GraphPINE: Graph importance propagation for interpretable drug response prediction

Yoshitaka Inoue[1,2,3], Tianfan Fu[4], Augustin Luna[2,3]

1 University of Minnesota, US, 2 National Library of Medicine, US,
3 National Cancer Institute, US, 4 Nanjing University, China

Introduction & Conclusion (TLDR)

Drug Response Prediction (DRP) is crucial in personalized medicine but remains challenging due to complex biological interactions. Existing GNN-based methods often lack interpretability and ignore known Drug-Target Interactions (DTIs). We propose **GraphPINE**, a GNN model incorporating prior biological knowledge through an **Importance Propagation Layer (IP Layer)**. This enables the model to learn from and explain predictions based on biologically meaningful gene networks.

GraphPINE achieves strong predictive performance (PR-AUC: **0.894**, ROC-AUC: **0.796**) and delivers interpretable insights.

- Integrates DTI-based initial importance with graph learning
- Highlights both known targets (e.g., **TOP1**) and novel candidates (e.g., **TOP1MT**, **TP53**)
- Expands relevant gene associations from **0.77% to 39.8%**, while preserving prior knowledge (**Cosine Sim. > 0.8**)

👉 A powerful and interpretable GNN for drug response prediction.

Method

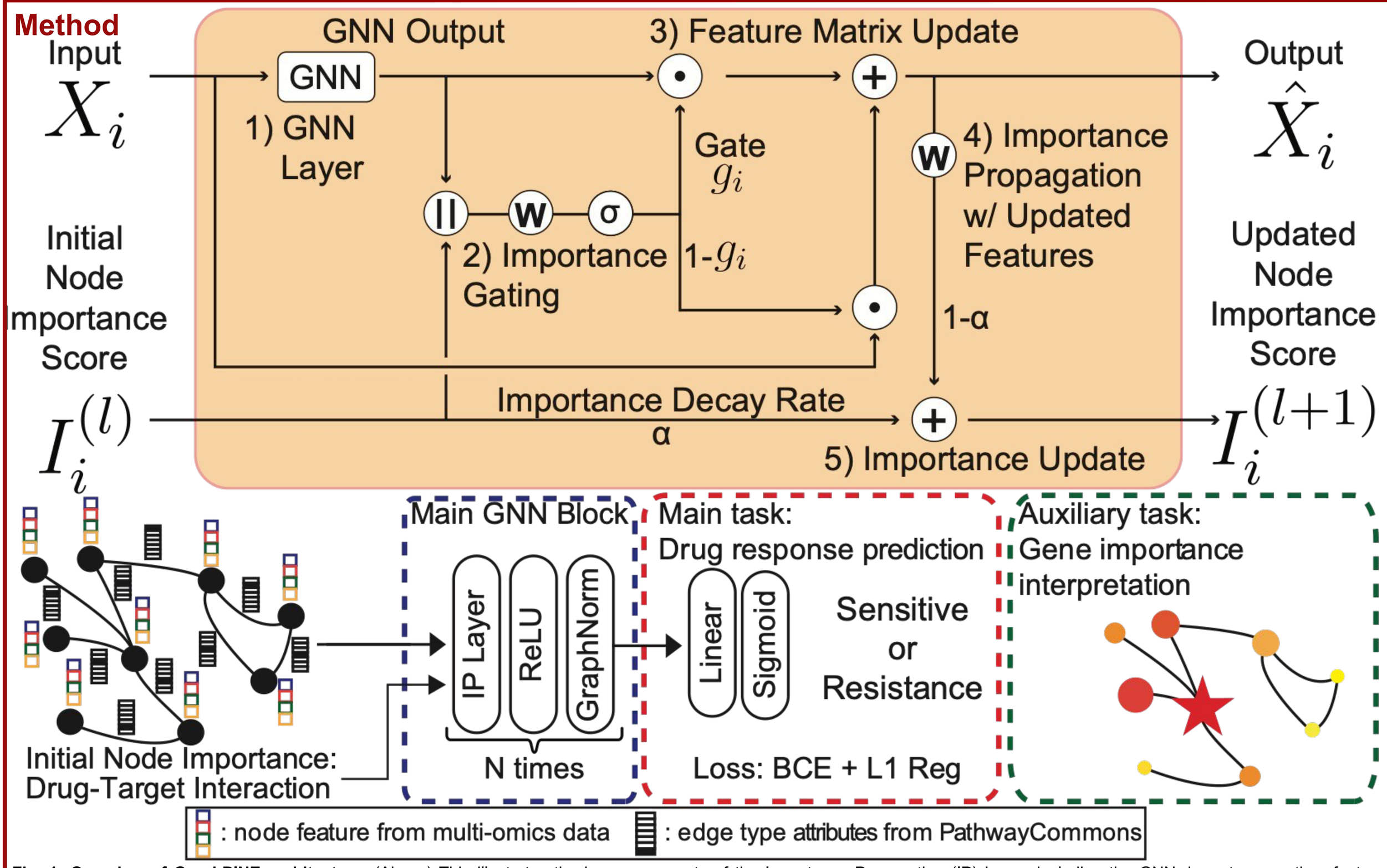

Fig. 1: Overview of GraphPINE architecture. (Above) This illustrates the key components of the Importance Propagation (IP) Layer, including the GNN, importance gating, feature updates with residual connections, importance propagation, and updates. The symbols represent the following operations: σ is the activation function, ⊙ is element-wise multiplication, × is multiplication, + is addition, W denotes weighted calculation with bias, || represents concatenation, and α is a hyperparameter for controlling importance. (Below) GraphPINE architecture. The model integrates multi-omics data (gene expression, copy number, methylation, mutation) from NCI60 with gene-gene interaction networks from PathwayCommons.

