameter σ_i.



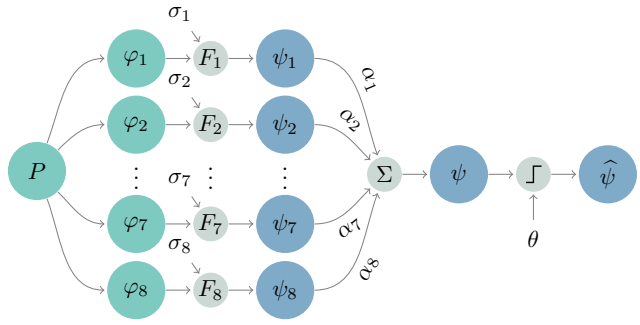
The properties of convolution, combined with the kernels invariance to rotations, guarantee that F_i are **equivariant** with respect to the group G of rigid motions of the space $\mathbb{R}^3$.

Moreover, by normalizing the kernels with respect to L^1 norm, F_i also acquire non-expansivity.

Model construction
Architecture

The eight GENEOs F_i are aggregated using convex combination with convex coefficients $\alpha_1, \ldots, \alpha_8$.

The network combines the intermediate outputs ψ_i into a global output $\psi : B \subset \mathbb{R}^3 \to [0, 1]$ which encodes the **likelihood** that a voxel belongs to a pocket.



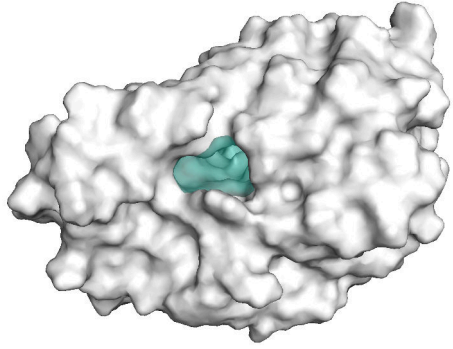
Model evaluation
Prediction

UNIVERSITÀ DEGLI STUDI DI MILANO

From the global output ψ the binary output $\widehat{\psi}$ is obtained by considering all the voxels x such that $\psi(x) \geq \theta$. By segmenting such set into connected components, we obtain the final prediction of GENEOnet.

Model evaluation
Prediction

From the global output ψ the binary output $\widehat{\psi}$ is obtained by considering all the voxels x such that $\psi(x) \geq \theta$. By segmenting such set into connected components, we obtain the final prediction of GENEOnet.



Model evaluation

Training

UNIVERSITÀ DEGLI STUDI DI MILANO

The model called GENEOnet (Bocchi, Frosini, et al., 2022; Bocchi, Frosini, et al., 2024) has 17 **learnable parameters** that can be estimated by optimizing a loss function of the binary prediction $\widehat{\psi}$ and the ground truth $\Pi\colon B \subseteq \mathbb{R}^3 \to \{0,1\}$ on a training set of data.

$$\mathcal{L}(\widehat{\psi}, \Pi) = \frac{|\widehat{\psi} \wedge \Pi| + \kappa|(\mathbf{1} - \widehat{\psi}) \wedge (\mathbf{1} - \Pi)|}{|\Pi| + \kappa|\mathbf{1} - \Pi|} \in [0,1]$$

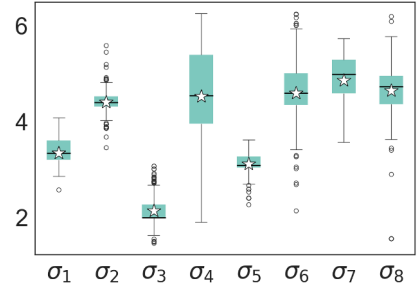
Model evaluation

Training

The model called GENEOnet (Bocchi, Frosini, et al., 2022; Bocchi, Frosini, et al., 2024) has 17 **learnable parameters** that can be estimated by optimizing a loss function of the binary prediction $\widehat{\psi}$ and the ground truth $\Pi\colon B \subseteq \mathbb{R}^3 \to \{0,1\}$ on a training set of data.

$$\mathcal{L}(\widehat{\psi},\Pi) = \frac{|\widehat{\psi} \wedge \Pi| + \kappa|(\mathbf{1}-\widehat{\psi}) \wedge (\mathbf{1}-\Pi)|}{|\Pi| + \kappa|\mathbf{1}-\Pi|} \in [0,1]$$

The optimization was repeated 200 times using random training sets of size 200. This analysis allows us to also evaluate the **sensitivity** of the coefficient's estimates.



Model evaluation
Scoring

UNIVERSITÀ DEGLI STUDI DI MILANO

For applications to virtual screenings is important to rank predictions by computing a pocket score.

The **score** is obtained as a weighed average of ψ restricted to one connected component of $\{x \in \mathbb{R}^3 \mid \psi(x) \geq \theta\}$.

Model evaluation
Scoring

UNIVERSITÀ DEGLI STUDI DI MILANO

For applications to virtual screenings is important to rank predictions by computing a pocket score.

The **score** is obtained as a weighed average of ψ restricted to one connected component of $\{x \in \mathbb{R}^3 \mid \psi(x) \geq \theta\}$.

We define the **overlap** between a predicted pocket $\widehat{\Pi}$ and the the true pocket Π:

$$\mathcal{V}(\widehat{\Pi}, \Pi) = \frac{|\Pi \wedge \widehat{\Pi}|}{|\Pi|}$$

we say that a predicted pocket $\widehat{\Pi}$ **hits** the ground truth Π if

$$\widehat{\Pi} = \arg\max_{C} \mathcal{V}(C, \Pi)$$

Model evaluation
Scoring

UNIVERSITÀ DEGLI STUDI DI MILANO

For applications to virtual screenings is important to rank predictions by computing a pocket score.

The **score** is obtained as a weighed average of ψ restricted to one connected component of $\{x \in \mathbb{R}^3 \mid \psi(x) \geq \theta\}$.

We define the **overlap** between a predicted pocket $\widehat{\Pi}$ and the the true pocket Π:

$$\mathcal{V}(\widehat{\Pi}, \Pi) = \frac{|\Pi \wedge \widehat{\Pi}|}{|\Pi|}$$

we say that a predicted pocket $\widehat{\Pi}$ **hits** the ground truth Π if

$$\widehat{\Pi} = \arg\max_{C} \mathcal{V}(C, \Pi)$$

Given a dataset $(P_t, \Pi_t)_{t=1}^n$ we can assess the scoring capabilities of a model $\mathcal{M}$ by computing:

$$\mathcal{H}_j = \frac{1}{n} \sum_{t=1}^{n} \Delta_j(\mathcal{M}(P_t), \Pi_t)$$

where, for every $j \geq 1$, $\Delta_j(\mathcal{M}(P_t), \Pi_t)$ is 1 if the j-th ranked pocket hits Π_t or 0 otherwise.

