Artificial Neural Networks for the Rapid Prediction of Possible Ferroptosis Inducers Using the GPx4 Enzyme

Jesus Antonio Alvarado-Huayhuaz[1], Miquéias Amorim Santos Silva[2], Elvis Mauricio Jimenez Peña[3], Joel Claudio Rengifo Maraví[4], Angella M. Cordova-Muñoz[5], Karina dos Santos Machado*[2]

1. Universidad Nacional de Ingeniería, Perú | 2. Universidade Federal do Rio Grande, Brasil | 3. Universidad Nacional Mayor de San Marcos, Perú | 4. Universidad Nacional Agraria La Molina, Perú | 5. Universidad Nacional de Trujillo, Perú.

karina.machado@furg.br

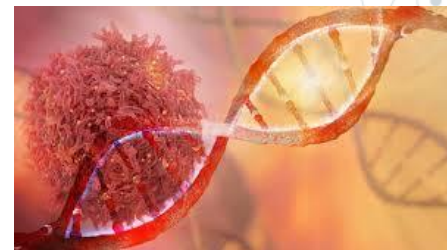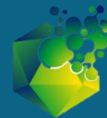*Natal, Brazil, August 27-29, 2024*

Outline

Cancer in the World

- The fight against cancer represents one of the most significant challenges in the field of global health;

- Cancer is a major contributor to global mortality, causing about 1 in every 6 deaths and affecting nearly every household.

- Globally, there were an estimated **20 million new cases of cancer and 9.7 million deaths from cancer in 2022** (WHO 2022)

- **With its increasing prevalence and the urgent need to discover effective treatments, the search for new drugs has become a research priority**

Rebecca L Siegel, Angela N Giaquinto, and Ahmedin Jemal. Cancer statistics, (2024). CA: a cancer journal for clinicians, 74(1):12–49, 2024.

Cancer fact sheet. World Health Organization; 2022.

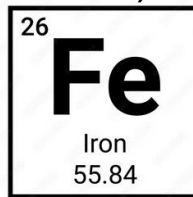# New treatments for cancer: Chelation Therapy and Ferroptosis

Chelation Therapies: remove metals from cells
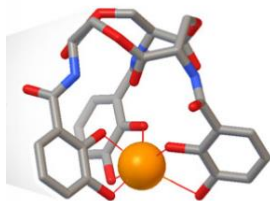
Enterobactin and Iron

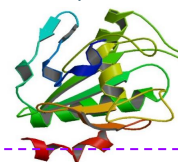Iron (Fe) essential to various cellular reactions (ferric ion)

26
Fe
Iron
55.84

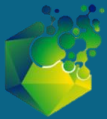Excess of Iron ⇒ participation in reactions that generates oxidative stress (OE)

OE can cause **Ferroptosis**
Type of programed cell death

One important system against OE is the GPx4 enzyme

PROPOSE: inhibits the GPx4 enzyme has emerged as a promising approach to therapy -> ferroptosis

Jie Li, Feng Cao, He-liang Yin, Zi-jian Huang, Zhi-tao Lin, Ning Mao, Bei Sun, and Gang Wang. Ferroptosis: past, present and future. Cell death & disease, 11(2):88, 2020

https://pdb101.rcsb.org/motm/193   https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964%2824%2900123-3/fulltext

GPx4 Enzyme

- The reactivation of cell death induced by the accumulation of iron-dependent lipid peroxides (Ferroptosis) is a potential mechanism for cancer treatment.

- One pathway to trigger Ferroptosis involves the inactivation of the glutathion-dependent peroxidase (GPx4) protein.
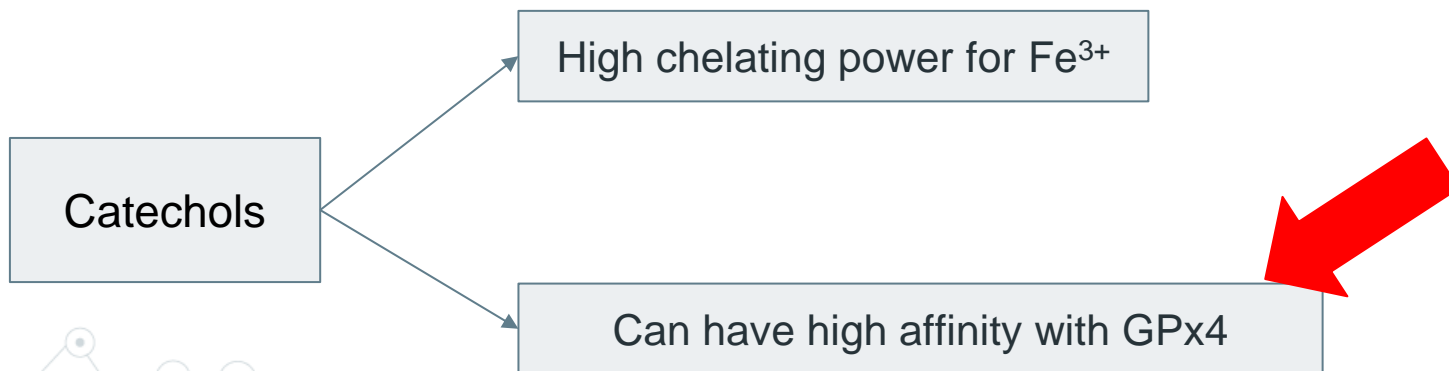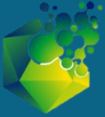

