World Academy of Science, Engineering and Technology
International Journal of Chemical and Molecular Engineering
Vol:18, No:03, 2024

Current Methods for Drug Property Prediction in the Real World

Authors : Jacob Green, Cecilia Cabrera, Maximilian Jakobs, Andrea Dimitracopoulos, Mark van der Wilk, Ryan Greenhalgh

Abstract : Predicting drug properties is key in drug discovery to enable de-risking of assets before expensive clinical trials and to find highly active compounds faster. Interest from the machine learning community has led to the release of a variety of benchmark datasets and proposed methods. However, it remains unclear for practitioners which method or approach is most suitable, as different papers benchmark on different datasets and methods, leading to varying conclusions that are not easily compared. Our large-scale empirical study links together numerous earlier works on different datasets and methods, thus offering a comprehensive overview of the existing property classes, datasets, and their interactions with different methods. We emphasise the importance of uncertainty quantification and the time and, therefore, cost of applying these methods in the drug development decision-making cycle. To the best of the author's knowledge, it has been observed that the optimal approach varies depending on the dataset and that engineered features with classical machine learning methods often outperform deep learning. Specifically, QSAR datasets are typically best analysed with classical methods such as Gaussian Processes, while ADMET datasets are sometimes better described by Trees or deep learning methods such as Graph Neural Networks or language models. Our work highlights that practitioners do not yet have a straightforward, black-box procedure to rely on and sets a precedent for creating practitioner-relevant benchmarks. Deep learning approaches must be proven on these benchmarks to become the practical method of choice in drug property prediction.

Keywords : activity (QSAR), ADMET, classical methods, drug property prediction, empirical study, machine learning
Conference Title : ICCBMC 2024 : International Conference on Chemical Biology and Medicinal Chemistry
Conference Location : Barcelona, Spain
Conference Dates : March 04-05, 2024