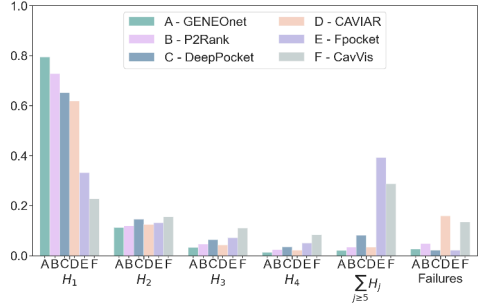
Model comparison

Scoring on BIND_TEST

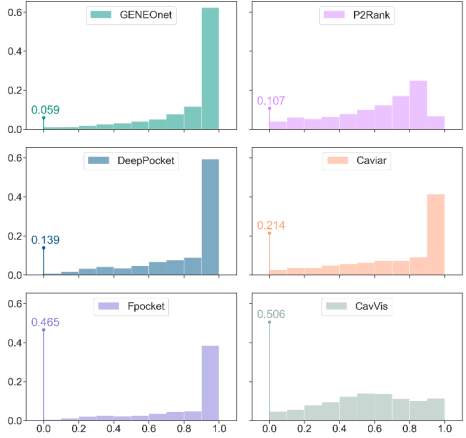
Method	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$	$\mathcal{H}_4$	$\sum_{j\geq 5}\mathcal{H}_j$	failures
GENEOnet	0.794	0.113	0.033	0.013	0.021	0.026
P2Rank	0.728	0.119	0.046	0.024	0.035	0.048
DeepPocket	0.651	0.146	0.064	0.035	0.082	0.022
CAVIAR	0.618	0.125	0.043	0.022	0.033	0.163
Fpocket	0.332	0.131	0.072	0.050	0.393	0.022
CavVis	0.228	0.156	0.110	0.084	0.287	0.135

Method	$\mathcal{T}_1$	$\mathcal{T}_2$	$\mathcal{T}_3$	$\mathcal{T}_4$	$\sum_{j\geq 1}\mathcal{T}_j$
GENEOnet	**0.794**	**0.907**	**0.940**	**0.953**	0.974
P2Rank	0.728	0.847	0.893	0.917	0.952
DeepPocket	0.651	0.797	0.861	0.896	**0.978**
CAVIAR	0.618	0.743	0.786	0.808	0.841
Fpocket	0.332	0.463	0.535	0.585	**0.978**
CavVis	0.228	0.384	0.494	0.578	0.865

$$\mathcal{T}_j = \sum_{i=1}^{j} \mathcal{H}_i$$

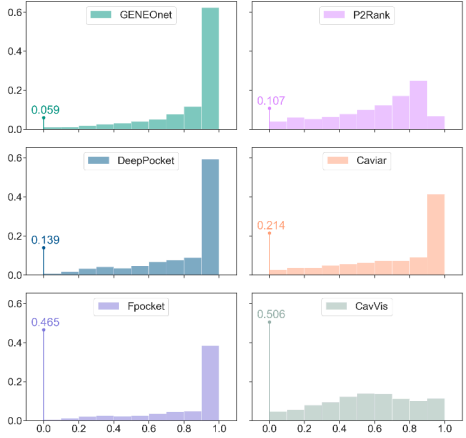


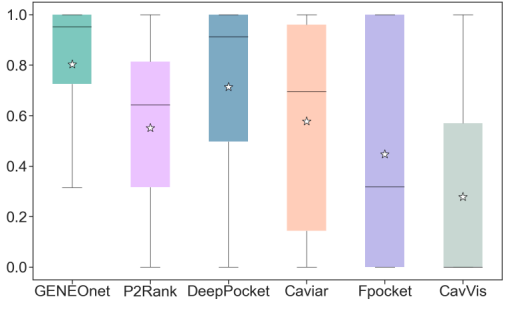
Model comparison
Overlap on BIND_TEST



Model comparison

Overlap on BIND_TEST





Model explainability

UNIVERSITÀ DEGLI STUDI DI MILANO

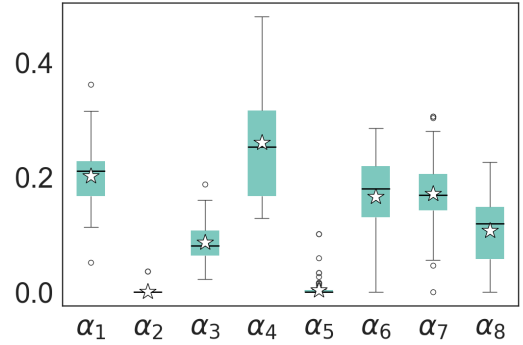
One of the great advantages of GENEOnet compared to other methods is its intrinsic **explainability** and **reliability** (Guidotti, Monreale, et al., 2021; Giudici, 2024).

Unit	Channel	σ_i	α_j	θ
1	Distance	3.110	0.362	0.756
2	Gravitational	5.197	0.002	
3	Electrostatic	2.561	0.054	
4	Lipophilic	4.678	0.338	
5	Hydrophilic	3.545	0.001	
6	Polar	6.166	0.185	
7	HB Acceptor	4.186	0.056	
8	HB Donor	3.908	0.001	

Model explainability

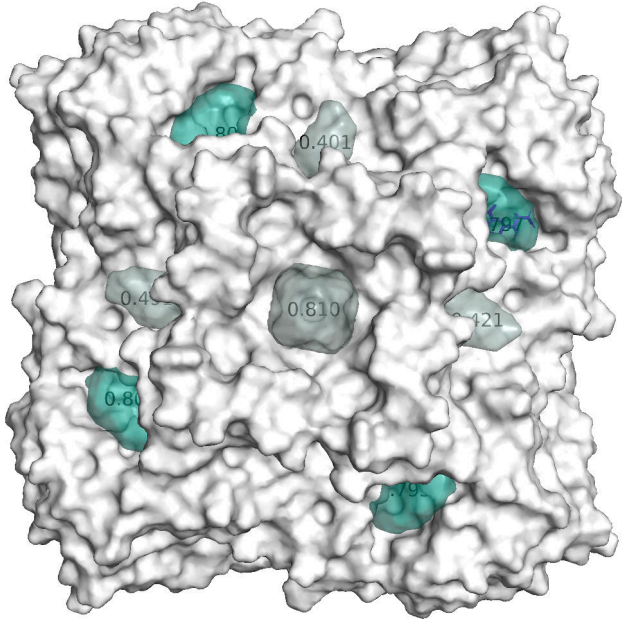
One of the great advantages of GENEOnet compared to other methods is its intrinsic **explainability** and **reliability** (Guidotti, Monreale, et al., 2021; Giudici, 2024).

Unit	Channel	σ_i	α_j	θ
1	Distance	3.110	0.362	0.756
2	Gravitational	5.197	0.002	
3	Electrostatic	2.561	0.054	
4	Lipophilic	4.678	0.338	
5	Hydrophilic	3.545	0.001	
6	Polar	6.166	0.185	
7	HB Acceptor	4.186	0.056	
8	HB Donor	3.908	0.001	



The Figure shows GENEOnet prediction for protein ID 2QWE which is made of **four symmetrical units** very similar to each other in their spatial arrangement.

This example shows equivariance and non-expansivity jointly in action.



GraphGENEOs

Graph GENEOs
Isomorphism problem

UNIVERSITÀ DEGLI STUDI DI MILANO

Given a pair of undirected and unweighted graphs $\Gamma_1 = (V_1, E_1)$ and $\Gamma_2 = (V_2, E_2)$ decide whether there exists a bijection $f\colon V_1 \to V_2$ such that $\{u, v\} \in E_1$ if and only if $\{f(u), f(v)\} \in E_2$.

L. Babai (Babai, 2019) proved that such problem is solvable in quasi-polynomial time[a]. Anyway, **no polynomial time** algorithm is currently known and it is also unclear if it is a NP-complete problem (most experts believe it to be an NP-intermediate problem).

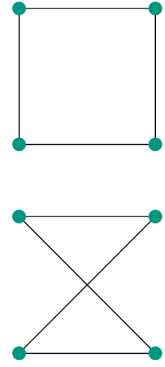
Graph GENEOs
Isomorphism problem

UNIVERSITÀ DEGLI STUDI DI MILANO

Given a pair of undirected and unweighted graphs $\Gamma_1 = (V_1, E_1)$ and $\Gamma_2 = (V_2, E_2)$ decide whether there exists a bijection $f\colon V_1 \to V_2$ such that $\{u, v\} \in E_1$ if and only if $\{f(u), f(v)\} \in E_2$.