GPx4

Ding, Y. et al (2021). Journal of hematology & oncology, 14, 1-21. | Fatonah, A., et al (2020). Biointerface Research in Applied Chemistry, 10(1), 4929–4933.
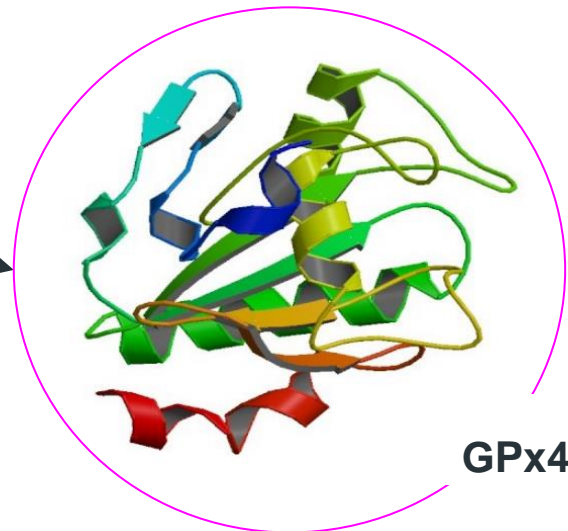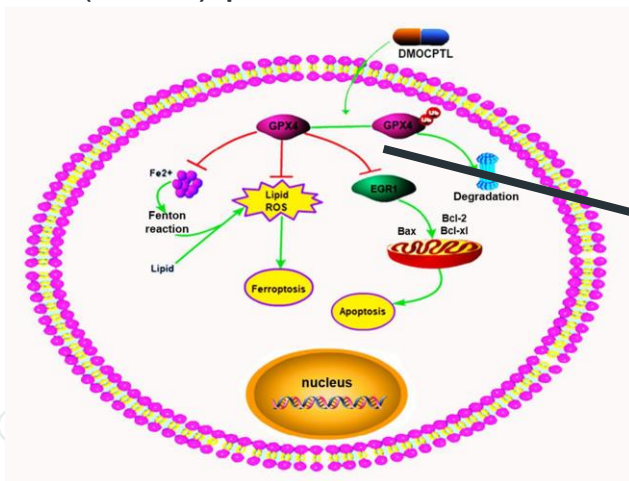
Siderophores - catechols

- Siderophores are low molecular weight organic molecules

- Its chelation power to Fe3+ has been studied for chelation in cancer therapy

- Catechols -> common iron-binding functional groups within siderophores

```
                    ┌─────────────────────────────────────┐
                 →  │  High chelating power for Fe³⁺        │
┌──────────────┐    └─────────────────────────────────────┘
│  Catechols   │
└──────────────┘    ┌─────────────────────────────────────┐
                 →  │  Can have high affinity with GPx4    │
                    └─────────────────────────────────────┘
```

Daniel James Raines, Thomas James Sanderson, AK Duhme-Klair, and Ellis Wilde. Siderophores. In Elsevier Reference Module in Chemistry, Molecular Sciences and Chemical Engineering. Elsevier, 2015. https://www.pnas.org/doi/full/10.1073/pnas.0630018100

Natural Products in Peru

- The first version of PeruNPDB, a compound database of NPs from Peru that includes 280 compounds from plant and animal sources.

Barazorda-Ccahuana, et al. (2023). Scientific Reports, 13(1), 7577.
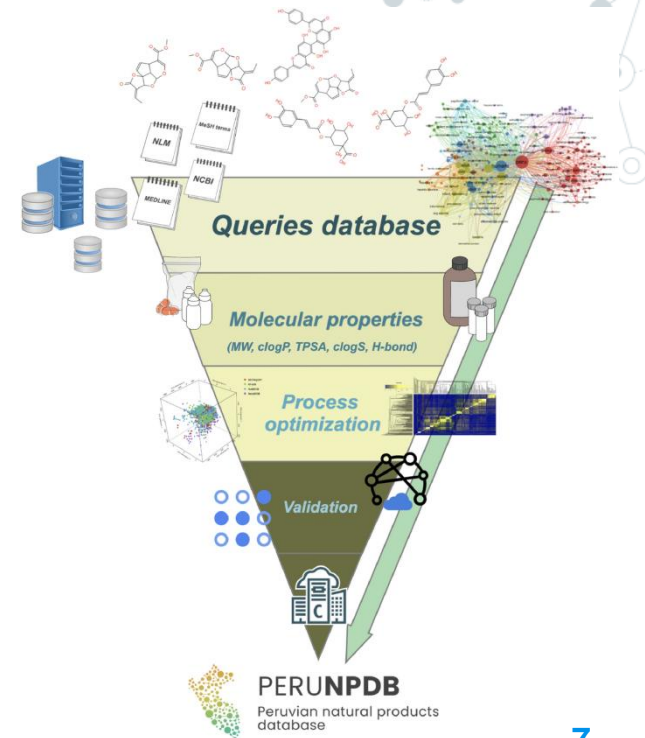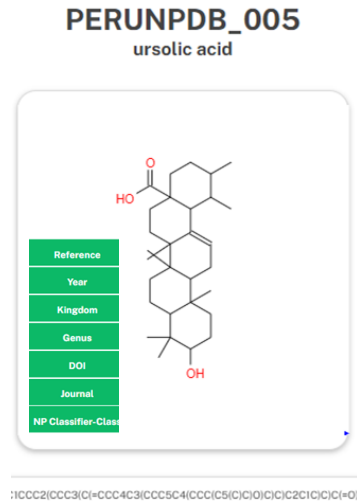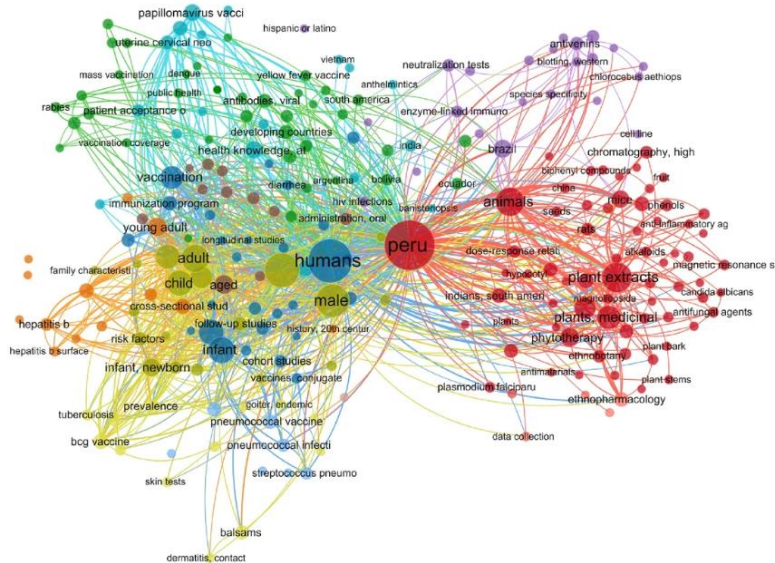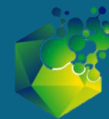
Molecular docking and Scoring Functions

