摘要：

Abstract

Time series anomalies can offer information rel-evant to critical situations facing various fields,from finance and aerospace to the IT, ecu-rity, and medical domains. However, detecting anomalies in time series data is particularly chal-lenging due to the vague definition of anomaliesand said data’s frequent lack of labels and highlycomplex temporal correlations. Current state-of-the-art unsupervised machine learning methods for anomaly detection suffer from scalability and portability issues, and may have high false posi-tive rates. In this paper, we propose TadGAN, an unsupervised anomaly detection approach built

onGenerativeAdversarialNetworks(GANs). To capture the temporal correlations of time series distributions, we use LSTM Recurrent Neural

Networks as base models for Generators and Critics. TadGAN is trained with cycle consis- tency loss to allow for effective time-series data reconstruction. We further propose several novel methods to compute reconstruction errors, as well as different approaches to combine recon-struction errors and Critic outputs to compute anomaly scores. To demonstrate the performance and generalizability of our approach, we test several anomaly scoring techniques and report the best-suited one. We compare our approach to 8 baseline anomaly detection methods on 11 datasets from multiple reputable sources such as NASA, Yahoo, Numenta, Amazon, and Twitter. The results show that our approach can effec- tively detect anomalies and outperform baseline methods in most cases (6 out of 11). Notably, our method has the highest averaged F1 score across all the datasets. Our code is open source and is available as a benchmarking tool.

时间序列异常可以为从金融、航空航天到IT、安全和医疗等各个领域的关键情况提供相关信息。然而，在时间序列数据中检测异常是非常具有挑战性的，因为异常的定义很模糊，并且数据经常缺乏标签和高度复杂的时间相关性。目前用于异常检测的最先进的无监督机器学习方法存在可扩展性和可移植性问题，可能有很高的假阳性率。在本文中，我们提出了TadGAN，一种基于generativeadversarialnetworks (GANs)的无监督异常检测方法。为了捕捉时间序列分布的时间相关性，我们使用LSTM递归神经网络作为生成器和批评者的基础模型。TadGAN使用周期一致性损失进行训练，以允许有效的时间序列数据重构。我们进一步提出了几种新的计算重构误差的方法，以及不同的方法来结合重构误差和Critic输出来计算异常分数。为了演示我们的方法的性能和通用性，我们测试了几种异常评分技术，并报告了最适合的一种。我们将我们的方法与来自多个著名来源的11个数据集的8种基线异常检测方法进行了比较，这些数据集包括NASA、Yahoo、Numenta、Amazon和Twitter。结果表明，我们的方法可以有效地检测异常，并在大多数情况下优于基线方法(6 / 11)。值得注意的是，我们的方法在所有数据集中拥有最高的平均F1得分。我们的代码是开源的，可以作为基准测试工具使用。

引言

最近时间观测数据的激增导致了从能源和金融到医疗保健和云计算等许多领域对时间序列异常检测的需求不断增加。时间序列异常定义为系统行为异常[8]的时间点或时间段。一般来说,有两种类型的异常:一个点异常是一个数据点,已经达到了一个不寻常的价值,而集体异常是一个连续的序列被认为异常的数据点作为一个整体,即使单个数据点可能不是不寻常的[8]。

时间序列异常检测的目的是分离时间序列内不同长度的异常子序列。最简单的检测技术之一是阈值，它检测超出正常范围[8]的数据点。然而，许多异常没有超过任何边界——例如，它们可能具有所谓的“正常”值，但在它们发生的特定时间是不正常的(即上下文异常)。这些异常更难识别，因为信号的背景往往不清楚[1,8]。

在阈值的基础上提出了各种统计方法，如统计过程控制(SPC)[39]，其中数据点如果不能通过统计假设检验，则被识别为异常。但这些方法无法处理多变量数据流，而且统计特征往往是未知的和多样的。

研究人员还研究了一些基于无监督机器学习的异常检测方法。一种流行的方法是将时间序列分割成一定长度的子序列(重叠或其他)，然后应用聚类算法来寻找离群值。另一种学习一种模型，预测或重建一个时间序列信号，并对真实值和预测或重建值进行比较。较高的预测或重建误差表明存在异常。

深度学习方法[22]具有很强的处理复杂时间相关性非线性的能力，具有优秀的学习能力。出于这个原因，他们做到了被用于许多时间序列异常检测方法[14,21,26]，包括Microsoft[30]等公司创建的工具。生成对抗网络(GANs)[16]也被证明在生成时间序列序列方面非常成功，表现优于最先进的基准[37]。这种方法的激增引出了一个问题:这些新的、复杂的方法真的比简单的基线统计方法表现得更好吗?评价新方法,我们使用11的数据集(真实和合成),集体有492信号与成千上万的已知异常设置基准测试系统(见细节部分5和表3)。我们实现5最近的深度学习技术介绍了从2016年到2019年,并将他们的表现与20世纪70年代的基准方法ARIMA进行了比较。虽然一些方法能够在50%的数据集上击败ARIMA，但有两种方法完全不能胜过ARIMA (c.f. Table 1)。