L. Babai (Babai, 2019) proved that such problem is solvable in quasi-polynomial time[a]. Anyway, **no polynomial time** algorithm is currently known and it is also unclear if it is a NP-complete problem (most experts believe it to be an NP-intermediate problem).

[a] Time complexity in the class $2^{\mathcal{O}((\log n)^C)}$



Graph GENEOs
Isomorphism tests

UNIVERSITÀ DEGLI STUDI DI MILANO

An **isomorphism test** is a procedure that, given a pair of graphs (Γ_1, Γ_2) gives at least a sufficient condition for non isomorphism.

1. A negative result means the two graphs are surely non isomorphic.
    1. **Exact tests:** A positive result implies the two graphs are isomorphic.

Graph GENEOs
Isomorphism tests

UNIVERSITÀ DEGLI STUDI DI MILANO

An **isomorphism test** is a procedure that, given a pair of graphs (Γ_1, Γ_2) gives at least a sufficient condition for non isomorphism.

1. A negative result means the two graphs are surely non isomorphic.
   1. **Exact tests:** A positive result implies the two graphs are isomorphic.
   2. **Inexact tests:** After a positive result the graphs can be isomorphic as well as non isomorphic.

Graph GENEOs
Isomorphism tests

UNIVERSITÀ DEGLI STUDI DI MILANO

An **isomorphism test** is a procedure that, given a pair of graphs (Γ_1, Γ_2) gives at least a sufficient condition for non isomorphism.

1. A negative result means the two graphs are surely non isomorphic.
    1. **Exact tests:** A positive result implies the two graphs are isomorphic.
    2. **Inexact tests:** After a positive result the graphs can be isomorphic as well as non isomorphic.
1. Exact tests are **slow** (no polynomial algorithm available).

UNIVERSITÀ DEGLI STUDI DI MILANO

An **isomorphism test** is a procedure that, given a pair of graphs (Γ_1, Γ_2) gives at least a sufficient condition for non isomorphism.

1. A negative result means the two graphs are surely non isomorphic.
   1. **Exact tests:** A positive result implies the two graphs are isomorphic.
   2. **Inexact tests:** After a positive result the graphs can be isomorphic as well as non isomorphic.

1. Exact tests are **slow** (no polynomial algorithm available).
2. Inexact tests are usually **fast** (can achieve linear complexity).

Mathematical setting

From now on we will consider **undirected** and **unweighed** graphs Γ with (at most) N nodes seen as subgraphs of the complete graph K_N as described in (Bocchi, Ferri, and Frosini, 2025).

1. As domain $X = \{\{v_i, v_j\} : i \neq j; i, j \in \mathbb{N}_N\}$.

Mathematical setting

From now on we will consider **undirected** and **unweighed** graphs Γ with (at most) N nodes seen as subgraphs of the complete graph K_N as described in (Bocchi, Ferri, and Frosini, 2025).

1. As domain $X = \{\{v_i, v_j\} : i \neq j; i, j \in \mathbb{N}_N\}$.
2. As functional space $\Phi = \{\varphi : X \to \{0, 1\}\}$.

Mathematical setting

UNIVERSITÀ DEGLI STUDI DI MILANO

From now on we will consider **undirected** and **unweighed** graphs Γ with (at most) N nodes seen as subgraphs of the complete graph K_N as described in (Bocchi, Ferri, and Frosini, 2025).

1. As domain $X = \{\{v_i, v_j\} \colon i \neq j; i, j \in \mathbb{N}_N\}$.
2. As functional space $\Phi = \{\varphi \colon X \to \{0, 1\}\}$.
3. As group $G = \{g \colon X \to X \mid g(\{v_i, v_j\}) = \{v_{\sigma(i)}, v_{\sigma(j)}\}\}$ where σ is a permutation of $\{1, \dots, N\}$.

Mathematical setting

From now on we will consider **undirected** and **unweighed** graphs Γ with (at most) N nodes seen as subgraphs of the complete graph K_N as described in (Bocchi, Ferri, and Frosini, 2025).

1. As domain $X = \{\{v_i, v_j\} : i \neq j; i, j \in \mathbb{N}_N\}$.
2. As functional space $\Phi = \{\varphi : X \to \{0, 1\}\}$.
3. As group $G = \{g : X \to X \mid g(\{v_i, v_j\}) = \{v_{\sigma(i)}, v_{\sigma(j)}\}\}$ where σ is a permutation of $\{1, \ldots, N\}$.
4. A graph Γ is uniquely represented by $\varphi \in \Phi$ denoted by $\Gamma \sim \varphi$.

Mathematical setting

From now on we will consider **undirected** and **unweighed** graphs Γ with (at most) N nodes seen as subgraphs of the complete graph K_N as described in (Bocchi, Ferri, and Frosini, 2025).

1. As domain $X = \{\{v_i, v_j\} : i \neq j; i, j \in \mathbb{N}_N\}$.
2. As functional space $\Phi = \{\varphi : X \to \{0, 1\}\}$.
3. As group $G = \{g : X \to X \mid g(\{v_i, v_j\}) = \{v_{\sigma(i)}, v_{\sigma(j)}\}\}$ where σ is a permutation of $\{1, \dots, N\}$.
4. A graph Γ is uniquely represented by $\varphi \in \Phi$ denoted by $\Gamma \sim \varphi$.
5. Γ_1, Γ_2 isomorphic $\iff$ it exists $g \in G$ such that $\varphi_1 = \varphi_2 \circ g$.

Mathematical setting

From now on we will consider **undirected** and **unweighed** graphs Γ with (at most) N nodes seen as subgraphs of the complete graph K_N as described in (Bocchi, Ferri, and Frosini, 2025).

1. As domain $X = \{\{v_i, v_j\} \colon i \neq j; i, j \in \mathbb{N}_N\}$.
2. As functional space $\Phi = \{\varphi \colon X \to \{0, 1\}\}$.
3. As group $G = \{g \colon X \to X \mid g(\{v_i, v_j\}) = \{v_{\sigma(i)}, v_{\sigma(j)}\}\}$ where σ is a permutation of $\{1, \dots, N\}$.
4. A graph Γ is uniquely represented by $\varphi \in \Phi$ denoted by $\Gamma \sim \varphi$.
5. Γ_1, Γ_2 isomorphic $\iff$ it exists $g \in G$ such that $\varphi_1 = \varphi_2 \circ g$.

We will also consider auxiliary graphs Λ seen as **subgraphs** of a smaller complete graph K_k with $k \ll N$.

1. As domain $Y = \{\{w_i, w_j\} \colon i \neq j; i, j \in \mathbb{N}_k\}$.
2. Two functional spaces: $\Psi_0 = \{\varphi \colon Y \to \{0, 1\}\}$ and $\Psi = \{\varphi \colon Y \to [0, 1]\}$.
3. A subgraph Λ is uniquely represented by $\psi_0 \in \Psi_0$ denoted by $\Lambda \sim \psi_0$.

