Some recent advances in the theory of equivariant non-expansive operators and its application to Machine Learning

Patrizio Frosini

Department of Mathematics, University of Bologna patrizio.frosini@unibo.it

ICMDD, Rajagiri School of Engineering & Technology February 1-3, 2024

Outline

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

Data can be often regarded as functions

Some examples of data that can be seen as functions:

- An electrocardiogram (a function from \mathbb{R} to \mathbb{R});
- A gray-level image (a function from \mathbb{R}^2 to \mathbb{R});
- A computerized tomography scan (a function from a helix to \mathbb{R}).

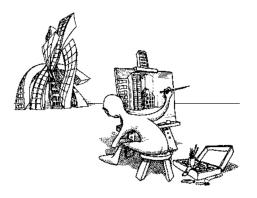






Data are processed by observers

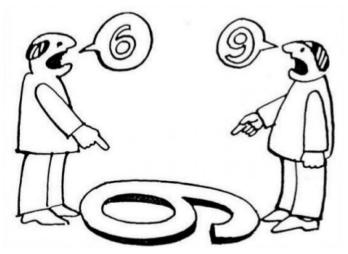
Data have no meaning if no observer elaborates them.



An observer is an agent that transforms data while respecting its symmetries.

Observers are variables in data analysis

Data interpretation strongly depends on the chosen observer:



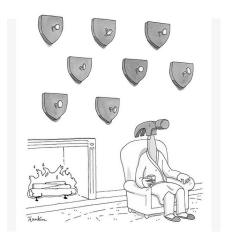
Our interest in data is greatly overrated

We are rarely directly interested in the data, but rather in how observers react to their presence.



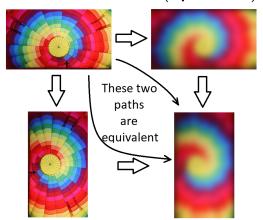
No data structure

Generally speaking, there is no structure in data. The structure of data is a projection of the structure of the observer.



Representing observers as equivariant operators

Observers are structures able to change data into other data, and usually do that by respecting some data symmetries, i.e., by commuting with some transformations (**equivariance**).



Representing observers as equivariant operators

As a first approximation, observers can be represented as **Group Equivariant Operators** (**GEOs**).

In this talk we will illustrate some results on the theory of **Group Equivariant Non-Expansive Operators** (**GENEOs**).

Why "non-expansive"?

Because

- observers are often assumed to simplify the metric structure of data in order to produce meaningful interpretations;
- 2. non-expansiveness guarantees good topological properties.

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

How could we represent observers?

machine intelligence

ARTICLES

https://doi.org/10.1038/s42256-019-0087-3

Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning

Mattia G. Bergomi¹, Patrizio Frosini¹, Daniela Giorgi¹ and Nicola Quercioli¹

We provide a general mathematical framework for group and set equivariance in machine learning. We define group equivariant non-expansive operators (GENEOs) as maps between function spaces associated with groups of transformations. We study the topological and metric properties of the space of GENEOs to evaluate their approximating power and set the basis for general strategies to initialize and compose operators. We define suitable pseudo-metrics for the function spaces, the equivariance groups and the set of non-expansive operators. We prove that, under suitable assumptions, the space of GENEOs is compact and convex. These results provide fundamental guarantees in a machine learning perspective. By considering isometry-equivariant non-expansive operators, we describe a simple strategy to select and sample operators. Thereafter, we show how selected and sampled operators can be used both to perform classical metric learning and to inject knowledge in artificial neural networks.

https://rdcu.be/bP6HV

All begins with the space of admissible functions

Let X be a nonempty set. Let Φ be a topological subspace of the set \mathbb{R}^X_b of all bounded functions φ from X to \mathbb{R} , endowed with the topology induced by the metric

$$D_{\Phi}(\varphi_1,\varphi_2):=\|\varphi_1-\varphi_2\|_{_{\infty}}.$$

We can see X as the space where we can make our measurements, and Φ as the space of all possible measurements. We will say that Φ is the set of admissible functions. In other words, Φ is the set of all functions from X to $\mathbb R$ that can be produced by our measuring instruments (or by other observers). For example, a gray-level image can be represented as a function from the real plane to the interval [0,1] (in this case $X=\mathbb R^2$).

