

DL lab5

PB18000149 吴越凡

June 2021

1 实验目的

在给定的三个图数据 Cora,Citeseer,PPI 上实现 GCN 算法去用于节点分类和链路分析。

2 实验细节

实验内容中的 GCN 是通过 torch 库进行修改得到。

模型首先是通过一个 GCN 映射到隐藏层上，然后经过 relu 操作之后，通过 dropout 后再经过一个 GCN。

3 实验结果

3.1 Cora

3.1.1 Base model

```
GCN(  
  (gc1): GraphConvolution (1433 -> 16)  
  (gc2): GraphConvolution (16 -> 7)  
)
```

Accuracy: 79.2%

Self_Loop	Activation	Accuracy
True	ReLU	83.4%
True	None	84.1%
False	None	77.5%
False	ReLU	79.8%

表 1: Self-Loop

3.1.2 Self-Loop

3.1.3 Layer number

没有自环

Layer number	Activation	Accuracy
2	None	79.2%
3	None	70.9%
4	None	74.3%
5	None	72.7%
2	ReLU	79.7%
3	ReLU	73.4%
4	ReLU	70.9%
5	ReLU	30.9%

表 2: Layer Number

3.1.4 DropEdge

从以上结果中，选择两层 + 自环 + ReLU

3.1.5 PairNorm

鉴于上述的情况，这里选择两层 + 自环 + 无 dropedge+ ReLU

DropEdge	Accuracy
0.1	79.9%
0.2	76.6%
0.3	66.7%
0.4	57.4%

表 3: DropEdge

Pairnorm_scale	Accuracy
0	79.1%
10	70.2%
20	66.6%

表 4: PairNorm

3.1.6 Activation

Activation	Accuracy
None	76.0%
ReLU	83.67%
Sigmoid	30.3%
Tanh	81.67%

表 5: Activation

3.2 Citeseer

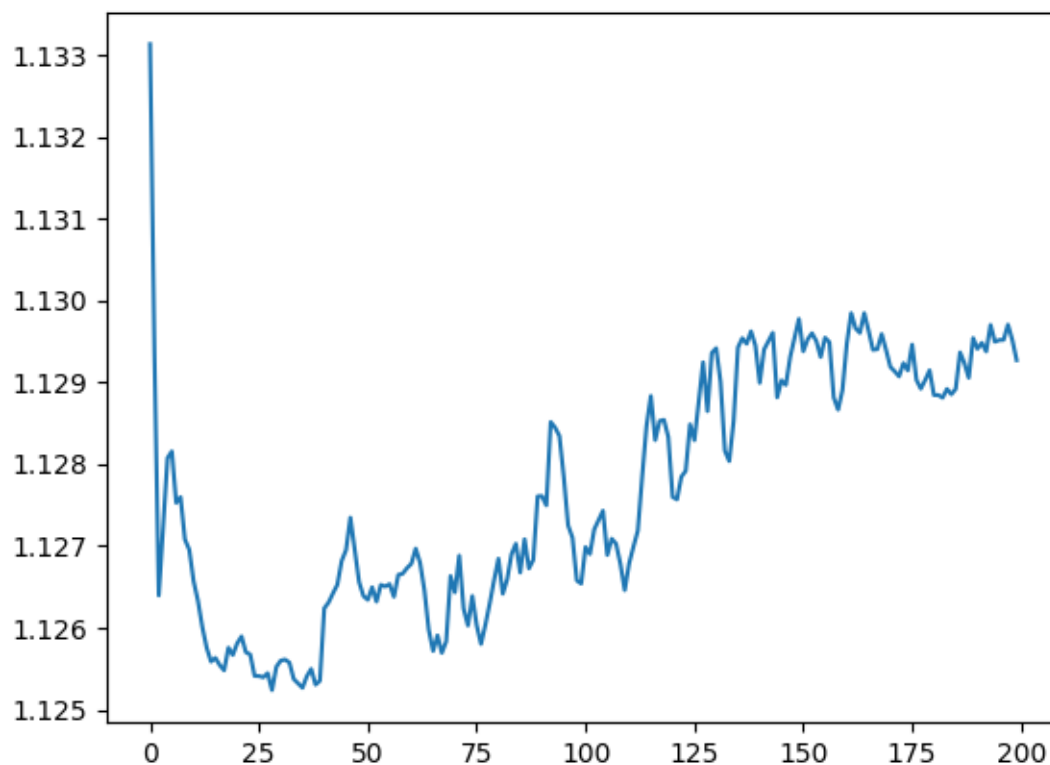
运用上面选出的参数：

两层 + 自环 + ReLU + 无 Dropedge + 无 PairNorm

Accuracy: 47.33%

3.3 PPI

accuracy= 0.7444 f1_val: 0.4003



4 总结

本次实验中，主要对 Cora 数据集上进行了比较实验，对比在自环、层数、DropEdge、PairNorm、激活函数等方面的变化对实验结果的影响。

- 自环的增加对模型效果的提升有帮助
- 关于层数方面，层数数量与模型效果总体成反比
- 关于 DropEdge 方面，DropEdge 的比例越高，模型的效果越差

- PairNorm 的情况，在两层的 GCN 上显示，PairNorm 的 scale 大小与模型效果成反比
- 在内部的 Activation 方面，结果显示使用 ReLU 的效果相对于另外的激活函数来说更好