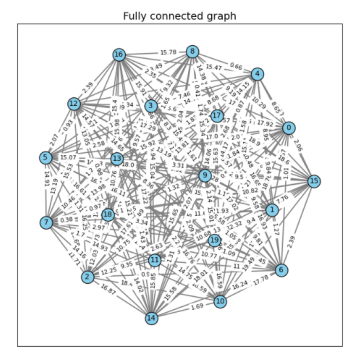
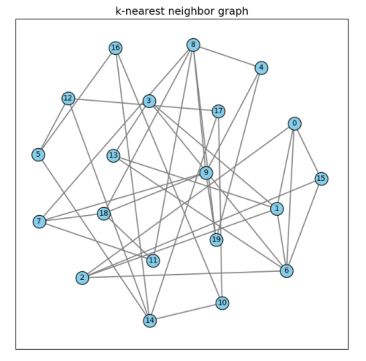
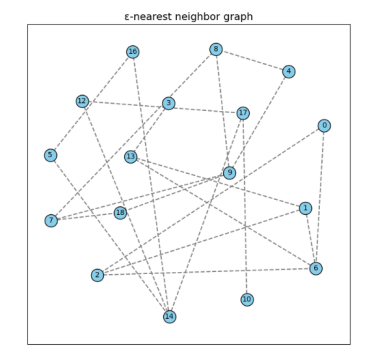
# 三种构图方式可视化

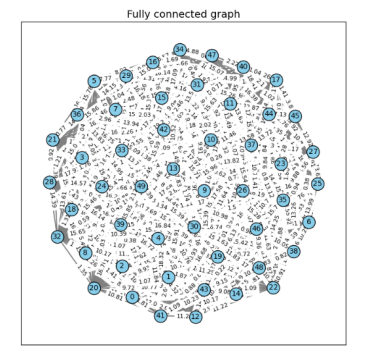
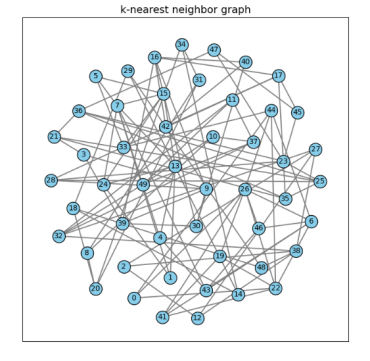
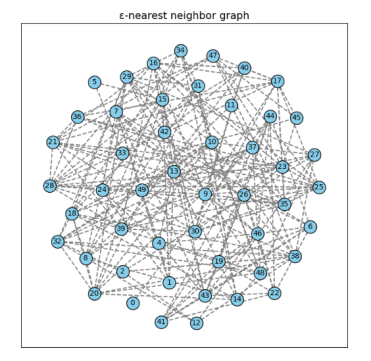
## 软件学院 人工智能 20221120286 许凯

我构建了动态图和静态图两种形式，gif文件无法插入，故放在了文件夹下，展现了图的动态构建形式，以下为静态图与三种构建的分析，动态文件参照文件夹下20样本点的图构建过程。

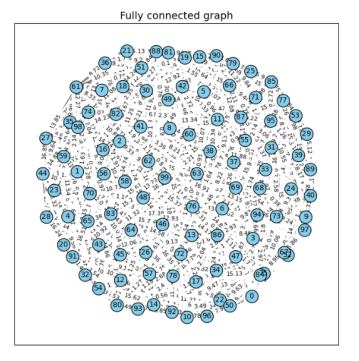
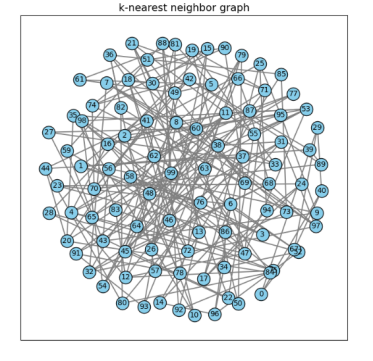
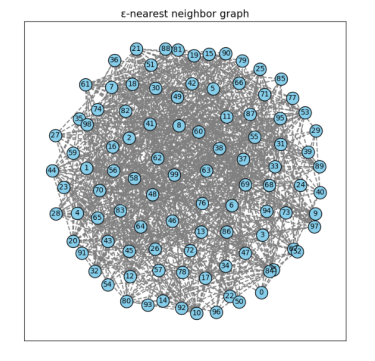
**样本点=20**