Perception pairs

Let us consider a group G of bijections $g: X \to X$ such that $\varphi \in \Phi \implies \varphi \circ g \in \Phi$ for every $\varphi \in \Phi$. We say that (Φ, G) is a perception pair.

The choice of a perception pair states which data can be considered as legitimate measurements (the functions in Φ) and which group represents the symmetries between data (the group G).

To proceed, we need to introduce suitable topologies on X and G. Before doing that, we recall that the initial topology $\tau_{\rm in}$ on X with respect to Φ is the coarsest topology on X such that every function φ in Φ is continuous.

A pseudo-metric on X

Let us define on *X* the pseudo-metric

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

 D_X induces a topology τ_{D_X} on X.

The use of D_X implies that we can distinguish two points only if a measurement exists, taking those points to different values.

Proposition

The topology τ_{D_X} is finer than the initial topology τ_{in} on X with respect to Φ . If Φ is totally bounded, then τ_{D_X} coincides with τ_{in} .

A pseudo-metric on X

The following properties are of use in our model.

Proposition

Every function in Φ is non-expansive, and hence continuous.

Proposition

If Φ is compact and X is complete, then X is compact.

In the following, we will usually assume that Φ is compact and X is complete (and hence compact).

Some magic happens: each bijection is an isometry

- $\operatorname{Bij}_{\Phi}(X) = \{ \text{bijections } g : X \to X \text{ s.t. } \Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi \};$
- Homeo $_{\Phi}(X) = \{\text{homeomorphisms } g: X \rightarrow X \text{ s.t. } \Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi\};$
- Iso_{Φ}(X) = {isometries $g: X \rightarrow X$ s.t. $\Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi$ }.

Proposition

$$\operatorname{Bij}_{\Phi}(X) = \operatorname{Homeo}_{\Phi}(X) = \operatorname{Iso}_{\Phi}(X).$$

A pseudo-metric on *G*

Let us now focus our attention on a subgroup G of $\operatorname{Homeo}_{\Phi}(X)$.

We can define a pseudo-metric D_G on G by setting

$$D_G(g_1,g_2) := \sup_{\varphi \in \Phi} D_{\Phi}(\varphi \circ g_1, \varphi \circ g_2).$$

Theorem

G is a topological group with respect to D_G and the action of G on Φ by right composition is continuous.

Theorem

If Φ is compact and G is complete, then G is compact.

GEOs and GENEOs

Each pair (Φ, G) with $G \subseteq \operatorname{Homeo}_{\Phi}(X)$ is called a *perception pair*.

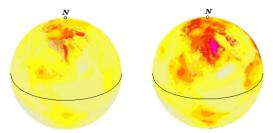
Let us assume that two perception pairs (Φ, G) , (Ψ, H) are given, and fix a group homomorphism $T: G \to H$.

Each function $F: \Phi \to \Psi$ such that $F(\varphi \circ g) = F(\varphi) \circ T(g)$ for every $\varphi \in \Phi, g \in G$ is called a *Group Equivariant Operator (GEO)* associated with the homomorphism T.

If F is also non-expansive (i.e., $D_{\Psi}(F(\varphi_1), F(\varphi_2)) \leq D_{\Phi}(\varphi_1, \varphi_2)$ for every $\varphi_1, \varphi_2 \in \Phi$), then F is called a *Group Equivariant Non-Expansive Operator (GENEO)* associated with the homomorphism T.

An example of GENEO

Let us assume to be interested in the comparison of the distributions of temperatures on a sphere, taken at two different times:



Let us also assume that only two opposite points N, S can be localized on the sphere.

An example of GENEO

Let us introduce two perception pairs $(\Phi, G), (\Psi, H)$ by setting

- $X = S^2$
- $\Phi = \text{set of 1-Lipschitz functions from } S^2 \text{ to a fixed interval } [a, b]$
- $G = \text{group of rotations of } S^2 \text{ around the axis } N S$ and
- Y =the equator S^1 of S^2
- $\Psi = \text{set of } 1\text{-Lipschitz functions from } S^1 \text{ to } [a,b]$
- $H = \text{group of rotations of } S^1$

An example of GENEO

This is a simple example of GENEO from (Φ, G) to (Ψ, H) :

- T(g) is the rotation $h \in H$ of the equator S^1 that is induced by the rotation g of S^2 , for every $g \in G$.
- $F(\varphi)$ is the function ψ that takes each point y belonging to the equator S^1 to the average of the temperatures along the meridian containing y, for every $\varphi \in \Phi$;

We can easily check that F verifies the properties defining the concept of group equivariant non-expansive operator with respect to the isomorphism $T: G \to H$.