Predict the best orientation and conformation of a ligand (small molecule) in a binding site of a receptor (Protein);

Estimate the binding affinity score with a scoring function
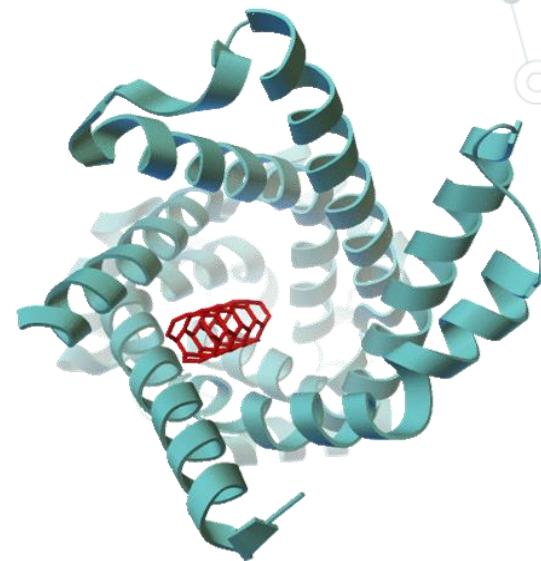
Molecular docking and Scoring Functions

PHYSICS BASED (Force Field)
Derived from mechanics force fields to calculate protein-ligand binding

KNOWLEDGE-BASED

- utilize knowledge derived from experimentally determined complexes
- obtains the score from summing pairwise statistical potentials between protein and ligand

EMPIRICAL

- Based on empirical observations;
- Calculates the affinity through the sum of terms that represent factors such as hydrogen bonding, van der Waals, eletrostactic and hydrophobic interactions;

MACHINE LEARNING

Trains scoring functions using features obtained from known protein-ligand binding experiments

Scoring Functions

Physics-based

DOCK: $E_{bind} = \sum_{i=1}^{L} \sum_{j=1}^{R} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right)$ (1)

Empirical

X-Score: $E_{bind} = w_0 + w_1 \Delta G_{vdW} + w_2 \Delta G_{Hbond} + w_3 \Delta G_{rot} + w_4 \Delta G_{hydro}$ (2)

Knowledge-based

PMF: $E_{bind} = \sum_{i=1}^{L} \sum_{j=1}^{R} -k_B T \ln[g(r)]$ (3)

Machine learning-based

RF-Score: $E_{bind} = f_{RF}(x_m)$

$x_{ij} = \sum_{i=1}^{L} \sum_{j=1}^{R} \Theta[d_{cutoff} - d_{ij}]$ (4)

