

图神经网络导论

图异常检测

周晟

浙江大学 软件学院

2022.12



课程内容

- ① 异常检测任务
- ② 基于分类的异常检测
- ③ 基于自编码器的异常检测
- ④ 基于对比学习的图异常检测
- ⑤ 其他异常检测方法



数据中的异常

异常的定义

An **outlier** is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism.

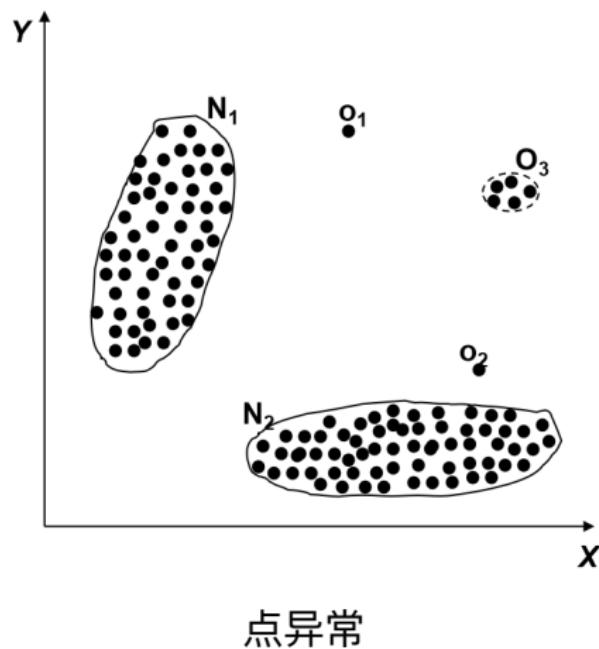
数据集中与大部分样本不一样的样本称为异常样本 (anomaly sample, abnormalities, deviants, outliers)。

异常检测的定义

离群点检测 (又称为异常检测) 是找出行为不同于预期的异常的过程。

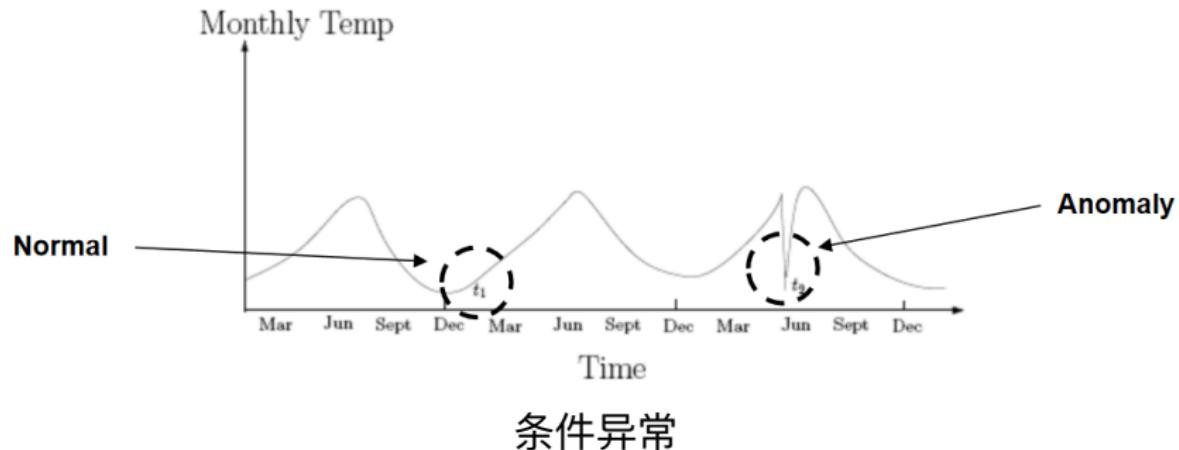
点异常

- 单个数据实例是异常的



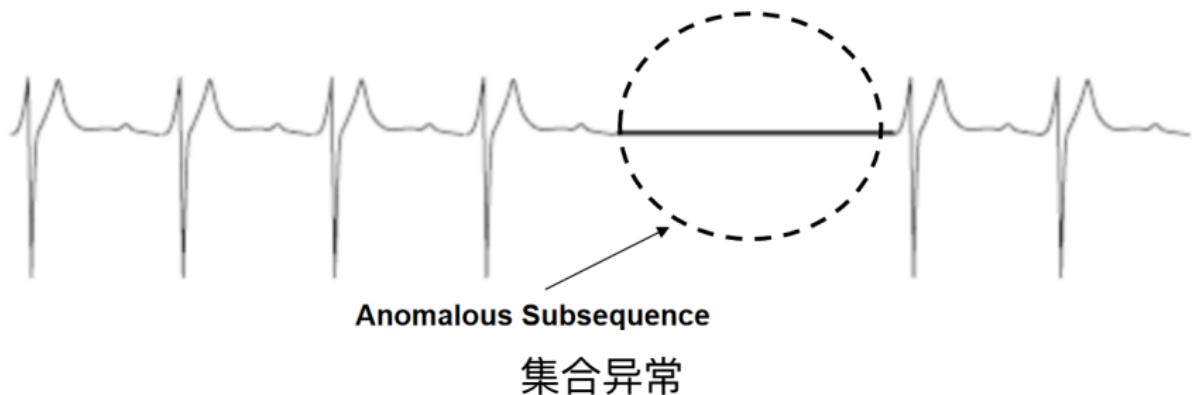
条件异常

- 单个数据实例在某个“条件下”是异常的
- 也称为“上下文异常 (Contextual Anomalies) ”

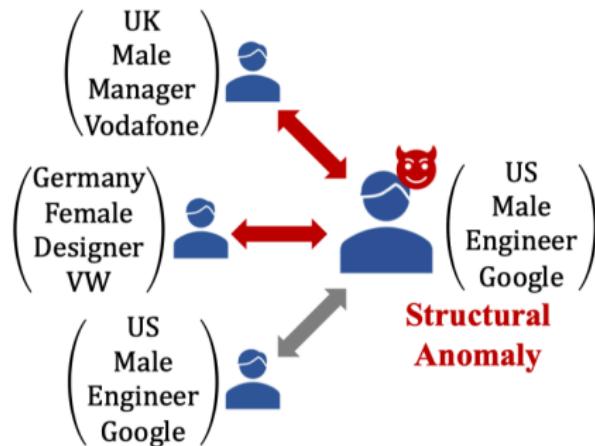


集合异常

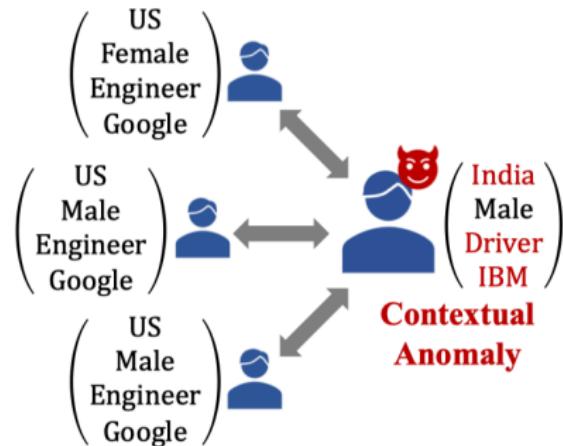
- 一个由相关数据实例构成的集合是异常的
- 数据实例间存在某种关系
 - 连续数据
 - 空间数据
 - 图数据
- 集合中的单个数据实例本身并不是异常



图上的异常



(a) Structural Anomaly



(b) Contextual Anomaly

图上的异常类型

不同异常类型的案例

- ① 超大金额的单笔交易（点异常）
- ② 发生在国外的小额交易（条件异常）
- ③ 发生在凌晨的小额交易（条件异常）
- ④ 连续十天购买同一件商品（集合异常）



真实大规模场景中往往同时包含多种类型的异常！



异常检测范式

① 有监督异常检测

- ① 标签获得困难
- ② 类别不平衡问题

② 半监督异常检测

- ① 数据分布偏移 (Distribution Shift)

③ 弱监督异常检测

- ① 异常标签噪声大
- ② 训练数据分布偏移 (Distribution Shift)

④ 无监督异常检测

- ① 缺少监督信号和数据分布信息
- ② 训练目标设计困难

Out-of-distribution (OOD) Detection



① 异常检测任务

② 基于分类的异常检测

③ 基于自编码器的异常检测

④ 基于对比学习的图异常检测

⑤ 其他异常检测方法



单分类问题

常见的分类方法：

- ① 单分类（只能定义正样本不能定义负样本）
- ② 二分类（邮件分类）
- ③ 多分类（图像分类）

单分类与二分类的区别

单分类问题中的训练样本只有一类，因此训练出的分类器将不属于该类的所有其他样本判别为“不是”即可，而不是由于属于另一类才返回“不是”的结果。

模型假设

寻找一个超平面将样本中的正例圈出来，预测就是用这个超平面做决策，在圈内的样本就认为是正样本。

- 无监督学习的方法，不需要训练集的标签。
- 如何在无监督场景下寻找划分的超平面以及寻找支持向量？
- SVDD(support vector domain description) 支持向量域描述