In plain words, our GENEO simplifies the data by transforming "temperature distributions on the earth" into "temperature distributions on the equator".

Two key results (and two good news for applications)

Let us assume that a homomorphism $T:G\to H$ has been fixed. Let us define a metric D_{GENEO} on $\text{GENEO}((\Phi,G),(\Psi,H))$ by setting

$$D_{\mathrm{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} D_{\Psi}(F_1(\varphi), F_2(\varphi)).$$

Theorem

If Φ and Ψ are compact, then GENEO($(\Phi, G), (\Psi, H)$) is compact with respect to D_{GENEO} .

Theorem

If Ψ is convex, then GENEO $((\Phi, G), (\Psi, H))$ is convex.

Two key observations (1)

• While the space of data is often non-convex (and hence averaging data does not make sense), the assumption of convexity of Ψ implies the convexity of the space of observers and allows us to consider the "average of observers".



Two key observations (2)

• Our main goal is to develop a good geometric and compositional theory to approximate an ideal observer. In our model, "to approximate an observer" means to look for a GENEO F that minimizes a suitable "cost function" c(F). The cost function quantifies the error that is committed by taking the GENEO F instead of the ideal observer. Since the space of GENEOs is compact and convex (under the assumption that the data spaces are compact and convex), if the cost function c(F) is strictly convex we have that there is one and only one GENEO that best approximates the ideal observer.

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

How can we build linear and nonlinear GENEOs?



ORIGINAL RESEARCH published: 15 February 2022 doi: 10.3389/frai.2022.786091



On the Construction of Group Equivariant Non-Expansive Operators *via* Permutants and Symmetric Functions

Francesco Conti 1,2, Patrizio Frosini 3,4,5,6* and Nicola Quercioli 3,7

¹ Department of Mathematics, University of Pisa, Pisa, Italy, ² Institute of Information Science and Technologies 'A. Faedo', National Research Council of Italy (CNR), Pisa, Italy, ² Department of Mathematics, University of Bologna, Bologna, Italy, ⁴ Alma Mater Research Center on Applied Mathematics, University of Bologna, Bologna, Italy, ⁴ Alma Mater Research Institute for Human-Centered Artificial Intelligence, University of Bologna, Bologna, Italy, ⁴ Research Centre on Bectronic Systems for the Information and Communication Technology, University of Bologna, Bologna, Italy, ⁴ ENEA Centro Ricerche Bologna, Bologna, Italy

https://www.frontiersin.org/articles/10.3389/frai.2022.786091/full

Elementary methods to build GENEOs

Proposition (Composition)

If
$$F_1 \in \text{GENEO}((\Phi, G), (\Psi, H))$$
 w.r.t. $T_1 : G \to H$ and $F_2 \in \text{GENEO}((\Psi, H), (\chi, K))$ w.r.t. $T_2 : H \to K$ then $F_2 \circ F_1 \in \text{GENEO}((\Phi, G), (\chi, K))$ w.r.t. $T_2 \circ T_1 : G \to K$.

Proposition (Image by a 1-Lipschitz function)

If
$$F_1, \ldots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$$
 w.r.t. $T : G \to H$, L is a 1-Lipschitz map from \mathbb{R}^n to \mathbb{R} , and $L^*(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$ (where L^* is the map induced by L), then $L^*(F_1, \ldots, F_n) \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T .

The next three statements follow from the last proposition.

Elementary methods to build GENEOs

Proposition (LATTICE OF GENEOS)

If $F_1, \ldots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T: G \to H$ and $\max(F_1, \ldots, F_n)(\Phi), \min(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$, then $\max(F_1, \ldots, F_n), \min(F_1, \ldots, F_n) \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T.

Proposition (Translation)

If $F \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T : G \to H$, and $F_b(\Phi) \subseteq \Psi$ for $F_b(\varphi) := F(\varphi) - b$, then $F_b \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T.

