Evaluation of Reinforcement Learning in Transformer-based Molecular Design

Jiazhen He[1*], Tibo Alessandro[1], Jon Paul Janet[1], Eva Nittinger[2], Christian Tyrchan[2], Werngard Czechtizkyr[2], Ola Engkvist[1,3]

[1]*Molecular AI, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden*
[2]*Medicinal Chemistry, Research and Early Development, Respiratory and Immunology (R&I), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden*
[3]*Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg, Sweden*
[*]*Correspondence: jiazhen.he@astrazeneca.com*

Supplementary material

Effect of learning rate

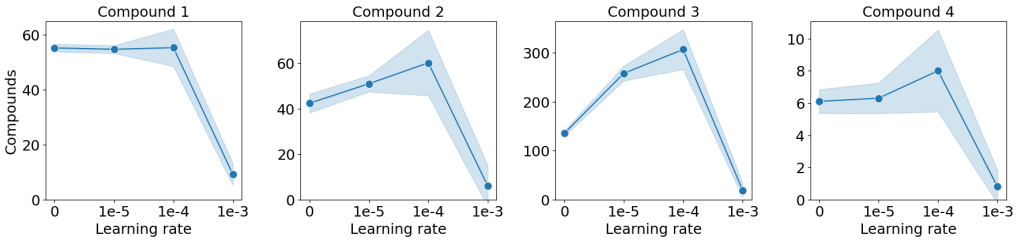
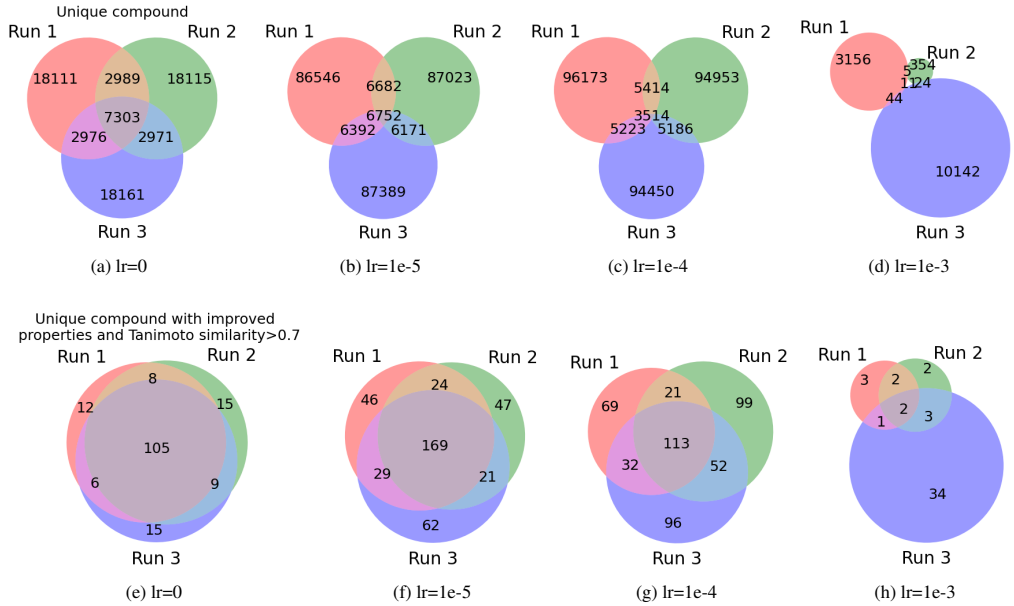


Figure S1: Molecular optimization task: effect of learning rates on the number of unique compounds with improved P(active) and QED, and Tanimoto similarity >0.7 when using RL_DF(cmp)_Sim. Results are mean values and ±1 standard deviation over 10 runs.



(a) lr=0 (b) lr=1e-5 (c) lr=1e-4 (d) lr=1e-3

(e) lr=0 (f) lr=1e-5 (g) lr=1e-4 (h) lr=1e-3

Figure S2: Molecular optimization task: overlap of three runs with varying learning rates (lr) on the unique compounds (a-d) and the unique compounds with improved P(active) and QED, and Tanimoto similarity above 0.7 compared to corresponding input molecule (e-h) produced by RL_DF(cmp)_Sim for compound 3.