Previous Work -> RFL-Score

9

Objective
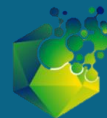
Proposes a GPx4 target-specific scoring function for molecular docking
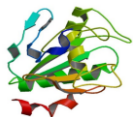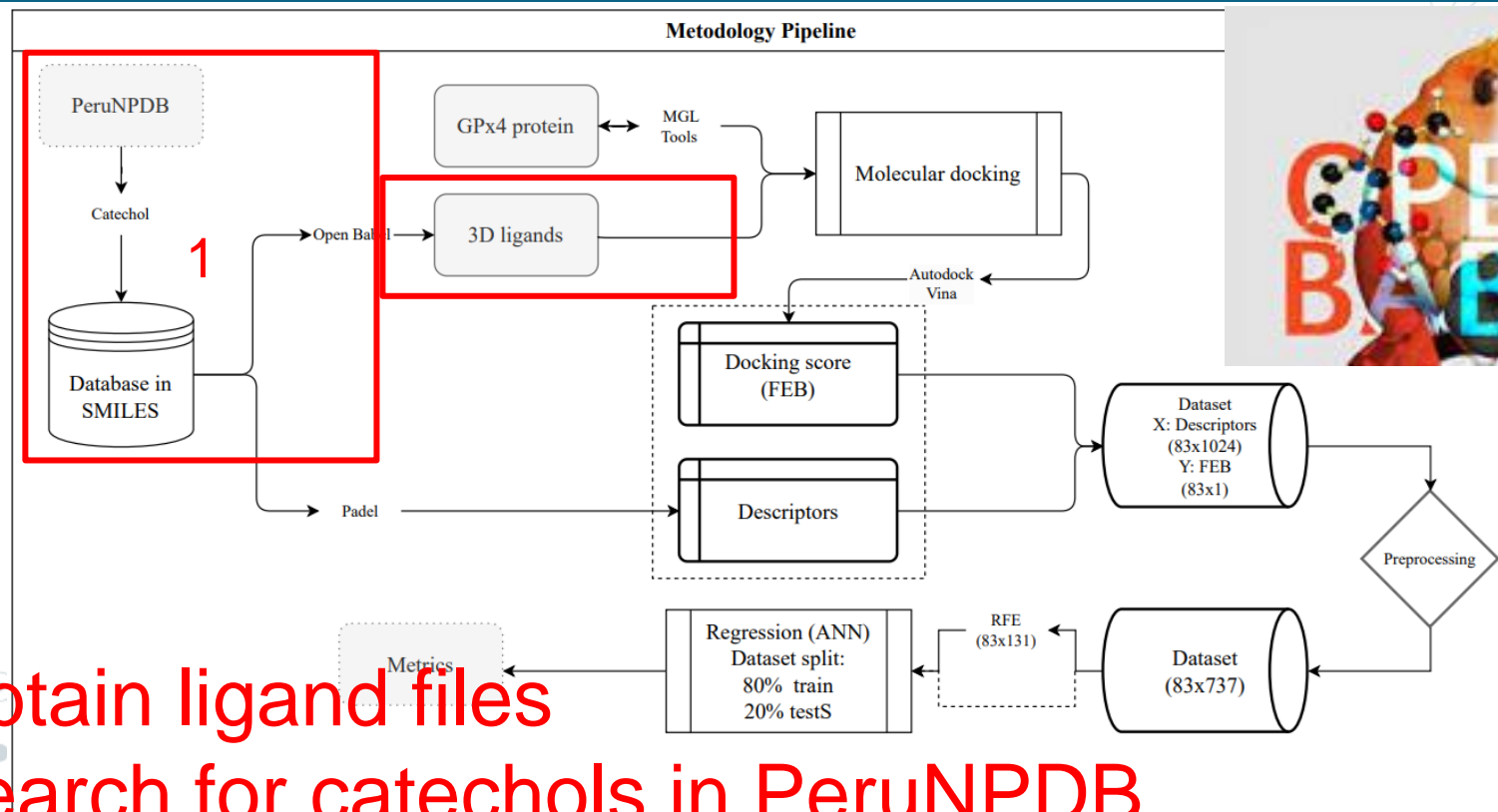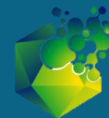
Training Set	ML algorithm	Validation

PERUNPDB

R2, RMSE, MSE

Available in: https://github.com/inefable12/ExpInSilico24
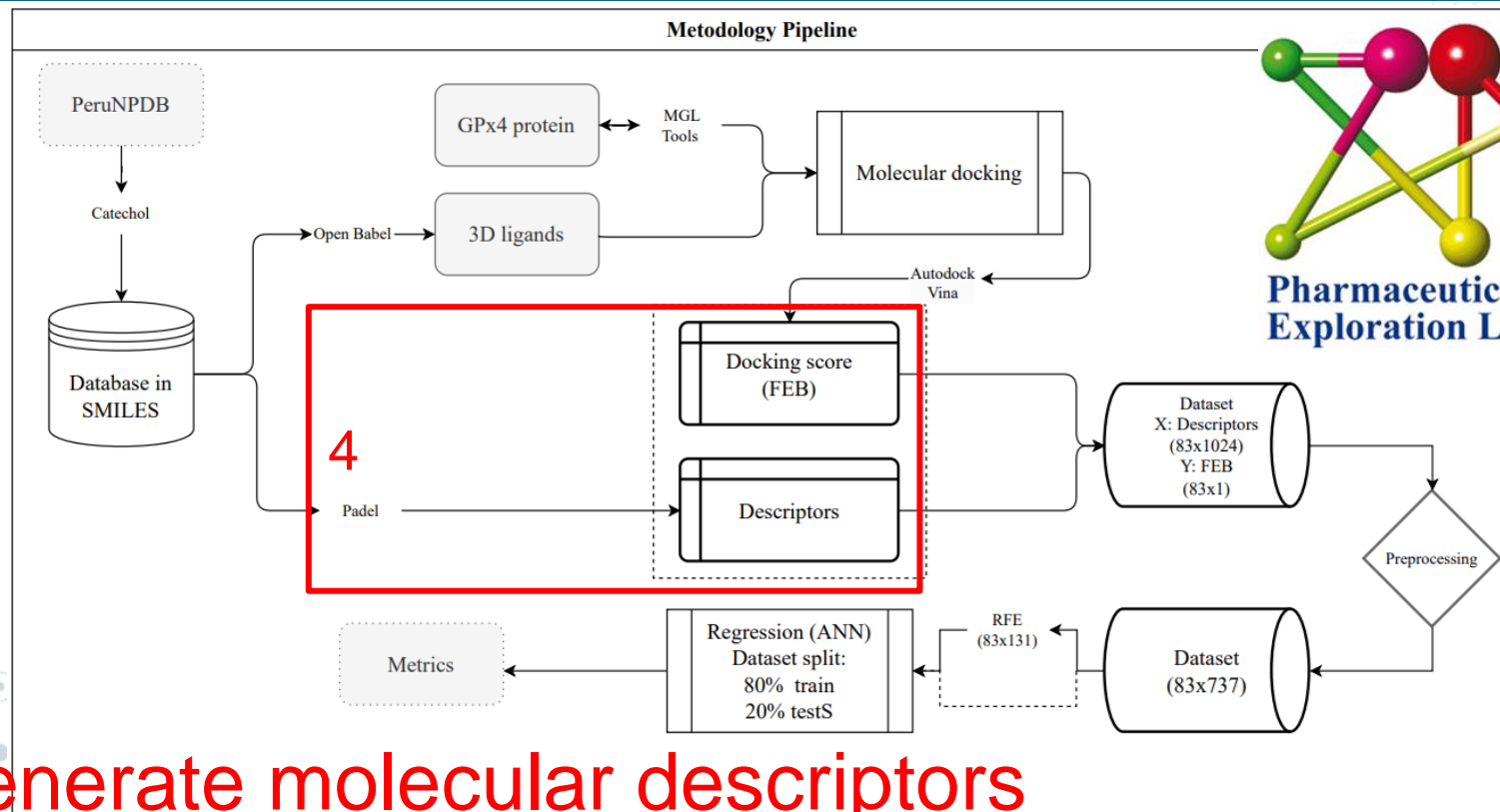
Methodology

