

网络表示学习第三次作业报告

- 姓名：甘云冲
- 学号：2101213081

本次作业需要使用GNN+分类head的模型结构进行节点分类任务，并且将GNN的输出作为表示利用t-SNE在二维空间内进行可视化，来查看GNN表示学习的最终效果。具体实现借助torch_geometric第三方库。

本次作业所采用的数据集与第二次相同，三个数据集的基本统计信息如下：

| | Cora | Citeseer | Amazon Photo |
|--------------------------------|-------|----------|--------------|
| Number of nodes | 2708 | 3264 | 7535 |
| Number of edges | 5278 | 4536 | 119082 |
| Average Degree | 3.898 | 1.779 | 31.61 |
| Average Clustering Coefficient | 0.241 | 0.145 | 0.410 |

在图神经网络模型的embedding之上直接做一个线性映射，将图神经网络学到的表征映射到分类的类别数量，作为模型的分头。以下参数在本次作业所有实验当中保持一致：

| Hyperparameters | Value |
|-----------------|-------|
| hidden_units | 128 |
| dropout | 0.5 |
| num_layers | 2 |
| early_stop | 50 |
| max_epochs | 300 |
| optimizer | Adam |
| learning_rate | 0.01 |

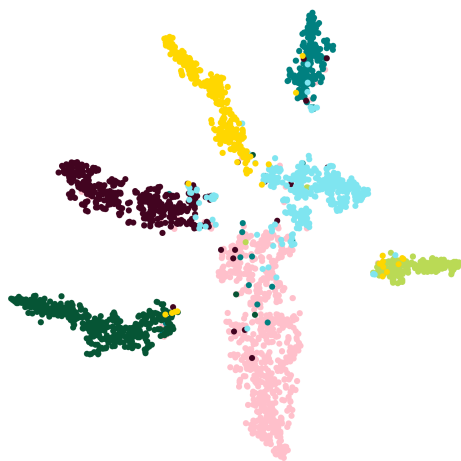
以下为各模型在不同数据集下的分类准确率，其中加粗部分为本次作业实验结果：

| Model | Cora | Citeseer | Amazon Photo |
|------------------|--------|----------|--------------|
| DeepWalk | 0.8429 | 0.6120 | 0.9085 |
| Node2Vec | 0.8429 | 0.6380 | 0.9131 |
| GCN | 0.8928 | 0.7684 | 0.9484 |
| GAT | 0.8909 | 0.7699 | 0.9529 |
| GIN | 0.8632 | 0.7193 | 0.9412 |
| GraphSAGE | 0.9039 | 0.7653 | 0.9510 |

可以发现，通过GNN学习所得到的表示，能够较DeepWalk和Node2Vec得到更高的分类准确率。

Cora

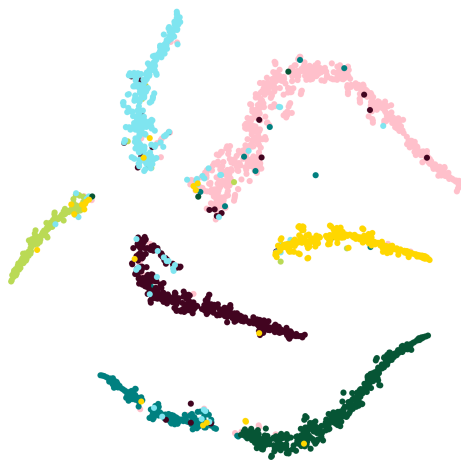
Cora GCN



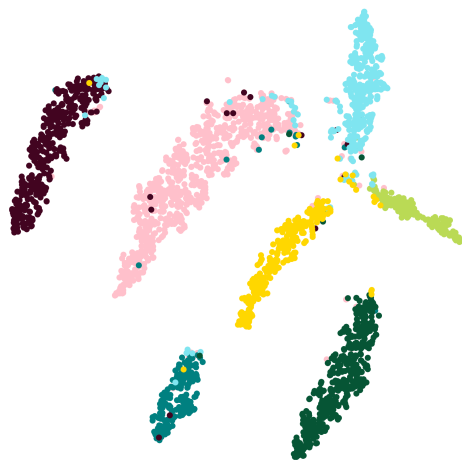
Cora GAT



Cora GIN



Cora SAGE



Citeseer

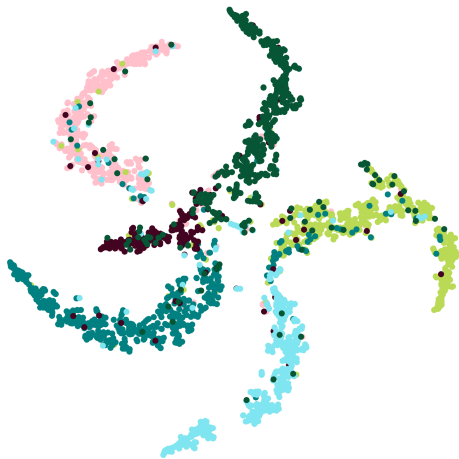
Citeseer GCN



Citeseer GAT



Citeseer GIN



Citeseer SAGE

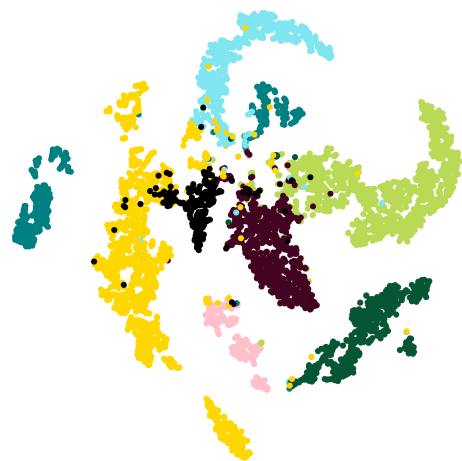


Amazon Photo

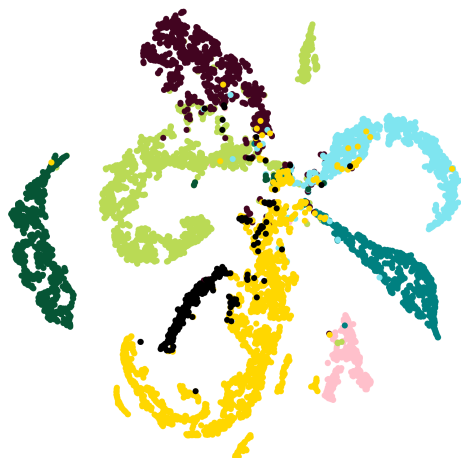
Amazon_photo GCN



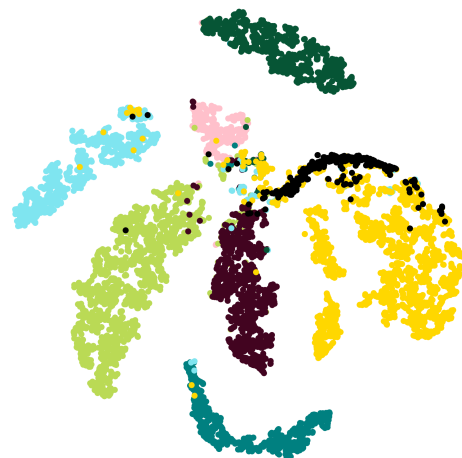
Amazon_photo GAT



Amazon_photo GIN



Amazon_photo SAGE



从可视化结果可以看出，图神经网络能够确实学到图节点上特征的关系，并且在二维平面上的投影明显相同类别的节点聚集在一块儿。