Subgraph permutants
Definition

UNIVERSITÀ DEGLI STUDI DI MILANO

Using subgraphs, we introduced the following set that proves to be a **generalized permutant**. Given $\Lambda \sim \psi_0$ as a subgraph of $\Gamma \sim \varphi$ is the set of functions from Y to X of the form:

$$H_\Lambda^\varphi = \{h\colon Y \to X \mid (\varphi \circ h)(\{w_i, w_j\}) = 1 \iff \psi_0(\{w_i, w_j\}) = 1 \\ \{w_i, w_j\} \in Y, h \text{ is injective}\}$$

Subgraph permutants

Definition

UNIVERSITÀ DEGLI STUDI DI MILANO

Using subgraphs, we introduced the following set that proves to be a **generalized permutant**. Given $\Lambda \sim \psi_0$ as a subgraph of $\Gamma \sim \varphi$ is the set of functions from Y to X of the form:

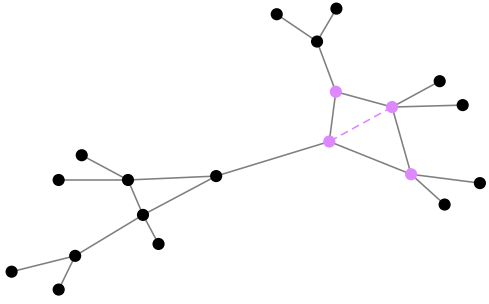
$$H_\Lambda^\varphi = \{h\colon Y \to X \mid (\varphi \circ h)(\{w_i, w_j\}) = 1 \iff \psi_0(\{w_i, w_j\}) = 1 \quad \{w_i, w_j\} \in Y, h \text{ is injective}\}$$

Proposition (G.B., M. Ferri, P. Frosini)

Any set H_Λ^φ is a generalized permutant for $T\colon G_\Gamma \to \{id_Y\}$.

Subgraph permutants

Example: Λ is a 4-cycle

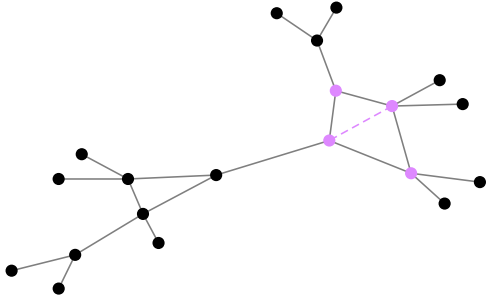


(a) $\Gamma_1 \quad H_\Lambda^{\varphi_1} = \emptyset$

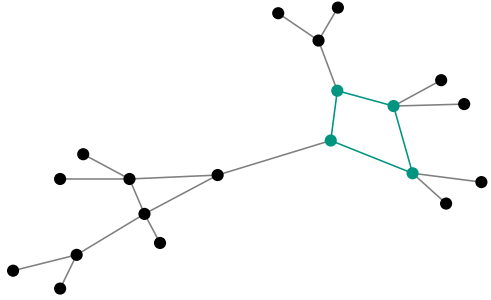
Subgraph permutants

Example: Λ is a 4-cycle

UNIVERSITÀ DEGLI STUDI DI MILANO



(a) $\Gamma_1 \quad H_\Lambda^{\varphi_1} = \emptyset$



(b) $\Gamma_2 \quad |H_\Lambda^{\varphi_2}| = 8$

Subgraph GENEOs

UNIVERSITÀ DEGLI STUDI DI MILANO

We proved this new proposition in order to admit the case in which the permutant is φ dependent:

Proposition (G.B., M. Ferri, P. Frosini)

The operator F_Λ defined as

$$F_\Lambda(\varphi) = \frac{1}{n_\Lambda |G_\Lambda|} \sum_{h \in H_\Lambda^\varphi} \varphi \circ h$$

is a linear GENEO between the pairs (Φ, G), $(\Psi, \{id_Y\})$ with respect to the trivial homomorphism $T : G \to \{id_Y\}$, where n_Λ is the number of occurrences of Λ in a complete graph with N nodes.

GENEO based isomorphism test

Theorem

UNIVERSITÀ DEGLI STUDI DI MILANO

Theorem (G.B., M. Ferri, P. Frosini)

Given two graphs $\Gamma_1 \sim \varphi_1$ and $\Gamma_2 \sim \varphi_2$, for every choice of $p \geq 1$ and $\Lambda_1, \ldots, \Lambda_p$ the map:

$$F_p(\varphi_1, \varphi_2) = \max_{j=1}^{p} F_{\{j\}}(\varphi_1, \varphi_2) \quad \text{where} \quad F_{\{j\}}(\varphi_1, \varphi_2) = \left\| \frac{1}{2} \left| F_{\Lambda_j}(\varphi_1) - F_{\Lambda_j}(\varphi_2) \right| \right\|_\infty$$

is a GENEO from $(\Phi \times \Phi, G \times G)$ to $(\mathbb{R}, \{id_S\})$ with respect to the trivial homomorphism $T\colon G \times G \to \{id_S\}$ ($\mathbb{R}$ is identified with the set of real-valued functions on a singleton S). Moreover, the followings hold true:

1. *If Γ_1 and Γ_2 are isomorphic then $F_p(\varphi_1, \varphi_2) = 0$.*
2. *$F_p(\varphi_1, \varphi_2) \neq 0$ then Γ_1 and Γ_2 are not isomorphic.*

GENEO based isomorphism test

Theorem

UNIVERSITÀ DEGLI STUDI DI MILANO

Theorem (G.B., M. Ferri, P. Frosini)

Given two graphs $\Gamma_1 \sim \varphi_1$ and $\Gamma_2 \sim \varphi_2$, for every choice of $p \geq 1$ and $\Lambda_1, \ldots, \Lambda_p$ the map:

$$F_p(\varphi_1, \varphi_2) = \max_{j=1}^{p} F_{\{j\}}(\varphi_1, \varphi_2) \quad \text{where} \quad F_{\{j\}}(\varphi_1, \varphi_2) = \left\| \frac{1}{2} \left| F_{\Lambda_j}(\varphi_1) - F_{\Lambda_j}(\varphi_2) \right| \right\|_\infty$$

is a GENEO from $(\Phi \times \Phi, G \times G)$ to $(\mathbb{R}, \{id_S\})$ with respect to the trivial homomorphism $T\colon G \times G \to \{id_S\}$ ($\mathbb{R}$ is identified with the set of real-valued functions on a singleton S). Moreover, the followings hold true:

1. *If Γ_1 and Γ_2 are isomorphic then $F_p(\varphi_1, \varphi_2) = 0$.*
2. *$F_p(\varphi_1, \varphi_2) \neq 0$ then Γ_1 and Γ_2 are not isomorphic.*

Thus $\operatorname{sgn}(F_p(\varphi_1, \varphi_2))$ provides an inexact isomorphism test !

Case study
r-regular graphs

UNIVERSITÀ DEGLI STUDI DI MILANO

How to choose the right subgraphs to obtain a **simple** and **efficient** test ?

We considered *r*-regular graphs, which are known to be hard to distinguish for methods like the Weisfeiler-Leman test (Weisfeiler and Leman, 1968) and (Message Passing) Graph Neural Networks (Scarselli, Gori, et al., 2009; Morris, Ritzert, et al., 2019).

Case study
r-regular graphs

UNIVERSITÀ DEGLI STUDI DI MILANO

How to choose the right subgraphs to obtain a **simple** and **efficient** test ?

We considered r-regular graphs, which are known to be hard to distinguish for methods like the Weisfeiler-Leman test (Weisfeiler and Leman, 1968) and (Message Passing) Graph Neural Networks (Scarselli, Gori, et al., 2009; Morris, Ritzert, et al., 2019).

A graph is called r-regular for $r \geq 2$ if every node has degree exactly equal to r.

