Chemical Language Models and Prediction Anatomy

Jürgen Bajorath[1,2,3,4]

[1]Department of Life Science Informatics and Data Science, B-IT
[2]Interdisciplinary Area AI in the Life Sciences,
Lamarr Institute for Machine Learning and Artificial Intelligence
[3]Chemical Biology & Medicinal Chemistry Unit, LIMES Institute
University of Bonn, Friedrich-Hirzebruch-Allee 5/6, D-53115 Bonn,
[4]Department of Neurobiology & Biophysics, School of Medicine,
University of Washington, Seattle, WA 98195.
E-mail: bajorath@bit.uni-bonn.de

Chemical language models (CLMs) are based on textual representations of molecular structure and often conditioned using context-dependent rules. Transformers have become preferred CLM architectures. Their performance is largely determined by the self-attention mechanism. A hallmark of CLMs is their ability to learn a variety of mappings of molecular representations and associated properties. The versatility of CLMs in addressing machine translation tasks and conditioning them through defined property constraints provides many new opportunities for generative molecular design. Transformer CLMs often produce promising results in off-the-beaten-path prediction scenarios. However, rationalizing predictions of transformers is challenging and a topical issue in explainable artificial intelligence (XAI). So far, transformer predictions have mostly been analyzed by determining attention weight distributions and attention flow, but other approaches are emerging. In this lecture, CLMs are presented for different applications in molecular design, for instance, the prediction of potent compounds or peptidomimetics using different input data. In addition, new model-agnostic XAI concepts are introduced. CLM predictions are dissected through scientific rationale-guided control calculations at different levels, taking principles from XAI and scientific theory into account. Going beyond individual models and predictions, the analysis illustrates the interplay of model explanation, interpretation, and causal reasoning in machine learning.