基于深度学习的方法面临的一个基本挑战是其拟合数据的显著能力。因此，它也可以拟合异常数据。例如，自动编码器使用L2目标函数，可以非常精确地拟合和重建数据，从而也拟合异常。另一方面，gan可能无法学习生成器完全捕获数据隐藏分布，从而导致误报。因此，我们发现有必要通过混合这两种方法来创建一种更微妙的方法。此外，在这些作品中，通常更多地强调的是深度学习模型本身。然而，正如我们在本文中所展示的，后处理步骤可以显著帮助减少假阳性的数量。

在这项工作中，我们介绍了一种新的GAN体系结构，TadGAN，用于时间序列域。我们使用TadGAN重建时间序列，并以上下文方式评估错误以识别异常。我们将探索基于Generators和Critics的输出计算异常分数的不同方法。。我们将我们的方法与几个著名的经典和基于深度学习的方法在11个时间序列数据集上进行基准测试。具体结果见表3。

本文的主要贡献如下:

•我们提出了一种新的基于无监督gan重构的时间序列数据异常检测方法。特别是，我们为时间序列到时间序列的映射引入了一个周期一致的GAN体系结构。

•我们确定了两个时间序列相似性度量，适合于评估原始序列和gan重建序列之间的上下文相似性。我们的新方法利用GAN的Generator和Critic来计算每个时间步的鲁棒异常得分。

•我们使用来自3个信誉良好的实体(NASA, Yahoo，和Numenta)的11个时间序列数据集进行了广泛的评估，表明我们的方法优于其他8个基线。我们进一步提供了使用gan对时间序列数据进行异常检测的几个见解。

•我们开发了一个时间序列异常检测的标杆系统。该系统是开源的，可以通过其他方法和数据集进行扩展。截至撰写本文时，基准测试已包括9个异常检测管道、13个数据集和2个评价机制。本文的其余部分结构如下。

第2节介绍了相关文献的概述。第3节介绍了我们的GAN模型的细节。然后，我们在第4节中描述了如何使用gan进行异常检测，然后在第5节中对我们提出的框架进行评价。最后，第6节总结了本文并报告了我们的主要发现。

2. 相关工作。

2.1时间序列数据的异常检测。

在过去的几年中，异常类型、数据类型和应用场景的丰富多样催生了一系列的检测方法[8,15,18,20]。其中最简单的方法是超限方法，它标记值超过某个阈值的区域[10,28]。虽然这些方法很直观，但它们不够灵活，无法检测上下文异常。为了克服这一更高级的技术，即:基于接近性，基于预测和基于重建，已被提出。

Proximity-based methods first use a distance measure to quantify similarity between objects – single data points for point anomalies, or fixed length sequences of data points for collective anomalies. Objects that are distant from others are considered as anomalies. This detection type can be further divided into distance-based methods, such as KNearest Neighbor (KNN) [3] – which use a given radius to define neighbors of an object and use the number of neighbors to determine an anomaly score – and densitybased methods, such as Local Outlier Factor (LOF) [7] and Clustering-Based Local Outlier Factor [19], which further consider the density of an object and that of its neighbors. There are two major drawbacks to applying proximitybased methods in time series data: (1) a priori knowledge about anomaly duration and the number of anomalies is required; (2) these methods are unable to capture temporal correlations.

基于近似值的方法首先使用距离度量来量化对象之间的相似性——单数据点用于点异常，或固定长度的数据点序列用于集体异常。距离较远的物体被认为是异常。这种检测可以进一步分为基于距离的方法,如KNearest邻居(资讯)[3]——使用一个给定的半径定义对象的邻居和使用邻居的数量来确定异常分数densitybased方法,如当地异常因素(LOF)[7]和Clustering-Based局部离群值因子[19],它进一步考虑了一个物体及其邻近物体的密度。在时间序列数据中应用基于近似的方法存在两个主要缺陷:(1)需要先验的异常持续时间和异常数目;(2)这些方法无法获取时间相关性。

Prediction-based methods learn a predictive model to fit the given time series data, and then use that model to predict future values. A data point is identified as an anomaly if the difference between its predicted input and the original input exceeds a certain threshold. Statistical models, such as ARIMA [29], Holt-Winters [29], and FDA [35], can serve this purpose, but are sensitive to parameter selection and often require strong assumptions and extensive domain knowledge about the data. Machine learning based approaches attempt to overcome these limitations. [1] introduce Hierarchical Temporal Memory (HTM), an unsupervised online sequence memory algorithm, to detect anomalies in streaming data. HTM encodes the current input to a hidden state and predicts the next hidden state. A prediction error is measured by computing the difference between the current hidden state and the predicted hidden state. Hundman et al. [21] propose Long Short Term Recurrent Neural Networks (LSTM RNNs), to predict future time steps and flag large deviations from predictions.

