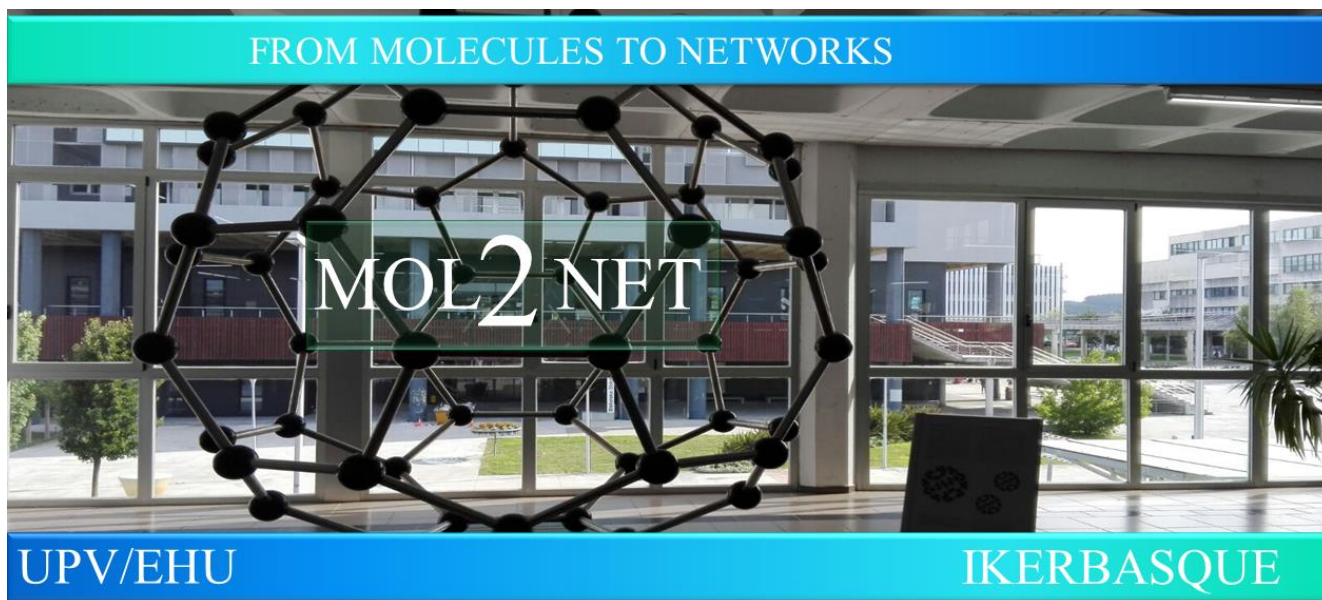




# MOL2NET'23, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 8th ed.



## Predicting Antimalarial Activity using Atomic Weight Vectors and Machine Learning

*Yoan Martínez López<sup>a</sup>, Juan A. Castillo Garit<sup>b</sup>, Gerardo M. Casanola-Martin<sup>c</sup>,  
Stephen J Barigye<sup>d</sup>, Oscar Martínez-Santiago<sup>e,f</sup>, Julio Madera Quintana<sup>a</sup>, Ansel  
Rodríguez González<sup>g</sup>, Jahiro Sutherland<sup>h</sup>*

<sup>a</sup> *Department of Computer Sciences, Faculty of Informatics, Camagüey University*

<sup>b</sup> *Universidad Tecnológica Metropolitana (UTEM), Santiago 8940577, Chile*

*Laboratorio de Bioinformática y Química Computacional, Universidad Católica del Maule, Talca,  
Chile*

<sup>c</sup> *Department of Coatings and Polymeric Materials, North Dakota State University*

<sup>d</sup> *Departamento de Química Física Aplicada, Facultad de Ciencias, Universidad Autónoma de Madrid  
(UAM), 28049 Madrid, Spain*

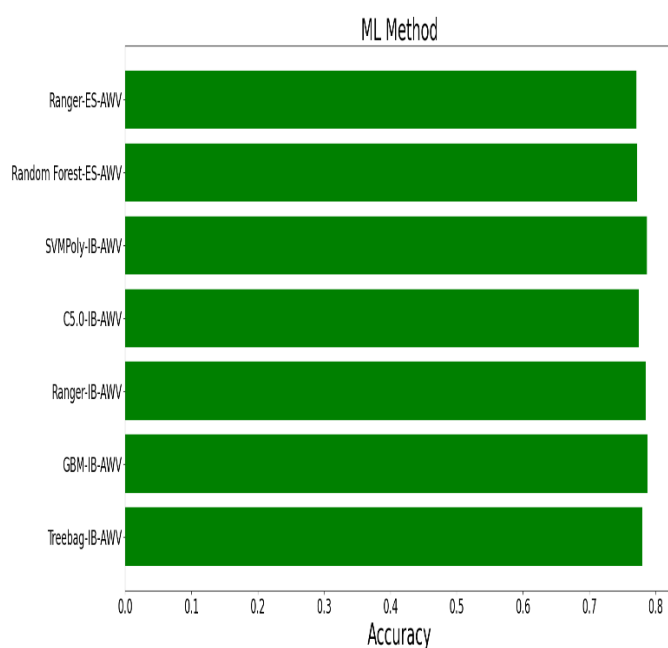
<sup>e</sup> *Alfa Vitamins Laboratories, Miami, Florida, 33166, USA*