One Class SVM

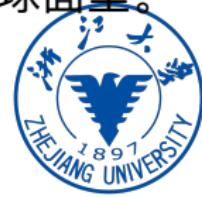
SVDD 的优化目标就是，求一个中心为 a , 半径为 R 的最小球面：

$$F(R, a, \xi_i) = R^2 + C \sum_i \xi_i$$

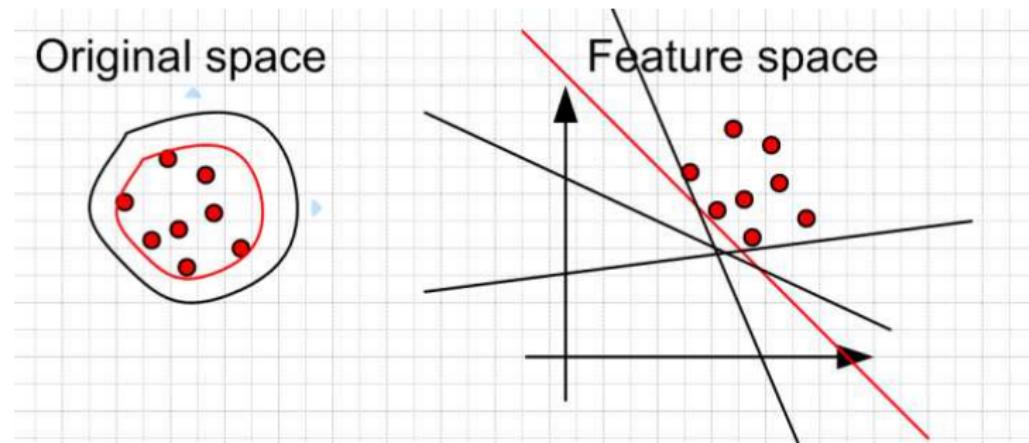
ξ_i 是球体的松弛变量，使得这个球面满足：

$$(x_i - a)^T (x_i - a) \leq R^2 + \xi_i \quad \forall i, \xi_i \geq 0$$

满足这个条件就是说要把 training set 中的数据点都包在球面里。



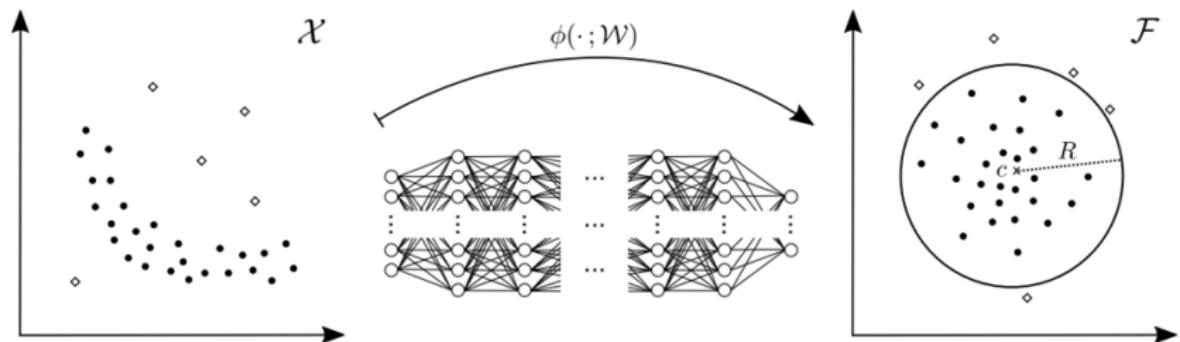
One Class SVM



寻找超平面与特征空间中的零点距离最大，并且将零点与所有的数据点分隔开。

Deep One-Class Classification(DeepSVDD)

采用深度学习的方法来实现传统的 One-Class SVM 算法



DeepSVDD[Ruff et al., 2018]¹



¹Deep One-Class Classification(ICML2018)

DeepSVDD

将样本学习到表征空间，并且学习出一个尽可能小的超平面来包裹住所有的训练样本

$$\min_{R, \mathcal{W}} R^2 + \frac{1}{\nu n} \sum_{i=1}^n \max \left\{ 0, \|\phi(\mathbf{x}_i; \mathcal{W}) - \mathbf{c}\|^2 - R^2 \right\} + \frac{\lambda}{2} \sum_{\ell=1}^L \|\mathbf{W}^\ell\|_F^2$$

- ① 第一项为超平面的半径
- ② 第二项对超平面外的点进行惩罚
- ③ 第三项利用正则化防止模型坍塌

这里允许训练样本中包含少量的异常样本， ν 控制异常比例

DeepSVDD

在实际场景中，训练集可以只包含正常样本，因此直接优化所有点到中心的距离

$$\min_{\mathcal{W}} \frac{1}{n} \sum_{i=1}^n \|\phi(\mathbf{x}_i; \mathcal{W}) - \mathbf{c}\|^2 + \frac{\lambda}{2} \sum_{\ell=1}^L \|\mathbf{W}^\ell\|_F^2$$

中心 c 的选择通常采用经验方法，对表征学习器随机初始化参数，取所有样本输出的均值作为中心，中心固定不变

异常分数：

$$s(\mathbf{x}) = \|\phi(\mathbf{x}; \mathcal{W}^*) - \mathbf{c}\|^2$$

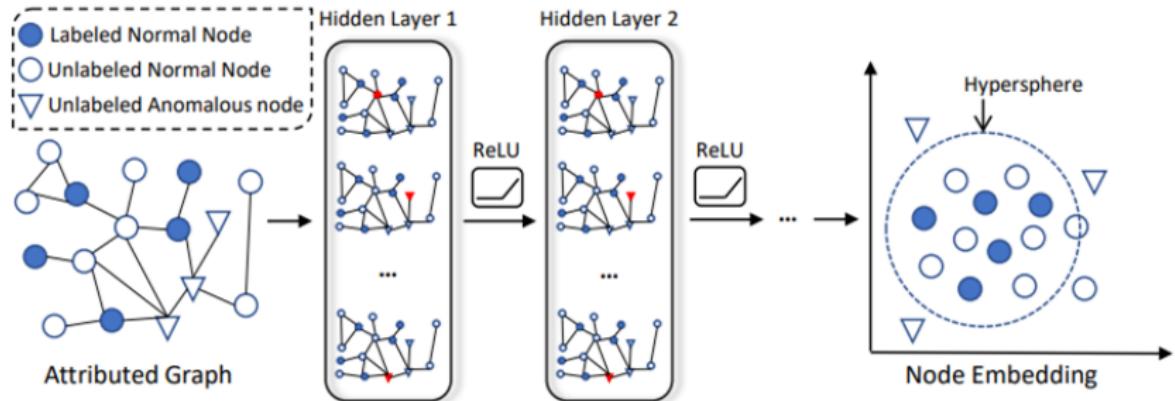


DeepSVDD

Table 1. Average AUCs in % with StdDevs (over 10 seeds) per method and one-class experiment on MNIST and CIFAR-10.

NORMAL CLASS	OC-SVM/ SVDD	KDE	IF	DCAE	ANOGAN	SOFT-BOUND. DEEP SVDD	ONE-CLASS DEEP SVDD
0	98.6 \pm 0.0	97.1 \pm 0.0	98.0 \pm 0.3	97.6 \pm 0.7	96.6 \pm 1.3	97.8 \pm 0.7	98.0 \pm 0.7
1	99.5 \pm 0.0	98.9 \pm 0.0	97.3 \pm 0.4	98.3 \pm 0.6	99.2 \pm 0.6	99.6 \pm 0.1	99.7 \pm 0.1
2	82.5 \pm 0.1	79.0 \pm 0.0	88.6 \pm 0.5	85.4 \pm 2.4	85.0 \pm 2.9	89.5 \pm 1.2	91.7 \pm 0.8
3	88.1 \pm 0.0	86.2 \pm 0.0	89.9 \pm 0.4	86.7 \pm 0.9	88.7 \pm 2.1	90.3 \pm 2.1	91.9 \pm 1.5
4	94.9 \pm 0.0	87.9 \pm 0.0	92.7 \pm 0.6	86.5 \pm 2.0	89.4 \pm 1.3	93.8 \pm 1.5	94.9 \pm 0.8
5	77.1 \pm 0.0	73.8 \pm 0.0	85.5 \pm 0.8	78.2 \pm 2.7	88.3 \pm 2.9	85.8 \pm 2.5	88.5 \pm 0.9
6	96.5 \pm 0.0	87.6 \pm 0.0	95.6 \pm 0.3	94.6 \pm 0.5	94.7 \pm 2.7	98.0 \pm 0.4	98.3 \pm 0.5
7	93.7 \pm 0.0	91.4 \pm 0.0	92.0 \pm 0.4	92.3 \pm 1.0	93.5 \pm 1.8	92.7 \pm 1.4	94.6 \pm 0.9
8	88.9 \pm 0.0	79.2 \pm 0.0	89.9 \pm 0.4	86.5 \pm 1.6	84.9 \pm 2.1	92.9 \pm 1.4	93.9 \pm 1.6
9	93.1 \pm 0.0	88.2 \pm 0.0	93.5 \pm 0.3	90.4 \pm 1.8	92.4 \pm 1.1	94.9 \pm 0.6	96.5 \pm 0.3
AIRPLANE	61.6 \pm 0.9	61.2 \pm 0.0	60.1 \pm 0.7	59.1 \pm 5.1	67.1 \pm 2.5	61.7 \pm 4.2	61.7 \pm 4.1
AUTOMOBILE	63.8 \pm 0.6	64.0 \pm 0.0	50.8 \pm 0.6	57.4 \pm 2.9	54.7 \pm 3.4	64.8 \pm 1.4	65.9 \pm 2.1
BIRD	50.0 \pm 0.5	50.1 \pm 0.0	49.2 \pm 0.4	48.9 \pm 2.4	52.9 \pm 3.0	49.5 \pm 1.4	50.8 \pm 0.8
CAT	55.9 \pm 1.3	56.4 \pm 0.0	55.1 \pm 0.4	58.4 \pm 1.2	54.5 \pm 1.9	56.0 \pm 1.1	59.1 \pm 1.4
DEER	66.0 \pm 0.7	66.2 \pm 0.0	49.8 \pm 0.4	54.0 \pm 1.3	65.1 \pm 3.2	59.1 \pm 1.1	60.9 \pm 1.1
DOG	62.4 \pm 0.8	62.4 \pm 0.0	58.5 \pm 0.4	62.2 \pm 1.8	60.3 \pm 2.6	62.1 \pm 2.4	65.7 \pm 2.5
FROG	74.7 \pm 0.3	74.9 \pm 0.0	42.9 \pm 0.6	51.2 \pm 5.2	58.5 \pm 1.4	67.8 \pm 2.4	67.7 \pm 2.6
HORSE	62.6 \pm 0.6	62.6 \pm 0.0	55.1 \pm 0.7	58.6 \pm 2.9	62.5 \pm 0.8	65.2 \pm 1.0	67.3 \pm 0.9
SHIP	74.9 \pm 0.4	75.1 \pm 0.0	74.2 \pm 0.6	76.8 \pm 1.4	75.8 \pm 4.1	75.6 \pm 1.7	75.9 \pm 1.2
TRUCK	75.9 \pm 0.3	76.0 \pm 0.0	58.9 \pm 0.7	67.3 \pm 3.0	66.5 \pm 2.8	71.0 \pm 1.1	73.1 \pm 1.2

