Improving Molecular Docking Accuracy for Binding Affinity Prediction in Virtual Screening: A Critical Review of Methods and Benchmarking Integrity

André Maio[1], Denner Evaristo dos Santos Silveira[2], Nathalie Esper[3], and Shani Yahav[4]

[1]Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
[2]Department of Chemistry, Université Paris Cité
[3]Faculty of Chemistry, Université de Strasbourg
[4]Department of Chemistry, Bar-Ilan University
Erasmus Mundus Joint Master Chemoinformatics+
Project Proposer: Pavel Polishchuk
Academic Supervisors: Olivier Taboureau, Alexey Orlov

Abstract

Molecular docking is a foundational tool in structure-based drug discovery, yet standard scoring functions struggle significantly with binding affinity prediction, typically achieving Pearson correlation coefficients of only 0.5–0.6 against experimental constants. This project critically reviews methodological advances published between 2022 and 2026 that aim to improve score-affinity correlation while maintaining throughput suitable for ultra-large chemical libraries. A paradigm-shifting finding in recent literature is the pervasiveness of data leakage; approximately 50% of the standard CASF-2016 test set shares high structural or ligand similarity with PDBbind training data, artificially inflating machine learning performance claims. Upon rigorous evaluation on leak-proof benchmarks, pure coordinate-based deep learning methods frequently exhibit limited generalization. Conversely, methods demonstrating robust, genuine performance include physics-informed machine learning architectures (e.g., PIGNet2, SCORCH), consensus scoring leveraging 3–5 diverse methods, and explicit water thermodynamics modeling. To translate these findings into practical application, we propose an evidence-based, tiered virtual screening workflow that balances computational cost with progressive accuracy improvements, culminating in recommendations for statistically rigorous experimental validation.