Metodology Pipeline

- PeruNPDB
- Catechol
- Database in SMILES
- Open Babel
- 3D ligands
- GPx4 protein
- MGL Tools
- Molecular docking
- Autodock Vina
- Docking score (FEB)
- Padel
- Descriptors
- Dataset X: Descriptors (83x1024) Y: FEB (83x1)
- Preprocessing
- Dataset (83x737)
- RFE (83x131)
- Regression (ANN) Dataset split: 80% train 20% testS
- Metrics

Methodology

1. Obtain ligand files
Search for catechols in PeruNPDB

Methodology

Metodology Pipeline

2. Obtain and prepare receptor

Methodology

3. Perform molecular docking simulations
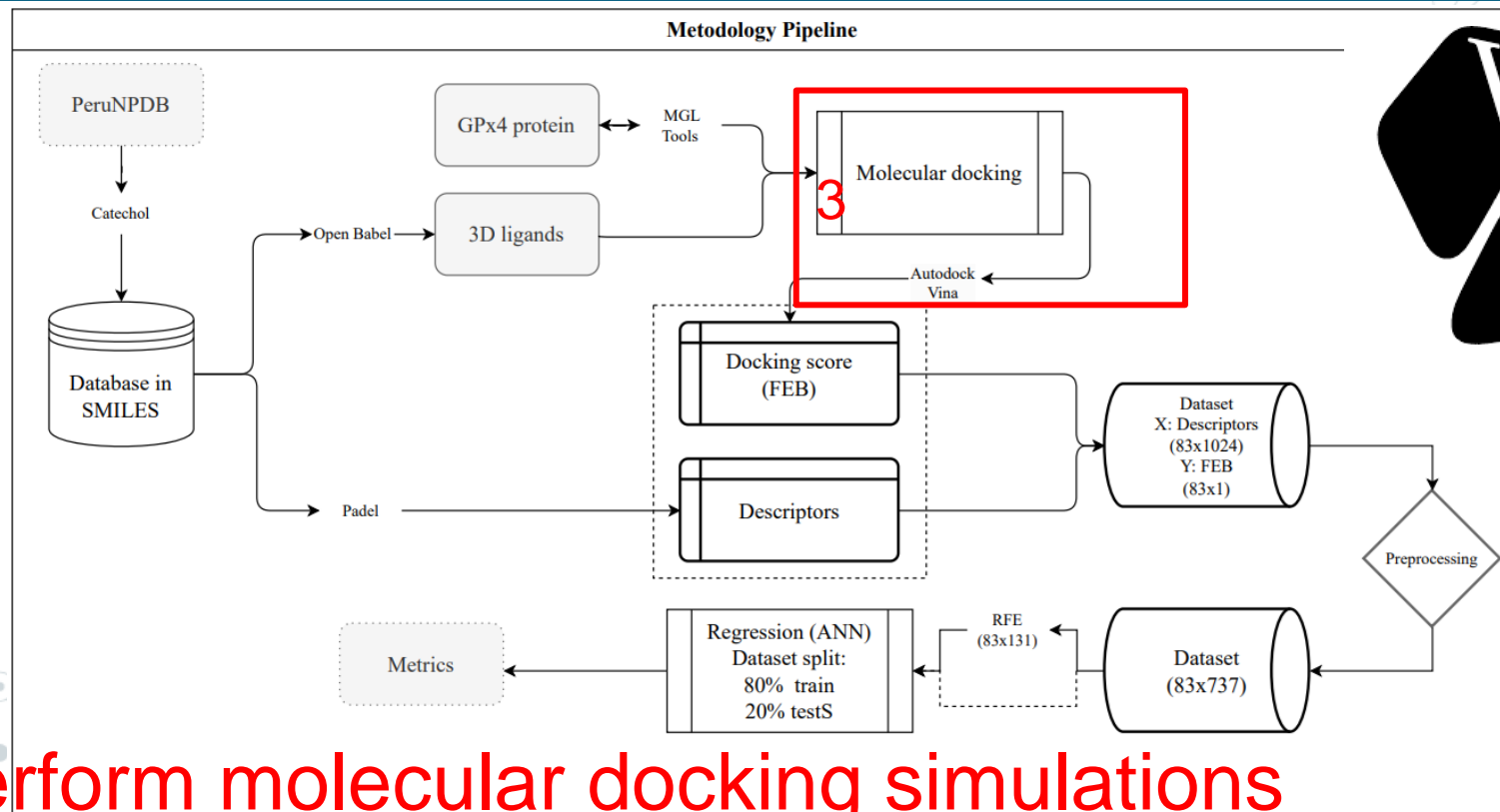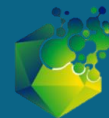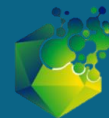
Methodology