<sup>f</sup> *Laboratorio de Bioinformática y Química Computacional, Universidad Católica del Maule, Talca,  
Chile*

<sup>g</sup> *Unidad de Transferencia Tecnológica, Centro de Investigación Científica y de Educación Superior  
de Ensenada*

<sup>h</sup> *Centro Regional Universitario de Colón. Universidad de Panamá*

## Graphical Abstract



## Abstract.

*Malaria is a disease caused by the Plasmodium parasite, which is transmitted through the bites of infected mosquitos. Only the Anopheles genus of mosquito can transmit malaria. The symptoms of this disease can include fever, vomiting, and headache. As millions of people are exposed to the threat of the Plasmodium parasite, it leads to millions of deaths annually. Therefore, there is a need to develop models for predicting compounds that can counteract this disease.*

*Objective: The primary objective of this research was to employ different techniques of machine learning on molecular descriptors obtained from Atomic Weight Vectors (AWV) and MD-LOVIs tool to predict the activity of potential antimalarial compounds.*

*Methods: Several machine learning techniques such as Ranger-ES-AWV (accuracy = 0.7714), Random Forest-ES-AWV (accuracy = 0.7718), SVMPoly-IB-AWV (accuracy = 0.787), C5.0-IB-AWV (accuracy = 0.7746), Ranger-IB-AWV (accuracy = 0.7854), GBM-IB-AWV (accuracy = 0.7882), and Treebag-IB-AWV (accuracy = 0.7798) were applied to predict the activity of antimalarial compounds.*

*Results: The results showed that the models obtained using machine learning techniques can be a powerful tool for predicting the activity of antimalarial compounds.*

*Conclusion: This study demonstrates the potential of machine learning techniques for predicting the activity of antimalarial compounds. These models can be used to identify new compounds with antimalarial properties and contribute to reducing the number of malaria-related deaths worldwide. The experiment file and dataset can be downloaded at: [https://github.com/cybervalient/DeepMD\\_AWV/blob/main/malaria.ipynb](https://github.com/cybervalient/DeepMD_AWV/blob/main/malaria.ipynb) and [https://github.com/cybervalient/DeepMD\\_AWV/blob/main/MD-Malaria.zip](https://github.com/cybervalient/DeepMD_AWV/blob/main/MD-Malaria.zip)*

## References

- Garit, J. A. C., Carmenate, Y. C. I., Torrens, F., & DiÁ, K. (2020). Aplicaciones y potencialidades de los métodos de diseño computacional en estudios ambientales y farmacocinéticos. *Anales de la Academia de Ciencias de Cuba*, 11(1), 811.
- Martínez-López, Y., Castillo-Garit, J. A., Casanola-Martin, G. M., Rasulev, B., Rodríguez-Gonzalez, A. Y., Martínez-Santiago, O., & Barigye, S. J. (2023). Exploring proteasome inhibition using atomic weighted vector indices and machine learning approaches. *Molecular Diversity*, 1-12.
- Martínez-López, Y., Marrero-Ponce, Y., Barigye, S. J., Teran, E., Martínez-Santiago, O., Zambrano, C. H., & Torres, F. J. (2020). When global and local molecular descriptors are more than the sum of its parts: Simple, But Not Simpler?. *Molecular Diversity*, 24(4), 913-932.
- Martínez-López, Y., et al., Prediction of aquatic toxicity of benzene derivatives using molecular descriptor from atomic weighted vectors. *Environmental Toxicology and Pharmacology*, 2017. 56: p. 314-321.
- Martínez-López, Y., et al., State of the Art Review and Report of New Tool for Drug Discovery. *Current Topics in Medicinal Chemistry*, 2017. 17: p. 1-20.
- Mariki, M., Mkoba, E., & Mduma, N. (2022). Combining clinical symptoms and patient features for malaria diagnosis: machine learning approach. *Applied Artificial Intelligence*, 36(1), 2031826.
- Saavedra Reyes, L. M. (2020). Aplicaciones biológicas de la teoría QSAR en el control del mosquito *Aedes aegypti* L (Doctoral dissertation, Universidad Nacional de La Plata).
- Tran, T., & Ekenna, C. (2023). Molecular Descriptors Property Prediction Using Transformer-Based Approach. *International Journal of Molecular Sciences*, 24(15), 11948.
- Zothantluanga, J. H., Chetia, D., Rajkhowa, S., & Umar, A. K. (2023). Unsupervised machine learning, QSAR modelling and web tool development for streamlining the lead identification process of antimalarial flavonoids. *SAR and QSAR in Environmental Research*, 34(2), 117-146.