Explainable Machine Learning based on Group Equivariant Non-Expansive Operators (GENEOs). Protein pocket detection: a case study

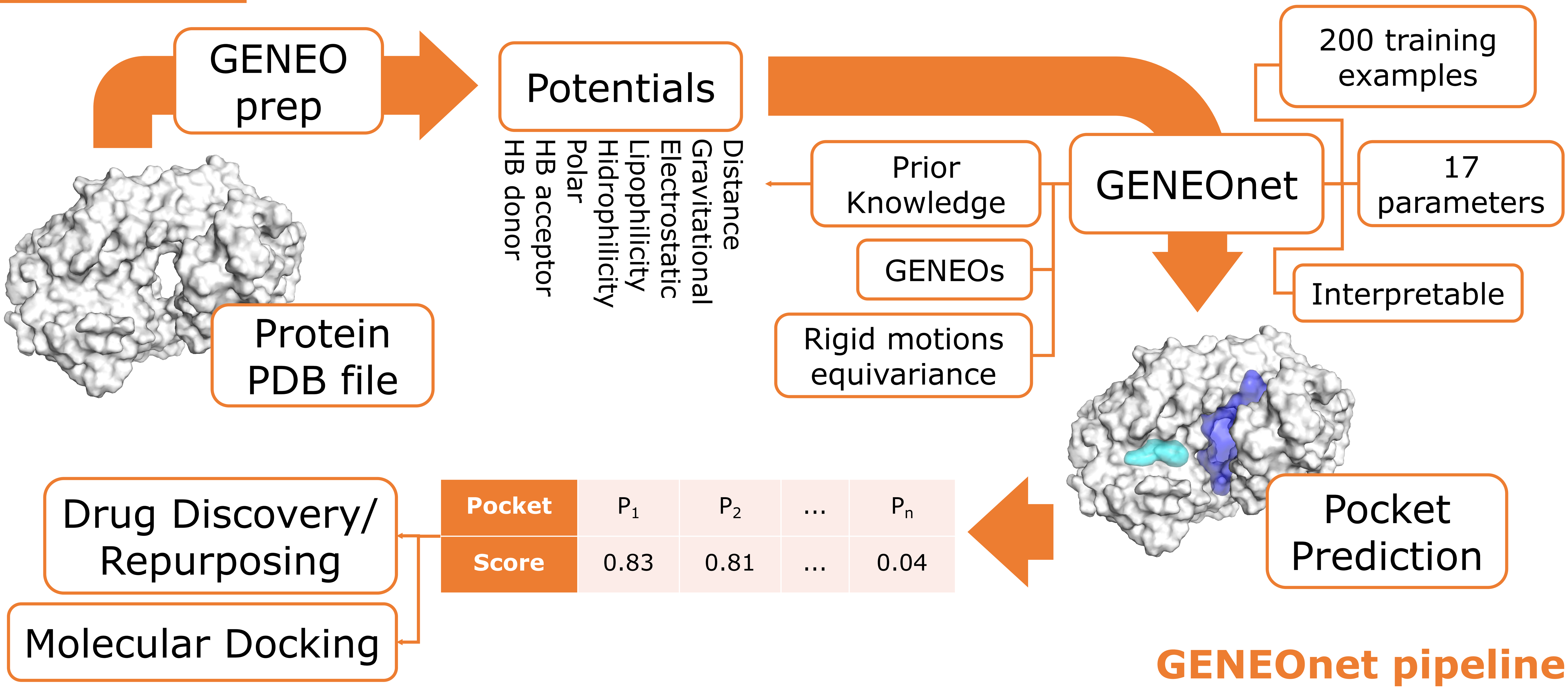
Giovanni Bocchi
Department of Environmental Science and Policy, University of Milan.

