[BIOINFO 2025]

MMDIMoE: A Mixture-of-Experts Framework for Multimodal Drug–Drug Interaction Prediction

Yelin Park[1], and Sangseon Lee[2,*]

[1] *Department of Computer Engineering, Inha University*
[2] *Department of Artificial Intelligence, Inha University*
* *Corresponding author: ss.lee@inha.ac.kr*

Drug–drug interactions (DDIs) pose critical risks in clinical treatment, as unexpected interactions may cause adverse effects or reduce therapeutic efficacy. Predicting DDIs is challenging because drug mechanisms involve heterogeneous and multimodal information, including chemical structures, pharmacological properties, targets, and enzymes. Conventional approaches relying on a single modality or simple fusion often fail to capture such complexity. We propose MMDIMoE, a novel DDI prediction framework that integrates multimodal relationships among drugs, chemical entities, and molecular substructures through a Mixture of Experts (MoE) architecture. First, modality subnetworks are constructed to model drug–modality relationships, and graph neural networks derive modality-specific embeddings. The four modality embeddings for each drug are concatenated into a unified representation, and embeddings of two drugs in a candidate pair are further concatenated to form the input to the prediction module. On top of this, the MoE assigns each expert to a distinct feature subspace, while a gating mechanism dynamically selects experts most relevant for each pair. The gating network makes instance-wise decisions, adapting to the characteristics of each input. To ensure efficiency and diversity, a top-k routing strategy is employed, where only the k most informative experts are activated for each prediction. This selective routing reduces redundancy, improves scalability, and enables the model to capture complementary information across modalities rather than uniform integration. Experiments on benchmark datasets demonstrate that MMDIMoE achieves consistent improvements across accuracy and AUPR. For unseen drug pairs, our model improves accuracy by 4.8% and AUPR by 4.9% compared to the SOTA baseline, highlighting its strong generalization ability. Analysis of expert utilization shows that drug pairs with similar mechanisms of action activate overlapping expert subsets, suggesting that the model learns mechanism-specific patterns useful for DDI prediction. In conclusion, MMDIMoE provides an effective and interpretable approach to multimodal data integration for DDI prediction.