One-Class Graph Neural Networks for Anomaly Detection in Attributed Networks(OCGNN)



OCGNN²直接将 DeepSVDD 的表征学习器替换为 GNN

²One-Class Graph Neural Networks for Anomaly Detection in Attributed Networks(Neural Comput & Applic 2021)

OCGNN

	Method	Cora	Citeseer	Pubmed
Raw Features	IForest	53.09 ± 0.03	46.33 ± 0.03	65.57 ± 0.02
	OCSVM	54.35 ± 0.02	57.05 ± 0.03	45.50 ± 0.01
	PCA	62.17 ± 0.01	58.10 ± 0.03	71.06 ± 0.01
	AE	62.17 ± 0.01	58.11 ± 0.03	71.05 ± 0.01
DeepWalk	IForest	57.87 ± 0.02	51.00 ± 0.03	60.73 ± 0.01
	OCSVM	52.10 ± 0.03	43.13 ± 0.02	60.22 ± 0.01
	PCA	55.90 ± 0.03	46.65 ± 0.02	61.66 ± 0.01
	AE	55.91 ± 0.03	46.42 ± 0.02	61.66 ± 0.01
DeepWalk+Raw Feat.	IForest	53.56 ± 0.04	45.55 ± 0.06	65.60 ± 0.02
	OCSVM	51.59 ± 0.03	42.95 ± 0.02	60.10 ± 0.01
	PCA	62.38 ± 0.02	57.96 ± 0.03	72.04 ± 0.01
	AE	62.39 ± 0.02	57.96 ± 0.03	71.91 ± 0.01
GAE based	GCN-AE	80.53 ± 0.05	59.52 ± 0.09	58.26 ± 0.02
	GAE [11]	60.15 ± 0.08	51.80 ± 0.03	54.27 ± 0.02
	Dom [7]	67.50 ± 0.25	62.44 ± 0.15	53.92 ± 0.04
Our OCGNNs	OC-GCN	73.25 ± 0.02	62.81 ± 0.01	54.53 ± 0.01
	OC-GAT	88.19 ± 0.02	79.06 ± 0.03	60.98 ± 0.01
	OC-SAGE	86.97 ± 0.04	85.62 ± 0.01	74.72 ± 0.03

OCGNN 实验效果



Subtractive Aggregation for Attributed Network Anomaly Detection(CIKM2021)

研究动机

图上的异常节点往往呈现出与邻居不同的模式，这种局部的便宜可以被用于发现异常

使用节点的表征与邻居表征的差值作为节点的特征

$$\begin{aligned} z_i &= \phi(\mathbf{x}_i; \mathbf{W}), \\ \mathbf{h}_i &= \sigma(z_i - \text{AGGREGATE}(z_j, \forall j \in \mathcal{N}_i^k)) \end{aligned}$$

进而使用 DeepSVDD 学习超平面：

$$\min_{\Theta} \frac{1}{n} \sum_{i=1}^n \|\mathbf{h}_i - \mathbf{c}\|_2^2 + \frac{\lambda}{2} \|\Theta\|_F^2$$



两种具体的聚合邻居方式

Mean aggregator:

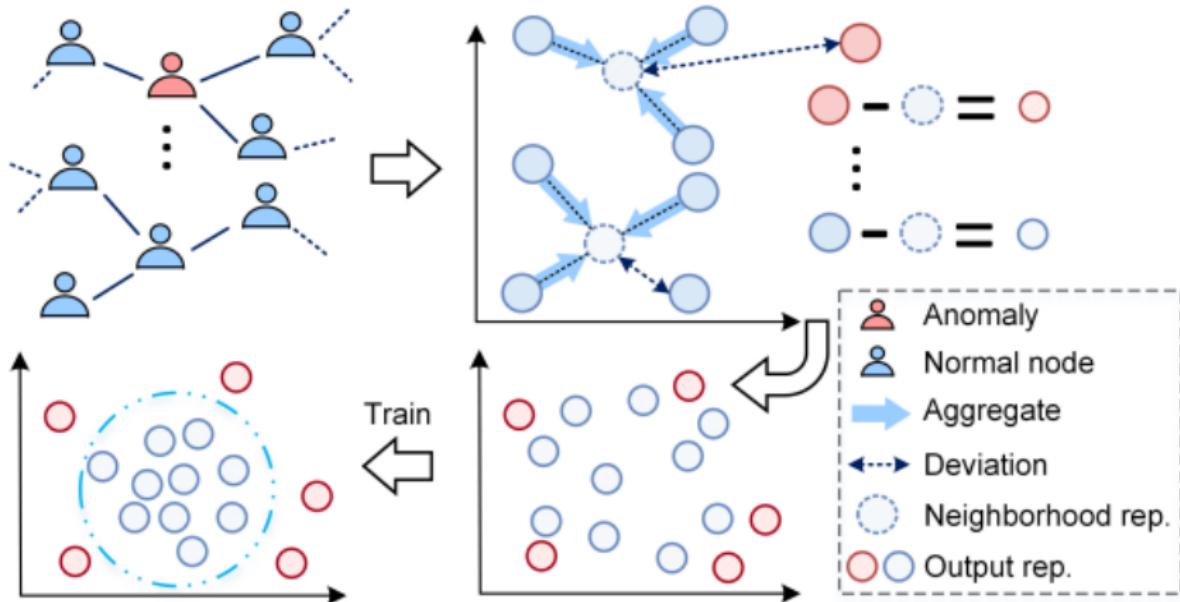
$$\mathbf{h}_i = \sigma \left(\mathbf{z}_i - \frac{1}{|\mathcal{N}_i^k|} \sum_{j \in \mathcal{N}_i^k} \mathbf{z}_j \right)$$

Attention aggregator:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^T [\mathbf{z}_i \oplus \mathbf{z}_j]))}{\sum_{j \in \mathcal{N}_i^k} \exp(\text{LeakyReLU}(\mathbf{a}^T [\mathbf{z}_i \oplus \mathbf{z}_j]))}$$
$$\mathbf{h}_i = \sigma \left(\mathbf{z}_i - \sum_{j \in \mathcal{N}_i^k} \alpha_{ij} \mathbf{z}_j \right)$$



AAGNN



- 1 异常检测任务
- 2 基于分类的异常检测
- 3 基于自编码器的异常检测
- 4 基于对比学习的图异常检测
- 5 其他异常检测方法



基于自编码器的异常检测

基本假设

自编码器的目标是最小化所有数据的重构损失，按照少数服从多数的原则，正常样本可以被更好地重构而异常样本则难以被完美重构。

模型结构

$$\begin{aligned}\mathbf{z} &= \phi_e(\mathbf{x}; \Theta_e), \hat{\mathbf{x}} = \phi_d(\mathbf{z}; \Theta_d) \\ \{\Theta_e^*, \Theta_d^*\} &= \arg \min_{\Theta_e, \Theta_d} \sum_{\mathbf{x} \in X} \|\mathbf{x} - \phi_d(\phi_e(\mathbf{x}; \Theta_e); \Theta_d)\|^2 \\ s_{\mathbf{x}} &= \|\mathbf{x} - \phi_d(\phi_e(\mathbf{x}; \Theta_e^*); \Theta_d^*)\|^2\end{aligned}$$



基于自编码器的异常检测

常用的 AutoEncoder 结构：

- ① Denoising AutoEncoder
- ② Sparse AutoEncoder
- ③ Contractive AutoEncoder
- ④ Variational AutoEncoder
- ⑤ Robust AutoEncoder
- ⑥ Masked AutoEncoder



