Interpretable Molecular Foundation Modeling via Functional Group Tokenization

Recent advances in deep learning have enabled the development of foundation models trained on large-scale molecular datasets, opening new avenues for mechanism-based compound modeling. However, most existing molecular foundation models are designed for tasks such as property prediction or molecular generation (e.g., ADME) and often lack interpretability, limiting their utility for downstream biomedical applications.

In this study, we present a novel foundation model designed to capture mechanistically relevant features of chemical compounds and enable the prediction of drug-induced biological responses. In this framework, compounds are decomposed into functional groups, and 3D atom pair maps are constructed. Each atom pair is encoded as a token by concatenating its value and contextual knowledge, and these tokens are jointly trained with a physicochemical property module. This architecture allows for modeling interactions among tokens and their associations with chemical properties, reflecting the systemic relationships between compound characteristics and their underlying functional groups.

The learned embeddings from model were fine-tuned for a variety of downstream tasks, including MoleculeNet benchmarks and DILI classification. Across tasks, model demonstrated performance comparable to or exceeding that of existing models, while providing interpretable embeddings grounded in functional group-level chemistry.

We aim to further extend model for inferring target modules and predicting compound-target interaction. As a foundation model, model holds promise for broad applications in mechanism-driven drug discovery, offering both generalizability and interpretability across molecular tasks.