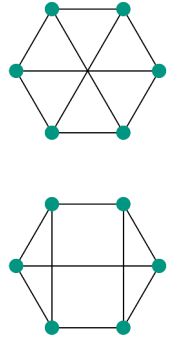
Case study
r-regular graphs

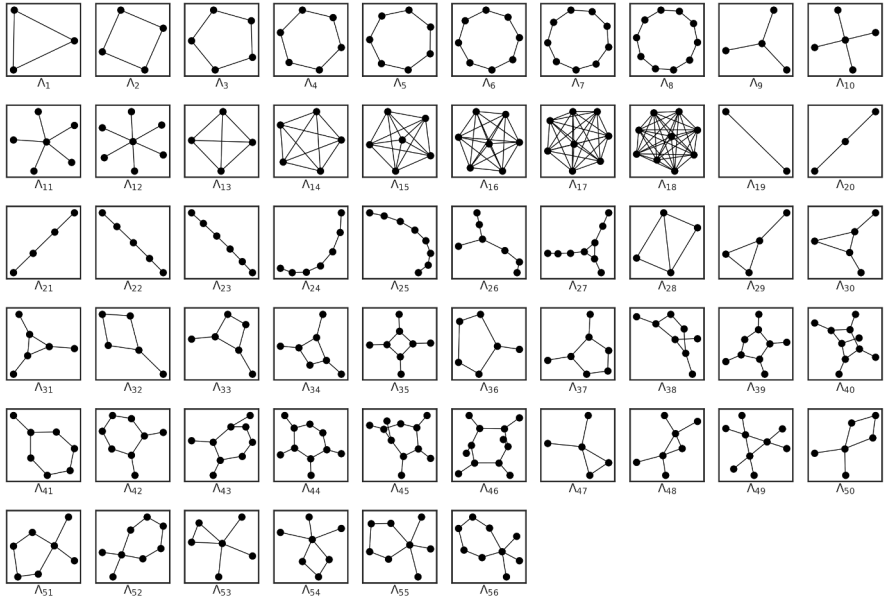
UNIVERSITÀ DEGLI STUDI DI MILANO

How to choose the right subgraphs to obtain a **simple** and **efficient** test ?

We considered r-regular graphs, which are known to be hard to distinguish for methods like the Weisfeiler-Leman test (Weisfeiler and Leman, 1968) and (Message Passing) Graph Neural Networks (Scarselli, Gori, et al., 2009; Morris, Ritzert, et al., 2019).

A graph is called r-regular for $r \geq 2$ if every node has degree exactly equal to r.




Λ_1
Λ_2
Λ_3
Λ_4
Λ_5
Λ_6
Λ_7
Λ_8
Λ_9
Λ_{10}
Λ_{11}
Λ_{12}
Λ_{13}
Λ_{14}
Λ_{15}
Λ_{16}
Λ_{17}
Λ_{18}
Λ_{19}
Λ_{20}
Λ_{21}
Λ_{22}
Λ_{23}
Λ_{24}
Λ_{25}
Λ_{26}
Λ_{27}
Λ_{28}
Λ_{29}
Λ_{30}
Λ_{31}
Λ_{32}
Λ_{33}
Λ_{34}
Λ_{35}
Λ_{36}
Λ_{37}
Λ_{38}
Λ_{39}
Λ_{40}
Λ_{41}
Λ_{42}
Λ_{43}
Λ_{44}
Λ_{45}
Λ_{46}
Λ_{47}
Λ_{48}
Λ_{49}
Λ_{50}
Λ_{51}
Λ_{52}
Λ_{53}
Λ_{54}
Λ_{55}
Λ_{56}

Evaluation of F_{Λ_j}

Single operator accuracy

UNIVERSITÀ DEGLI STUDI DI MILANO

To select the most promising operators to distinguish non isomorphic graphs, we first computed the distinguishing power of each operator $F_{\{j\}}$ alone on a random sample of {3,4,5}-regular graphs:

$$A\Big(F_{\{j\}}\Big) = \frac{1}{m}\sum_{k=1}^{m} \mathbb{1}\Big(\mathrm{sgn}(F_{\{j\}}((\varphi_1,\varphi_2)_k)) = y_k\Big)$$

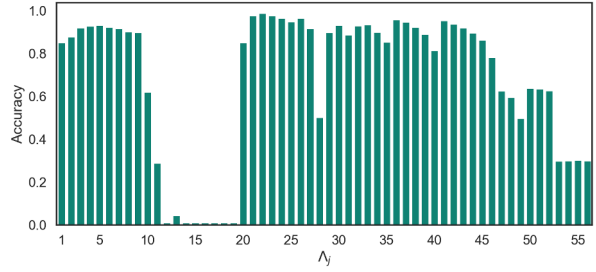
Evaluation of F_{Λ_j}

Single operator accuracy

UNIVERSITÀ DEGLI STUDI DI MILANO

To select the most promising operators to distinguish non isomorphic graphs, we first computed the distinguishing power of each operator $F_{\{j\}}$ alone on a random sample of {3,4,5}-regular graphs:

$$A\Big(F_{\{j\}}\Big) = \frac{1}{m}\sum_{k=1}^{m} \mathbb{1}\Big(\text{sgn}(F_{\{j\}}((\varphi_1,\varphi_2)_k)) = y_k\Big)$$



Selection of Λ_j

Selection algorithm

UNIVERSITÀ DEGLI STUDI DI MILANO

Algorithm 1: Forward Selection

Data: $\Lambda_1, \ldots, \Lambda_p$, $((\varphi_1, \varphi_2)_k)_{k=1}^m$

Result: $S \subseteq \{\Lambda_1, \ldots, \Lambda_p\}$

$I \leftarrow 1$; $i_I \leftarrow \underset{t \in \{1,\ldots,p\}}{\arg\max}\, A\Big(F_{\{t\}}\Big)$;

while $\underset{t \in \{1,\ldots,p\} \setminus \{i_1,\ldots,i_I\}}{\max} A\Big(F_{\{i_1,\ldots,i_I\} \cup \{t\}}\Big) > A\Big(F_{\{i_1,\ldots,i_I\}}\Big)$ **do**

  $i_{I+1} \leftarrow \underset{t \in \{1,\ldots,p\} \setminus \{i_1,\ldots,i_I\}}{\arg\max} A\Big(F_{\{i_1,\ldots,i_I\} \cup \{t\}}\Big)$;

end

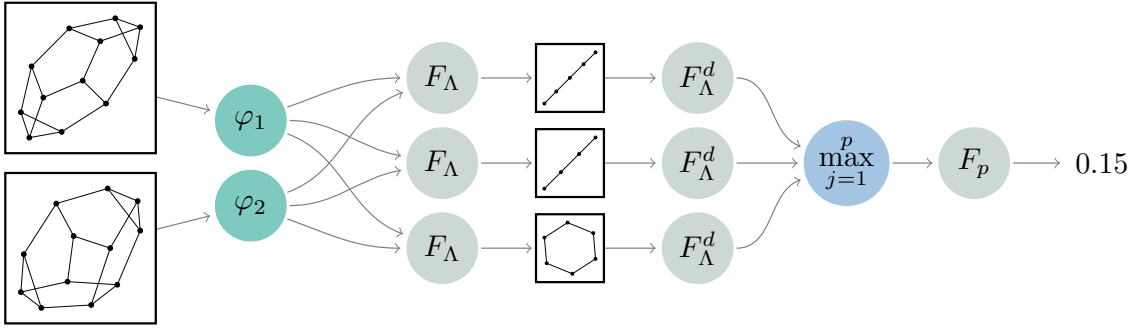
$S = \{\Lambda_{i_1}, \ldots, \Lambda_{i_I}\}$;

Selection of Λ_j

Model architecture

UNIVERSITÀ DEGLI STUDI DI MILANO

After running the selection algorithm we found that maximum accuracy could be achieved with only three operators F_{22}, F_{21} and F_4:



Comparison
Accuracy and execution time

UNIVERSITÀ DEGLI STUDI DI MILANO

Finally we tested the network against other SOTA methods, both exact and inexact, considering both **expressivity** and **time efficiency**.