Outlier Detection with Robust Deep AutoEncoders(KDD 2017)

研究动机

- ① 自编码器在训练过程中容易受到异常样本的影响
- ② Robust Principal Component Analysis (RPCA) 也是一种降维方法，但是为异常样本做了专门的优化

RPCA 将数据矩阵 X 拆分为低秩的矩阵 L 和一个稀疏矩阵 S :

$$X = L + S$$

矩阵分解的过程可以理解为如下的优化目标:

$$\begin{aligned} & \min_{L,S} \|L\|_* + \lambda \|S\|_1 \\ \text{s.t. } & \|X - L - S\|_F^2 = 0 \end{aligned}$$

Robust Deep AutoEncoders(RDA)

$$X = L_D + S$$

L_D 是指能被 AutoEncoder 重构的特征， S 包含了难以被 AutoEncoder 重构的噪声和异常。

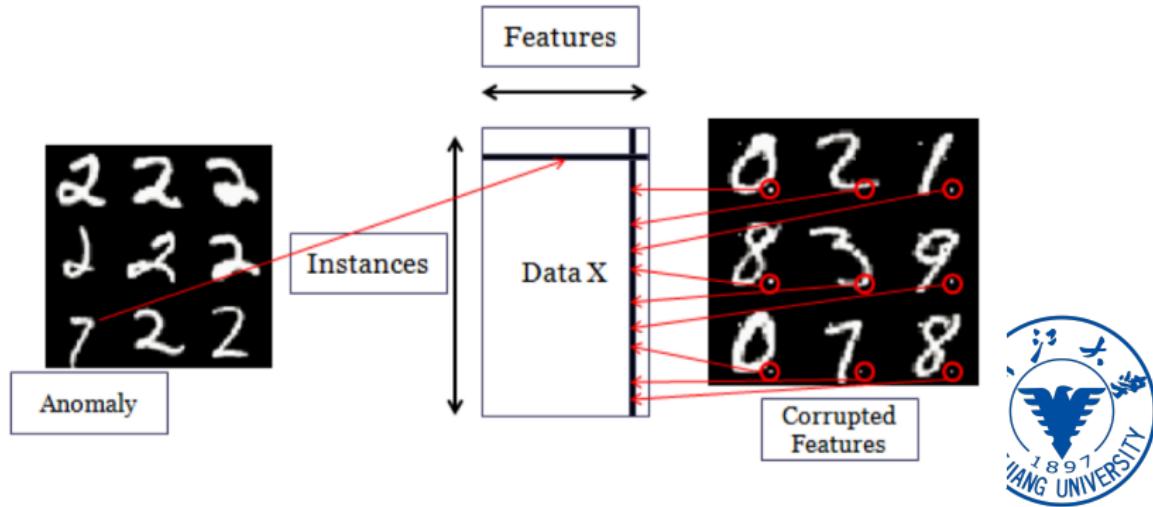
$$\begin{aligned} \min_{\theta} & \|L_D - D_{\theta}(E_{\theta}(L_D))\|_2 + \lambda \|S\|_1 \\ \text{s.t. } & X = L_D + S = 0 \end{aligned}$$



Robust Deep AutoEncoders(RDA)

Group Anomalies

- ① 许多样本共享一个相同的特征维度（系统噪声而不是异常）
- ② 一个样本中异常的特征应当相对确定



Robust Deep AutoEncoders(RDA)

$\mathcal{L}_{2,1}$ norm

\mathcal{L}_2 norm 作用于所有特征维度, \mathcal{L}_1 norm 作用于所有的样本

$$\|X\|_{2,1} = \sum_{j=1}^n \|x_j\|_2 = \sum_{j=1}^n \left(\sum_{i=1}^m |x_{ij}|^2 \right)^{1/2}$$

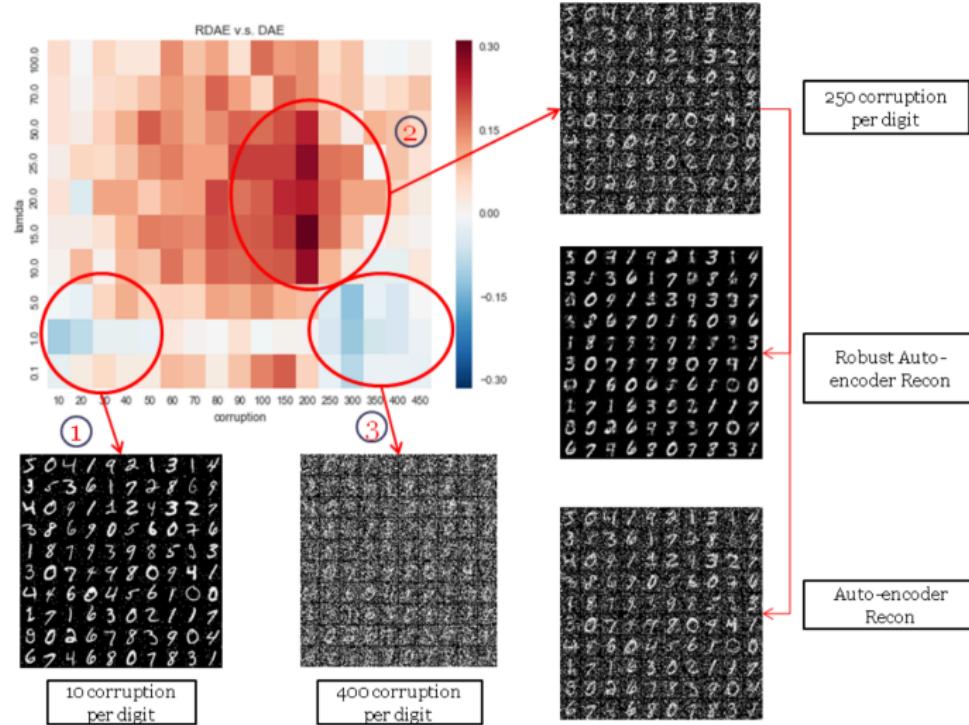
$$\min_{\theta, S} \|L_D - D_\theta(E_\theta(L_D))\|_2 + \lambda \|S\|_{2,1}$$

$$\min_{\theta, S} \|L_D - D_\theta(E_\theta(L_D))\|_2 + \lambda \|S^T\|_{2,1}$$

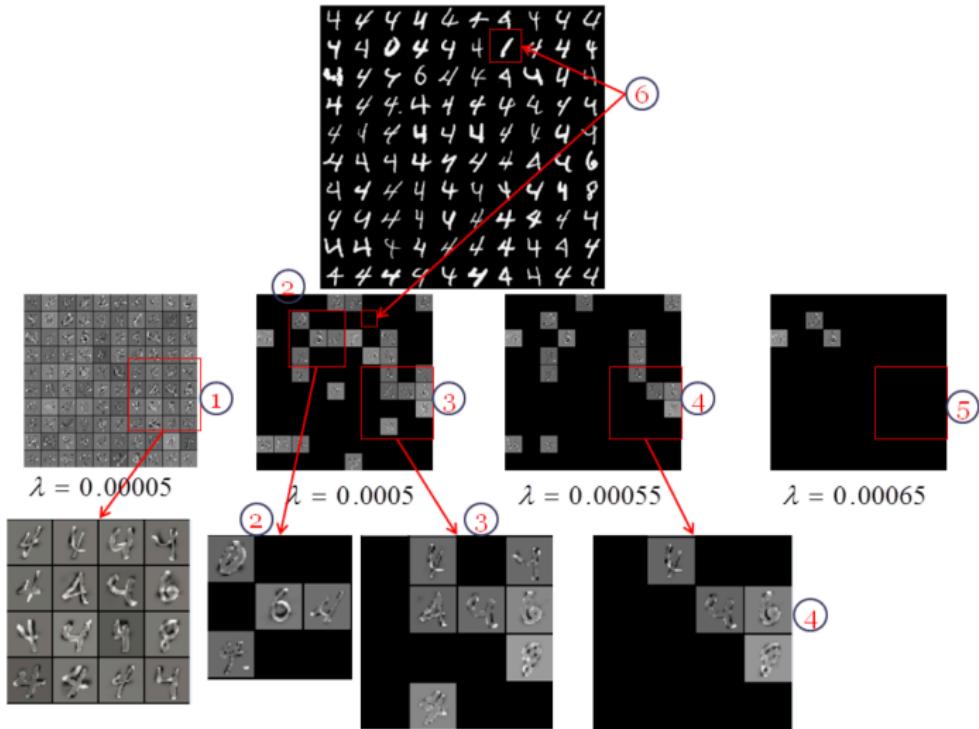
$$\text{s.t. } X - L_D - S = 0$$



Robust Deep AutoEncoders(RDA)



Robust Deep AutoEncoders(RDA)



Deep Anomaly Detection on Attributed Networks(Dominant)

研究动机

使用图神经网络对结构和属性进行统一编码，进而提升异常捕获的能力。使用重构损失作为异常分值。

核心模块³:

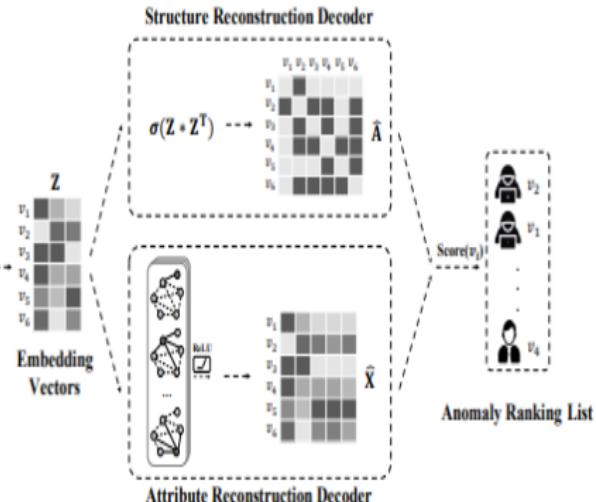
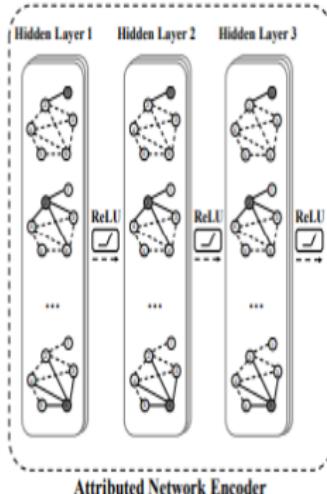
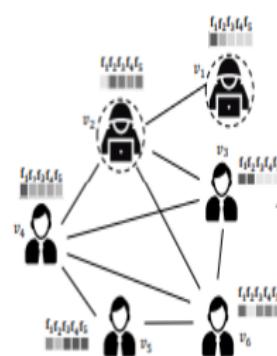
$$\hat{\mathbf{A}} = \text{sigmoid}(\mathbf{Z}\mathbf{Z}^T)$$

$$\hat{\mathbf{X}} = f_{\text{Relu}}(\mathbf{Z}, \mathbf{A} \mid \mathbf{W}^{(3)})$$

$$\begin{aligned}\mathcal{L} &= (1 - \alpha)\mathbf{R}_S + \alpha\mathbf{R}_A \\ &= (1 - \alpha)\|\mathbf{A} - \hat{\mathbf{A}}\|_F^2 + \alpha\|\mathbf{X} - \hat{\mathbf{X}}\|_F^2,\end{aligned}$$

³Deep Anomaly Detection on Attributed Networks(SDM 2019)

Dominant



Dominant 模型架构

DOMINANT

Precision@K												
	BlogCatalog				Flickr				ACM			
K	50	100	200	300	50	100	200	300	50	100	200	300
LOF	0.300	0.220	0.180	0.183	0.420	0.380	0.270	0.237	0.060	0.060	0.045	0.037
Radar	0.660	0.670	0.550	0.416	0.740	0.700	0.635	0.503	0.560	0.580	0.520	0.430
ANOMALOUS	0.640	0.650	0.515	0.417	0.790	0.710	0.650	0.510	0.600	0.570	0.510	0.410
DOMINANT	0.760	0.710	0.590	0.470	0.770	0.730	0.685	0.593	0.620	0.590	0.540	0.497
Recall@K												
	BlogCatalog				Flickr				ACM			
K	50	100	200	300	50	100	200	300	50	100	200	300
LOF	0.050	0.073	0.120	0.183	0.047	0.084	0.120	0.158	0.005	0.010	0.015	0.018
Radar	0.110	0.223	0.367	0.416	0.082	0.156	0.282	0.336	0.047	0.097	0.173	0.215
ANOMALOUS	0.107	0.217	0.343	0.417	0.087	0.158	0.289	0.340	0.050	0.095	0.170	0.205
DOMINANT	0.127	0.237	0.393	0.470	0.084	0.162	0.304	0.396	0.052	0.098	0.180	0.248

Table 2: Performance of different anomaly detection methods w.r.t. precision@ K and recall@ K .

Dominant 实验结果

AnomalyDAE: Dual autoencoder for anomaly detection on attributed networks

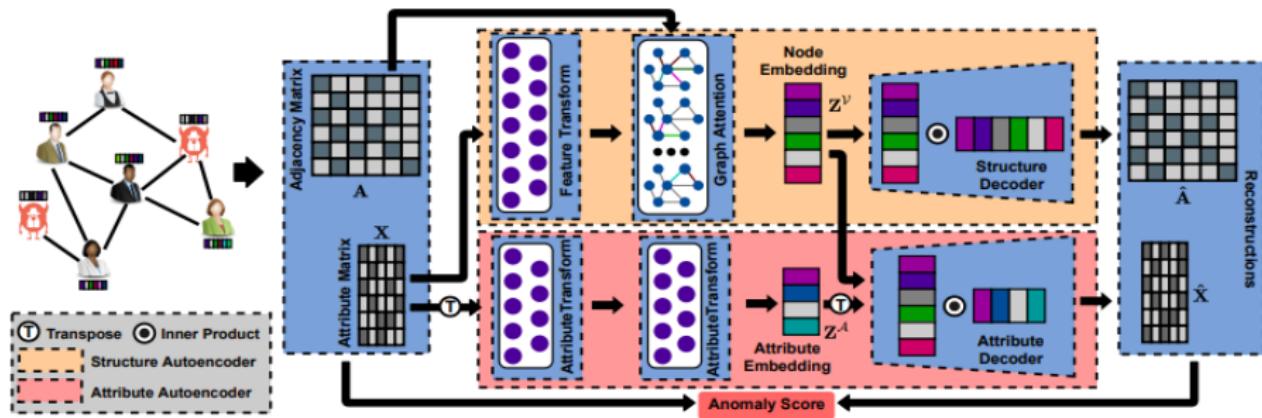


Fig. 1. The framework of the proposed AnomalyDAE.

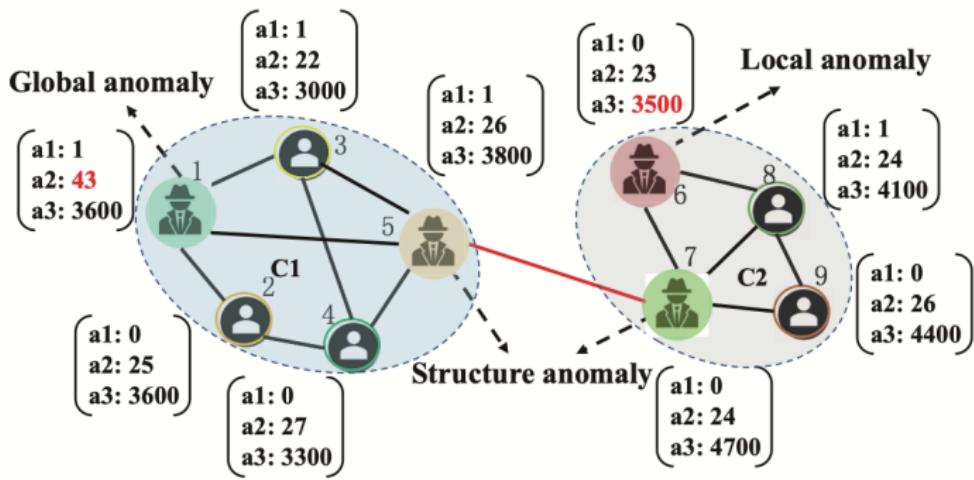
AnomalyDAE⁴与 Dominant 的区别在于使用了两个编码器

⁴AnomalyDAE: Dual autoencoder for anomaly detection on attributed networks(ICASSP)

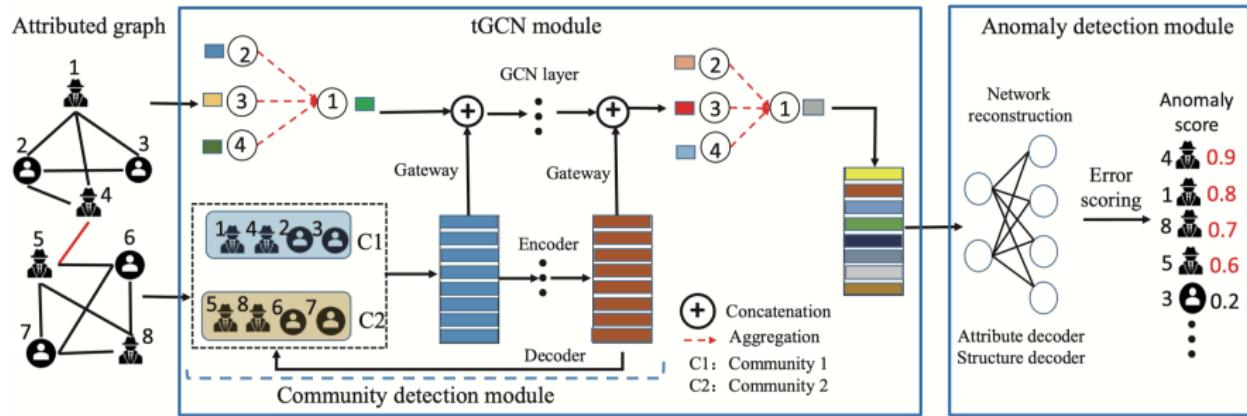
ComGA: Community-Aware Attributed Graph Anomaly Detection(WSDM 2022)

研究动机

真实网络中的异常往往包含多种粒度的异常，而不同异常的表现形式往往不同。



ComGA



社区发现：基于自编码器的模块度分解

$$L_{res} = \|\mathbf{B} - \hat{\mathbf{B}}\|_F^2 = \sum_{i=1}^n \left\| \mathbf{b}_i - \hat{\mathbf{b}}_i \right\|_2^2$$

