

GuacaMol Benchmarking Models for De Novo Molecular Design

Segler, M.¹

¹BenevolentAI London UK.

Recently, generative models based on deep neural networks have been proposed to perform de-novo design, that is to directly generate molecules with required property profiles by virtual design-make-test cycles [1,2]. Neural generative models can learn to produce diverse and synthesisable molecules from large datasets, for example by employing recurrent neural networks, which makes them simpler to set up and potentially more powerful than established de novo design approaches relying on hand-coded rules or fragmentation schemes.

Even though dozens of different models have been proposed in the last few months, comparison studies against strong baselines and established models have only seldom been performed. Here, we present a collection of benchmarks to assess the performance of ligand-based de novo-design approaches [3]. The benchmark tasks encompass measuring the fidelity of the models to reproduce the property distribution of the training sets, and a variety of single- and multi-objective optimisation tasks meaningful for drug design.

We found that generative models perform well in the task of learn to generate diverse and valid molecules. Our results also show that graph-based genetic algorithms slightly outperform neural models in terms of optimization performance, indicating potential room for improvement for neural generative models.

The benchmark code and pretrained models are available under github.com/BenevolentAI/guacamol or installable via “pip install guacamol”. With our benchmark, we hope to contribute to a broader discussion in the chemoinformatics community on how to assess the quantitative and qualitative performance of de novo-design algorithms.

References

- [1] Segler et al., ACS Cent. Sci., 2018, 4, 120–131
- [2] Gómez-Bombarelli et al, ACS Cent. Sci., 2018, 4, 268–276
- [3] Brown, Fiscato, Segler, Vaucher, *J. Chem. Inf. Model.* 2019, 1096-1108