Comparison

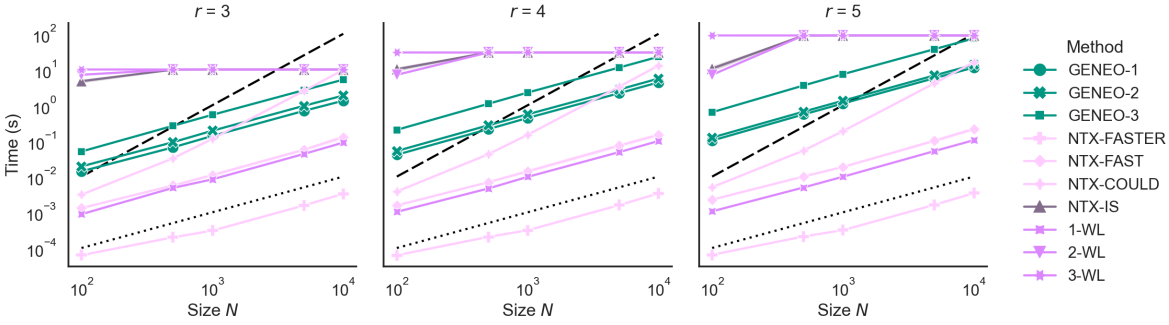
Accuracy and execution time

UNIVERSITÀ DEGLI STUDI DI MILANO

Finally we tested the network against other SOTA methods, both exact and inexact, considering both **expressivity** and **time efficiency**.

Method	*N*					
	100		1000		10000	
	Time	Acc	Time	Acc	Time	Acc
GENEO-1	0.014	0.990	0.138	0.986	1.338	0.984
GENEO-2	0.019	0.996	0.193	0.994	1.866	0.994
GENEO-3	0.051	**1.000**	0.539	**1.000**	5.202	**1.000**
NTX-FASTER	0.000	0.000	0.000	0.000	0.003	0.000
NTX-FAST	0.001	0.719	0.011	0.758	0.128	0.754
NTX-COULD	0.003	0.719	0.114	0.758	9.747	0.754
NTX-IS	4.719	0.691	10.000	0.000	10.000	0.000
1-WL	0.001	0.000	0.008	0.000	0.090	0.000
2-WL	7.134	0.000	10.000	0.000	10.000	0.000
3-WL	10.040	0.000	10.000	0.000	10.000	0.000

Comparison
Execution times plot



Thank you for your attention!

“It’s still magic even if you know how it’s done.”

(Terry Pratchett — A Hat Full of Sky)

References

[1] G. Bocchi, S. Botteghi, et al. “On the finite representation of linear group equivariant operators via permutant measures”. In: *Annals of Mathematics and Artificial Intelligence* 91.4 (2023), pp. 465–487. ISSN: 1012-2443. DOI: 10.1007/s10472-022-09830-1.

[2] G. Bocchi, M. Ferri, and P. Frosini. “A novel approach to graph distinction through GENEOs and permutants”. In: *Scientific Reports* (2025). (in press). eprint: 2406.08045.

[3] G. Bocchi, P. Frosini, et al. “A geometric XAI approach to protein pocket detection”. In: *Joint Proceedings of the xAI 2024 Late-breaking Work, Demos and Doctoral Consortium co-located with the 2nd World Conference on eXplainable Artificial Intelligence (xAI-2024), Valletta, Malta, July 17-19, 2024*. Vol. 3793. 2024, pp. 217–224. URL: https://ceur-ws.org/Vol-3793/paper%5C_28.pdf.

[4] G. Bocchi, P. Frosini, et al. “GENEOnet: A new machine learning paradigm based on Group Equivariant Non-Expansive Operators. An application to protein pocket detection”. preprint at arXiv. 2022. arXiv: 2202.00451.

Other references I

[5] F. Ahmad, M. Ferri, and P. Frosini. “Generalized Permutants and Graph GENEOs”. In: *Machine Learning and Knowledge Extraction* 5.4 (2023), pp. 1905–1920. DOI: 10.3390/make5040092.

[6] L. Babai. “Canonical form for graphs in quasipolynomial time: preliminary report”. In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. 2019, pp. 1237–1246.

[7] M. G. Bergomi, P. Frosini, et al. “Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning”. In: *Nature Machine Intelligence* 1.9 (2019), pp. 423–433. DOI: 10.1038/s42256-019-0087-3.

[8] F. Camporesi, P. Frosini, and N. Quercioli. “On a New Method to Build Group Equivariant Operators by Means of Permutants”. In: *Machine Learning and Knowledge Extraction, CD-MAKE 2018*. Vol. 11015. Lecture Notes in Computer Science. 2018, pp. 265–272. DOI: 10.1007/978-3-319-99740-7_18.

[9] T. S. Cohen and M. Welling. “Group Equivariant Convolutional Networks”. In: *33rd International Conference on Machine Learning*. Ed. by M. Balcan and K. Weinberger. Vol. 48. 2016.

Other references II

[10] F. Conti, P. Frosini, and N. Quercioli. “On the Construction of Group Equivariant Non-Expansive Operators *via* Permutants and Symmetric Functions”. In: *Frontiers in Artificial Intelligence* 5 (2022). DOI: 10.3389/frai.2022.786091.

[11] Z. Galil, C. M. Hoffmann, et al. “$O(n^3 \log n)$ deterministic and an $O(n^3)$ Las Vegas Isomorphism test for trivalent graphs.”. In: *Journal of the ACM* 34.3 (1987), pp. 5l3–53l. DOI: 10.1145/28869.28870.

[12] P. Giudici. “Safe machine learning”. In: *Statistics* 58.3 (2024), pp. 473–477. DOI: 10.1080/02331888.2024.2361481.

[13] R. Guidotti, A. Monreale, et al. “Principles of Explainable Artificial Intelligence”. In: *Explainable AI Within the Digital Transformation and Cyber Physical Systems: XAI Methods and Applications*. 2021, pp. 9–31. ISBN: 978-3-030-76409-8. DOI: 10.1007/978-3-030-76409-8_2.

Other references III

[14] E. M. Luks. “Isomorphism of graphs of bounded valence can be tested in polynomial time”. In: *Journal of Computer and System Sciences* 25.1 (1982), pp. 42–65. ISSN: 0022-0000. DOI: 10.1016/0022-0000(82)90009-5.

[15] C. Morris, M. Ritzert, et al. “Weisfeiler and Leman Go Neural: Higher-Order Graph Neural Networks”. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 33.1 (2019), pp. 4602–4609. DOI: 10.1609/aaai.v33i01.33014602.

[16] F. Scarselli, M. Gori, et al. “The Graph Neural Network Model”. In: *IEEE Transactions on Neural Networks* 20.1 (2009), pp. 61–80. DOI: 10.1109/TNN.2008.2005605.

[17] M. Weiler and G. Cesa. “General E(2)-Equivariant Steerable CNNs”. In: *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates, Inc., 2019.

[18] M. Weiler, F. A. Hamprecht, and M. Storath. “Learning Steerable Filters for Rotation Equivariant CNNs”. In: *Conference on Computer Vision and Pattern Recognition (CVPR 2018)*. 2018.

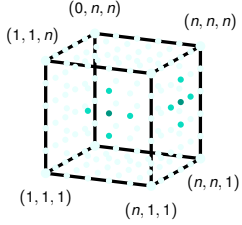
Other references IV

[19] B. Weisfeiler and A. Leman. *Weisfeiler and Leman original paper English translation*. https://www.iti.zcu.cz/wl2018/pdf/wl_paper_translation.pdf. 1968.