社区信息与局部信息的融合：

$$\tilde{\mathbf{Z}}_{l-1} = \mathbf{Z}_{l-1} + \mathbf{H}_{l-1}$$

社区信息与局部信息的约束：

$$L_{gui} = KL(\mathbf{Z} || \mathbf{H})$$

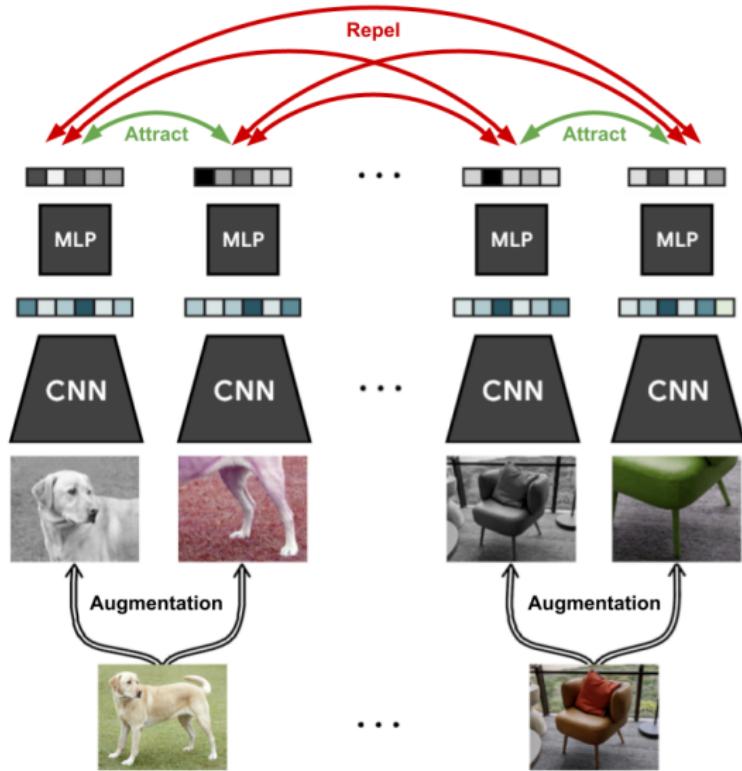
整体优化目标：

$$L_{rec} = (1 - \alpha) \|\mathbf{A} - \hat{\mathbf{A}}\|_F^2 + \alpha \|\mathbf{X} - \hat{\mathbf{X}}\|_F^2.$$

$$L = L_{res} + L_{gui} + L_{rec}$$

- 1 异常检测任务
- 2 基于分类的异常检测
- 3 基于自编码器的异常检测
- 4 基于对比学习的图异常检测
- 5 其他异常检测方法

对比学习



基于对比学习的异常检测

基本假设

对比学习是一个利用分类模型捕获样本的变换不变性的表征学习方法。

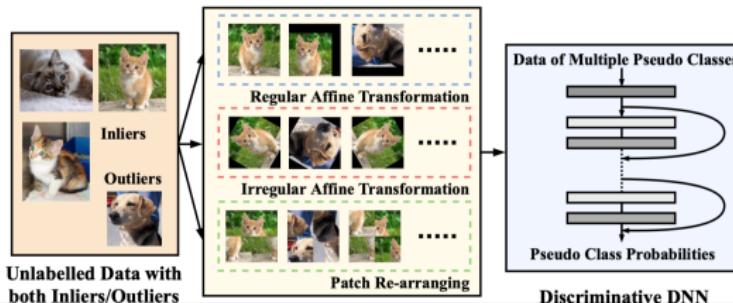
基本思路：

- ① 分类模型的置信度（正常样本可以被准确分类，异常样本无法准确分类）
- ② 变换不变的能力（正常样本多种变换间差异不大，异常样本多种变换间差异较大）

Effective End-to-end Unsupervised Outlier Detection via Inlier Priority of Discriminative Network(NIPS2019)

数据增强

- ① 旋转
- ② 翻折
- ③ 平移
- ④ 区块打乱



表征学习对正常样本的偏好

Motivation

- ① 类别不平衡状态下，有监督训练会在训练过程中偏向于捕获规模较大的类的信息
- ② 正常样本在训练过程中会提供更强的梯度方向指引（模型优化方向），且与异常样本有显著差异

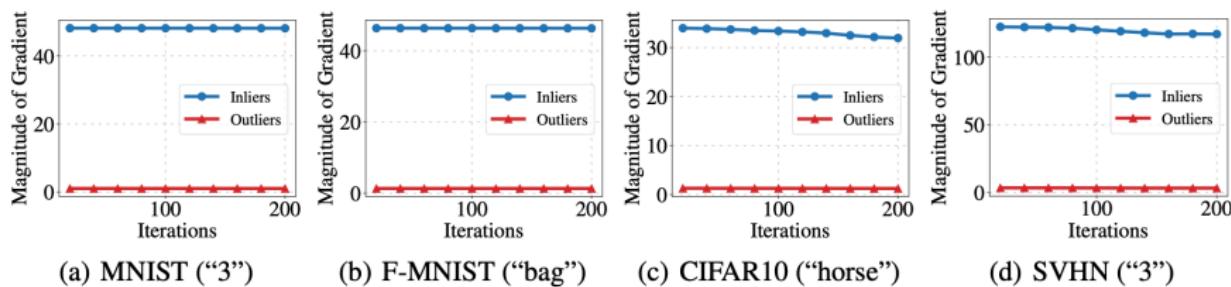


Figure 2: Inliers and outliers' gradient magnitude on example cases of benchmark datasets during SSD training. The class used as inliers is in brackets.

训练过程中的梯度对比

异常分值

Pseudo Label based Score (PL):

$$S_{pl}(\mathbf{x}) = \frac{1}{K} \sum_{y=1}^K P^{(y)} (\mathbf{x}^{(y)} | \boldsymbol{\theta})$$

Maximum Probability based Score (MP):

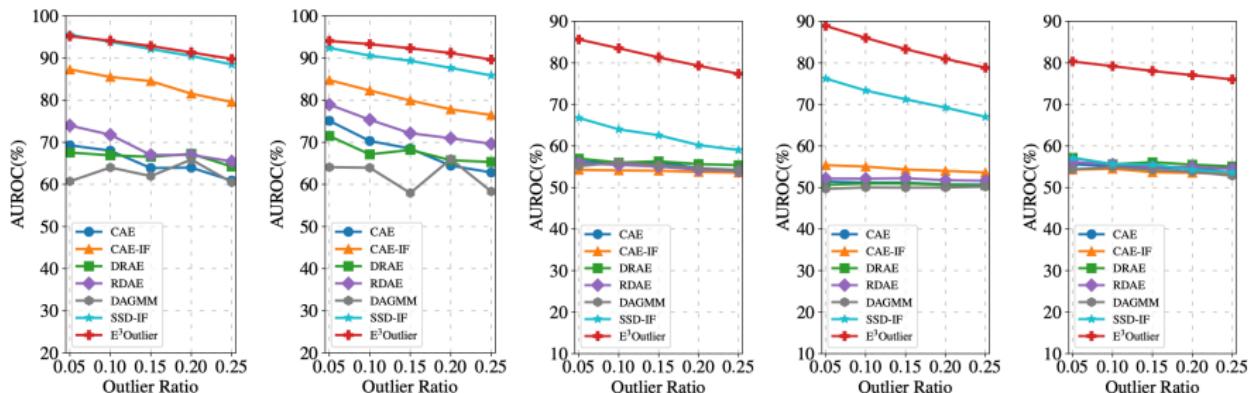
$$S_{mp}(\mathbf{x}) = \frac{1}{K} \sum_{y=1}^K \max_t P^{(t)} (\mathbf{x}^{(y)} | \boldsymbol{\theta})$$

Negative Entropy based Score (NE)

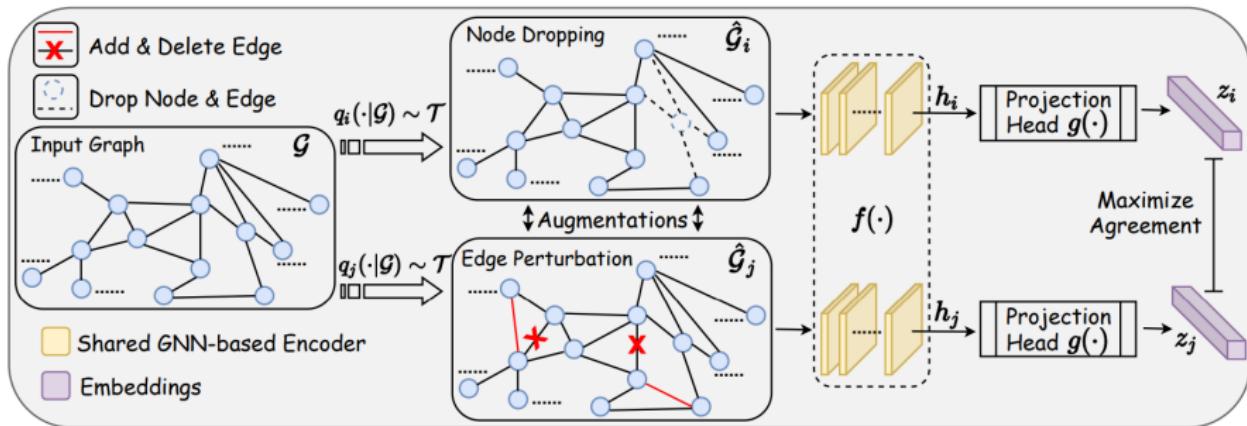
$$S_{ne}(\mathbf{x}) = \frac{1}{K} \sum_{y=1}^K \sum_{t=1}^K P^{(t)} (\mathbf{x}^{(y)} | \boldsymbol{\theta}) \log (P^{(t)} (\mathbf{x}^{(y)} | \boldsymbol{\theta}))$$

