

日期: /

## GNN Augmentation and Training

图增强技术的理念是, 原始输入图不一定完全呈现计算图。

使用图增强技术的动机:

1) 原始图节点可能缺工特征

2) 原始图可能存在不良结构:

① 过于稀疏: 低效的消息传递

② 过于稠密: 昂贵的消息传递

③ 过大: 计算图不能适配 GPU

## [Reasons for Feature Augmentation on Graphs]

1) 原始图可能并不具有节点特征, 以种情况下可以为节点分配特征:

	(常数特征)	(独热编码特征)
Expressive power	<b>Constant node feature</b> Medium. All the nodes are identical, but GNN can still learn from the graph structure 	<b>One-hot node feature</b> High. Each node has a unique ID, so node-specific information can be stored 
Inductive learning (Generalize to unseen nodes)	High. Simple to generalize to new nodes: we assign constant feature to them, then apply our GNN	Low. Cannot generalize to new nodes: new nodes introduce new IDs, GNN doesn't know how to embed unseen IDs
Computational cost	Low. Only 1 dimensional feature	High. $O( V )$ dimensional feature, cannot apply to large graphs
Use cases	Any graph, inductive settings (generalize to new nodes)	Small graph, transductive settings (no new nodes)

日期： /

2) 某些图结构对 GNN 来说难以学习，因此我们希望将其编码至节点特征中。例如 GNN 无法学习节点所在环中的节点数。

### [稀疏图的增强方式]

- 1) 加入虚拟边：使用虚拟边连接 2-hop 邻居节点  
(使用  $A + A^2$  代替原始邻接矩阵  $A$ )
- 2) 加入虚拟节点：虚拟节点与图中所有节点均有连接  
这使得图中任意两个节点的距离不多于 2-hop

### [稠密图的增强方式]

使用特殊的采样策略，从特定的邻居子集中接收消息