Application to dice classification

Data

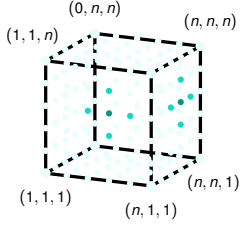
The space X models a discretization of a cube into a grid



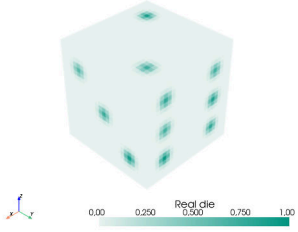
Application to dice classification

Data

The space X models a discretization of a cube into a grid



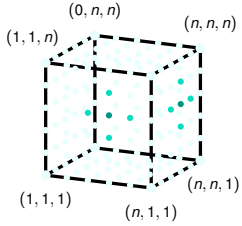
Φ represents scans of dice: **real** (opposite faces sum to 7)...



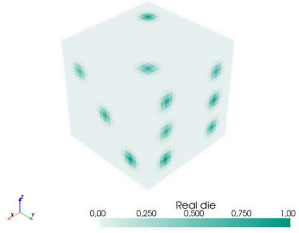
Application to dice classification

Data

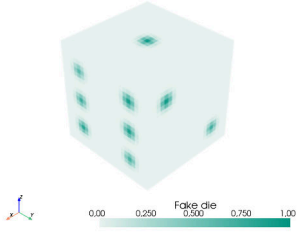
The space X models a discretization of a cube into a grid



Φ represents scans of dice: **real** (opposite faces sum to 7)...



... And **fake** (opposite faces never sum to 7).



Application to dice classification
Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

We identified three permutants H_i for the group G of orientation-preserving isometries of the cube, such permutants define an equal number of GENEOs

$$F_i(\varphi) = \frac{1}{|H_i|} \sum_{h \in H_i} \varphi \circ h^{-1} \quad i \in \{1, 2, 3\}$$

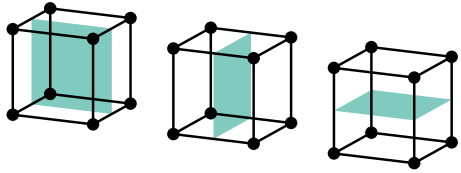
Application to dice classification
Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

We identified three permutants H_i for the group G of orientation-preserving isometries of the cube, such permutants define an equal number of GENEOs

$$F_i(\varphi) = \frac{1}{|H_i|} \sum_{h \in H_i} \varphi \circ h^{-1} \quad i \in \{1, 2, 3\}$$

Planes parallel to one of the faces ...



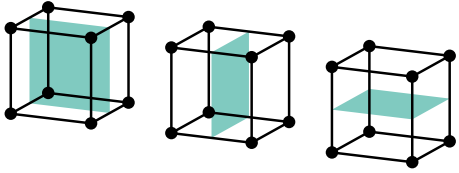
Application to dice classification
Operators

UNIVERSITÀ DEGLI STUDI DI MILANO

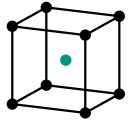
We identified three permutants H_i for the group G of orientation-preserving isometries of the cube, such permutants define an equal number of GENEOs

$$F_i(\varphi) = \frac{1}{|H_i|} \sum_{h \in H_i} \varphi \circ h^{-1} \quad i \in \{1, 2, 3\}$$

Planes parallel to one of the faces ...



... central symmetry of cube's center ...



Application to dice classification

Network and results

UNIVERSITÀ DEGLI STUDI DI MILANO

We combined these three GENEOs to obtain a small network by using convex combination

$$F_\alpha(\varphi) = \sum_{i=1}^{3} \alpha_i F_i(\varphi)$$

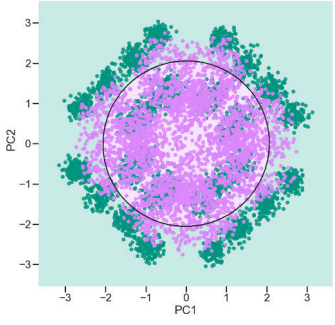
We used F_α to **classify** data between real/fake comparing to a similar approach without using GENEOs.

Application to dice classification

Network and results

UNIVERSITÀ DEGLI STUDI DI MILANO

Without using F_α



We combined these three GENEOs to obtain a small network by using convex combination

$$F_\alpha(\varphi) = \sum_{i=1}^{3} \alpha_i F_i(\varphi)$$

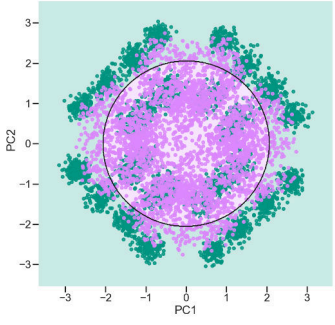
We used F_α to **classify** data between real/fake comparing to a similar approach without using GENEOs.

Application to dice classification

Network and results

UNIVERSITÀ DEGLI STUDI DI MILANO

Without using F_α

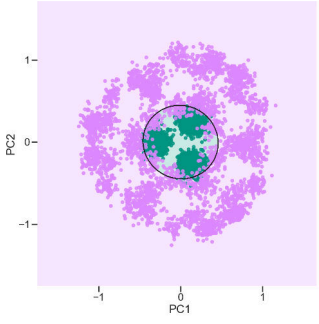


We combined these three GENEOs to obtain a small network by using convex combination

$$F_\alpha(\varphi) = \sum_{i=1}^{3} \alpha_i F_i(\varphi)$$

We used F_α to **classify** data between real/fake comparing to a similar approach without using GENEOs.

Using F_α



Connections between GENEOs and TDA

UNIVERSITÀ DEGLI STUDI DI MILANO

GENEOs and Topological Data Analysis (TDA) are indeed deeply interconnected:

1. TDA helps in the comparison of GENEOs, it makes available an efficiently computable pseudo-metric between GENEOs, which are usually not easily compared using d_{GENEO} (Bergomi, Frosini, et al., 2019).
2. The extension of one-parameter TDA to multiparameter TDA can be based on a suitable GENEO (Cerri, Ethier, and Frosini, 2019)
3. The use of GENEOs allows one to restrict the invariance of TDA to suitable invariance groups (Frosini and Jabłoński, 2016)

How to introduce a topology on X

UNIVERSITÀ DEGLI STUDI DI MILANO

In principle the set X initially has **no topological structure**, anyway, to introduce GENEOs is necessary to consider $\mathrm{Homeo}(X)$. Indeed, once fixed the space Φ of admissible functions, a topology on X can be derived using the pseudo-metric

$$D_X(x_1, x_2) = \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|$$

The natural pseudo distance

UNIVERSITÀ DEGLI STUDI DI MILANO

An important concept in TDA is that of natural pseudo-distance associated to the group G:

Definition

The pseudo-distance $\mathfrak{D}_G \colon \Phi \times \Phi \to \mathbb{R}$

$$\mathfrak{D}_G(\varphi_1, \varphi_2) = \inf_{g \in G} ||\varphi_1 - \varphi_2 \circ g||_\infty$$

It is called the natural pseudo-distance associated with a group G acting on Φ.

Often $\mathfrak{D}_G$ represents the **ground truth** of data comparison. Indeed in the part concerning GraphGENEOs we had that:

$$\mathfrak{D}_G(\varphi_1, \varphi_2) = 0 \iff \Gamma_1 \sim \varphi_1, \Gamma_2 \sim \varphi_2 \text{ are isomorphic}$$
$$\mathfrak{D}_G(\varphi_1, \varphi_2) = 1 \iff \Gamma_1 \sim \varphi_1, \Gamma_2 \sim \varphi_2 \text{ are non isomorphic}$$

Permutants VS Group convolution

Since H must be finite, the provides a practical and possibly efficient way of computing GENEOs.

$$F_H(\varphi) = \begin{cases} \frac{1}{|H|} \sum_{h \in H} \varphi \circ h & \text{if } |H| > 0 \\ \mathbf{0} & \text{otherwise} \end{cases}$$

No approximation needed and the smallest H the fasted F_H can be computed.

