Few-shot Learning from Biomedical Network

- with a focus on predicting emerging drugs interactions

Dr. Quanming Yao

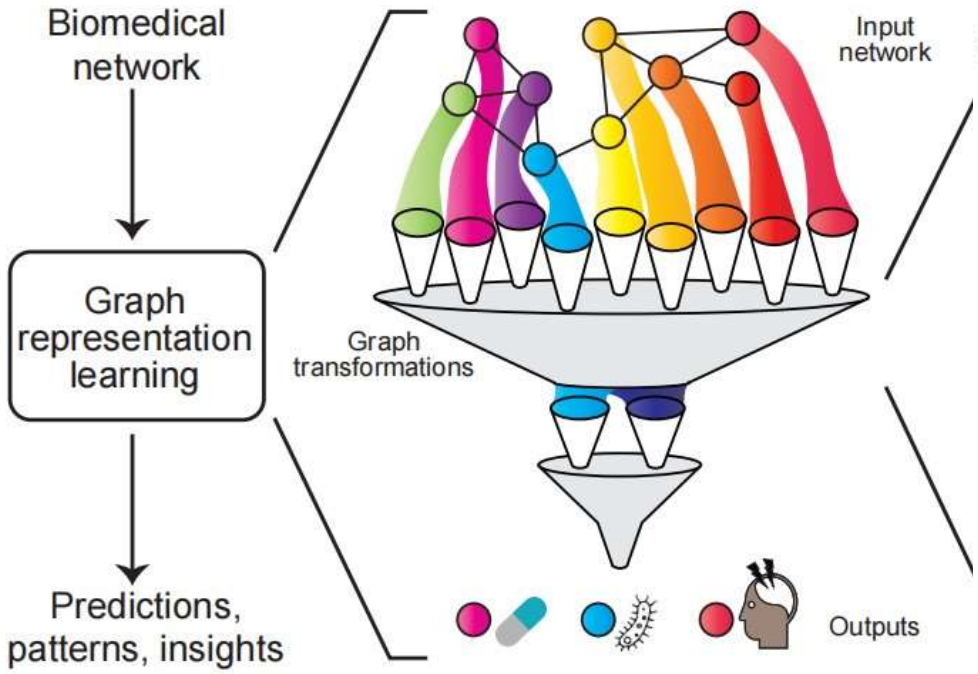
Assistant professor, EE Tsinghua

qyaoaa@tsinghua.edu.cn

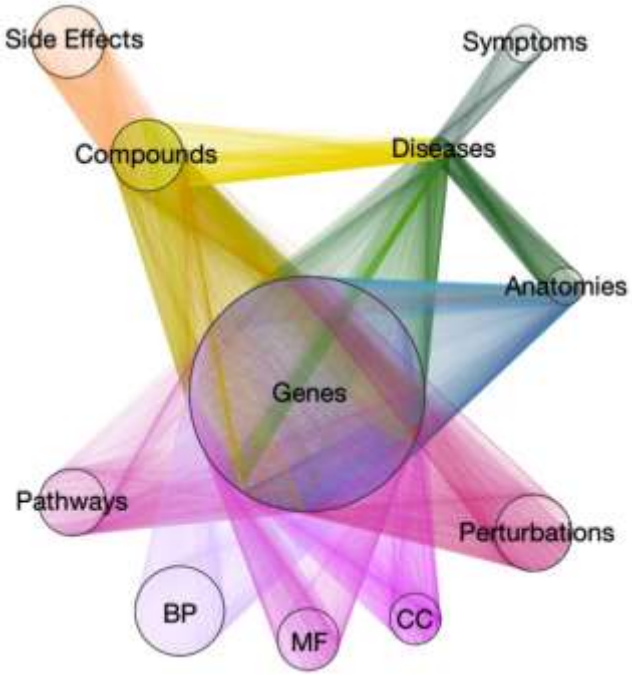
https://lars-group.github.io/index.html

2024/01/10

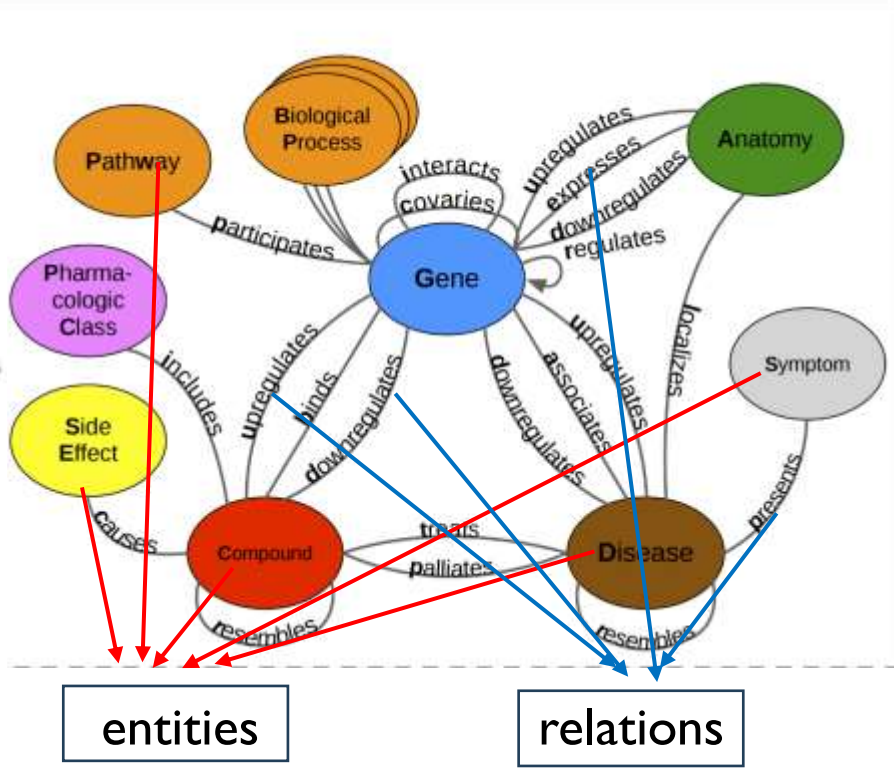
Biomedical Network (BN)



Representation learning for networks in biology and medicine.

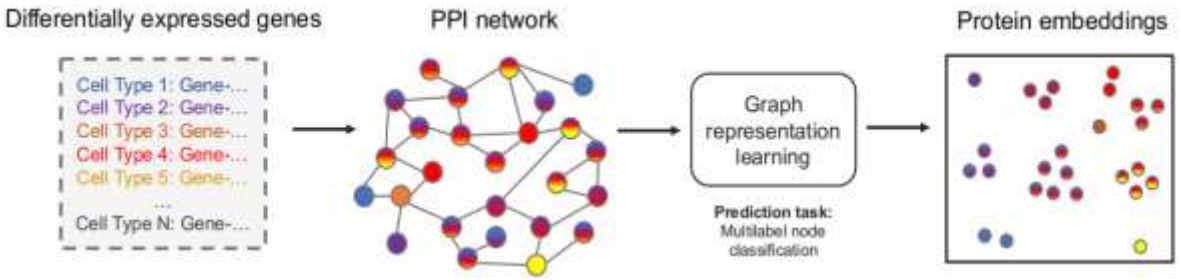


The network visualized

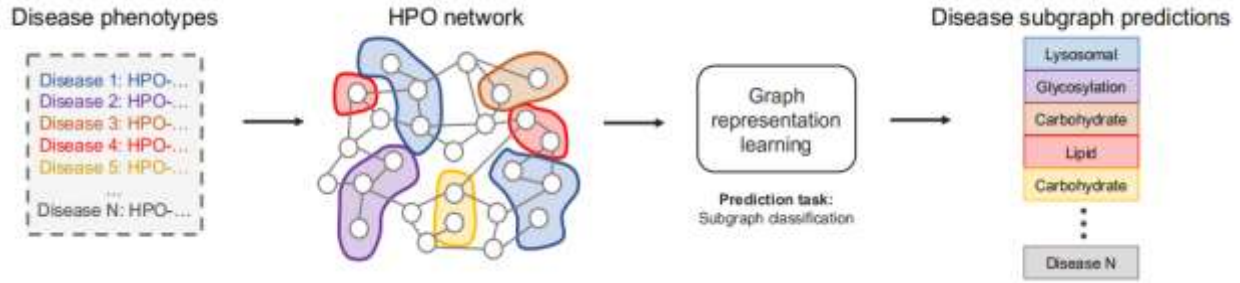


The metagraph, a schema of the network types.

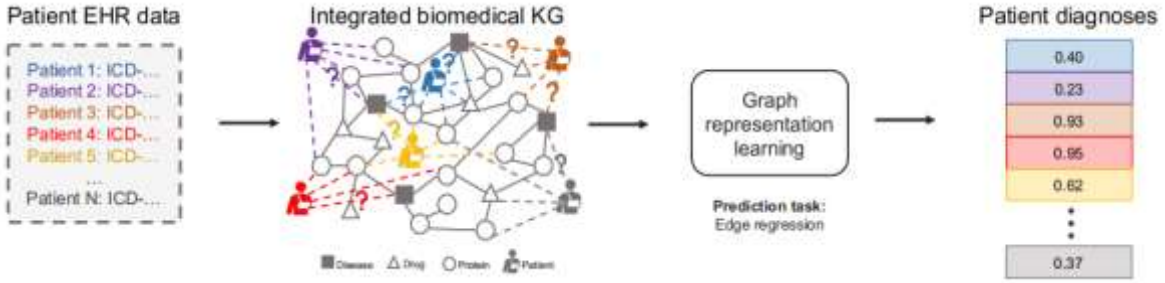
Important Applications



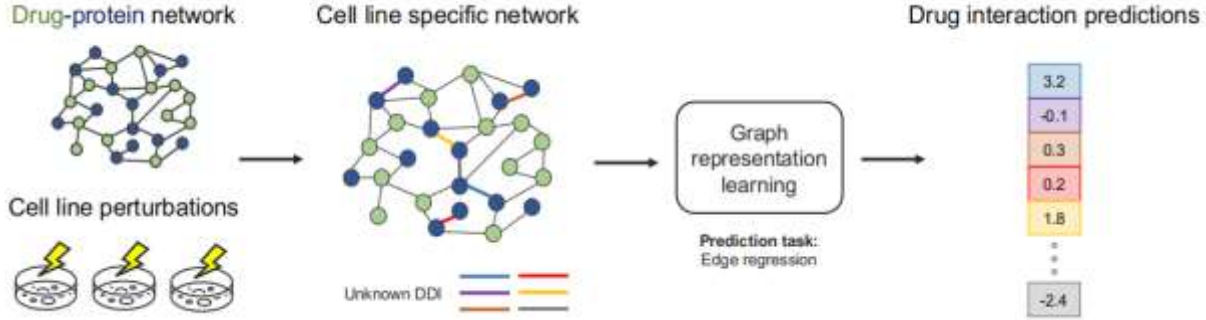
Cell-type aware protein representation learning