4. Generate molecular descriptors

Methodology
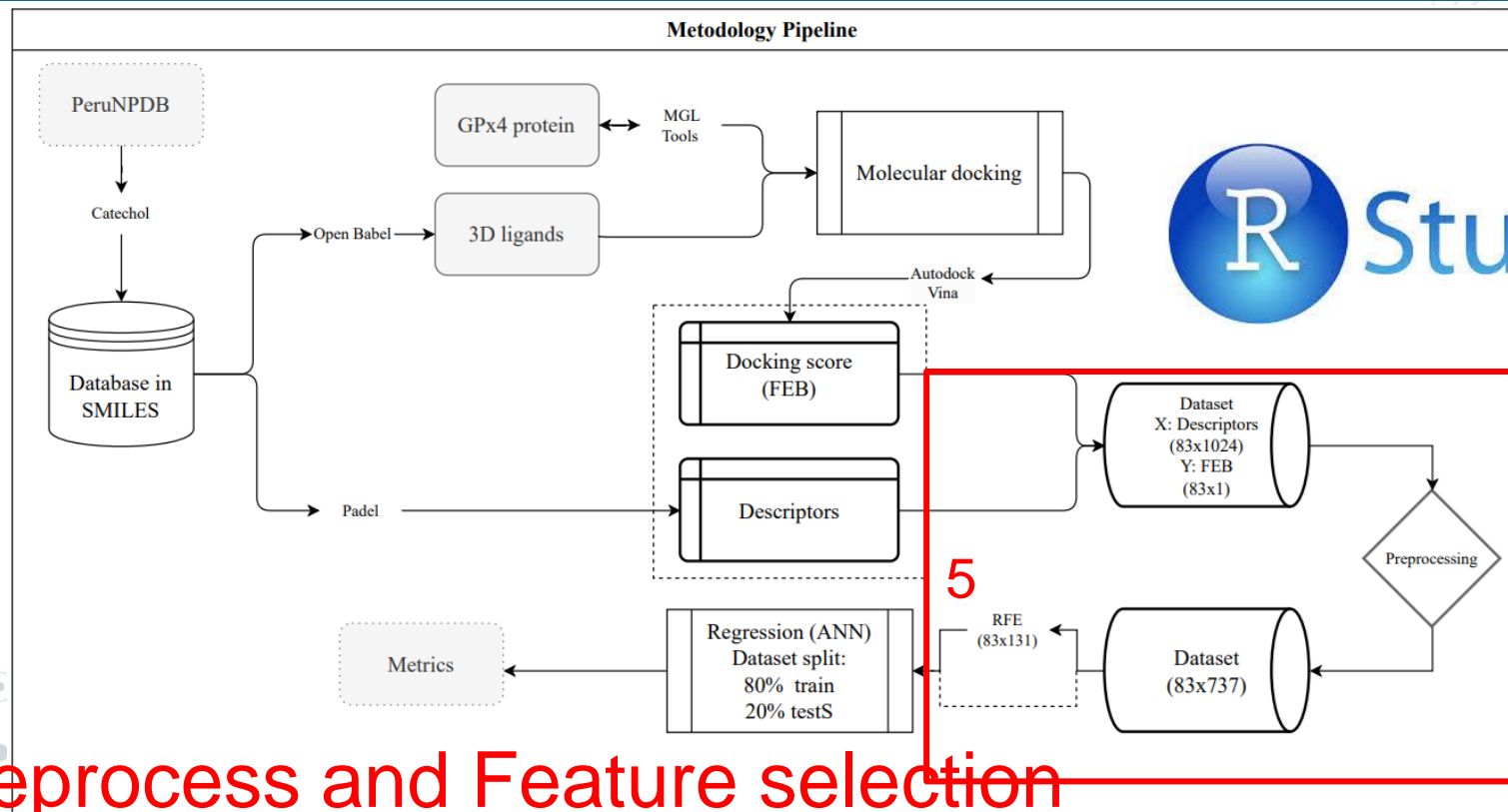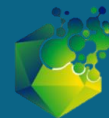
5. Preprocess and Feature selection

