# 问题

在run\_er\_debugl中

使用前面几张图作为训练DQN的数据,然后用最后一张图训练+求出种子 这样需要区分list中的种子是不是代表一个或者一类图吗?,这个list中的图有什么关联, what if list中只给出一张图,对一张图进行多次训练会获得更好的结果吗?

---这个模型是具有泛化性的还是针对一个图只有一个模型有效?大概想清楚了

sdne中的训练数据和目标编码图像的关系,是不是具有相同特征的图像(如小世界特征)则可以用来当做训练数据

如何查找研究的graph数据,有没有可以生成graph的工具或方法,爬取微信?

关键问题:模型中采用的将无向图转换为有向图的并给边赋予权重 ---发现在无向图中一条边总是存储 double: a-b和b-a均存在,但是在sdne中的模型图中值存a-b或者b-a,即原来的图像就是有向图数据,这里会不会对影响力计算方式带来某些问题? --按理说不会,因为无向图最终也是转换为有向图进行计算的,这里需要关注一下,可能还需重新训练sdne模型

# 考虑

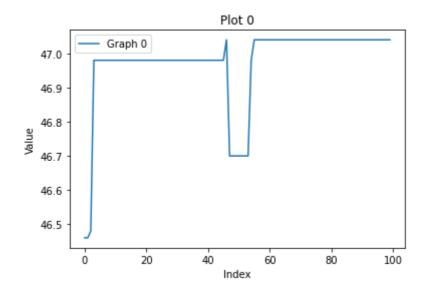
TODO: 考虑如果转为SDNE编码之后是否还需要增加节点度这一列信息

DQN模型训练的时候是将节点是否作为种子也当成一个特征值,这样会如何影响归一化?如果不这样的话如何将是否为种子节点编码到特征中去?---key:了解编码模型的原理

# 实验结果

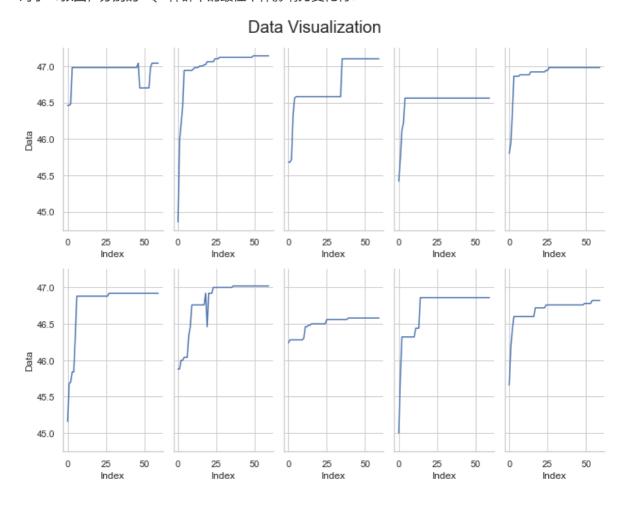
图中展示的是100个种群当中最好个体的变化情况,很多时候出现训练过程中最好的个体得分不变,但 是平均分会有变化(大体上升)

对GN图像进行line编码训练过程影响力变化情况,图1:



下降原因:可能是训练的DQN参数随机替换的时候把pop[0]给替换掉了?

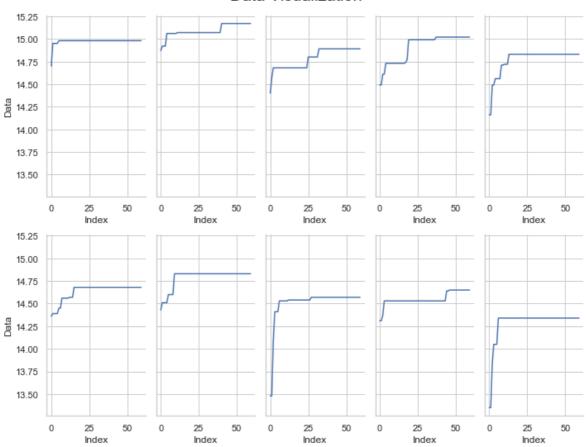
对于10张图,分别的DQN种群中的最佳个体影响力变化有:



Karate\_graph\_transform 34节点图像进行100轮进化:

line编码

# **Data Visualization**



### influence:

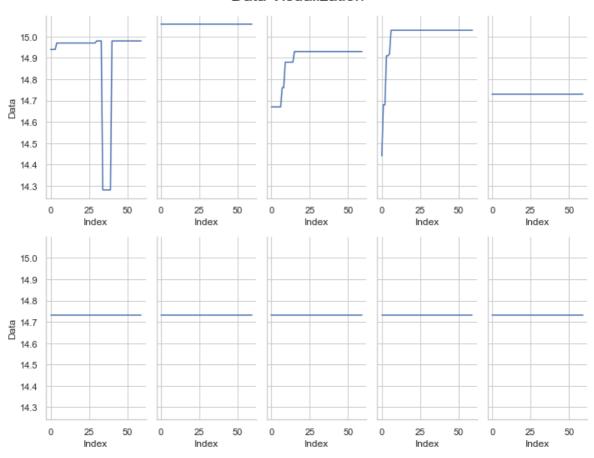
14.98000000000008 15.170000000000007 14.89000000000008 15.02000000000007 14.830000000000007 14.68000000000000 14.8300000000000 14.57000000000007

14.6500000000001 14.34000000000012

14.340000000000012

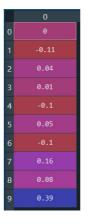
sdne编码

## **Data Visualization**

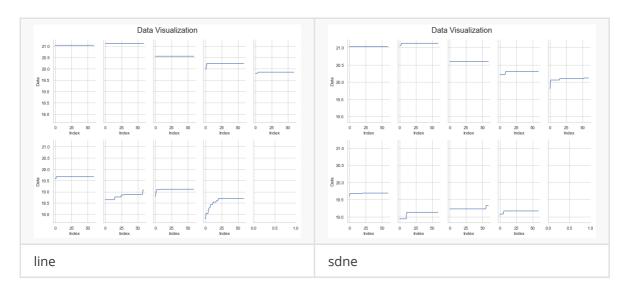


#### influence:

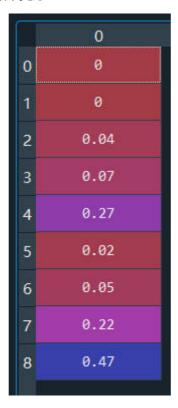
10个graph最终的influence 求差(SDNE结果减去LINE结果):



尝试调低学习率为0.0001,每次DQN参数更新后更换8个个体,得到图像最优值变化很小,但是sdne由于line似乎可以验证



差值: 正数代表sdne编码所取得的效果更好



一个实验结果: sdne编码的最大影响-line编码得到的最大影响

# 关键疑惑点:

代码里面会创建100个DQN参数作为种群,然后用种群中最好的一个DQN参数作为这一轮的结果。

现在有一点我感觉有点奇怪:往往训练很多轮,其余不是最优的DQN参数会得到改进,但是最好的DQN参数一直没有变化(如上面图中最大影响力变化图中最优影响力始终不变)

### 猜想原因:

这就代表有时候随机产生的DQN参数选取的种子达到了最优效果,但是我看到达到相同最大影响力时选择的种子集有时候会不同,因此可能是由于图比较简单导致。

还可以尝试继续增加训练轮数: (目前是训练100轮)

## reword的算法

# 找相关论文、数据

- 1.代码改进
- 2.研究背景与实际应用 (为什么研究该问题,有什么实际价值)
- 3.数据搜集(搜集什么网络适用于该问题研究or自己创建一个图网)
- 4.文献调研(了解最新进展用的方法,能否应用到该问题中)