Assessing Methods and Obstacles in Chemical Space Exploration

Shawn Reeves[1], Benjamin DiFrancesco[1], Vijay Shahani[1], Stephen MacKinnon[1], Andreas Windemuth[1], and Andrew Brereton[1]

[1]Cyclica, Inc.

August 17, 2020

Abstract

Benchmarking the performance of generative methods for drug design is complex and multifaceted. In this report, we propose a separation of concerns for de novo drug design, categorizing the task into three main categories: generation, discrimination, and exploration. We demonstrate that changes to any of these three concerns impacts benchmark performance for drug design taks. In this report we present Deriver, an open-source Python package that acts as a modular framework for molecule generation, with a focus on integrating multiple generative methods. Using Deriver, we demonstrate that changing parameters related to each of these three concerns impacts chemical space traversal significantly, and that the freedom to independently adjust each is critical to real-world applications having conflicting priorities. We find that combining multiple generative methods can improve optimization of molecular properties, and lower the chance of becoming trapped in local minima. Additionally, filtering molecules for drug-likeness (based on physicochemical properties and SMARTS pattern matching) before they are scored can hinder exploration, but can improve the quality of the final molecules. Finally, we demonstrate that any given task has an exploration algorithm best suited to it, though in practice linear probabilistic sampling generally results in the best outcomes, when compared to Monte Carlo sampling or greedy sampling. We intend that Deriver, which is being made freely available, will be helpful to others interested in collaboratively improving existing methods in de novo drug design centered around inheritance of molecular structure, modularity, extensibility, and separation of concerns.

Hosted file

Reevesetal2020.pdf available at https://authorea.com/users/351613/articles/476178-assessing-methods-and-obstacles-in-chemical-space-exploration