Methodology
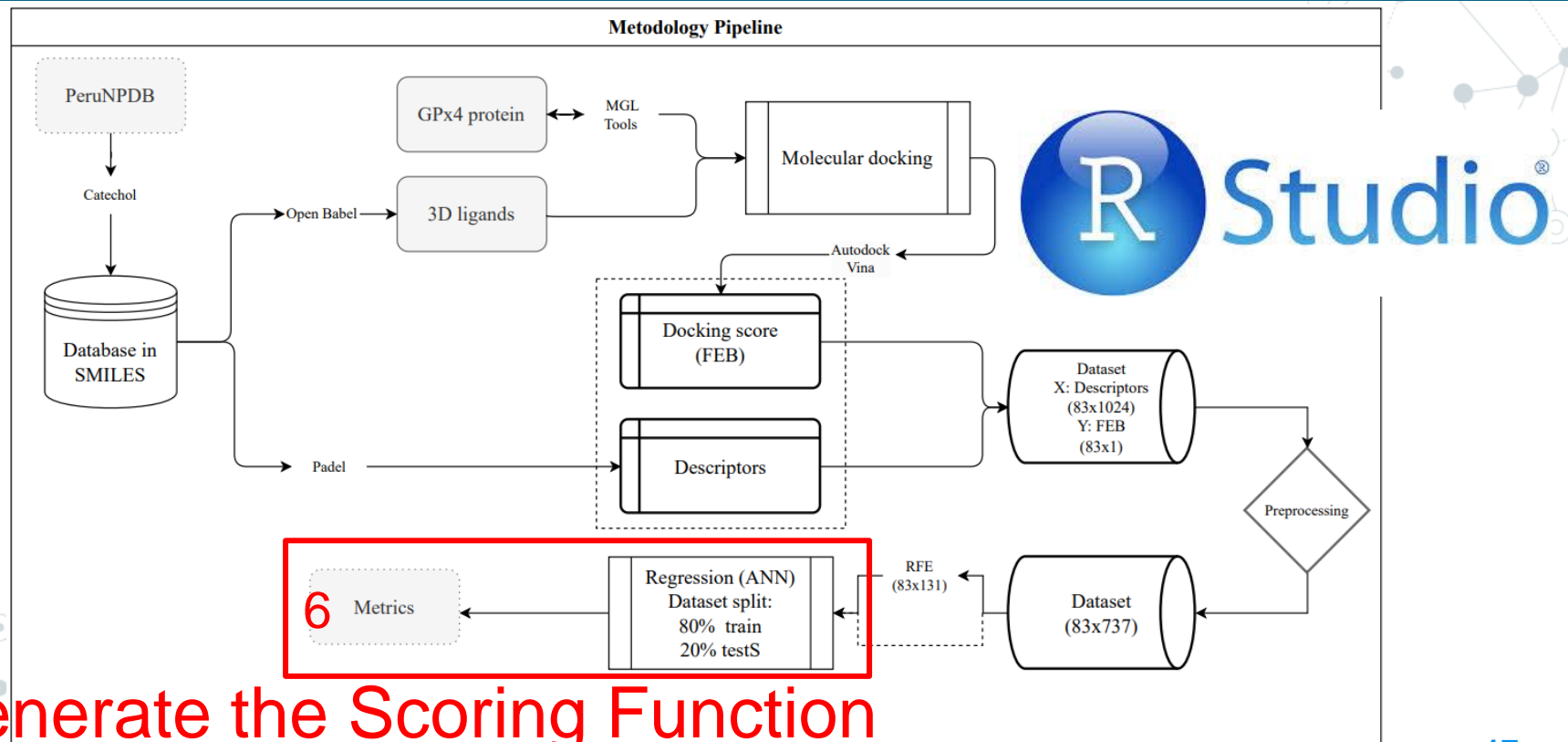
6. Generate the Scoring Function

Results | Molecules
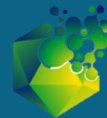
PERUNPDB
Peruvian Natural Products Database

280 Natural Products
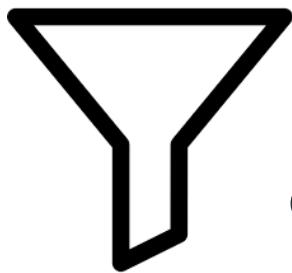
Functional group catechol

83 molecules (SMILES)

CC1CCC2(CCC3(C(=CCC4C3(CCC5C4(CCC(C5(C)C)O)C)C)C2C1C)C)C(=O)O
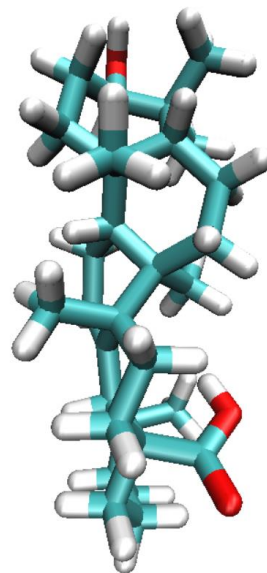
Molecules minimized and protonated

18
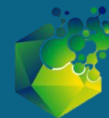
Results ⇒ Dataset Descriptors -> 1,024
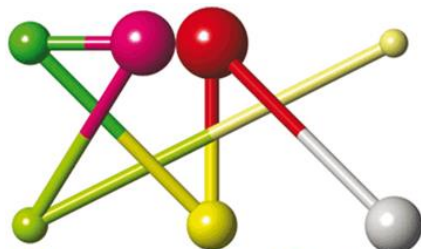
1,024 descriptors