****

**样本点=50**



**样本点=100**



# 实验总结：不同节点数的图构建可视化分析

## 1. 概括

通过 ε-近邻图、k-近邻图、全连通图 三种方式构建图结构，并在不同样本点规模（20、50、100）下进行可视化，观察它们的结构变化、连通性特点以及计算复杂度

样本点数量：分别取 20、50、100 个随机样本点进行实验

## 图构建方式：

ε-近邻图（Epsilon-nearest neighbor graph）：连接欧式距离小于 ε 的点

k-近邻图（k-Nearest neighbor graph）：每个点与最近的 k 个点相连

全连通图（Fully Connected Graph）：所有点互相连接，边权重为点之间的距离

动画展示方式：动态变化 ε 值和 k 值，观察图的逐步演化

## 可视化的图的分析

### ε-近邻图

| 样本点数 | 连接情况 | 结构特征 | 可视化效果 |
| --- | --- | --- | --- |
| 20 | 形成多个小的连通分量，逐步合并 | 结构较松散，ε 增大时逐渐连通 | 逐步从孤立点聚合成连通图 |
| 50 | 形成多个小团簇，部分点仍未连通 | 需要更大的 ε 才能形成完整连通分量 | 结构逐步由稀疏变密集 |
| 100 | 初始时很多孤立点，ε 需较大值才能连通 | 局部连通性强，但全局连通较慢 | 逐步出现大规模聚集结构 |

小样本时 ε 控制得当可以形成合理的局部连通结构样本量增大时，必须适当增大 ε，才能保证整体连通性，过大的 ε 会导致图过于密集，信息冗余

### k-近邻图

| 样本点数 | 连接情况 | 结构特征 | 可视化效果 |
| --- | --- | --- | --- |
| 20 | k=1 时很多点仍未连通，k=5+ 时形成较好结构 | 结构比 ε 近邻图更均匀 | 逐步增加 k 时图变得稠密 |
| 50 | k=3+ 时大部分点已连通，但仍存在部分弱连通区域 | 结构紧凑，局部群集明显 | 图结构稳定增长，局部团簇突出 |
| 100 | k=5+ 时基本完全连通 | 结构类似小世界网络 | 图整体呈现较均匀的分布 |

k 近邻图相比 ε 近邻图更稳定，k 值相同时不同规模样本的连通性相似k过小如 k=1）时，可能存在孤立点；k 过大时，趋向全连通图，在大规模数据中，k近邻比ε近邻更可控，且能避免过度连通

### 全连通图

| 样本点数 | 连接情况 | 结构特征 | 可视化效果 |
| --- | --- | --- | --- |
| 20 | 所有点相连 | 计算简单，但信息冗余 | 图结构密集 |
| 50 | 计算复杂度显著上升 | 需要适当筛选边 | 图显示较拥挤 |
| 100 | 计算和可视化开销大 | 过度连通导致信息冗余 | 图几乎变成一团黑色 |

全连通图适合 小规模数据，但不适合大规模数据，计算复杂度 O(n²)，当样本数增大时 计算成本急剧上升。

## 综合对比

| 图类型 | 适用场景 | 计算开销 | 适用数据规模 | 主要特点 |
| --- | --- | --- | --- | --- |
| ε-近邻图 | 适合局部结构分析，聚类可视化 | 依赖 ε 值调整 | 中等 | 结构松散，逐步连通 |
| k-近邻图 | 适合构建稳定邻接关系 | 依赖 k 值调整 | 适中 | 结构均衡，可控性强 |
| 全连通图 | 适用于小规模完整图分析 | O(n²) 计算量大 | 小规模 | 过度连接，信息冗余 |

小样本（20 点）：三种方法都能可视化，但全连通图仍显得冗余

中等样本（50 点）：k 近邻和 ε 近邻 能较好保持结构，全连通图计算成本较高

大规模（100 点）：k 近邻最优，ε 近邻需要调优，全连通图不可取