Disease classification using subgraphs



Integration of health data into knowledge graphs to predict patient diagnoses



Cell-line specific prediction of interacting drug pairs

Needs Few-shot Learning

Drug discovery process

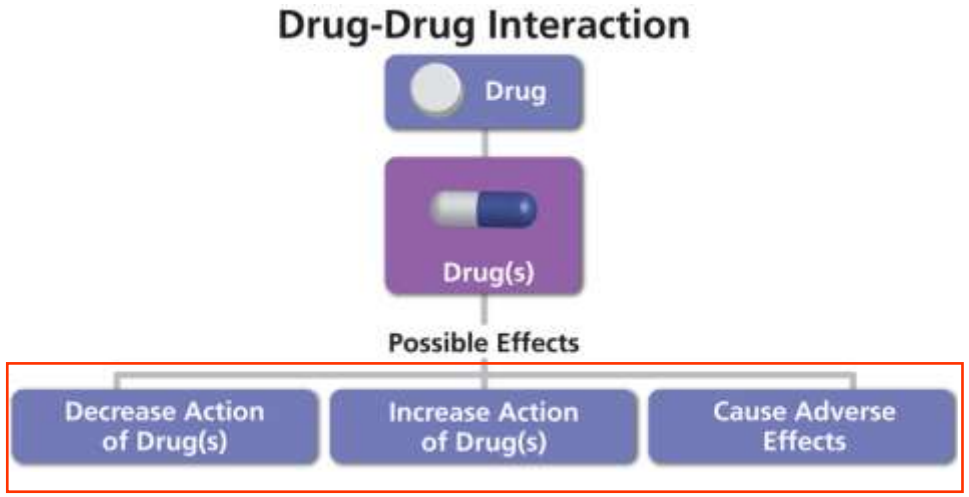
- Time-consuming but necessary.
- Hard to obtain clinical data.
- Data sparsity.

- Limited experimental data
- Heterogeneous data
- Lack of supervision data

→ Weakly supervised learning from biomedical network.

- Here we focus on drug-drug interaction prediction.

Emerging Drug-Drug Interaction (DDI) Prediction



Adverse effects!

S. No.	Drugs Interaction Combination	Frequency	Outcome
1	Ceftriaxone + Calcium Gluconate	6	Precipitation of ceftriaxone-calcium salt
2	Furosemide + Amikacin	5	Potentiate the risk of oto- and nephrotoxicity
3	Atracurium + Amikacin	3	Severe and/or prolonged respiratory depression
4	Omeprazole + Clopidogril	2	Decreased effectiveness of Clopidogril
5	Aspirin + Clopidogril	2	Increased platelet inhibition effect

Clinical combination use is common.

Top 5 major drug interaction combinations and their outcomes

6.7% of hospitalized patients have a serious adverse drug reaction with a fatality rate of 0.32%.

--U.S. Food and Drug Administration

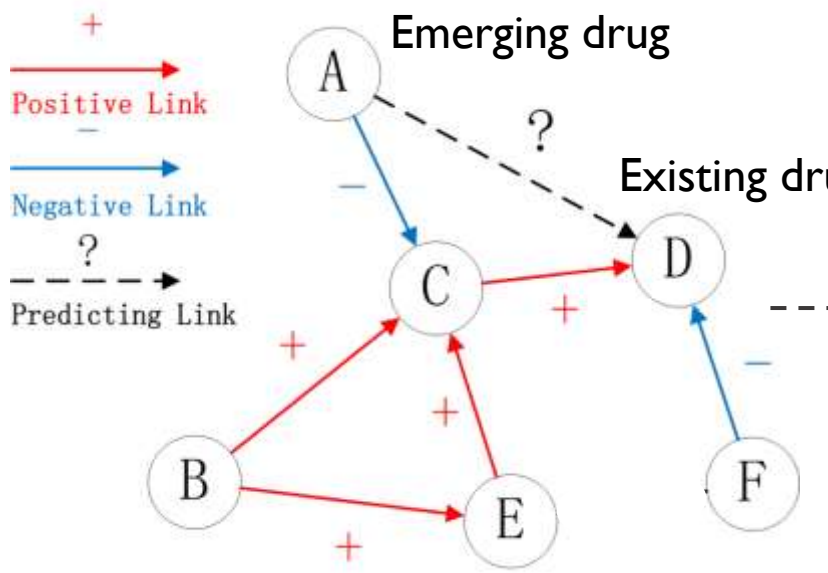
data sparsity v.s. data hungry

DDI prediction

deep learning

These problems are even more serious for emerging drugs.

DDI Task



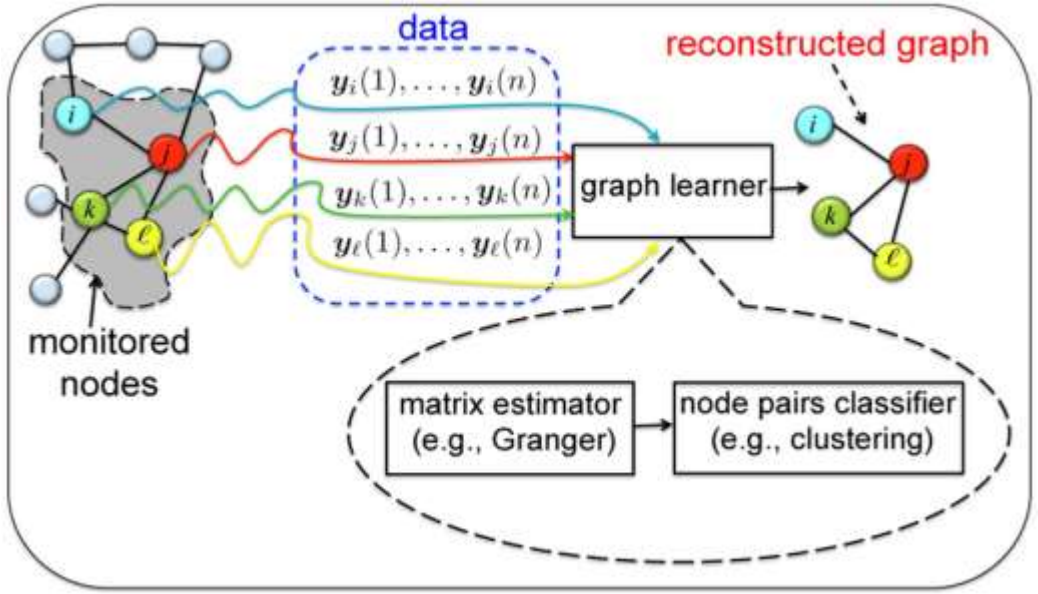
- Input:
  - Fingerprint features (f_u) of a given emerging drug u.
  - Drug drug interaction (DDI) network.
  - Biomedical network.
- Output:
  - Interaction type between a given emerging drug and a given existing drug.

(Emerging drug, ? , Drug)

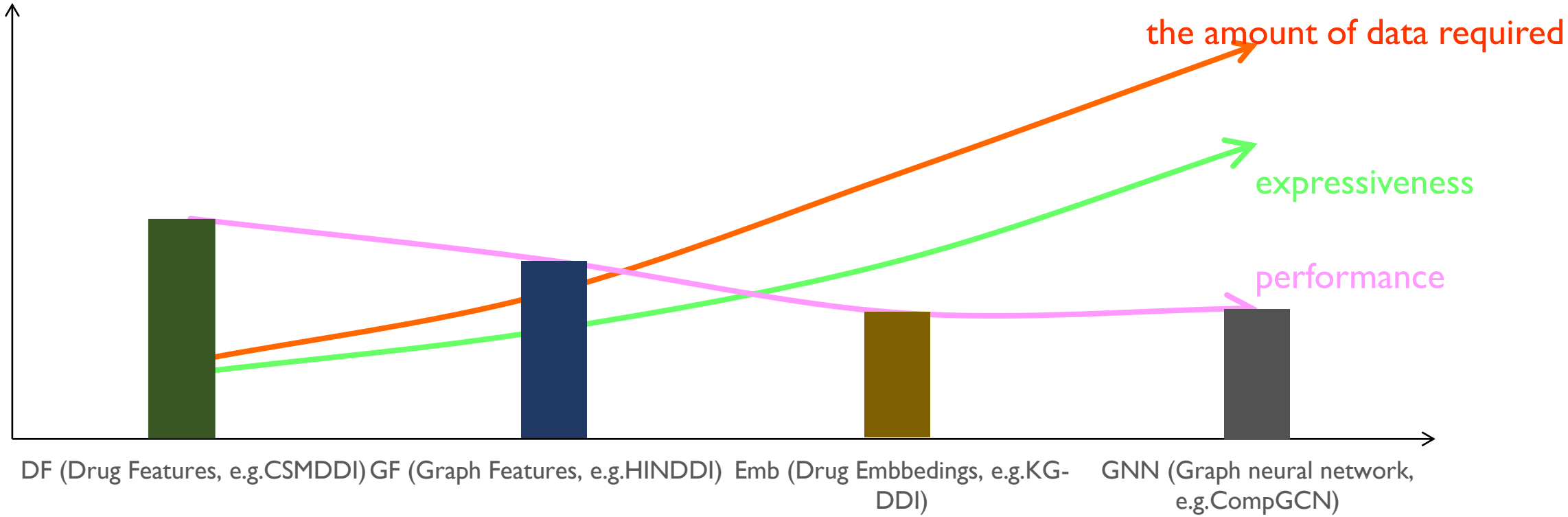
Link prediction on biomedical network (BN)

transform

A graph learning problem



Related Works



learn a function to map drugs' attributes to representations

extract meta-paths on the BN and predict interactions

graph embbeding

fuse multiple types of information

Existing deep learning methods do not perform well, as they require large amounts of data to train their over-parameterized models.

Our contributions

Emerging Drug Interaction Prediction Enabled by Flow-based Graph Neural Network with Biomedical Network.
Y. Zhang, Q. Yao, L. Yue, X. Wu, Z. Zhang, Z. Lin, Y. Zheng
Nature Computational Science. 2023

- Develop the first effective deep learning method for emerging drug-drug interaction (DDI) prediction.
- Propose a variant of Graph Neural Network (EmerGNN) by integrating the relevant biomedical concepts.
- Extensive experiments with interpretable learned concepts on biomedical network (BN).
- Possible applications for patient care and drug development processes.

Intuition

Drug name

【药品名称】
通用名称：布洛芬颗粒
英文名称：Ibuprofen Granules
汉语拼音：Buluofen Keli

Ibuprofen Granules

Element

【成份】
本品每包含布洛芬0.2克。辅料为：糊精、蔗糖、甘露醇。

Traits

【性状】
本品为白色或黄色颗粒，气芳香；味甜。

【作用类别】
本品为解热镇痛类非处方药药品。

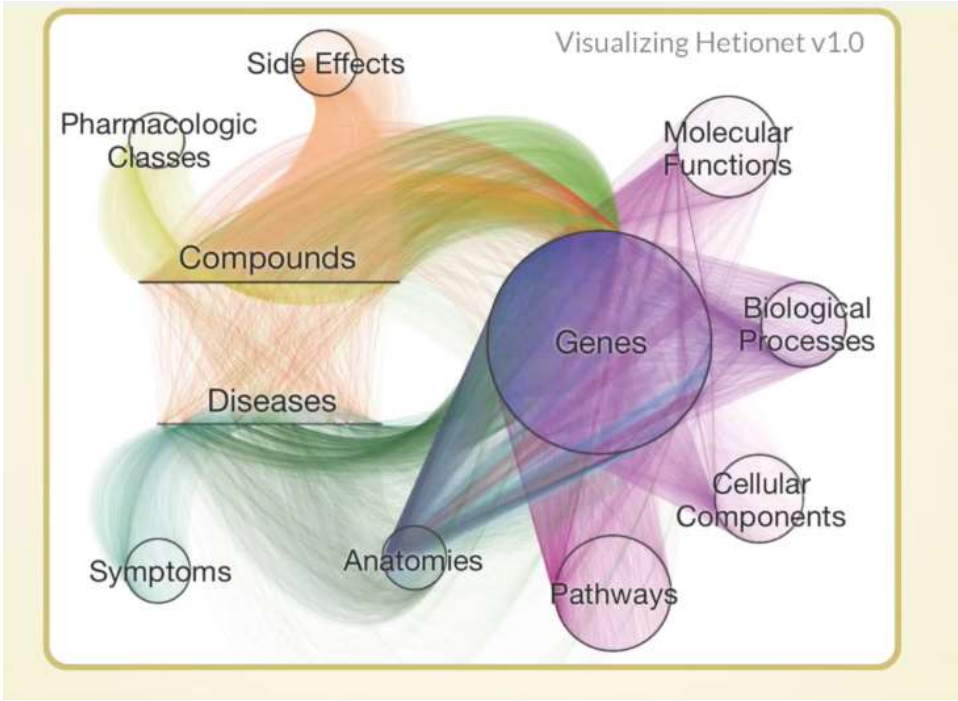
Indications

【适应症】
用于缓解轻至中度疼痛如头痛、关节痛、偏头痛、牙痛、肌肉痛、感冒或流行性感冒引起的发热。

Relieve fever and headache.

DDI

【药物相互作用】
1. 本品与其他解热、镇痛、抗炎药物同用时可增加胃肠道的副作用，并可能导致溃疡。
2. 本品与肝素、双香豆素等抗凝药同用时，可导致凝血酶原时间延长，增加出血倾向。
3. 本品与地高辛、甲氨蝶呤、口服降血糖药物同用时，能使这些药物的血药浓度增高，不宜同用。



- Medication instructions

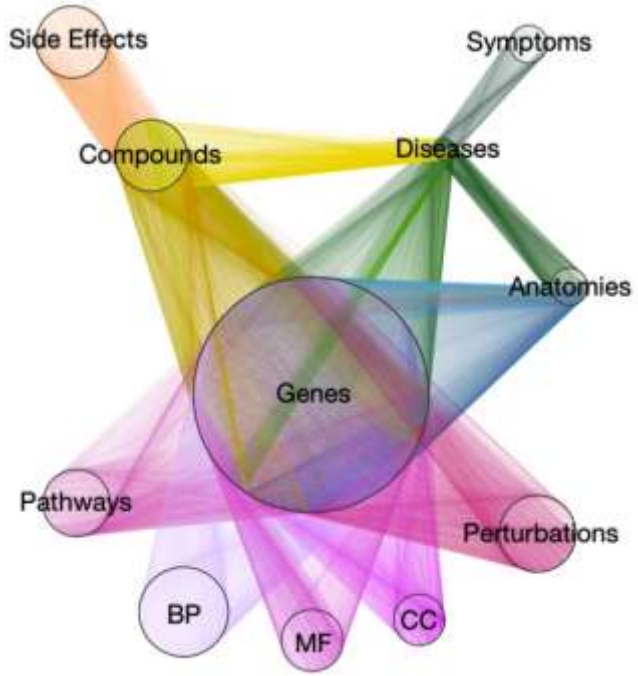
Emerging drugs often share the same entities with existing drugs, like the same targeted genes or diseases.

- How to extract relevant information?

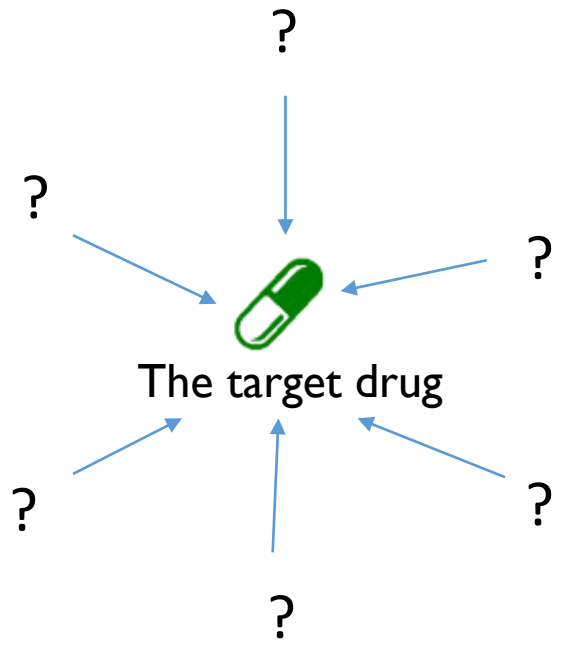
Need an effective and efficient method.

Challenges

- However, properly utilizing biomedical networks can be challenging as these networks are not specially developed for emerging drugs.
- The mismatch of objectives can lead the machine learning models to learn distractive knowledge.



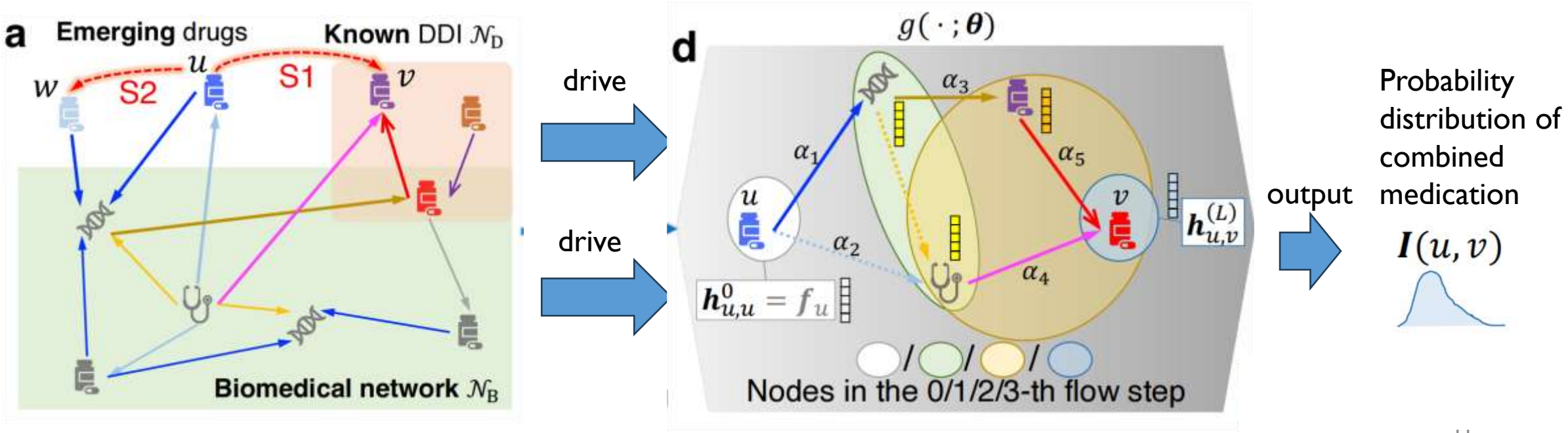
How to utilize biomedical network



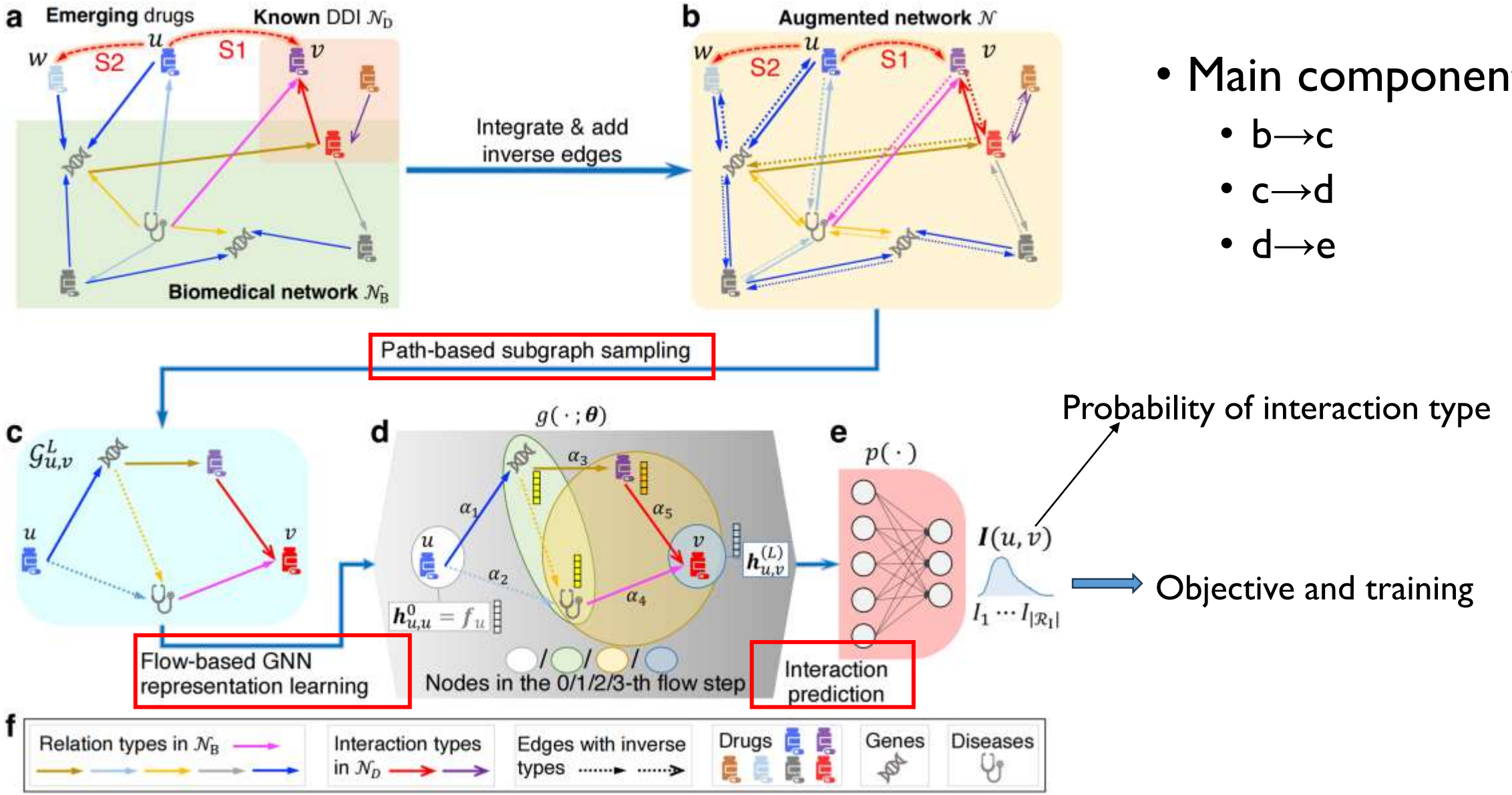
Architecture of EmerGNN

Main idea:

- Construct a subgraph to extract knowledge related to emerging drugs.
- Set attention weight for edges to highlight the important paths and design GNN.

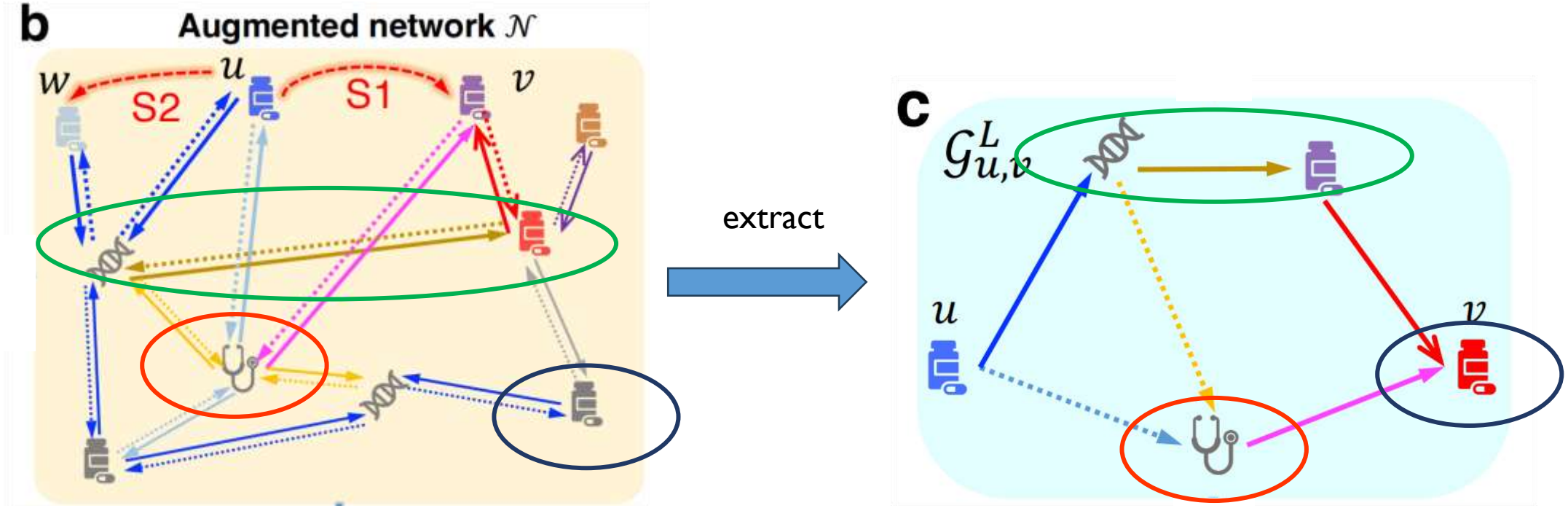


Framework



- Main components
  - b→c
  - c→d
  - d→e

Subgraph extraction



u (emerging drug) and v (existing drug) share some same entities, like genes, side effects and compounds.

Attention weight design

- Attention weight

Sigmoid function

$$\alpha_r^{(l)} = \sigma\left(\left(w_r^{(l)}\right)^{\mathrm{T}} [f_u; f_v]\right)$$

Relation weight The fingerprints of drugs to be predicted

Message function:

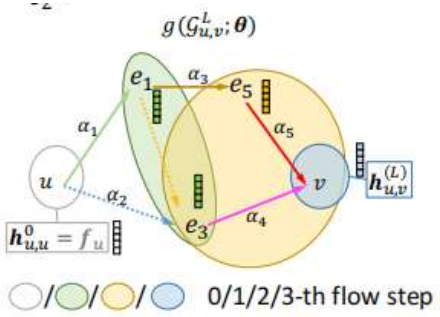
$$\phi\left(h_{u,e'}^{(l-1)}, h_r^{(l)}\right) = \alpha_r^{(l)} \cdot \left(h_{u,e'}^{(l-1)} \odot h_r^{(l)}\right)$$

Message passing:

$$h_{u,e}^{(l)} = \delta\left(W^{(l)} \Sigma_{e' \in \mathcal{V}_{u,v}^{l-1}} \left(h_{u,e}^{(l-1)} + \phi\left(h_{u,e'}^{(l-1)}, h_r^{(l)}\right)\right)\right)$$

- Weight the different types of relations on the biomedical network.
- The edges with larger weights on the paths are helpful for interpretation.

Applied to Construct GNN.



- After iterating for L steps, we can obtain the representation $h_{u,v}^{(L)}$.

Objective and training

- Bi-directional representations:

$$l(u,v) = W_{rel}[h_{u,v}^{(L)}; h_{v,u}^{(L)}]$$

- Two datasets: DrugBank (multi-class) and TWOSIDES (multi-label).

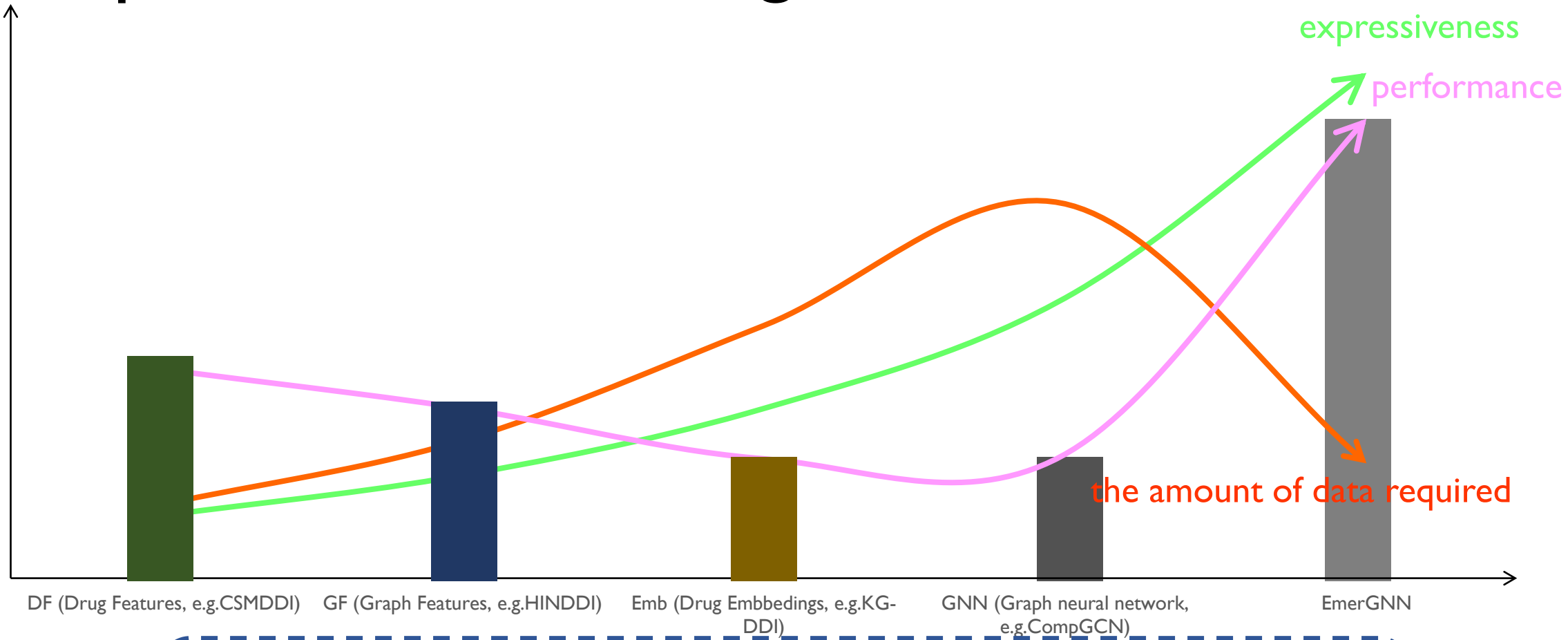
Dataset	Probability of interaction type i	Loss function
DrugBank	$I_i(u,v) = \frac{\exp(l_i(u,v))}{\Sigma_{j \in \mathcal{R}_1} \exp(l_j(u,v))}$	$\mathcal{L}_{DB} = -\Sigma_{(u,i,v) \in \mathcal{N}_{D-train}} y_i(u,v) log I_i(u,v)$
TWOSIDES	$I_i(u,v) = \frac{1}{1+\exp(-l_i(u,v))}$	$\mathcal{L}_{TS} = -\Sigma_{(u,i,v) \in \mathcal{N}_{D-train}} log I_i(u,v) + \Sigma_{(u',v') \in \mathcal{N}_i} \log(1 - I_i(u',v'))$

- Parameters: $\theta = \{W_{rel}, \{W^{(l)}, h_r^{(l)}, w_r^{(l)}\}_{l=1,\ldots,L, r \in \mathcal{R}}\}$

Technical differences

GNN method	Subgraph	Message passing
GCN	None	$h_i^{(l+1)} = \sigma(b^{(l)} + \Sigma_{j \in \mathcal{N}(i)} \frac{1}{c_{ji}} h_j^{(l)} W^{(l)})$
GAT	Can only aggregate 1-hop neighbors	$h_i^{(l+1)} = \Sigma_{j \in \mathcal{N}(i)} \alpha_{i,j} W^{(l)} h_j^{(l)}$
GraphSage	None	$h_i^{(l+1)} = ReLU(U^l Concat\left(h_i^l, Mean_{j \in \mathcal{N}(i)} h_j^{(l)}\right))$
CompGCN	None	$h_i^{(l+1)} = f(\Sigma_{(u,r) \in \mathcal{N}(i)} W_{\lambda(r)}^l \phi(h_u^l, h_r^l))$
Decagon	None	$h_i^{(l+1)} = \phi(\Sigma_r \Sigma_{j \in \mathcal{N}_r^i} c_r^{ij} W_r^{(l)} h_j^{(l)} + c_r^i h_i^{(l)})$
KGNN	Multiple hops neighbors	
SumGNN	Enclosing subgraph	$h_i^{(l)} = ReLU(W_{self}^l h_i^{(l-1)} + b_v^{(l)})$
EmerGNN	Path-based subgraph	$h_{u,e}^{(l)} = \delta\left(W^{(l)} \Sigma_{e' \in \mathcal{V}_{u,v}^{l-1}} \left(h_{u,e}^{(l-1)} + \phi\left(h_{u,e'}^{(l-1)}, h_r^{(l)}\right)\right)\right)$

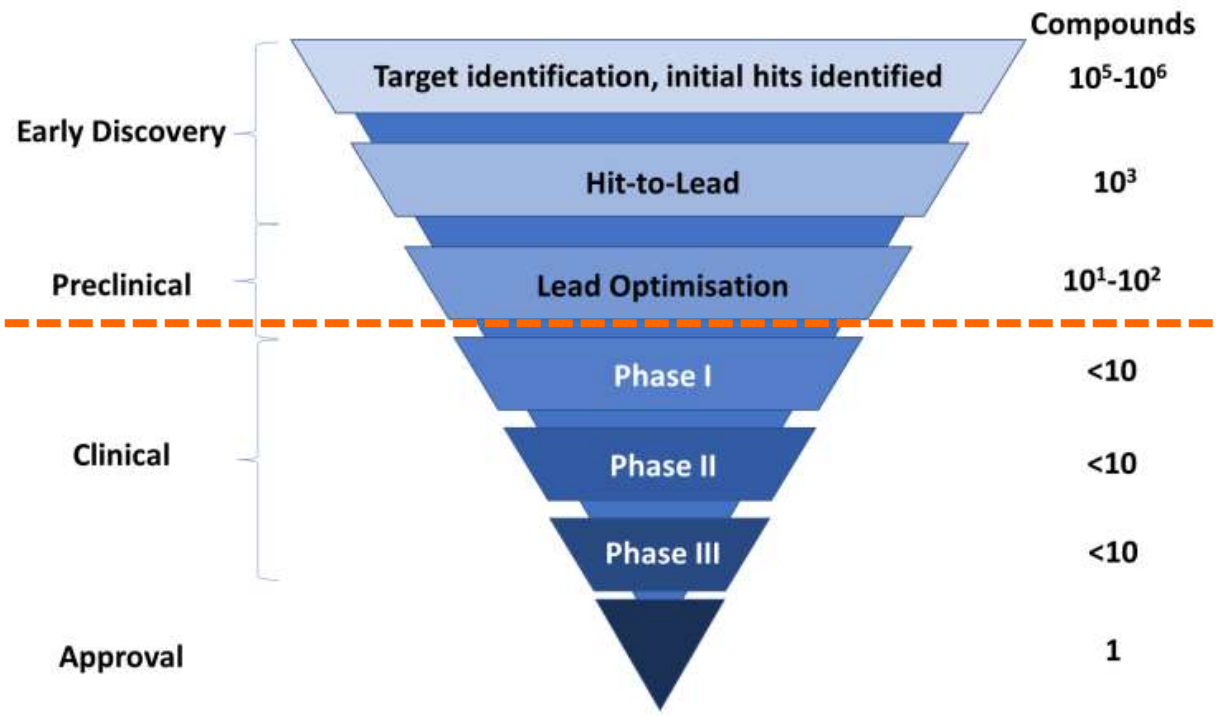
Comparison with Existing Works



Why can we solve data sparsity v.s. data hungry?

- Introduce biomedical network and extract essential information due to attention mechanism based on specially designed GNN.

Dataset



DrugBank

- Multi-class
- Interactions
- Pre-IPO
- Relation type: increase or decrease the expression of the metabolite, protein or gene etc..

TWOSIDES

- Multi-label
- Side effects
- Post-market
- Relation type: anaemia, nausea, pain, etc..

Statistics	$\lvert\mathcal{V}_{\mathrm{D}}\rvert$	$\lvert\mathcal{R}_{\mathrm{I}}\rvert$	$\lvert\mathcal{N}_{\text{D-train}}\rvert$	$\lvert\mathcal{N}_{\text{D-valid}}\rvert$	$\lvert\mathcal{N}_{\text{D-test}}\rvert$
Drugbank	1,710	86	134,641	19,224	38,419
TWOSIDES	604	200	177,568	24,887	49,656

Statistics of common DDI datasets used

Baseline

Method	Baseline
MLP	Learn a multi-layer perception (denoted as MLP) that mapped the fingerprints of drugs to their interaction type.
GAT	Use attention networks to aggregate neighborhood information in DDI network.
CSMDDI	Learn a function to map drugs' attributes to representations for DDI prediction in a cold-start setting.
HIN-DDI	Extract meta-paths on the biomedical network and predicts the interaction type based on the meta-paths.
MSTE	Predict interactions by learning drug embeddings.
CompGCN	Use GNN to learn high-order embeddings of entities from their neighbors in a knowledge Graph.
Decagon	Have similar model structure as CompGCN, but only use three types of entities, i.e., drug, protein and disease.
KGNN	Sample and aggregate neighborhoods for each node from their local receptive via GNN and with external KG, which achieves the state-of-the-art result on binary DDI prediction problem.
SumGNN	Use GNN to summarize knowledge in the subgraphs covering the drug pairs.

Experiments

(S1): DDI prediction between emerging drug and existing drug.

Datasets		DrugBank			TWOSIDES		
Type	Methods	F1-Score	Accuracy	Kappa	PR-AUC	ROC-AUC	Accuracy
DF	MLP (Rogers and Hahn, 2010)	21.1±0.8	46.6±2.1	33.4±2.5	81.5±1.5	81.2±1.9	76.0±2.1
	Similarity (Vilar et al, 2014)	43.0±5.0	51.3±3.5	44.8±3.8	56.2±0.5	55.7±0.6	53.9±0.4
	CSMDDI (Liu et al, 2022)	45.5±1.8	62.6±2.8	55.0±3.2	73.2±2.6	74.2±2.9	69.9±2.2
	STNN-DDI (Yu et al, 2022)	39.7±1.8	56.7±2.6	46.5±3.4	68.9±2.0	68.3±2.6	65.3±1.8
GF	HIN-DDI* (Tanvir et al, 2021)	37.3±2.9	58.9±1.4	47.6±1.8	81.9±0.6	83.8±0.9	79.3±1.1
Emb	MSTE (Yao et al, 2022)	7.0±0.7	51.4±1.8	37.4±2.2	64.1±1.1	62.3±1.1	58.7±0.7
	KG-DDI* (Karim et al, 2019)	26.1±0.9	46.7±1.9	35.2±2.5	79.1±0.9	77.7±1.0	60.2±2.2
GNN	CompGCN* (Vashishth et al, 2019)	26.8±2.2	48.7±3.0	37.6±2.8	80.3±3.2	79.4±4.0	71.4±3.1
	Decagon* (Zitnik et al, 2018)	24.3±4.5	47.4±4.9	35.8±5.9	79.0±2.0	78.5±2.3	69.7±2.4
	KGNN* (Lin et al, 2020)	23.1±3.4	51.4±1.9	40.3±2.7	78.5±0.5	79.8±0.6	72.3±0.7
	SumGNN* (Yu et al, 2021)	35.0±4.3	48.8±8.2	41.1±4.7	80.3±1.1	81.4±1.0	73.0±1.4
	DeepLGF* (Ren et al, 2022)	39.7±2.3	60.7±2.4	51.0±2.6	81.4±2.1	82.2±2.6	72.8±2.8
	EmerGNN*	**62.0**±2.0	**68.6**±3.7	**62.4**±4.3	**90.6**±0.7	**91.5**±1.0	**84.6**±0.7
p-value		8.9E-7	0.02	0.02	1.6E-6	6.0E-8	3.5E-5

- Metrics

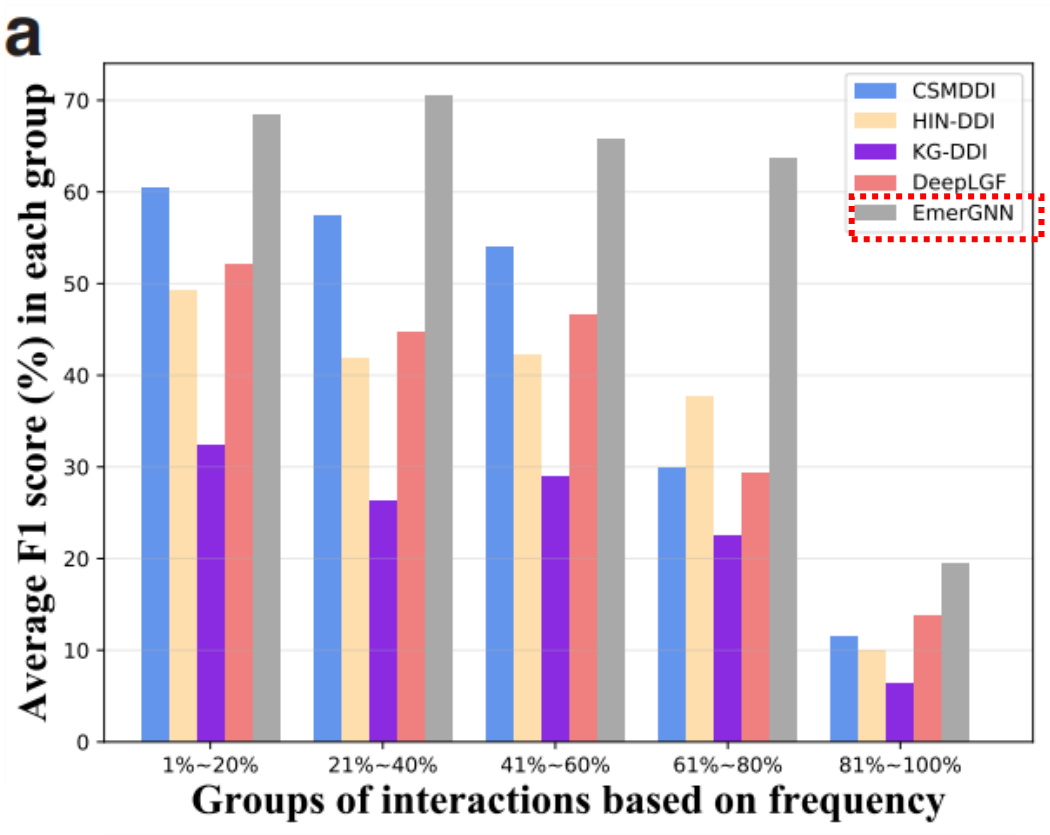
S1 setting

- F1-score (macro) (primary)
- Accuracy
- Cohen's Kappa (Cohen, 1960)

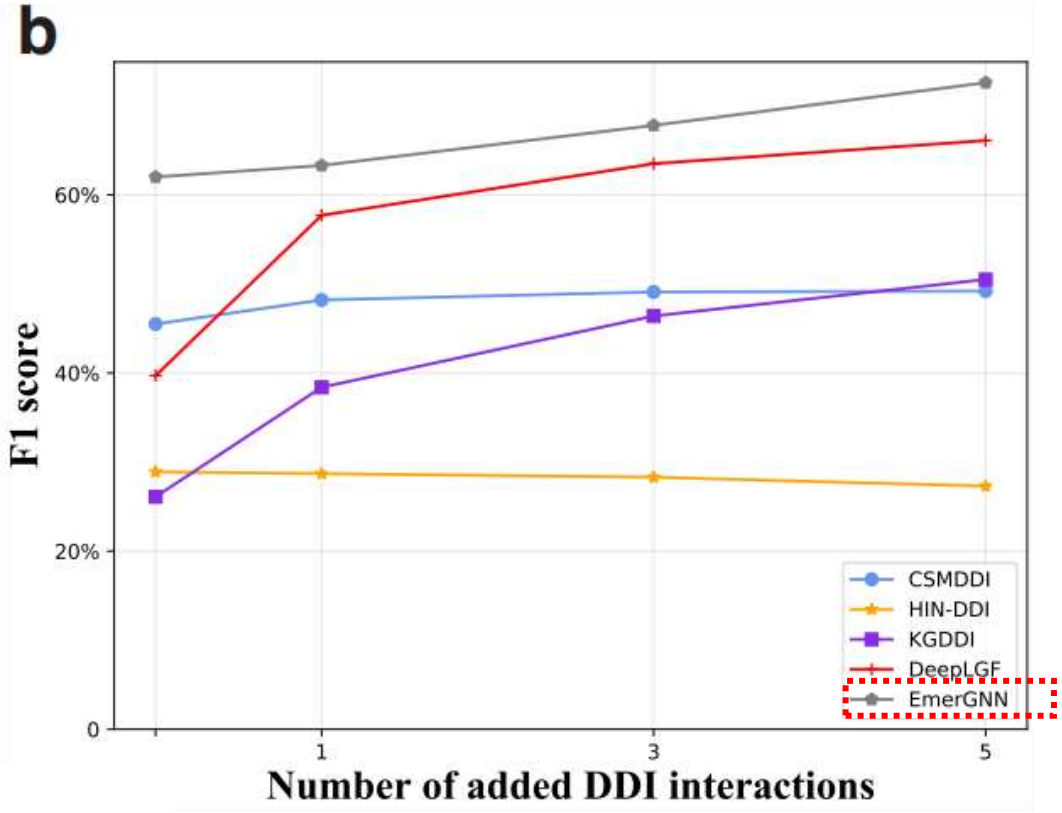
(run on a single Nvidia RTX 3090 GPU with 24GB memory)

Overall, EmerGNN significantly outperforms all the compared methods as indicated by the small p-values.

Performance



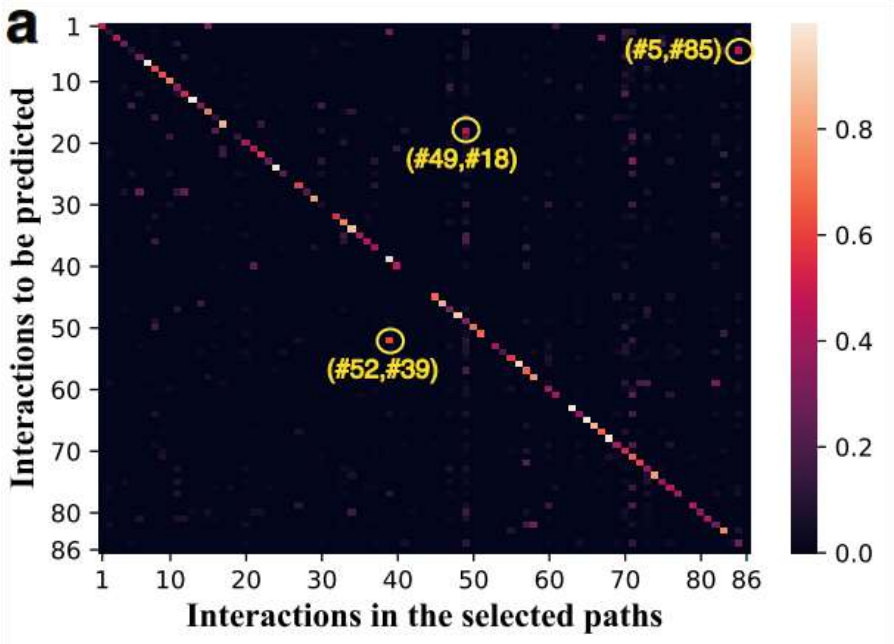
- EmerGNN outperforms the baselines in all frequencies.



Supplement interactions to emerging drugs.

- EmerGNN has increased performance with more interactions added and is still the best over all the compared methods.

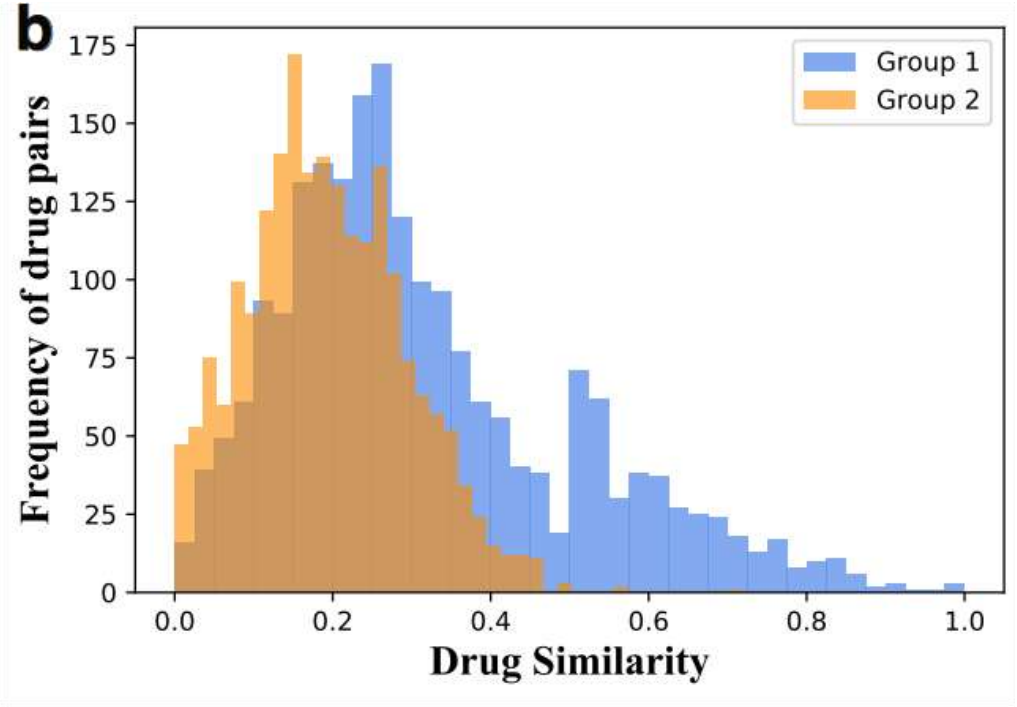
Analyzing the learned paths



- Correlations between interaction i_{pred} and interaction i on the paths in the subgraphs

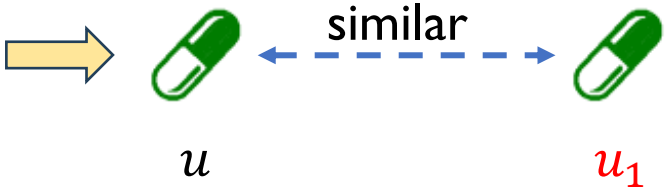
Observation:

- The diagonal elements are dominant.
- The paths with large attention weights are likely to go through drug u_1 which has $i_1 = i_{pred}$.
- Yellow cycles: strongly correlated pairs of interactions.

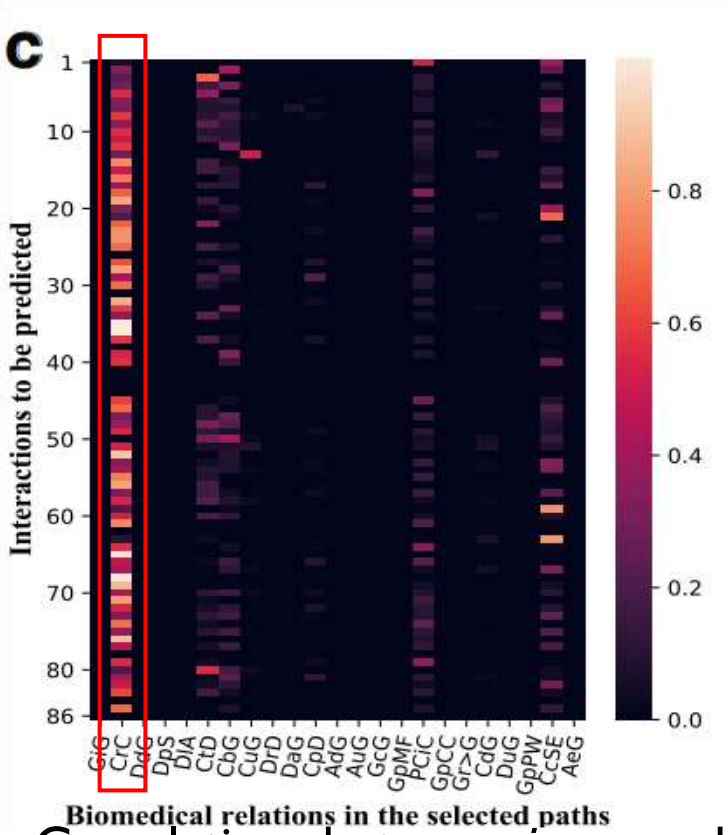


- Group 1: (u, u_1)

  Group 2: (u, u_2) (u_2 is a random drug)



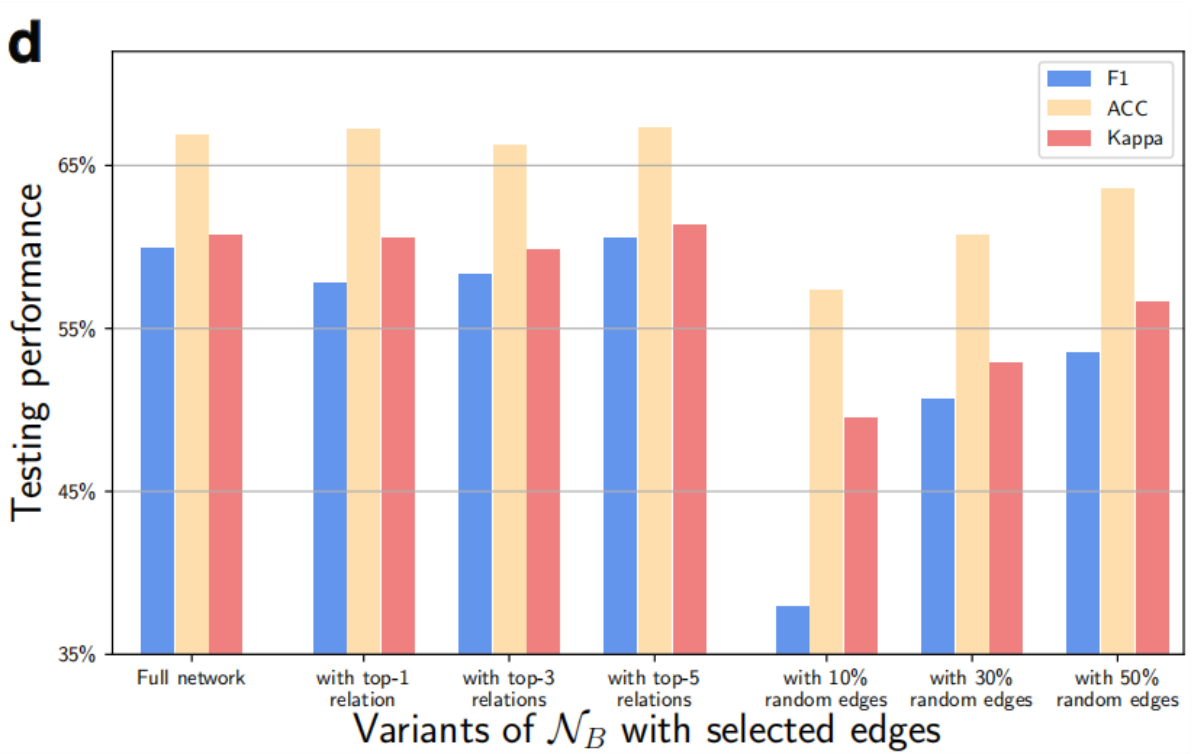
Analyzing the biomedical relation types



- Correlations between i_{pred} and biomedical relations r

Observation:

- Most frequent relation type: CrC (the drug resembling relation)

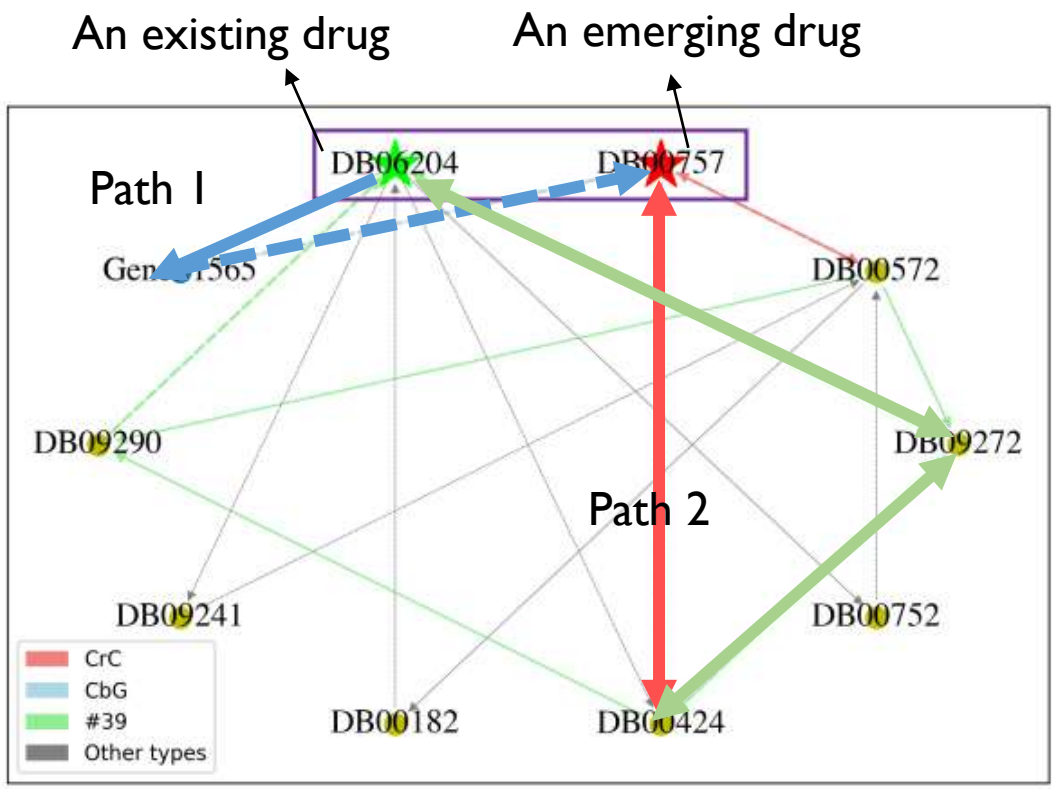


- Only use full network/ top-1 relation (CrC)/ top-3 relations/ …/10% random edges/ 30% random edges/ …

Observation:

- EmerGNN can select important and relevant relations in BN.

Interpretable

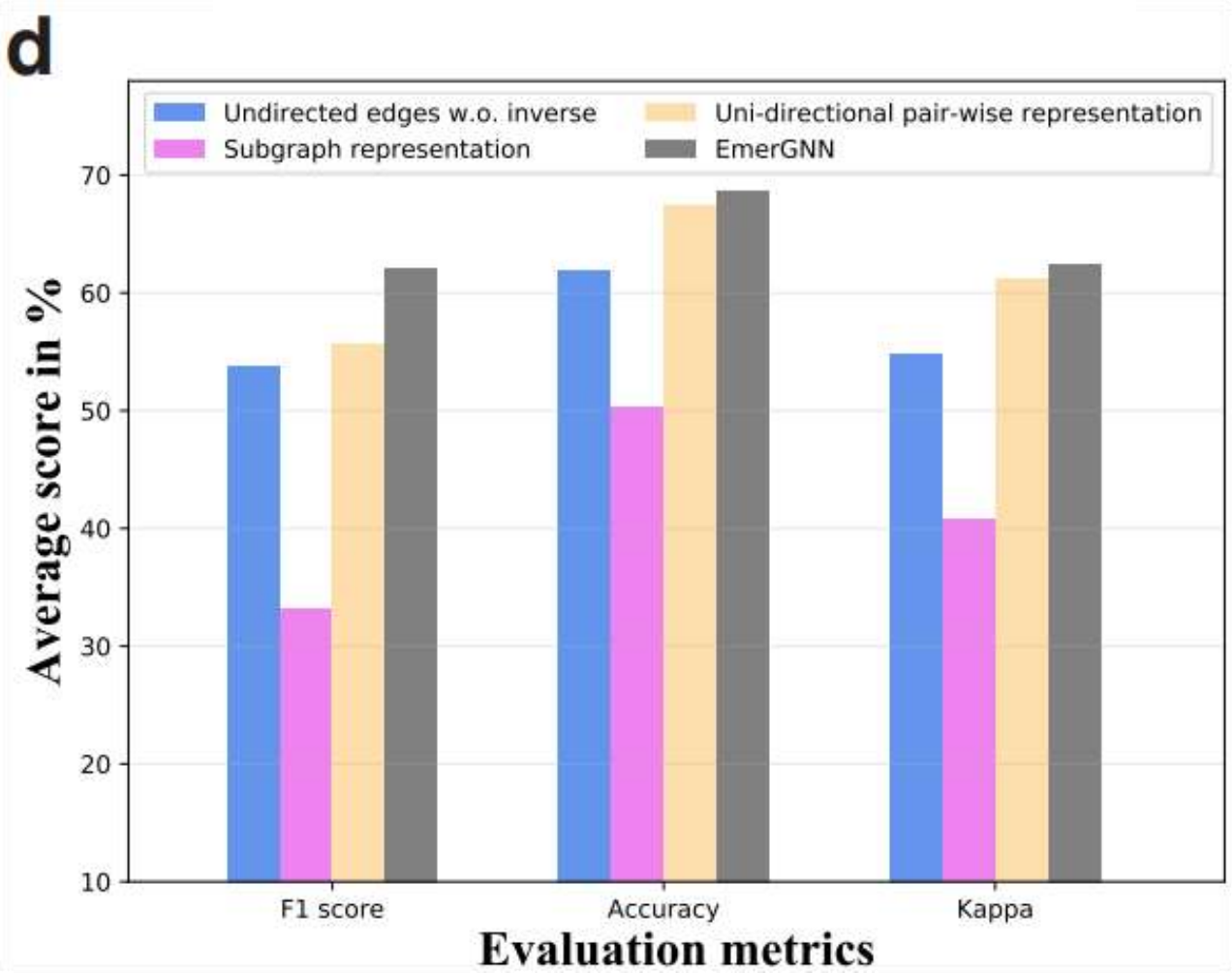


- Select top ten paths between u and v according to edges' attention weights.

Target: Tapentadol (DB06204) may decrease the analgesic activity of Dolasetron (DB00757).
Path1 (0.6666): Tapentadol $\xrightarrow{\text{binds}}$ CYP2D6 (P450) $\xrightarrow{\text{binds\_inv}}$ Dolasetron
Explanation: Tapentadol can binds the P450 enzyme CYP2D6 (Gene::1565), which is vital for the metabolism of many drugs like Dolasetron (Estabrook, 2003). In addition, Binding of drug to plasma proteins is reversible, and changes in the ratio of bound to unbound drug may lead to drug-drug interactions (Kneip et. al. 2008).
Path2 (0.8977): Dolasetron $\xrightarrow{\text{resembles}}$ Hyoscyamine $\xrightarrow{\text{\#39:↑ constipating}}$ Eluxadoline $\xrightarrow{\text{\#39\_inv}}$ Tapentadol
Explanation: Dolasetron is similar to drug Hyoscyamine (DB00424). Hyoscyamine and Tapentadol can get some connection since they will both increase the constipating activity of Eluxadoline (DB09272). As suggested by Liu and Wittbrodt (2022), reversing opioid-induced constipation often causes the unwanted side effect of analgesia reversal.

EmerGNN can find important paths for emerging DDI.

Effectiveness of GNN architecture design

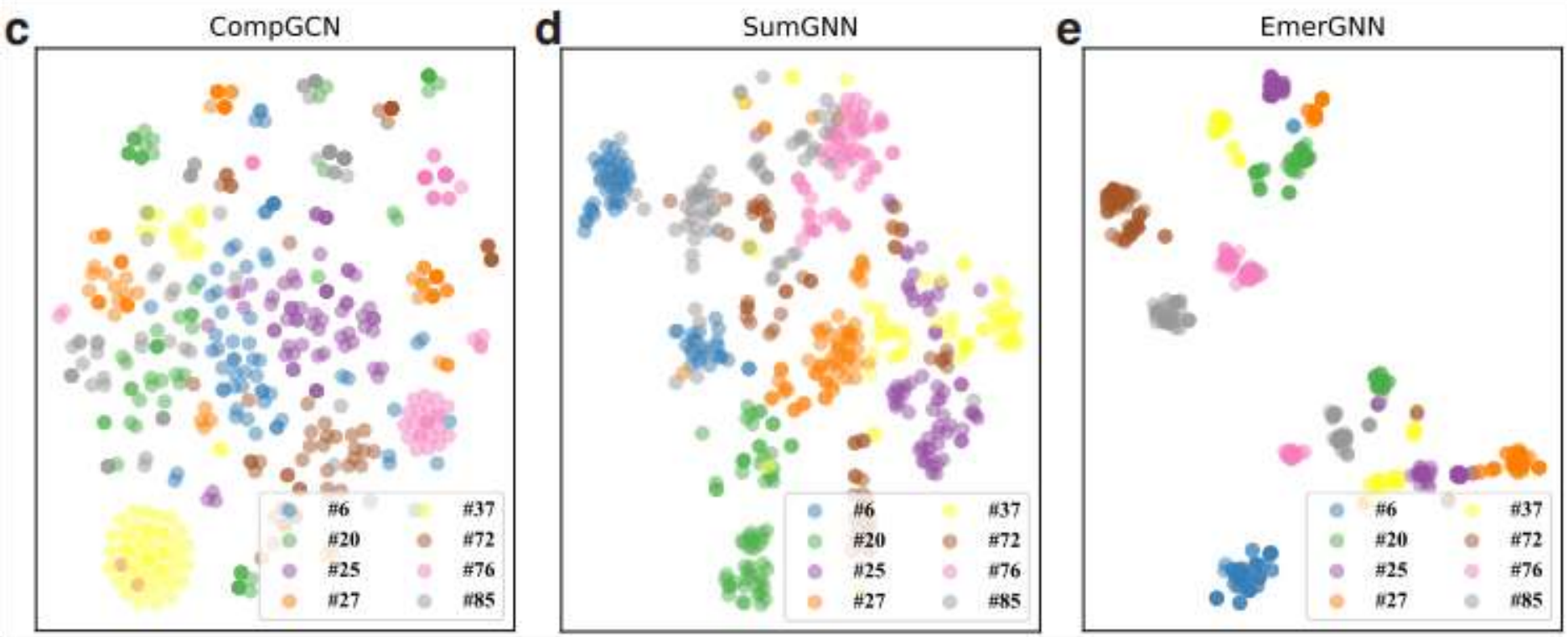


- Performance comparison of different GNN architecture designs.

Conclusion:

- The flow-based GNN architecture is more effective than any other variants of GNN architectures.
- Our design is effective.

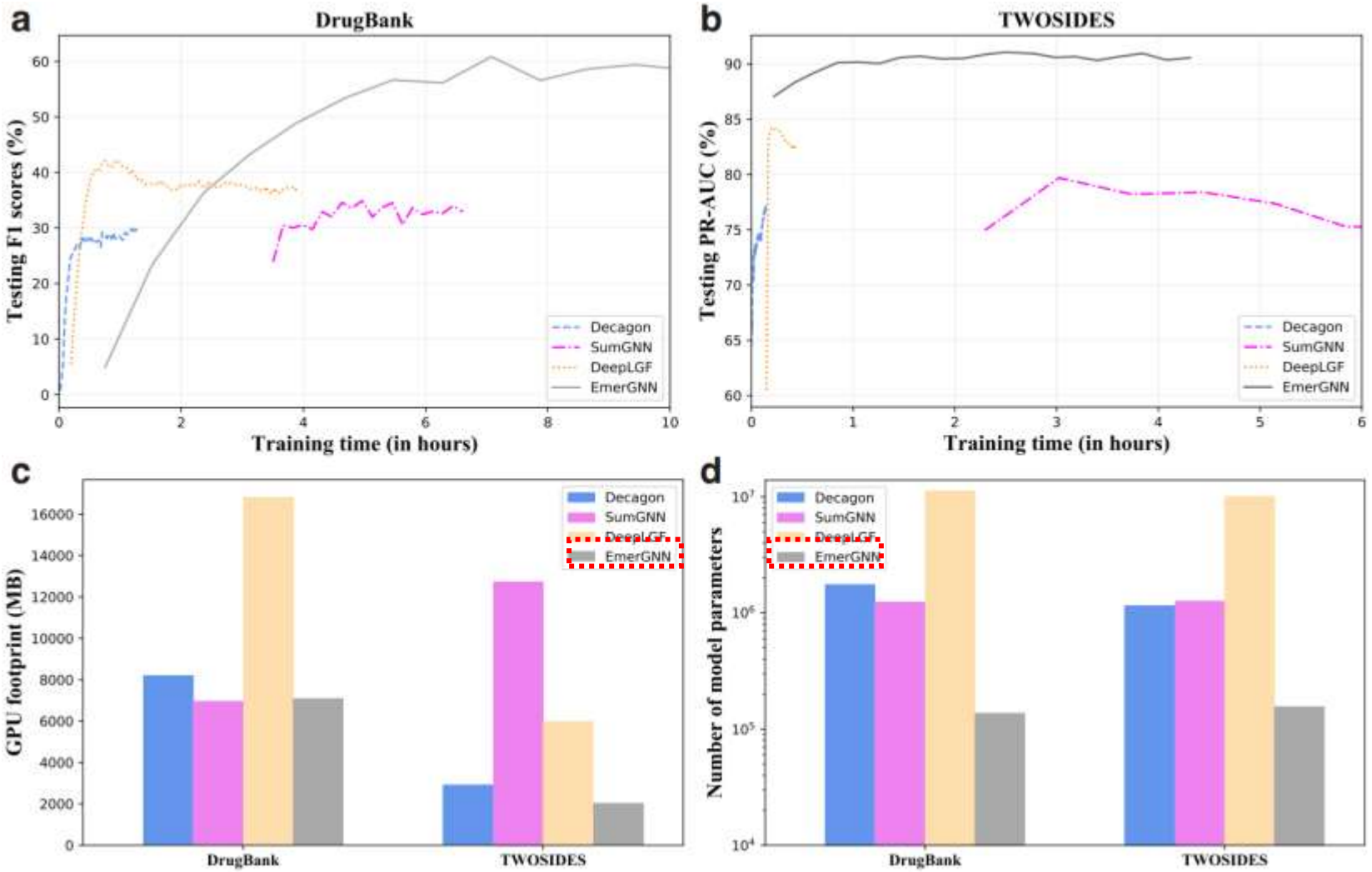
Discrimination ability of GNN



- Dot: interaction (u, i, v)
- Different colors: different interaction types

EmerGNN can separate the different interaction types better.

Complexity analysis



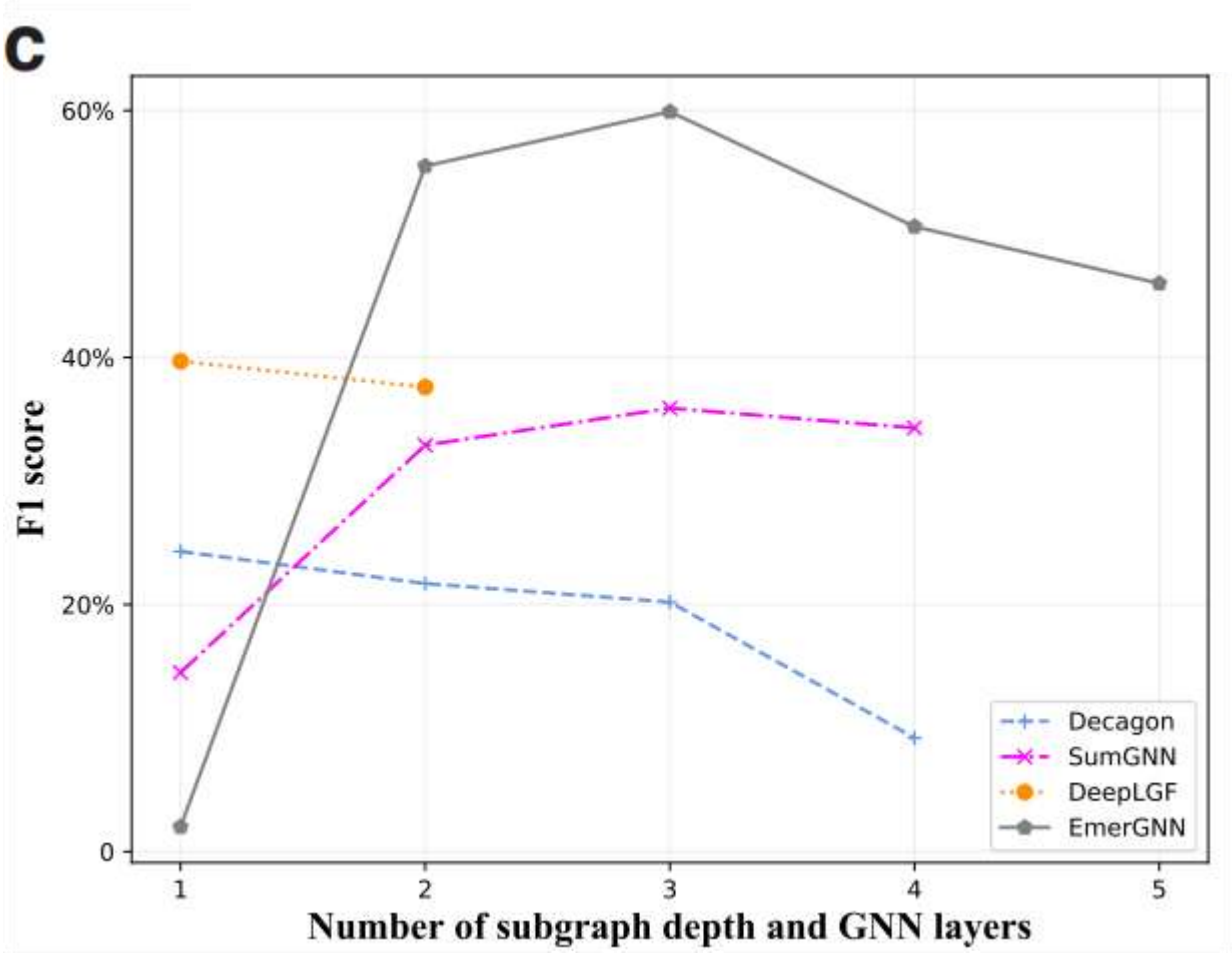
(a)

- EmerGNN can achieve higher accuracy although it takes a long time.
- But compared to clinical development, this is acceptable.

(b)

- EmerGNN is memory and parameter efficient.
- Smaller subgraphs and only relying on biomedical concepts.

Changing the length L



- The value of L determines the maximum number of hops of neighboring entities that the GNN-based models can visit.
- $L = 3$ is optimal for EmerGNN considering both the effectiveness and computation efficiency.

Summary of EmerGNN

- Outperforms existing methods in emerging DDI prediction.
- Exploit the rich knowledge in existing large biomedical networks for low-data scenarios.
- Customize small subgraphs and a flow-based GNN architecture to effectively extract essential information.
- May contribute to improving patient care and more efficient drug development processes.

Recent works in our group

- Emerging Drug Interaction Prediction Enabled by Flow-based Graph Neural Network with Biomedical Network. *Nature Computational Science*. 2023 [covered in this talk]
- Accurate and interpretable drug-drug interaction prediction enabled by knowledge subgraph learning. *Nat. Com. (Medicine)*. 2023
- Bilinear Scoring Function Search for Knowledge Graph Learning. *TPAMI*. 2023
- Relation-aware Ensemble Learning for Knowledge Graph Embedding. *EMNLP*. 2023
- Automated 3D Pre-Training for Molecular Property Prediction. *KDD*. 2023
- KGTuner: Efficient Hyper-parameter Search for Knowledge Graph Learning. *ACL*. 2022
- Knowledge Graph Reasoning with Relational Digraph. *WebConf*. 2022
- Property-Aware Relation Networks for Few-Shot Molecular Property Prediction. *NeurIPS*. 2021
- Generalizing from a Few Examples: A Survey on Few-Shot Learning. *CSUR*. 2020.

collaborators


Yongqi Zhang


Ling Yue


Yaqing Wang


Huan Zhao


Yefeng Zheng


James Kowk

Thanks & QA.