Result

Methods	ROC-AUC	PR-AUC	Precision	Specificity
Random Forest	0.79	**0.89**	0.73	0.63
LightGBM	0.79	0.87	0.77	0.46
MLP	0.75	0.84	0.72	0.27
DeepDSC	0.71	0.78	**0.81**	0.60
MOFGCN	0.49	0.67	0.65	**0.90**
GraphPINE	**0.80**	**0.89**	0.72	0.55

Table. 1 Predictive performance comparison. Results show averages of 5 independent runs. ROC-AUC: Receiver Operating Characteristic Area Under the Curve, PR-AUC: Precision-Recall Area Under the Curve, MLP: Multi Layer Perceptron.

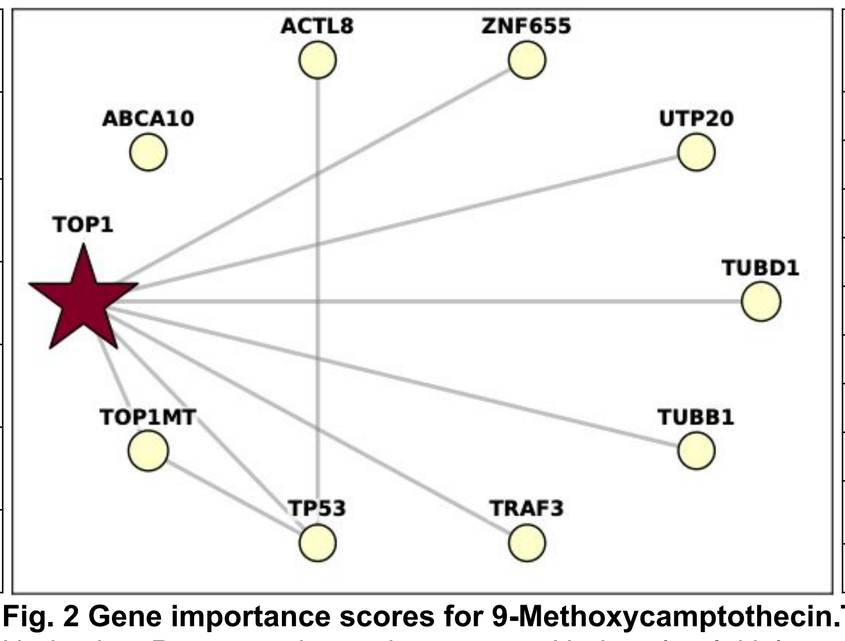

Fig. 2 Gene importance scores for 9-Methoxycamptothecin. Node size: Propagated gene importance. Node color: Initial DTI score. Lines: Protein-Protein interactions. ☆: Known DTI target. ○: Unknown DTI protein.

Rank	Initial Importance	Gene	PMID	Relationship
1	1	TOP1	29312794...	Target
2	-	TOP1MT	24890608...	Indirect
3	-	TUBD1	-	-
4	-	ZNF655	-	-
5	-	UTP20	-	-
6	-	TUBB1	-	-
7	-	ACTL8	-	-
8	-	ABCA10	10606239	Indirect
9	-	TRAF3	-	-
10	-	TP53	12082016...	Indirect

Table. 2 Top 10 predicted important genes for 9-Methoxycamptothecin and supporting literature. (-) indicates no initial DTI; (...) denotes multiple sources. Target: Direct drug-binding proteins. Indirect: Genes involved in drug response pathways without direct binding.

ArXiv

GitHub

This research was supported in part by the Division of Intramural Research (DIR) of the National Library of Medicine (NLM), National Institutes of Health (NIH) (ZIALM240126).

Metric	Value
Cosine sim.	0.87
Spearman corr.	0.82
Rank changes	90.42%
Avg. shift	±67.02
Max up	946
Max down	-932

Table 3: Differences in Node (Gene) Ranks Before and After Propagation. Cosine sim.: Cosine similarity between initial/propagated importance rank. Spearman corr.: Spearman Rank correlation between the initial/propagated importance rank. Rank changes: The percentage of genes whose ranks changed after propagation. Avg. shift: The average rank shift. Max up/down: Maximum upward/downward rank mobility