E^3 Outlier

Dataset	ρ	CAE	CAE-IF	DRAE	RDAE	DAGMM	SSD-IF	E^3 Outlier
MNIST	10%	68.0/92.0/32.9	85.5/97.8/49.0	66.9/93.0/30.5	71.8/93.1/35.8	64.0/92.9/26.6	93.8/99.2/ 68.7	94.1/99.3/67.5
	20%	64.0/82.7/40.7	81.5/93.6/57.2	67.2/86.6/42.5	67.0/84.2/43.2	65.9/86.4/41.3	90.5/97.3/71.0	91.3/97.6/72.3
F-MNIST	10%	70.3/94.3/29.3	82.3/97.2/40.3	67.1/93.9/25.5	75.3/95.8/31.7	64.0/92.7/30.3	90.6/98.5/68.6	93.3/99.0/75.9
	20%	64.4/85.3/36.8	77.8/92.2/49.0	65.7/86.9/36.6	70.9/89.2/41.4	66.0/86.7/43.5	87.6/95.6/71.4	91.2/97.1/78.9
CIFAR10	10%	55.9/91.0/14.4	54.1/90.2/13.7	56.0/90.7/14.7	55.4/90.7/14.0	56.1/91.3/15.6	64.0/93.5/18.3	83.5/97.5/43.4
	20%	54.7/81.6/25.5	53.8/80.7/25.3	55.6/81.7/26.8	54.2/81.0/25.7	54.7/81.8/26.3	60.2/85.0/28.3	79.3/93.1/52.7
SVHN	10%	51.2/90.3/10.6	55.0/91.4/11.9	51.0/90.3/10.5	52.1/90.6/10.8	50.0/90.0/19.3	73.4/95.9/22.0	86.0/98.0/36.7
	20%	50.7/80.2/20.7	54.0/82.0/22.4	50.6/80.4/20.5	51.8/80.9/21.1	50.0/79.9/29.6	69.2/89.5/33.7	81.0/93.4/47.0
CIFAR100	10%	55.2/91.0/14.5	54.5/90.7/13.8	55.6/90.9/15.0	55.8/90.9/15.0	54.9/91.1/14.2	55.6/91.5/13.0	79.2/96.8/33.3
	20%	54.4/81.7/25.6	53.5/80.9/25.1	55.5/81.8/27.0	54.9/81.5/26.5	53.8/81.5/24.7	54.3/82.1/23.4	77.0/92.4/46.5



图对比学习



图对比学习沿用经典对比学习架构，
并在数据增强和表征学习器上进行针对性设计。

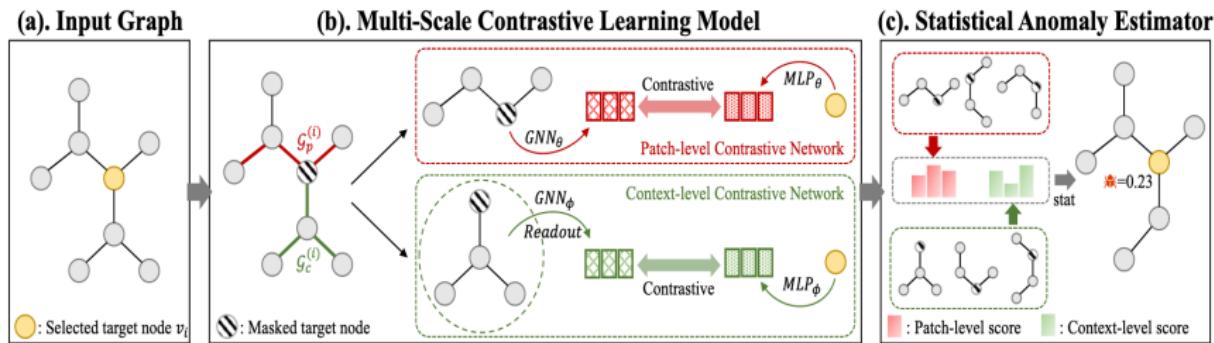
核心思路：

- ① 正负样本对对比作为异常分数
- ② 节点与邻居的对比作为异常分数

ANEMONE: Graph Anomaly Detection with Multi-Scale Contrastive Learning

研究动机

现有的图异常检测方法通常只考虑单个尺度的视图对比，因此会丢失不同尺度之间的信息，进而限制异常检测的效果。



Multi-scale Contrastive Learning

对比“节点-节点”和“节点-子图”两个尺度，其中子图通过随机游走采样，并掩码中心节点的属性，避免任务设计上存在漏洞，模型学到捷径解。

- patch-level (i.e., node versus node) agreement

$$\mathcal{L}_p = -\frac{1}{2n} \sum_{i=1}^n \left(\log(s_p^{(i)}) + \log(1 - \tilde{s}_p^{(i)}) \right).$$

- context-level (i.e., node versus ego-net) agreement

$$\mathcal{L}_c = -\frac{1}{2n} \sum_{i=1}^n \left(\log(s_c^{(i)}) + \log(1 - \tilde{s}_c^{(i)}) \right).$$

ANEMONE

statistical anomaly estimator: 推断异常分数，考虑到对比学习的正负样本和数据增强的随机性，设计了启发式的异常分数。

- 基础分数通过采样正负样本，计算对比分数的差值

$$b_{view,j}^{(i)} = \tilde{s}_{view,j}^{(i)} - s_{view,j}^{(i)},$$

- 提点的 trick：在基础分数之上计算均值和方差

$$\bar{b}_{view}^{(i)} = \sum_{j=1}^R b_{view,j}^{(i)} / R,$$

$$y_{view}^{(i)} = \bar{b}_{view}^{(i)} + \sqrt{\sum_{j=1}^R \left(b_{view,j}^{(i)} - \bar{b}_{view}^{(i)} \right)^2 / R},$$

Anomaly Detection on Attributed Networks via Contrastive Self-Supervised Learning

