# 机器学习 第六次作业

# 题目 1. 列举四个图卷积神经网络解决的问题,并简述这些问题相应的训练数 据和优化目标是什么样的。

问题一: 社交网络机器人检测

### 训练数据:

一、基于图论的方法(通常是通过关注用户之间的社交关系以及行为相似性)而构建的社交网络图

# 优化目标:

- 一、针对图中的具有异常边的节点进行社交机器人的检测
- 二、找出那些由于社交机器人的不寻常的行为以及行为的相似性而形成的不符合常理 的特殊子图,实现对社交机器人进行联合检测

# 问题二:蛋白质分类

# 训练数据

- 一、不同蛋白质不同结果的数据
- 二、已知蛋白质结构的各种不同集合

# 优化目标

- 一、蛋白质结构预测
- 二、提高预测结构与真实结构的接近程度

# 问题三: 推荐系统

## 训练数据

1. 推荐系统中大部分的信息具有图结构,例如社交关系、知识图谱、user-item 交互 组成的二部图(bipartite graph)、序列中的 item 转移图,这些图都可以作为训练 数据

2.

# 优化目标

不同的场景往往有不同的目标:

- 1. 电商场景中,希望能够在优化 GMV 的基础上提高点击率,从而提高用户的粘性;
- 2. 在信息流场景中,希望提高用户点击率的基础上提高用户关注,点赞,评论等 行为,营造更好的社区氛围从而提高留存

### 问题四: 机器视觉

#### 训练数据

1. 根据不同的任务会有不同的数据集。例如要对猫狗分类,那么就需要有大量的猫狗的图片。

优化目标:

- 1. 分类, 亦即如猫、狗、茄子、香蕉、飞机、自行车等物体之检测、分类
- 2. 尽可能地将物体区分出来,判断准确

题目 2. 已知一个图的邻接矩阵为 A,设基于邻接矩阵的图嵌入给出的嵌入矩阵 为 Z,证明 $ZZ^T$ 是矩阵 A 的低阶秩近似。

图嵌入的优化目标为 $L = ||ZZ^T - A||_F^2$ 

而我们知道根据奇异值分解,任意一个矩阵M可以分解为r个简单矩阵(秩为 1)  $\sigma_i u_i v_i^T$ 之和,且有 $\sigma_i \geq \sigma_{i+1} \geq 0$  ( $i = 1, 2, \dots, r$ )。

这几个简单矩阵按照重要性排序来表示矩阵M,则有 $M=U\Sigma V^T=\sigma_1u_1v_1^T+\cdots+\sigma_ru_rv_r^T$ 。若我们用秩为k< r的矩阵N来最优近似矩阵M,记秩为 k 的矩阵N为 $N_k$ ,则 $N_k=\sum_{i=1}^k\sigma_iu_iv_i^T$ 。这就是矩阵低秩近似定理。

而在图嵌入的优化目标中 $L = ||ZZ^T - A||_F^2$ ,我们要求L的最小值,就是求矩阵 $ZZ^T$ 和邻接矩阵A之差范数的最小值。那么范数的一个目标就是为了度量矩阵的相似度。那么我们如何来求解两个矩阵之差范数的最小值呢?

令e表示单位向量,则我们定义矩阵M和N的相似度为,它们对于同一单位向量变换之后的向量之差范数即||M-N||=||Me-Ne||=||(M-N)e||。

根据奇异值分解 $M-N=U\Sigma V^T=\sum_{i=1}^r\sigma_iu_iv_i^T$ , 我们知道

$$(M-N)e = \sum_{i=1}^{r} (\sigma_i v_i^T e) u_i$$

由于 $u_i$ 正交,所以有

$$\begin{aligned} \left| |(M - N)e| \right| &= ||\sum_{i=1}^{r} (\sigma_i v_i^T e) u_i|| \\ &= \sqrt{\left(\sum_{i=1}^{r} (\sigma_i v_i^T e)^2 \le \sqrt{\left(\sum_{i=1}^{r} (\sigma_1 v_i^T e)^2 \le \sigma_1 ||e|| = \sigma_1}\right)^2} \\ &= \sigma_1 \sqrt{\left(\sum_{i=1}^{r} v_i^T e)^2 \le \sigma_1 ||e|| = \sigma_1} \end{aligned}$$

所以矩阵之差范数 $||M-N||=\sigma_1$ ,根据矩阵低秩相似定理, $||M-N_k||=\sigma_k$ ,而对于其他秩为 k 的矩阵N',均有 $||M-N'||\geq ||M-N_k||$ 。

所以当我们将上述过程应用到 $||ZZ^T - A||_F^2$ 的求解当中,我们便可以知道我们求解的矩阵 $ZZ^T$ 为矩阵A的低阶秩近似。但此处仍然有一个不同之处,在于 $ZZ^T$ 是一个对称矩阵。而在前述过程中都为提到M和N为对称矩阵。此处的原因在于,在给定一个无向图进

行图嵌入时,其邻接矩阵为一个对称矩阵,所以其低阶秩近似矩阵也为对称矩阵,且有若  $A = U\Sigma U^T = \sum_{i=1}^r \sigma_i u_i u_i^T$ ,则 $Z_i = \sqrt{\sigma_i} u_i$ 。

题目 3. 写出图神经网络消息传递公式,并指出公式中每个变量的含义,以及 那些变量是可学习变量;指出基础的图神经网络消息传递方法,Kipf 方法, GraphSAGE 方法的不同点。

图神经网络消息传递公式:

$$\begin{cases} h_{v}^{0} = [F]_{v} \\ h_{v}^{k} = \sigma \left( W_{k} \sum_{u \in N(v)} \frac{h_{u}^{k-1}}{|N(u)|} + B_{k} h_{v}^{k-1} \right), k = 1, 2, \dots, K \\ z_{k} = h_{v}^{k} \end{cases}$$

公式中变量的含义:

- 1.  $h_{\nu}^{k}$ 表示节点 $\nu$ 在时刻k的特征表示
- 2. K表示消息传递需要进行K个时间步
- 3. F表示特征矩阵, $[F]_v$ 表示特征矩阵中对应于节点v中的那一行(列)
- 4. N(v)表示节点v的邻居节点集合
- 5. |N(v)|表示节点v的邻居集合的大小
- 6.  $W_k, B_k$ 需要学习的参数,用于权衡从邻居节点学到的信息和上一状态中的信息对当前要计算的状态表示的影响
- 7.  $z_k$  节点在消息传递结束后的最终表示
- 8. σ消息传递函数

可学习变量:  $W_k$ ,  $B_k$ 

基础方法:使用 $W_k$ ,  $B_k$ 权衡从邻居节点学到的信息和上一状态中的信息对当前要计算的状态表示的影响。

Kipf 方法: 在图中执行参数化的消息传递操作,被建模为谱图卷积的一阶近似值。将从邻居节点学到的信息和上一状态中的信息对当前状态的影响一视同仁,只学习参数  $W_k$ 。

GraphSAGE 方法: GraphSAGE 同时利用节点特征信息和结构信息得到 Graph Embedding 的映射,相比之前的方法,之前都是保存了映射后的结果,而 GraphSAGE 保存了生成 embedding 的映射,可扩展性更强,对于节点分类和链接预测问题的表现也比较突出。GraphSAGE 是一个 inductive(归纳式)框架,能够高效地利用节点的属性信息对新节点生成 embedding