DrugResponse-GNN

DrugResponse-GNN: Cross-Panel Pathway-Bottleneck Graph Neural Networks for Drug Sensitivity Prediction
Sally Yepes (2025)

---

1. Problem

Pharmacogenomic resources such as CCLE, GDSC, NCI-60, and CellMinerCDB offer unprecedented depth in linking molecular features to drug response. Yet current models often fail to generalize across panels: a predictor trained on CCLE collapses on GDSC. This lack of robustness undermines biomarker discovery and clinical translation. Existing GNN models (drGAT, DRPreter) rely on post-hoc attention, which is unstable and insufficient to yield reproducible mechanisms.

---

2. Innovation

DrugResponse-GNN introduces **pathway-bottleneck graph neural networks** that enforce interpretability by design. Multi-omic features are encoded as nodes in a prior-knowledge graph (protein–protein interactions, pathways, regulatory links). Information must flow through curated pathway/TF bottlenecks before reaching the drug-response head. This ensures that predictions reflect pathway-level reasoning and can be attributed to stable biological modules. The framework directly integrates **cross-panel generalization** as a first-class evaluation, not an afterthought.

---

3. Validation

The model will be trained and tested across multiple panels: CCLE, GDSC, NCI-60, CTRP, and CellMinerCDB. Predictive accuracy will be compared to elastic net, random forests, and GNN baselines. Novel metrics include **cross-panel validation** (e.g., train on CCLE, test on GDSC) and **attribution stability** under seed variation and bootstrapping. Biological case studies will focus on PI3K, EGFR, and PARP inhibitors, testing whether the model recovers known resistance subnetworks and proposes novel candidates.

---

4. Contribution

DrugResponse-GNN redefines pharmacogenomic modeling by elevating **cross-panel validation** to a methodological standard. It provides not just predictions but reproducible explanations of drug sensitivity, grounded in pathways and subnetworks. Scientifically, it will reveal conserved mechanisms of response across datasets, informing biomarker

development and therapeutic stratification. Methodologically, it positions pathway-bottleneck GNNs as a **generalizable, interpretable framework** for pharmacogenomics, bridging the gap between machine learning pipelines and translational oncology.