GENEOs

A Group Equivariant Non-Expansive Operator F is a map between two functional spaces $\Phi = \{\varphi: X \to \mathbb{R}\}$ and $\Psi = \{\psi: Y \to \mathbb{R}\}$ that, given two subgroups $G \trianglelefteq \mathrm{Homeo}(X)$ and $H \trianglelefteq \mathrm{Homeo}(Y)$ and a fixed group homomorphism $T: G \to H$, has two properties:

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

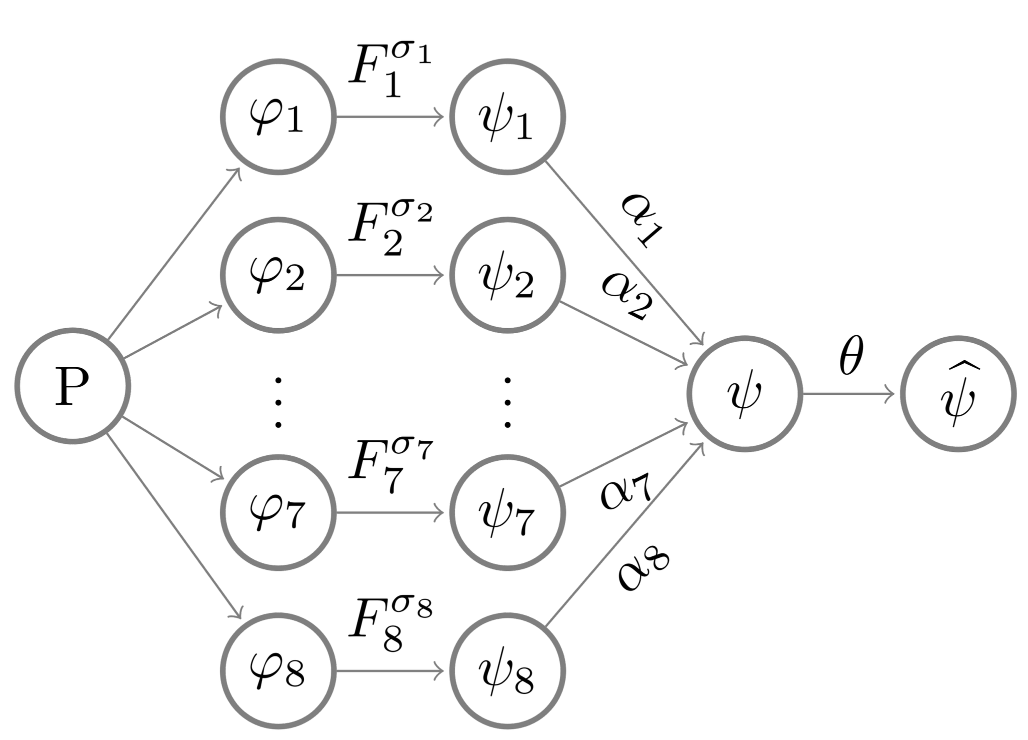
The problem



Pocket	P_1	P_2	...	P_n
Score	0.83	0.81	...	0.04

GENEOnet pipeline

GENEOnet



Results

Method	T_1	T_2	T_3	H_{1+}
GENEOnet	0.792	0.905	0.941	0.975
P2Rank	0.728	0.847	0.892	0.952
DeepPocket	0.652	0.798	0.860	0.978
CAVIAR	0.616	0.739	0.783	0.837
Fpocket	0.331	0.462	0.534	0.978
CavVis	0.224	0.376	0.483	0.842

Conclusions

GENEOs are powerful mathematical tools for **Explainable Machine Learning.**

1. Very few trainable parameters.
2. Intrepretability of the parameters.
3. Can incorporate Prior Knowledge.
4. Equivariance by design.
5. Fewer data are necessary for training.
6. Lower computational complexity compared to similar deep networks.

Joint work with:
Alessandra Micheletti, Patrizio Frosini, Alessandro Pedretti, Carmine Talarico, Filippo Lunghini, Andrea Beccari and Carmen Gratteri.

Reach me at:
giovanni.bocchi1@unimi.it

GENEOS GENEOnet