GraphChem: A Soft Graph Mixture of Experts Based Adversarial Variational Spectral Graph Autoencoder for De Novo Drug Design & Other Pharmacological Objectives

Moitra, Agnij (School: Amity International School - Saket)

Traditional drug discovery is costly, slow, and has low FDA approval rates. Deep learning models such as GAN, VAE, & reinforcement learning (RL) have emerged as a solution, using SMILES or graphs to create drugs tailored to target molecules/ligands, or employing heuristics like QED & Lipinski's rule. Additionally, molecule vector representations enable various downstreamed transfer learning (TL). However, existing works face the following limitations: generate small molecules, no implicit way of using information from scientific literature, use on-policy RL. Thereby, GraphChem introduces a novel DL method—soft graph mixture of experts (SGMoE)—using multi-headed spectral graph transformers, which is stacked in an adversarial VAE for generating latent vectors. Finally the vector representation of the molecule combined with embeddings from Galactica AI & ChemLLM are passed to a Deep Q Learning model for generating molecules. This had a 149% improvement on 39 curated drug design tasks, and 41% improvement on 37 downstreamed TL tasks for virtual screening, molecular property prediction (MPP using MoleculeNet & DOCKSTRING) & retrosynthesis (USPTO chemical reactions). The drug discovery dataset includes targets like JAK2, MET, ESR1, HIV protease etc which are relevant for bone marrow blood, breast, lung cancer, & various other diseases. Further, ablation studies demonstrated the significance & rationale of using SGMoE, scientific literature from Galactica, and other contributions of this work. In conclusion, GraphChem can drastically cut costs & time in medication development, enhancing accessibility to life-saving treatments worldwide & advancing global health equity and outcomes. Furthermore, computational MPP is more humane & non-invasive when compared to animal testing.

Awards Won: