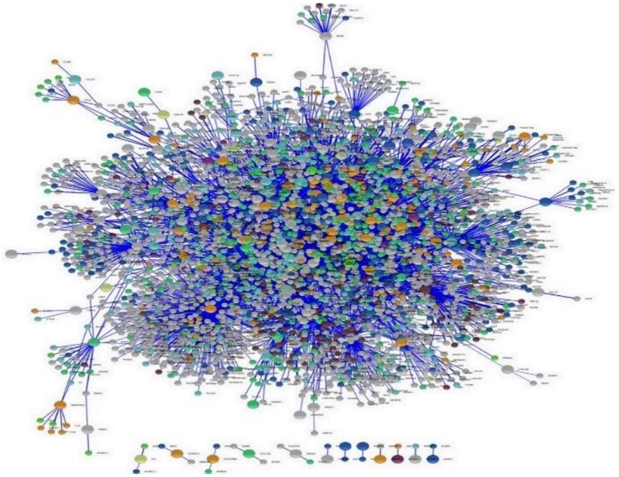
Disentangling Invariant Subgraph via Variance Contrastive Estimation under Distribution Shifts

Haoyang Li, Xin Wang, Xueling Zhu, Weigao Wen, Wenwu Zhu

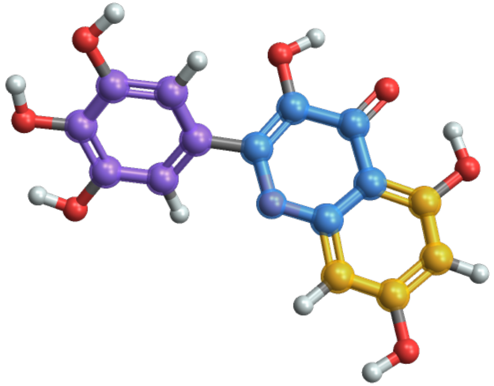
Graph Structured Data is Everywhere

Social Networks

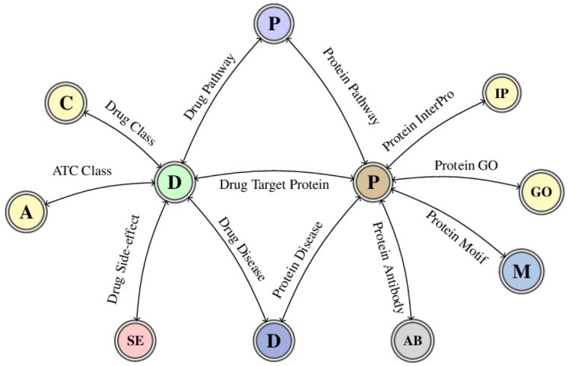
Biology Networks



Molecular Graphs

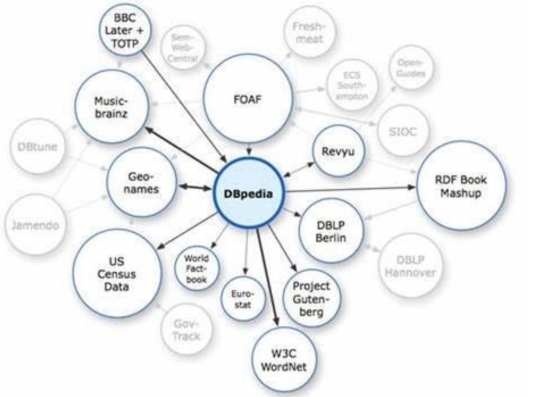


Knowledge Graphs



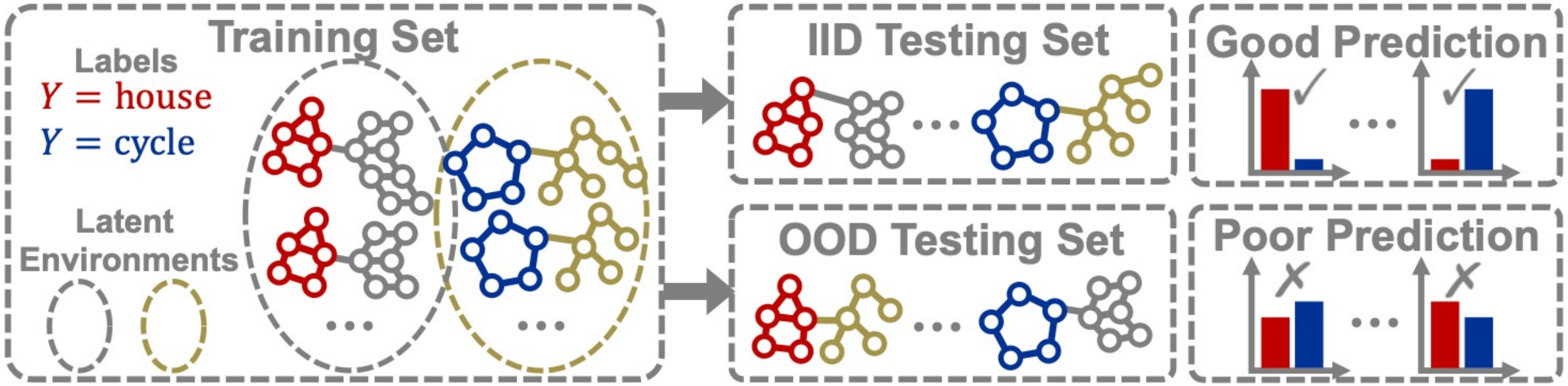
Traffic Networks

Information Networks

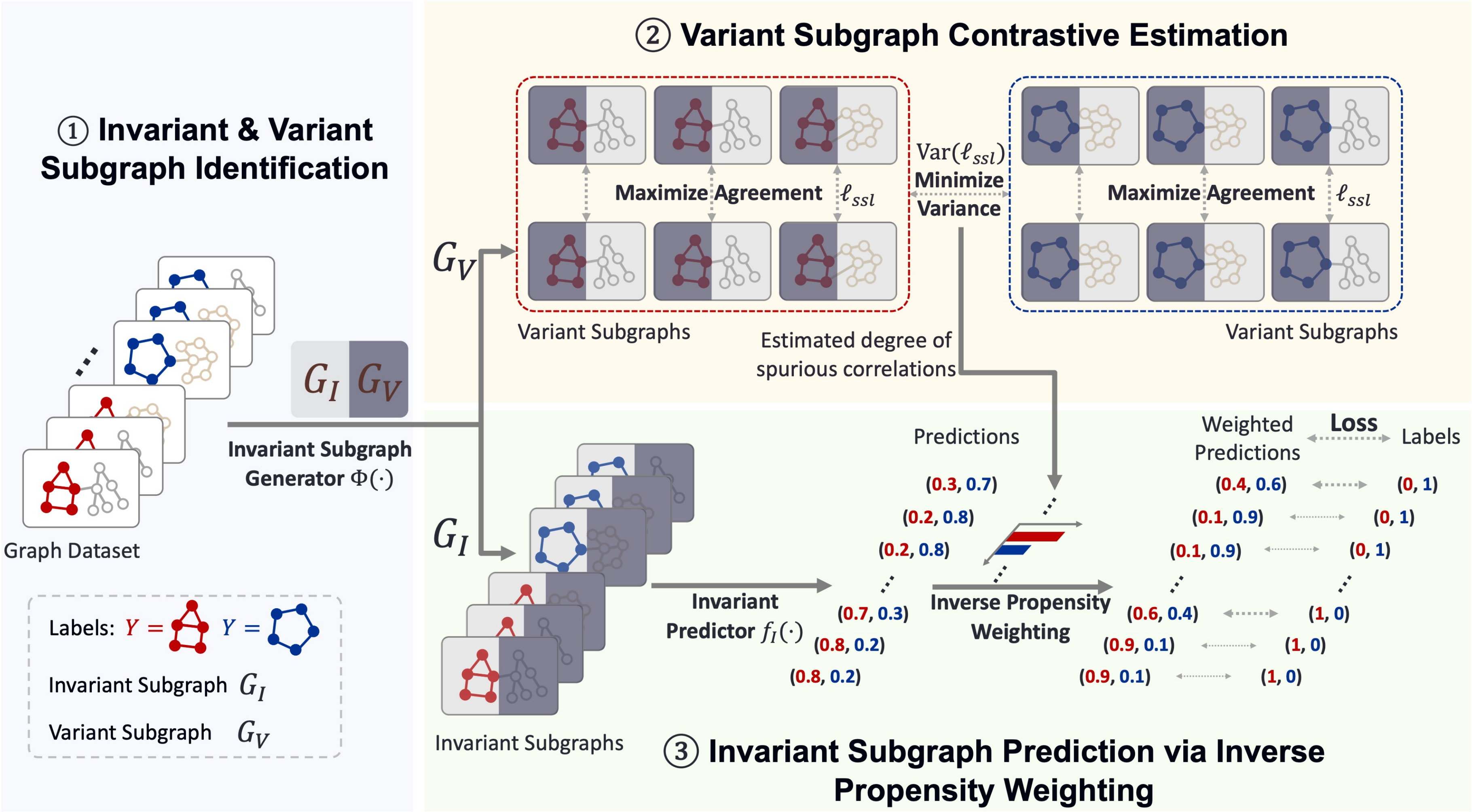


Graph Neural Networks

- **Today's graph neural networks (GNNs)**
  - GNNs built upon the in-distribution (I.D.) hypothesis fail to generalize to out-of-distribution (OOD) environments.
  - Most graph invariant learning methods heavily rely on the predefined or automatically generated environment labels, i.e., multiple training environments.
  - However, the environment labels are unavailable in most scenarios and directly annotating or generating environment labels is impractical or inaccurate.



Model Framework



Theoretical Analysis

- **Theoretical Analysis**
  - Our method can disentangle the ground-truth invariant and variant subgraphs which is a significant step towards OOD generalized predictions.

Theorem 1. *Denote the optimal invariant subgraph generator* Φ^* *that disentangles the ground-truth invariant subgraph* G_I^* *and variant subgraph* G_V^* *given the input graph* G, *where* G_I^* *satisfies Assumption 1 and denote the complement as* $G_V^* = G \backslash G_I^*$. *Assume the second variance term of Eq. (5) is minimized, we have that the first contrastive loss term is minimized iff the invariant subgraph generator* Φ *equals* Φ^*.

Experimental Results

- **Results on real-world benchmarks**

Table 1. Experimental results (%) of our method and baselines. The evaluation metric is accuracy for CMNIST, CFashion, and CKuzushiji, and ROC-AUC for MOLSIDER and MOLHIV. ± denotes the standard deviation. The best results are in bold for each row. Our **VIVACE** outperforms the baselines in all comparisons, indicating its superiority against graph distribution shifts.

Dataset	**Bias**	**Methods**							
	r	GCN	GIN	FactorGCN	DiffPool	DIR	LDD	DisC	**VIVACE**
CMNIST	0.8	50.43±4.13	57.75±0.78	72.30±1.18	73.79±0.02	9.98±0.33	64.95±1.22	82.60±0.93	**82.71±0.74**
	0.9	28.97±4.40	36.78±5.55	62.35±5.07	66.45±0.78	9.96±0.23	56.65±2.18	78.14±2.14	**79.46±1.87**
	0.95	13.50±1.38	16.04±1.14	42.50±4.91	47.12±1.04	10.03±0.27	46.83±2.88	63.47±5.65	**64.72±4.61**
CFashion	0.8	63.60±0.53	64.25±0.46	61.23±1.11	62.82±0.53	13.02±1.92	63.85±1.17	66.85±1.11	**67.09±1.23**
	0.9	57.22±0.93	58.03±0.40	53.50±1.29	57.50±0.39	12.80±1.67	64.30±0.89	65.33±4.70	**65.38±4.18**
	0.95	47.69±0.42	49.74±0.60	45.78±2.40	50.86±0.20	11.98±1.41	62.28±0.48	63.93±1.50	**63.96±1.27**
CKuzushiji	0.8	38.45±1.10	41.83±0.78	42.87±1.19	45.46±0.65	10.35±0.32	42.38±0.33	55.53±2.29	**55.58±1.87**
	0.9	28.35±0.79	30.09±0.87	32.35±2.79	36.18±0.19	10.72±0.27	38.75±0.49	48.13±2.59	**48.15±1.91**
	0.95	20.70±0.88	21.18±1.63	23.87±0.12	27.45±0.26	10.59±0.46	33.08±0.59	36.63±1.73	**37.01±1.67**
MOLSIDER		59.62±1.82	57.61±1.48	53.32±1.75	60.21±1.55	57.74±1.63	58.83±1.62	59.31±1.87	**62.15±1.10**
MOLHIV		76.13±1.01	75.63±1.41	57.18±1.54	76.32±1.48	77.05±0.57	76.91±1.81	76.97±1.03	**78.11±0.82**

Best OOD generalization performances

Thanks!

Disentangling Invariant Subgraph via Variance Contrastive Estimation under Distribution Shifts