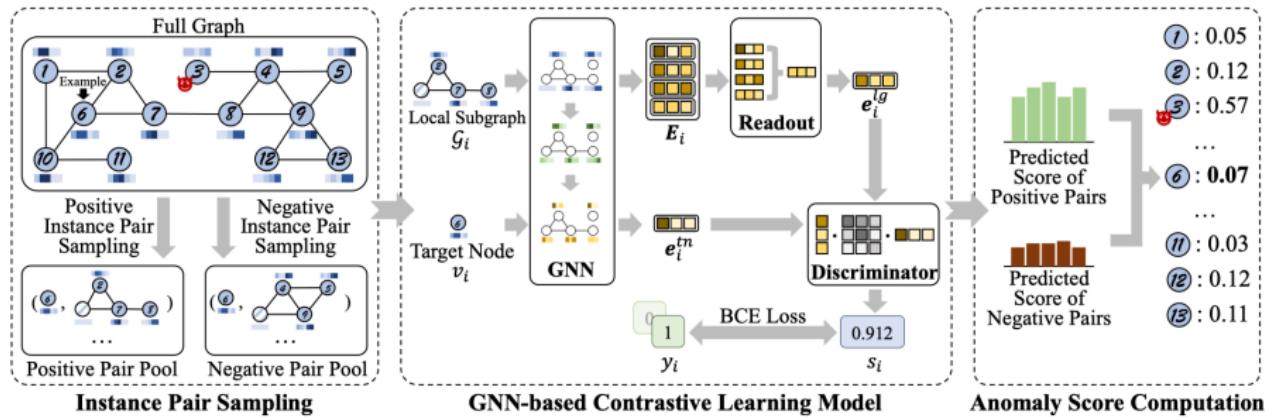
研究动机

由于同质性 (Homophily) 的存在，节点和其周围节点会表现出相似的属性或标签，而与周围节点模式不一致的节点则为异常节点。

核心思路：Node-Subgraph 对比学习框架，通过节点与邻居子图的一致性作为异常的指标。

- ① Random Walk 邻居子图生成
- ② Node VS Random Walk 邻居子图对比学习

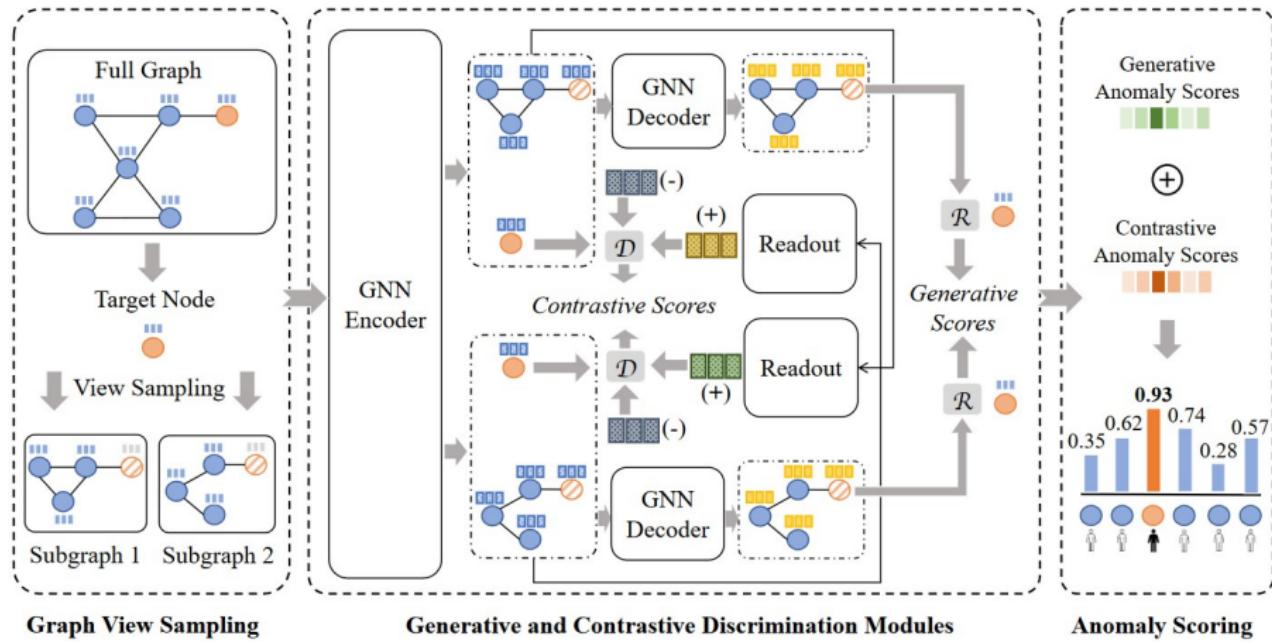
CoLA



异常分值：负例对比分数和正例对比分数之间的差值

$$f(v_i) = s_i^{(-)} - s_i^{(+)}$$

Generative and Contrastive Self-Supervised Learning for Graph Anomaly Detection: SL-GAD

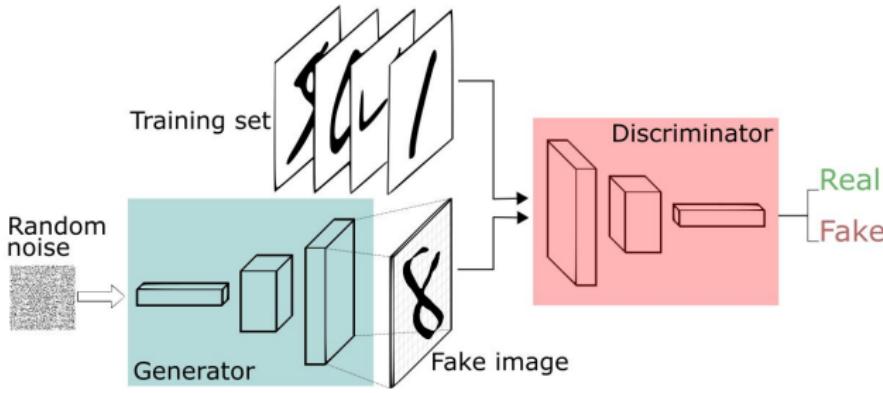


基于生成对抗网络的异常检测算法

生成对抗网络

生成对抗网络 Generative Adversarial Network 由生成器和判别器组成，两个网络相互对抗、不断调整参数，最终目的是使判别网络无法判断生成网络的输出结果是否真实。

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_X} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_Z} [\log(1 - D(G(\mathbf{z})))]$$



GANomaly

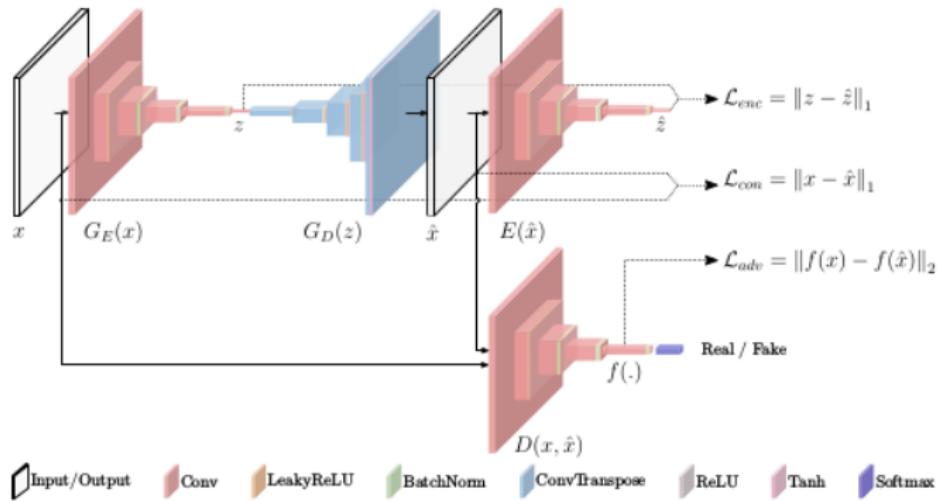


Figure 2: Pipeline of the proposed approach for anomaly detection.

GANomaly 网络结构

- 在训练阶段，整个模型均是通过正常样本做训练。也就是编码器 $GE(x)$ ，解码器 $GD(z)$ 和重构编码器 $E(\hat{x})$ ，都是适用于正常样本的
- 当模型在测试阶段接受到一个异常样本，理论上此时模型的编码器，解码器将不适用于异常样本，此时得到的编码后潜在变量 z 和重构编码器得到的潜在变量 \hat{z} 的差距是大的。这个差距记为：

$$\mathcal{A}(\mathcal{X}) = \|G_E(x) - E(G(x))\|_1$$

通过设定阈值 ϕ ，一旦 $A(x) > \phi$ 模型就认定送入的样本 x 是异常数据。

- 在训练阶段，整个模型均是通过正常样本做训练。也就是编码器 $GE(x)$ ，解码器 $GD(z)$ 和重构编码器 $E(\hat{x})$ ，都是适用于正常样本的
- 当模型在测试阶段接受到一个异常样本，理论上此时模型的编码器，解码器将不适用于异常样本，此时得到的编码后潜在变量 z 和重构编码器得到的潜在变量 \hat{z} 的差距是大的。这个差距记为：

$$\mathcal{A}(\mathcal{X}) = \|G_E(x) - E(G(x))\|_1$$

通过设定阈值 ϕ ，一旦 $A(x) > \phi$ 模型就认定送入的样本 x 是异常数据。

基于生成对抗网络的异常检测算法

优点

- ① GAN 作为最经典的深度生成模型之一，可以广泛用于生成与真实数据相似的样本。而难以从潜在空间生成的样本可能是异常样本。
- ② GAN 经过多年发展，已有大量成熟的模型可用于异常检测。

缺点

- ① GAN 模型的训练相对困难，容易出现模型坍塌等问题
- ② 当待检测数据较为复杂时，GAN 很容易生成与大部分样本不同的样本。异常数据集容易进一步加剧 GAN 模型的训练。
- ③ 基于 GAN 的异常检测模型本质上还是训练 GAN，而不是异常检测。

基于聚类的方法

研究动机

聚类和异常检测是最具代表性的两个无监督任务，两者假设类似且彼此依赖。

基于聚类的异常检测算法的假设

- ① 正常的数据会呈现一定的聚类分布，而异常数据不会属于任何类。
- ② 正常数据和他们最近的聚类中心比较接近，异常数据与聚类中心距离较远。
- ③ 正常数据属于大且稠密的类，异常数据不属于任何类或者属于小的稀疏的类。

基于聚类的异常检测算法

1. 对数据进行聚类，并把不在类中的数据作为异常数据。

- ① DBSCAN
- ② ROCK
- ③ SNN clustering

依赖于聚类的质量，效果难以保证

2. 对于每个数据，计算它到最近的聚类中心的距离，并作为异常分值

Two-step method

3. Cluster-Based Local Outlier Factor (CBLOF)

聚类大小 + 到聚类中心的距离

基于概率分布的异常检测模型

先验假设

许多真实数据是由概率分布生成的，而异常数据则是与整体数据的概率分布不同。

模型假设

正常样本出现在一个分布的高概率密度区域，而异常样本则出现在低概率密度区域。

通过学习数据集的概率密度参数，估计出样本所处的概率密度空间，并估计置信度。

基于概率分布的异常检测模型

基本流程

- ① 为数据选择一个概率模型
- ② 根据概率模型选择一个概率阈值
- ③ 计算观测到每个样本的概率
- ④ 将低于阈值的样本作为异常样本

常见的概率模型：高斯分布，泊松分布（Poisson Distribution）

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

基于距离的异常方法

基本假设

异常数据通常距离正常的数据较远。

异常的分值可以用样本到它的邻居的距离来定义。

优点

- ① 简单直接
- ② 无监督，数据驱动

缺点

- ① 依赖特征
- ② 难以处理高维数据
- ③ 结果难以解释
- ④ 对数据的密度敏感

总结与展望

当前图异常检测面临的主要困难：

- ① 异常假设的一致性
- ② 异常模式的动态性
- ③ 图神经网络的有效性
- ④ 异常结果的可解释性

References I

-  Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., Müller, E., and Kloft, M. (2018). Deep one-class classification. In *International conference on machine learning*, pages 4393–4402. PMLR.