Proposition (Convex combination)

If $F_1, \ldots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T: G \to H$, $(a_1, \ldots, a_n) \in \mathbb{R}^n$ con $\sum_{i=1}^n |a_i| \le 1$ and $F_{\Sigma}(\Phi) \subseteq \Psi$ for $F_{\Sigma}(\phi) := \sum_{i=1}^n a_i F_i(\phi)$, then $F_{\Sigma} \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T.

Permutant measures

Let us consider the set $\Phi = \mathbb{R}^X \cong \mathbb{R}^n$ of all functions from a finite set $X = \{x_1, \dots, x_n\}$ to \mathbb{R} , and a subgroup G of the group $\operatorname{Bij}(X)$ of all permutations of X.

Definition

A finite (signed) measure μ on $\operatorname{Bij}(X)$ is called a *permutant measure* with respect to G if every <u>subset</u> H of $\operatorname{Bij}(X)$ is measurable and μ is invariant under the conjugation action of G (i.e., $\mu(H) = \mu(gHg^{-1})$ for every $g \in G$).

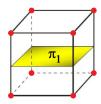
Proposition

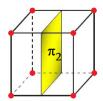
If μ is a permutant measure with respect to G, then the map $F_{\mu}: \mathbb{R}^{X} \to \mathbb{R}^{X}$ defined by setting $F_{\mu}(\varphi) := \sum_{h \in \operatorname{Bij}(X)} \varphi h^{-1} \mu(h)$ is a linear GEO. If $\sum_{h \in \operatorname{Bij}(X)} |\mu(h)| \leq 1$, then $F_{\mu}(\varphi)$ is a GENEO.

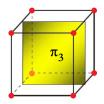
An example of permutant measure

Let us consider the set X of the vertices of a cube in \mathbb{R}^3 , and the group G of the orientation-preserving isometries of \mathbb{R}^3 that take X to X. Let π_1, π_2, π_3 be the three planes that contain the center of mass of X and are parallel to a face of the cube. Let $h_i: X \to X$ be the orthogonal symmetry with respect to π_i , for $i \in \{1,2,3\}$.

We can now define a permutant measure μ on the group $\mathrm{Bij}(X)$ by setting $\mu(h_1) = \mu(h_2) = \mu(h_3) = c$, where c is a positive real number, and $\mu(h) = 0$ for any $h \in \mathrm{Bij}(X)$ with $h \notin \{h_1, h_2, h_3\}$.







Building GENEOs by permutant measures

It is interesting to observe that the set PM(G) of permutant measures with respect to G is a lattice. Indeed, if $\mu_1, \mu_2 \in PM(G)$, then the measures μ', μ'' on Bij(X), respectively defined by setting $\mu'(h) := \min\{\mu_1(h), \mu_2(h)\}$ and $\mu''(h) := \max\{\mu_1(h), \mu_2(h)\}$, still belong to PM(G). Moreover, if $\mu \in PM(G)$ then $|\mu| \in PM(G)$. Furthermore, PM(G) is closed under linear combination. Therefore, PM(G) has a natural structure of real vector space.

We stress that when the group G becomes larger and larger the lattice PM(G) becomes smaller and smaller.

In other words, the theory of permutant measures becomes more and more useful as the groups get bigger and bigger.

Building GENEOs by permutant measures

The method for building GENEOs based on permutant measures can be generalized by replacing the arithmetic mean with another symmetric function. We can indeed show that when a symmetric function and a permutant for the equivariance group G are available, we can easily build a (non-linear) GENEO with respect to G. First of all, let us recall the concept of permutant, which is equivalent to the one of permutant measure uniformly distributed on its support.

Definition

We say that a subset $H \subseteq \operatorname{Bij}(X)$ is a **permutant** for G if either $H = \emptyset$ or $gHg^{-1} = H$ for every $g \in G$.

Note that a subset H of Bij(X) is a permutant for G if and only if H is a union of orbits for the conjugation action of G on Bij(X).