基于预测的方法学习一个预测模型来拟合给定的时间序列数据，然后使用该模型来预测未来的值。如果数据点的预测输入和原始输入之间的差异超过某个阈值，则该数据点被识别为异常。统计模型，如ARIMA[29]、Holt-Winters[29]和FDA[35]，可以用于此目的，但对参数选择很敏感，通常需要强大的假设和关于数据的广泛领域知识。基于机器学习的方法试图克服这些限制。[1]引入了分层时间记忆(HTM)，一种无监督在线序列记忆算法，用于检测流数据中的异常。HTM将当前输入编码为隐藏状态，并预测下一个隐藏状态。通过计算当前隐藏状态和预测隐藏状态之间的差异来测量预测误差。huntman等人提出了长期短期循环神经网络(LSTM RNNs)，以预测未来的时间步长和标记与预测的大偏差。

Reconstruction-based methods learn a model to capture

the latent structure (low-dimensional representations) of

the given time series data and then create a synthetic re-

construction of the data. Reconstruction-based methods

assume that anomalies lose information when they are

mapped to a lower dimension space, thereby cannot be

effectively reconstructed; thus, high reconstruction errors

suggest high chances of being anomalies.

Principal Component Analysis (PCA) [31], a

dimensionality-reduction technique, can be used to

reconstruct data, but it is limited to linear reconstruction

and requires data to be highly correlated and to follow a

Gaussian distribution [9]. More recently deep learning

based techniques have been investigated. These include the

use of Auto-Encoder (AE) [2], V ariational Auto-Encoder

(V AE) [2] and LSTM Encoder-Decoder [27].

基于重构的方法学习一个模型来捕获给定时间序列数据的潜在结构(低维表示)，然后创建数据的合成重构。基于重构的方法假定异常在映射到低维空间时丢失了信息，因此不能有效地重构;因此，高重建误差意味着高概率的异常。主成分分析(PCA)[31]是一种降维技术，可用于数据重构，但仅限于线性重构，要求数据高度相关，服从高斯分布[9]。最近，基于深度学习的技术得到了研究。这些包括使用自动编码器(AE) [2]， V可变自动编码器(V AE)[2]和LSTM编码器-解码器[27]。

However, without proper regularization, these

reconstruction-based methods may easily get overfit-

ted, resulting in low performance. In this work, we

propose to use adversarial learning to allow for time series

reconstruction. We introduce an intuitive approach for

regularizing reconstruction errors. The trained Generators

can be directly used to reconstruct more concise time series

data – thereby providing more accurate reconstruction

errors – while the Critics can offer scores as a powerful

complement to the reconstruction errors when computing

an anomaly score.

然而，如果没有适当的正则化，这些基于重构的方法容易出现过拟合，导致性能低下。在这项工作中，我们建议使用对抗学习允许时间序列重建。我们引入一种直观的方法来正则化重构误差。经过训练的生成器可以直接用于重构更简洁的时间序列数据——从而提供更准确的重构错误——而批评者可以提供得分，作为计算异常得分时重构错误的有力补充。

2.2. Anomaly Detection Using GANs. Generative adversarial networks can successfully perform many image-related tasks, including image generation [16], image translation [41], and video prediction [36], and researchers have recently demonstrated the effectiveness of GANs for anomaly detection in images [11, 34].

Adversarial learning for images. Schlegl et al. [33] use the Critic network in a GAN to detect anomalies in medical images. They also attempt to use the reconstruction loss as an additional anomaly detection method, and find the inverse mapping from the data space to the latent space. This mappingis doneina separate step, afterthe GAN is trained. However, Zenati et al. [38] indicate that this method has proven impractical for large data sets or real-time applications. They propose a bi-directional GAN for anomaly detection in tabular and image data sets, which allows for simultaneous training of the inverse mapping through an encoding network.

2.2。基于gan的异常检测。

生成式对抗网络可以成功地执行许多与图像相关的任务，包括图像生成[16]、图像转换[41]和视频预测[36]，最近有研究证明了生成式对抗网络在图像异常检测方面的有效性[11,34]。

图像的对抗性学习。Schlegl等[33]利用GAN中的Critic网络检测医学图像中的异常。他们还尝试使用重建损失作为一种额外的异常检测方法，并找到从数据空间到潜在空间的逆映射。这个映射是一个单独的步骤，在GAN被训练之后。然而，Zenati等人[38]指出，这种方法已被证明不适用于大数据集或实时应用。他们提出了一种双向GAN，用于表格和图像数据集的异常检测，它允许通过编码网络同时训练逆映射。

The idea of training both encoder and decoder networks was developed by Donahue et al. [12] and Dumoulin et al. [13], who show how to achieve bidirectional GANs by trying to match joint distributions. In an optimal situation, the joint distributions are the same, and the Encoder and Decoder must be inverses of each other. A cycle-consistent GAN was introduced by Zhu et al. [41], who have two networks try to map into opposite dimensions, such that samples can be mapped from one space to the other and vice versa.

