

Machine Learning Approaches for Predicting ADMET Properties of Novel Compounds

Lucas Mayert*

Department of Cheminformatics and Molecular Modeling, ETH Zurich, Zurich, Switzerland

*Corresponding author: Lucas Mayert, Department of Cheminformatics and Molecular Modeling, ETH Zurich, Zurich, Switzerland, E-mail: luc.mayert@eth.ch

Received date: January 02, 2025, Manuscript No. ipchi-25-20771; **Editor assigned date:** January 04, 2025, PreQC No. ipchi-25-20771 (PQ); **Reviewed date:** January 18, 2025, QC No. ipchi-25-20771; **Revised date:** January 24, 2025, Manuscript No. ipchi-25-20771 (R); **Published date:** January 31, 2025, DOI: 10.36648/2470-6973.9.01.239

Citation: Mayert L (2025) Machine Learning Approaches for Predicting ADMET Properties of Novel Compounds. Chem inform Vol.9.No.01: 239.

Introduction

The discovery and development of novel therapeutic compounds is an inherently complex and resource-intensive process, with high rates of attrition in later stages of clinical trials. Among the critical bottlenecks in drug development are pharmacokinetic and toxicological considerations collectively described as ADMET-Absorption, Distribution, Metabolism, Excretion, and Toxicity. A drug candidate with promising therapeutic activity can still fail due to poor oral bioavailability, off-target toxicity, or rapid metabolic clearance. Traditionally, evaluating ADMET properties has relied on in vitro assays and animal testing, both of which are costly, time-consuming, and ethically constrained. The emergence of Machine Learning (ML) approaches, fueled by large-scale chemical databases and computational advances, has revolutionized the predictive landscape of ADMET profiling. By capturing complex, non-linear relationships between molecular descriptors and pharmacological outcomes, ML models hold the potential to identify liabilities early, reduce experimental burden, and accelerate the delivery of safe and effective drugs to the market [1].

Description

The foundation of machine learning-based ADMET prediction lies in the availability of diverse and well-curated datasets. Public repositories such as ChEMBL, PubChem BioAssay, and ToxCast, along with proprietary pharmaceutical datasets, provide structural and experimental information on thousands of molecules. Using these resources, researchers derive molecular descriptors-including physicochemical features (e.g., lipophilicity, molecular weight, polar surface area), structural fingerprints, and 3D conformational data-that serve as input for ML models. Classical algorithms such as Random Forests (RF), Support Vector Machines (SVM), and Gradient Boosting Machines (GBM) have been extensively applied to ADMET classification tasks, offering interpretable predictions of solubility, permeability, and toxicity endpoints. More recently, deep learning models such as Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs) have demonstrated superior performance by learning directly from molecular structures without relying solely on pre-computed descriptors [2].

Among the most significant contributions of ML to ADMET prediction is its role in absorption and distribution modeling. Absorption, particularly oral bioavailability, is influenced by solubility, intestinal permeability, and transporter interactions. ML models trained on high-throughput permeability assays (e.g., Caco-2 cell data) or solubility datasets can predict oral drug absorption with remarkable accuracy, often outperforming traditional rule-based methods like Lipinski's "Rule of Five." Similarly, distribution profiles, including plasma protein binding and blood-brain barrier penetration, have been modeled using ensemble learning techniques and neural networks. These predictions are particularly valuable for Central Nervous System (CNS) drugs, where blood-brain barrier permeability is a crucial determinant of efficacy. By integrating physicochemical properties and molecular interaction networks, ML provides a holistic approach to estimating compound disposition, enabling early-stage filtering of candidates with poor pharmacokinetic potential [3].

Metabolism and excretion, two other pillars of ADMET, are highly amenable to ML-based approaches. Predicting metabolism involves identifying potential sites of enzymatic modification, particularly by cytochrome P450 (CYP) enzymes. ML models trained on CYP-substrate and inhibitor datasets have been successful in classifying molecules as substrates, inhibitors, or inducers, as well as pinpointing metabolic "hotspots." Such insights are vital to anticipating drug-drug interactions and optimizing metabolic stability. Excretion, which is governed by renal clearance and biliary elimination, has also been modeled through ML algorithms that incorporate structural features and transporter interactions. In this context, ML not only predicts clearance rates but also identifies molecular determinants of elimination pathways. By providing early insights into metabolic and excretory liabilities, ML enables medicinal chemists to rationally modify compounds, extending half-life and reducing the risk of adverse interactions. Perhaps the most challenging and critical ADMET property to predict is toxicity. Adverse effects, ranging from hepatotoxicity and cardiotoxicity to genotoxicity, are a leading cause of clinical trial failures and post-market drug withdrawals [5].

Conclusion

Machine learning approaches are rapidly transforming ADMET prediction into a powerful and indispensable tool for modern drug discovery. By leveraging vast chemical datasets and advanced computational algorithms, ML provides early, accurate, and scalable assessments of pharmacokinetic and toxicological properties, significantly reducing reliance on costly experimental screening. The integration of ML into ADMET workflows allows for the rapid triaging of candidate molecules, rational design modifications, and the prioritization of compounds with optimal safety and efficacy profiles. Nevertheless, challenges remain, including data heterogeneity, limited representation of rare toxicities, and the need for interpretable models that can guide chemical design. Future progress will likely involve hybrid approaches that integrate ML with systems pharmacology, multi-omics datasets, and high-throughput experimentation, creating a feedback loop between computational prediction and experimental validation. As these advances unfold, machine learning will not only streamline the identification of viable drug candidates but also enhance the safety and efficiency of therapeutic development, ultimately accelerating the translation of scientific discoveries into clinical solutions.

Acknowledgement

None.

Conflict of Interest

None.

References

1. Shang Y, Hao Q, Jiang K, He M, Wang J (2020). Discovery of heterocyclic carbohydrazide derivatives as novel selective fatty acid amide hydrolase inhibitors: Design, synthesis and anti-neuroinflammatory evaluation. *Bioorg Med Chem Lett* 30: 127118.
2. Morse GD, Fischl MA, Shelton MJ, Cox SR, Driver M, et al. (1997). Single-dose pharmacokinetics of delavirdine mesylate and didanosine in patients with human immunodeficiency virus infection. *Antimicrob Agents Chemother* 41: 169-174.
3. Kashtoh H, Baek KH (2022). Recent updates on phytoconstituent alpha-glucosidase inhibitors: An approach towards the treatment of type two diabetes. *Plants* 11: 2722.
4. Thakor PM, Patel JD, Patel RJ, Chaki SH, Khimani AJ, et al. (2024). Exploring new schiff bases: Synthesis, characterization and multifaceted analysis for biomedical applications. *ACS Omega* 9: 35431-35448.
5. Harohally NV, Cherita C, Bhatt P, Anu Appaiah KA (2017). Antiaflatoxigenic and antimicrobial activities of Schiff bases of 2-hydroxy-4-methoxybenzaldehyde, cinnamaldehyde and similar aldehydes. *J Agric Food Chem* 65: 8773-8778.