Building GENEOs by permutant measures

Let $\mathscr{S}: \mathbb{R}^n \to \mathbb{R}$ be a symmetric function. If $H = \{h_i\}_{i=1}^n$ is a non-empty permutant for $G \subseteq \operatorname{Bij}_{\Phi}(X)$, then we can define an operator $\mathscr{S}_H \colon \Phi \to \mathbb{R}^X$ by setting, for any $\varphi \in \Phi$,

$$\mathscr{S}_{H}(\varphi) := \mathscr{S}(\varphi \circ h_{1}, \ldots, \varphi \circ h_{n}),$$

where $\mathscr{S}(\varphi \circ h_1, \dots, \varphi \circ h_n)(x) := \mathscr{S}((\varphi \circ h_1)(x), \dots, (\varphi \circ h_n)(x))$ for every $x \in X$.

Proposition

 \mathscr{S}_{H} is a GEO from (Φ,G) to (\mathbb{R}^{X},G) with respect to the identity homomorphism $\mathrm{id}_{G}:G\to G$. If the restriction of \mathscr{S} to $\mathrm{Im}(\Phi)^{n}$ is non-expansive, then \mathscr{S}_{H} is a GENEO from (Φ,G) to (\mathbb{R}^{X},G) with respect to id_{G} .

How can we represent linear GENEOs?

Annals of Mathematics and Artificial Intelligence https://doi.org/10.1007/s10472-022-09830-1

On the finite representation of linear group equivariant operators via permutant measures

Giovanni Bocchi¹ · Stefano Botteghi² · Martina Brasini² · Patrizio Frosini² D · Nicola Ouercioli³

Accepted: 26 December 2022

© The Author(s) 2023

https://rdcu.be/c5Obw

Representation Theorem for linear GENEOs

The following theorem strengthens our previous result about building linear GENEOs via permutant measures.

Theorem (Representation Theorem for linear GENEOs)

Let us assume that $G \subseteq \operatorname{Bij}(X)$ transitively acts on the finite set X and that F is a map from \mathbb{R}^X to \mathbb{R}^X . The map F is a linear GENEO from \mathbb{R}^X to \mathbb{R}^X with respect to the identical homomorphism $\operatorname{id}_G\colon g\mapsto g$ if and only if a permutant measure μ with respect to G exists, such that $F(\phi)=\sum_{h\in\operatorname{Bij}(X)}\phi h^{-1}$ $\mu(h)$ for every $\phi\in\mathbb{R}^X$, and $\sum_{h\in\operatorname{Bij}(X)}|\mu(h)|\leq 1$.

Impulsive noise reduction by GENEOs





Article

A Probabilistic Result on Impulsive Noise Reduction in Topological Data Analysis through Group Equivariant Non-Expansive Operators

Patrizio Frosini *,† 10, Ivan Gridelli † and Andrea Pascucci

https://www.mdpi.com/1099-4300/25/8/1150

Partially equivariant GENEOs



TYPE Methods
PUBLISHED 21 December 2023
DOI 10.3389/frai 2023 1272619

A topological model for partial equivariance in deep learning and data analysis

Lucia Ferrari¹, Patrizio Frosini¹, Nicola Quercioli^{2*} and Francesca Tombari³

¹Department of Mathematics, University of Bologna, Bologna, Italy, ²Department of Electrical, Electronic, and Information Engineering (DEI) and Willab-National Laboratory for Wireless Communications, National Inter-University Consortium for Telecommunications (CNIT), University of Bologna, Bologna, Italy, ³Department of Mathematics, Royal Institute of Technology (KTH), Stockholm, Sweden

Generalized permutants and Graph GENEOs





Article

Generalized Permutants and Graph GENEOs

Faraz Ahmad, Massimo Ferri * D and Patrizio Frosini

ARCES and Department of Mathematics, University of Bologna, 40126 Bologna, Italy; faraz.ahmad2@unibo.it (F.A.); patrizio.frosini@unibo.it (P.F.)

* Correspondence: massimo.ferri@unibo.it

https://www.mdpi.com/2504-4990/5/4/92

A current line of research

We are presently studying this problem:

How can we extend the theory of GENEOs to probability spaces of signals?

P. Cascarano, P. Frosini, N. Quercioli, A. Saki,

On the geometric and Riemannian structure of the spaces of group
equivariant non-expansive operators,

https://arxiv.org/pdf/2103.02543.pdf

The key role of observers in data analysis

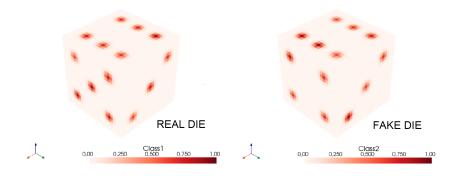
Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

What happens when we apply GENEOs to our data?