Preprocessing
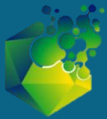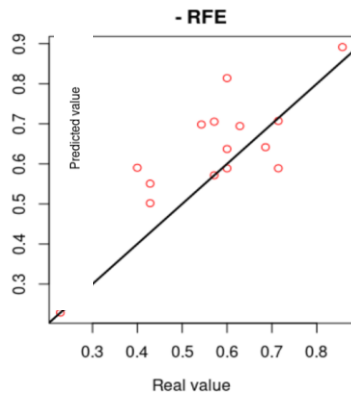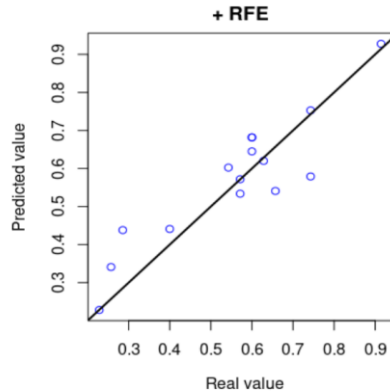
737 descriptors
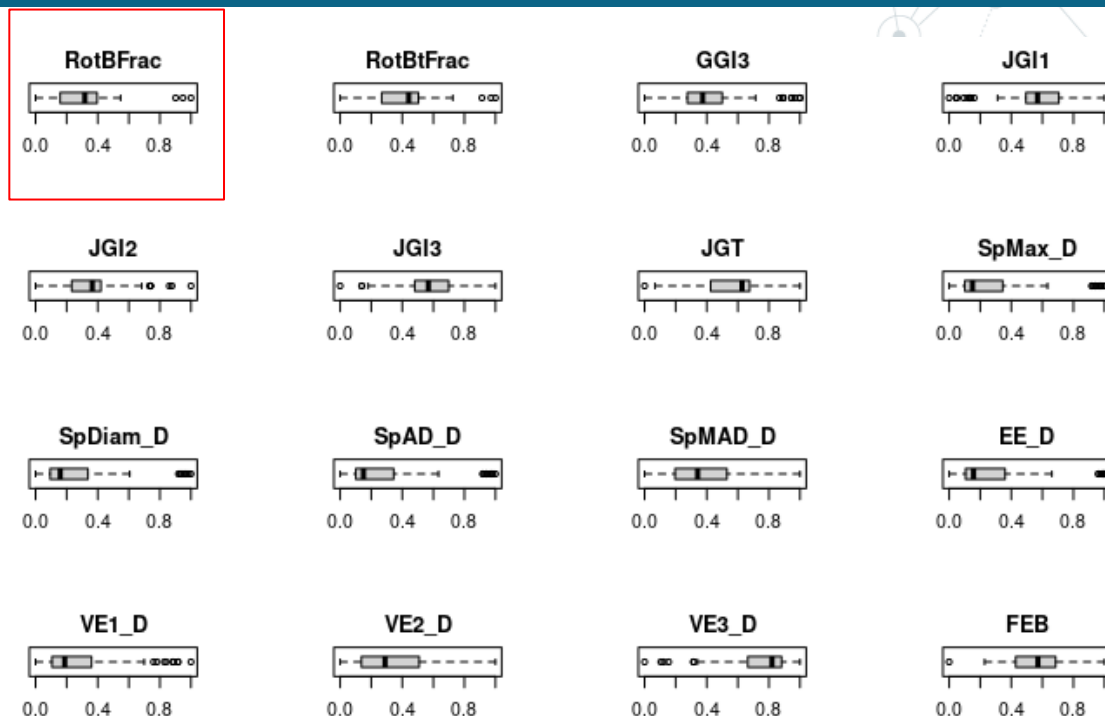
Results ⇒ Feature selection

Minimum RMSE 131 descriptors

Results ⇒ Metrics of the proposed scoring function
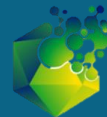

Input dataset: 83 (instances) X 131 descriptors + target (FEB)

Artificial Neural Network (ANN) algorithm

	R2	RMSE	MSE	MAE
Train	0.9894	0.0186	0.0003	0.0042
Test	0.6292	0.1062	0.0112	0.0809
Train + RFE	0.9881	0.0189	0.0003	0.0057
Test + RFE	0.8267	0.0787	0.0061	0.0596

Conclusion

★ Chelation therapy, particularly in relation to ferroptosis and the inhibition of Glutathione peroxidase 4 (GPx4), has emerged as a significant area of investigation

★ The proposed methodology and results can serve as a starting point for the search for catechols on others natural products databases as well as search for new inhibitors for GPx4

Future Work: Include more molecules in the input dataset (consider other databases); evaluate more ML algorithms; train new target-specific functions for other target proteins.

Available in: https://github.com/inefable12/ExpInSilico24

Acknowledgments
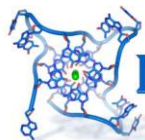
Thanks!

Contact information: Karina dos Santos Machado

E-mail: karina.machado@furg.br

COMBI-Lab, Grupo de Biología Computacional, Centro de Ciências Computacionais, Universidade Federal do Rio Grande, Rio Grande, RS, Brasil.

Natal, Brazil, August 27-29, 2024

@combilabfurg