

CHEMINFOICD3-03: Cheminfo. Int. Conf.
Drug Des. & Discov. Alfnas, Brazil-München,
Germany-Ch. Hill, Durham, USA, 2021



Predictive models as a useful tool for preclinical assay optimization in antimalarial compounds.

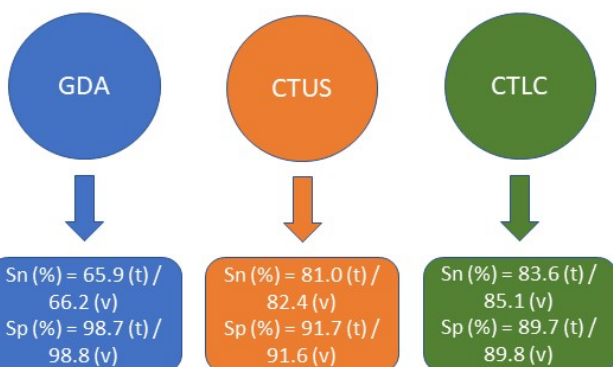
Viviana F. Quevedo-Tumaili ^{a, b}, Bernabe Ortega-Tenezaca ^{a, b}

^a RNASA-IMEDIR, Computer Science Faculty, University of A Coruña, 15071, A Coruña, Spain.

^b Universidad Estatal Amazónica UEA, Puyo, Pastaza, Ecuador.

Graphical Abstract

Modelos PTML



Abstract.

In this study, three Perturbation Theory Machine Learning (PTML) models were created to optimize preclinical assays on antimalarial compounds of the parasitic species of the genus *Plasmodium falciparum*. Between General Discriminant Analysis (GDA), Classification Tree with Univariate Splits (CTUS) and Classification Tree with Linear Combinations (CTLC). The PTML-CTLC presented the best performance with a Sensitivity percentage

equal to 83.6 for the training data set and 85.1 for validation; for specificity with a percentage of 89.8 for training and 89.7 for validation. The PTML-CTLC model has significant variables that could be a good option for pharmaceutical companies to optimize preclinical testing processes.

References

1. Kalanon, M.; McFadden, G. I., Malaria, Plasmodium falciparum and its apicoplast. *Biochem Soc Trans* 2010, 38, 775-82.
2. DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W., Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* 2016, 47, 20-33.
3. Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P., ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res* 2012, 40, D1100-7.
4. Quevedo-Tumaili, V. F.; Ortega-Tenezaca, B.; Gonzalez-Diaz, H., Chromosome Gene Orientation Inversion Networks (GOINs) of Plasmodium Proteome. *J Proteome Res* 2018, 17, 1258-1268.
5. Gonzalez-Diaz, H.; Arrasate, S.; Gomez-SanJuan, A.; Sotomayor, N.; Lete, E.; Besada-Porto, L.; Ruso, J. M., General Theory for Multiple Input-Output Perturbations in Complex Molecular Systems. 1. Linear QSPR Electronegativity Models in Physical, Organic, and Medicinal Chemistry. *Current Topics in Medicinal Chemistry* 2013, 13, 1713-1741.
6. Martinez, S. G.; Tenorio-Borroto, E.; Barbabosa Pliego, A.; Diaz-Albiter, H.; Vazquez-Chagoyan, J. C.; Gonzalez-Diaz, H., PTML Model for Proteome Mining of B-cell Epitopes and Theoretic-Experimental Study of Bm86 Protein Sequences from Colima Mexico. *J Proteome Res* 2017.
7. Vasquez-Dominguez, E.; Armijos-Jaramillo, V. D.; Tejera, E.; Gonzalez-Diaz, H., Multioutput Perturbation-Theory Machine Learning (PTML) Model of ChEMBL Data for Antiretroviral Compounds. *Mol Pharm* 2019, 16, 4200-4212.
8. Whittle, M.; Gillet, V. J.; Willett, P.; Loesel, J., Analysis of data fusion methods in virtual screening: similarity and group fusion. *J Chem Inf Model* 2006, 46, 2206-19.
9. Whittle, M.; Gillet, V. J.; Willett, P.; Loesel, J., Analysis of data fusion methods in virtual screening: theoretical model. *J Chem Inf Model* 2006, 46, 2193-205.
10. Weininger, D., SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Modeling* 1988, 28 31-36.