Methods belonging to the sphere of Geometric Deep Leaning (Weiler and Cesa, 2019; Cohen and Welling, 2016; Weiler, Hamprecht, and Storath, 2018) may obtain equivariant operators using group convolution:

$$\psi(u) = [K *_G \varphi](u) = \int_G K(v^{-1}u)\varphi(v) d\mu_G(v)$$

However, when the group is infinite, it needs to be approximated...

Model comparison
Datasets

UNIVERSITÀ DEGLI STUDI DI MILANO

To train, select and compare GENEOnet we employed two data sources

1. The PDBbind database http://pdbbind.org.cn/, which offers a well-curated collection of experimentally determined binding affinity data for protein-ligand complexes stored within the Protein Data Bank. From PDBbind we retrieved 12295 complexes that were divide into BIND_VAL (around 25%) and BIND_TEST (around 75%).
2. The Protein Data Bank https://www.rcsb.org/ is the largest biomolecular structure repository, it processes, checks, annotates, and releases new structures daily. From the PDB we retrieved 41519 complexes denoted as WILD.

The loss function $\mathcal{L}$

The hyperparameter κ

UNIVERSITÀ DEGLI STUDI DI MILANO

When $\kappa = 0$

$$\mathcal{L}(\widehat{\psi}, \Pi) = \frac{|\widehat{\psi} \wedge \Pi|}{|\Pi|} = \mathcal{V}(\widehat{\psi}, \Pi)$$

the loss function coincides with the overlap metric. This would produce huge pockets since the only things that matters is to include the ground truth in the prediction.

When $\kappa = 1$

$$\mathcal{L}(\widehat{\psi}, \Pi) = \frac{|\widehat{\psi} \wedge \Pi| + |(\mathbf{1} - \widehat{\psi}) \wedge (\mathbf{1} - \Pi)|}{|\Pi| + |\mathbf{1} - \Pi|}$$

the loss function coincides with the proportion of correctly classified voxels between cavity and non-cavity. This would favor the correct identification of non-cavity voxels, which are the vast majority, due to a label imbalance problem.

$\mathcal{H}_j$ with more than one true pocket

When the dataset of proteins is such that each protein has only **one** true pocket:

$$\mathcal{H}_j = \frac{1}{N}\sum_{t=1}^{N} \Delta_j(\mathcal{M}(P_t), \Pi_t)$$

for every $j \geq 1$.

Otherwise when proteins may have **multiple** replicas of the true pocket, for example due to symmetries:

$$\mathcal{H}_n = \frac{1}{N}\sum_{t=1}^{N}\sum_{j=1}^{n} \Delta_j(\mathcal{M}(P_t), \Pi_t)$$

$$\mathcal{H}_{n+j} = \frac{1}{N}\sum_{t=1}^{N} \Delta_{n+j}(\mathcal{M}(P_t), \Pi_t)$$

for every $j \geq 1$, where n is the number of replicas for each protein.

Overlap VS DVO

UNIVERSITÀ DEGLI STUDI DI MILANO

To asses GENEOnet performances we considered as primary metric the **overlap** of a predicted pocket $\widehat{\Pi}$ with the true pocket Π

$$\mathcal{V}(\widehat{\Pi}, \Pi) = \frac{|\Pi \wedge \widehat{\Pi}|}{|\Pi|}$$

In the literature is also possible to find another metric called **Discretized Volumetric Overlap** (DVO)

$$DVO(\widehat{\Pi}, \Pi) = \frac{|\Pi \wedge \widehat{\Pi}|}{|\Pi \vee \widehat{\Pi}|}$$

Graph permutants

UNIVERSITÀ DEGLI STUDI DI MILANO

Proposition

The set H_τ defined as

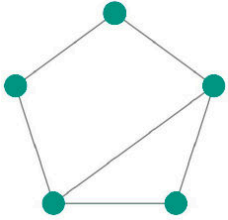
$$H_\tau = \{g \in G: \text{ induced by a permutation } \sigma \text{ of } \{1, \dots, N\} \text{ of kind } \tau\}$$

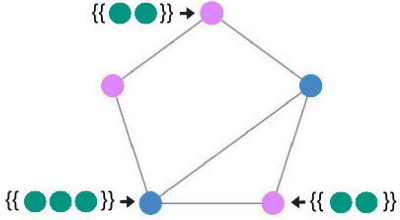
is a permutant for the pairs (Φ, G), (Φ, G) with respect to the trivial homomorphism $T = id_G$ for every possible choice of kind τ.

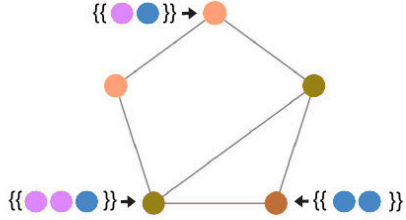
Considering $N = 6$:

τ	0	2	2,2,2	3,3	3	2,2	2,4	4	6	2,3	5	$\lvert G \rvert$
$\lvert H_\tau \rvert$	1	15	15	40	40	45	90	90	120	120	144	720

Weisfeiler Leman coloring test (1-WL)







Almost equivalence of 1-WL and GNNs

UNIVERSITÀ DEGLI STUDI DI MILANO

Algorithm 2: 1-WL

Data: $\Gamma = (V, E, \mathbf{X})$
Result: $\mathbf{c}_\Gamma$
$\mathbf{c}_v^{(0)} \leftarrow \text{HASH}(\mathbf{x}_v) \quad \forall\, v \in V;$
for $l \in \{1, \dots, L\}$ **do**
  $\mathbf{c}_v^{(l)} = \text{HASH}(\mathbf{c}_v^{(l-1)}, \{\{\mathbf{c}_u^{(l-1)} : u \in \mathcal{N}(v)\}\}) \quad \forall\, v \in V;$
end
$\mathbf{c}_\Gamma = \text{HISTOGRAM}(\{\{\mathbf{c}_v^{(L)} : v \in V\}\});$

Algorithm 3: GNNs

Data: $\Gamma = (V, E, \mathbf{X})$
Result: $\mathbf{h}_\Gamma$
$\mathbf{h}_v^{(0)} = \mathbf{x}_v \quad \forall\, v \in V;$
for $l \in \{1, \dots, L\}$ **do**
  $\mathbf{a}_v^{(l)} = f_{\text{aggregate}}^{(l)}(\{\{\mathbf{h}_u^{(l-1)} : u \in \mathcal{N}(v)\}\}) \quad \forall\, v \in V;$
  $\mathbf{h}_v^{(l)} = f_{\text{update}}^{(l)}(\mathbf{h}_v^{(l-1)}, \mathbf{a}_v^{(l)}) \quad \forall\, v \in V;$
end
$\mathbf{h}_\Gamma = f_{\text{readout}}(\{\{\mathbf{h}_v^{(L)} : v \in V\}\});$

Testing r-regular graphs

UNIVERSITÀ DEGLI STUDI DI MILANO

Some special families of graphs can be tested in polynomial time:

- Trees
- Planar graphs
- etc.

In particular, also r-regular graphs can be tested in polynomial time:

1. For 3-regular graphs
   1. A bound of $\mathcal{O}(n^5)$ was proven in (Luks, 1982).
   2. A bound of $\mathcal{O}(n^3 \log n)$ was proven in (Galil, Hoffmann, et al., 1987).
2. For r-regular graphs a general bound of $n^{\mathcal{O}(r)}$ can be established.