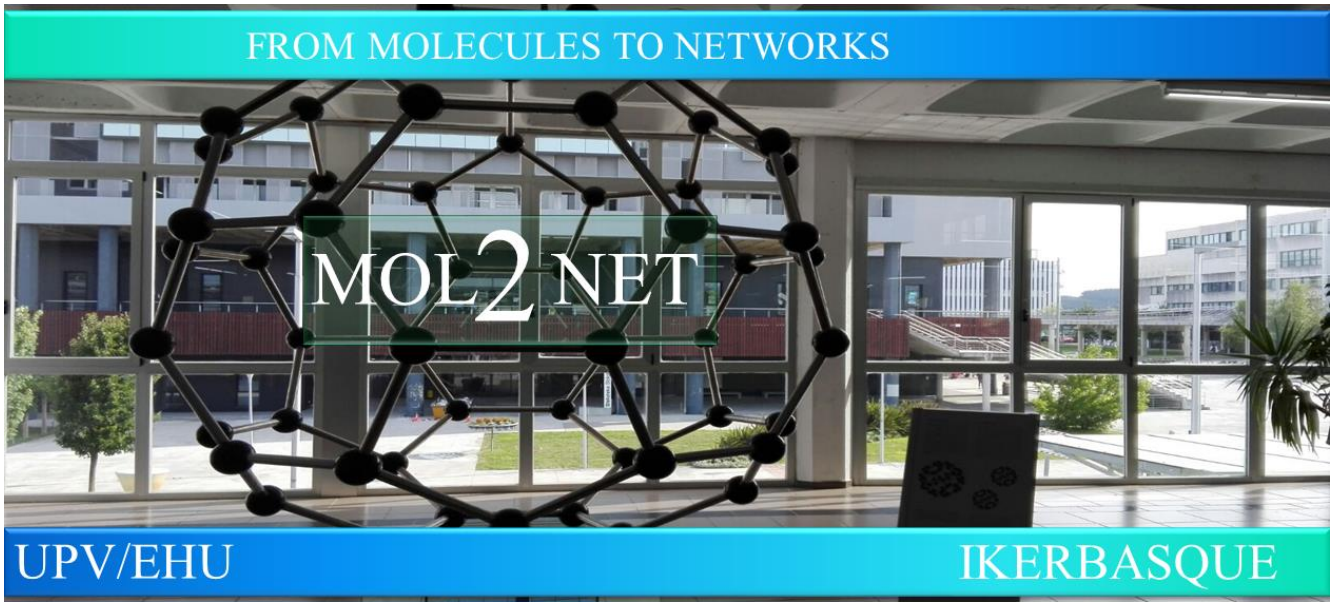
MDPI sciforum

MOL2NET'21, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 7th ed.

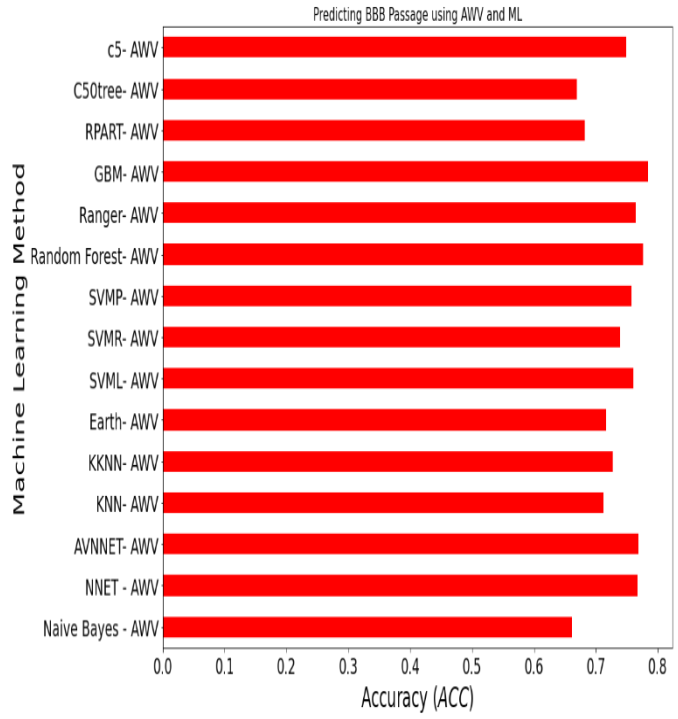


Predicting Blood-Brain Barrier Passage using Atomic Weighted Vector and Machine Learning

Yoan Martínez López[a], Yanaima Jauriga Ortiz[a], Ansel Rodríguez González[b], Juan A. Castillo Garit[c] , Gerardo M. Casanola-Martin[d] , Julio Madera Quintana[a]

[a] *Department of Computer Sciences, Faculty of Informatics, Camagüey University*
[b] *Unidad de Transferencia Tecnológica, Centro de Investigación Científica y de Educación Superior de Ensenada*
[c] *Unidad de Toxicología Experimental, Universidad de Ciencias Médicas de Villa Clara*
[d] *Department of Coatings and Polymeric Materials, North Dakota State University*

Graphical Abstract	**Abstract.**
	The blood-brain barrier (BBB) is a highly selective permeability barrier that separates circulating blood from brain extracellular fluid in the central nervous system (CNS). This barrier allows the passage of water, some gases, and lipid-soluble molecules by passive diffusion, as well as the selective transport of molecules such as glucose and amino acids that are crucial for neuronal function. In this research, we present an exploratory study, where several



machine learning techniques are applied to predict blood-brain barrier passage by applying molecular descriptors based on atomic vectors, obtained by MD-LOVIs software. Several techniques such as KNN-AWV(ACC= 0.712), AVNNET-AWV(ACC=0.768), Random Forest-AWV(ACC=0.776) and GBM-AWV(ACC=0.784) obtained good prediction performance. The results show that machine learning techniques are powerful tools for the prediction of this activity. The experiment file and dataset can be downloaded at: https://github.com/cybervalient/DeepMD_AWV/blob/main/BBB-AWV-ML.R and https://github.com/cybervalient/DeepMD_AWV/blob/main/BBB%20caret.rar

References

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.

Castillo-Garit, J. A., Casanola-Martin, G. M., Le-Thi-Thu, H., & Barigye, S. J. (2017). A simple method to predict blood-brain barrier permeability of drug-like compounds using classification trees. Medicinal Chemistry, 13(7), 664-669.

Svetnik, V., et al., Random forest: a classification and regression tool for compound classification and QSAR modeling. Journal of chemical information and computer sciences, 2003. 43(6): p. 1947-1958. • Liaw, A. and M. Wiener, Classification and regression by randomForest. R news, 2002. 2(3): p. 18-22.

Gomatam, A., Joseph, B., Gawde, U., Raikuvar, K., & Coutinho, E. (2022). Predicting toxicity of endocrine disruptors and blood–brain barrier permeability using chirality-sensitive descriptors and machine learning. Computational Toxicology, 24, 100240.

Witten, I.H.; Frank, E., Data Mining: Practical Machine Learning Tools and Techniques. Second ed.; Morgan Kaufmann Publishers: 2005.

Y. Freund and R.E. Schapire (1997) "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, 55(1):119-139.

J.H. Friedman (2001). "Greedy Function Approximation: A Gradient Boosting Machine," Annals of Statistics 29(5):1189-1232.

J.H. Friedman (2002). "Stochastic Gradient Boosting," Computational Statistics and Data Analysis 38(4):367-378.