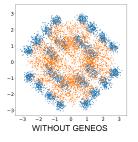
An example of use: comparison between real dice and fake dice.

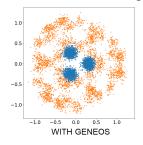


(Experiment and computations by Giovanni Bocchi)

What happens to data when we apply GENEOs?

We produced 10000 dice (a training set of size 7000 and a test set of size 3000), then we applied PCA to the test set and to the test set transformed by a suitable GENEO, optimized on the training set:





For each die the first two principal components are plotted. Blue points are associated with **real dice**, while orange ones with **fake dice**. The GENEO we use was built by a convex combination of 3 GENEOs defined by permutant measures.

A real application: finding pockets in proteins

GENEOnet: A new machine learning paradigm based on Group Equivariant Non-Expansive Operators. An application to protein pocket detection.

Giovanni Bocchi ¹, Patrizio Frosini ², Alessandra Micheletti ¹, Alessandro Pedretti ³ Carmen Gratteri ⁴, Filippo Lunghini ⁵, Andrea Rosario Beccari ⁵ and Carmine Talarico ⁵

https://arxiv.org/ftp/arxiv/papers/2202/2202.00451.pdf

¹ Department of Environmental Science and Policy, Università degli Studi di Milano

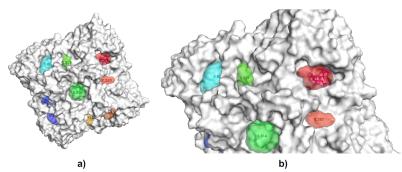
² Department of Mathematics, Università degli Studi di Bologna

³ Department of Pharmaceutical Sciences, Università degli Studi di Milano

⁴ Dipartimento di Scienze della Salute, Università degli Studi "Magna Græcia di Catanzaro"

⁵ Dompé Farmaceutici SpA

A real application: finding pockets in proteins

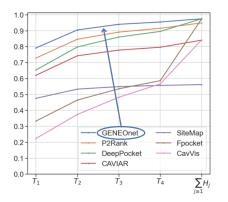


Model predictions for protein 2QWE. In Figure a) the global view of the prediction is shown, where different pockets are depicted in different colors and are labelled with their scores. In Figure b) the zoomed of the pocket containing the ligand is shown.

The search for the pockets was carried out by identifying an optimal GENEO in the convex hull of 8 GENEOs (each focused on a particular property of the pockets).

A real application: finding pockets in proteins

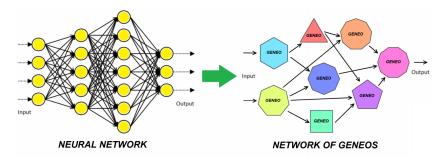
Here are the results of our experiments:



Please note that GENEOnet uses 17 parameters, while a CNN such as DeepPocket requires 665122 parameters.

The main point in our approach

In perspective, we are looking for a good compositional theory for building efficient and transparent networks of GENEOs. Some preliminary experiments suggest that replacing neurons with GENEOs could make deep learning more transparent and interpretable and speed up the learning process.



GENEOs and Machine Learning

For more details about the use of GENEOs in Machine Learning:

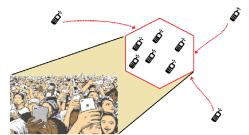


- A. Micheletti, A new paradigm for artificial intelligence based on group equivariant non-expansive operators, In: EMS Magazine, Online First, 24 April 2023.
- https://ems.press/content/serial-article-files/27673

Current research projects

CNIT / WiLab - Huawei Joint Innovation Center (JIC)

Project on GENEOs for 6G





Current research projects



Horizon Europe (HORIZON)
Call: HORIZON-CL4-2023-HUMAN-01-CNECT

Project: 101135775 — PANDORA

Funding: approximately 9 million euros.

Task 3.3 - Leveraging domain knowledge for explainable learning: This task aims to investigate the use of domain knowledge in the development of explainable AI models. Tools like GENEOs for applications in TDA and ML and new theoretical methods of GENEOs for explainable AI will be used.