Adversarial learning for time series. Prior GAN-related work has rarely involved time series data, because the complex temporal correlations within this type of data pose significant challenges to generative modeling. Three works published in 2019 are of note. First, to use GANs for anomaly detection in time series, Li et al. [24] propose using a vanilla GAN model to capture the distribution of a multivariate time series, and use the Critic to detect anomalies. Another approach in this line is BeatGAN [40], which is a Encoder and Decoder GAN architecture that allows for the use of the reconstruction error for anomaly detection in heartbeat signals. More recently, Y oon et al. [37] propose a time series GAN which adopts the same idea but introduces temporal embeddings to assist network training. However, their work is designed for time series representation learning instead of anomaly detection. To the best of our knowledge, we are the first introducing cycle-consistentGAN architecturesfortime series data, such that Generators can be directly used for time series reconstructions. Besides, we systematically investigate how to utilize Critic and Generator outputs for anomaly score computation. A complete framework of time series anomaly detection is introduced to work with GANs.

训练编码器和解码器网络的想法是由Donahue等人[12]和Dumoulin等人[13]开发的，他们展示了如何通过匹配联合分布来实现双向gan。在最优情况下，联合分布是相同的，编码器和解码器必须互为倒数。Zhu等人[41]引入了周期一致的GAN，他们有两个网络试图映射到相反的维度，这样样本可以从一个空间映射到另一个空间，反之亦然。

时间序列的对抗性学习。之前与gan相关的工作很少涉及时间序列数据，因为这类数据中复杂的时间相关性对生成模型提出了重大挑战。2019年发表的三部作品值得注意。首先，为了使用GAN来检测时间序列中的异常，Li等人[24]提出使用香草GAN模型来捕获一个多元时间序列的分布，并使用Critic来检测异常。这一行中的另一种方法是BeatGAN[40]，它是一种编码器和解码器GAN架构，允许在心跳信号中使用重构错误进行异常检测。最近，Y oon等人提出了时间序列神经网络采用了相同的思想，但引入了时间嵌入来辅助网络训练。然而，他们的工作是为了时间序列表示学习而不是异常检测。

据我们所知，我们是第一个为时间序列数据引入周期一致性gan架构的人，这样生成器就可以直接用于时间序列重构。此外，我们还系统地研究了如何利用Critic和Generator输出进行异常评分计算。介绍了一种完整的时间序列异常检测框架。

## 3. Adversarial Learning for time seriesreconstruction

The purpose of our GAN is to reconstruct the input time series: xi→ E(xi) → G(E(xi)) ≈ ˆ xi. However, training the GAN with the adversarial losses (i.e., Wasserstein losses) alone cannot guarantee mapping individual input xi to a desired output ziwhich will be further mapped back to ˆ xi. To reduce the possible mapping function search space, we adapt cycle consistency loss to time series reconstruction. It was first introduced by Zhu et al. [41] for image translation tasks. We train the generative network E and G with the adapted cycle consistency loss by minimizing the L2 norm of the difference between the original and the reconstructed samples:

我们的GAN的目的是重构输入时间序列:xi→E(xi)→G(E(xi))≈ˆxi。然而，仅用对位损失(即瓦瑟斯坦损失)来训练GAN并不能保证将单个输入xi映射到期望的输出ziz, ziz将进一步映射回ˆxi。为了减少可能的映射函数搜索空间，我们将周期一致性损失引入到时间序列重构中。它最初是由Zhu等人引入用于图像平移任务的。我们通过最小化原始样本和重建样本之间差值的L2范数来训练生成网络E和G，并具有自适应的周期一致性损失:



Considering that our target is anomaly detection, we use L2 norm instead of L1 norm (the one used by Zhu et al. [41] for image translation) to emphasize the impacts of anomalous values. In our preliminary experiments, we observed that adding the backward consistency loss (i.e., Ez∼Pz[kz − E(G(z))k2]) did not show an improved performance.

考虑到我们的目标是异常检测，我们使用L2范数代替L1范数(Zhu et al.[41]用于图像平移的L1范数)来强调异常值的影响。在我们的初步实验中，我们发现添加反向一致性损失(即Ez ~ Pz[kz−E(G(z))k2])并没有显示出性能的改善。