

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,500

Open access books available

176,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Anomaly Detection in Time Series: Current Focus and Future Challenges

*Farrukh Arslan, Aqib Javaid, Muhammad Danish Zaheer Awan and Ebad-ur-Rehman*

## Abstract

Anomaly detection in time series has become an increasingly vital task, with applications such as fraud detection and intrusion monitoring. Tackling this problem requires an array of approaches, including statistical analysis, machine learning, and deep learning. Various techniques have been proposed to cater to the complexity of this problem. However, there are still numerous challenges in the field concerning how best to process high-dimensional and complex data streams in real time. This chapter offers insight into the cutting-edge models for anomaly detection in time series. Several of the models are discussed and their advantages and disadvantages are explored. We also look at new areas of research that are being explored by researchers today as their current focuses and how those new models or techniques are being implemented in them as they try to solve unique problems posed by complex data, high-volume data streams, and a need for real-time processing. These research areas will provide concrete examples of the applications of discussed models. Lastly, we identify some of the current issues and suggest future directions for research concerning anomaly detection systems. We aim to provide readers with a comprehensive picture of what is already out there so they can better understand the space – preparing them for further development within this growing field.

**Keywords:** anomaly detection, anomaly detection in time series, high dimensional data, big data, current focus and future challenges, machine learning, deep learning, forecasting, real time

## 1. Introduction

Time series data mining is becoming increasingly important due to the advances in technology which have allowed us to collect and store large amounts of structured temporal data. A wide range of tasks can be performed with this time series data such as classification [1], clustering [2], forecasting [3] and outlier detection [4–6]. Extracting meaningful insights from this data opens up new opportunities for research across diverse areas.

Anomaly detection in time series is a critical task with significant implications in numerous fields, including finance [7], healthcare [8], and security [9]. Identifying and analyzing outliers in time-series data is a critically important operation for obtaining meaningful insights [6]. As described in a seminal paper, there are two types of univariate time-series outliers - type I and type II [10, 11]. Whilst type I outlier events occur individually, type II can be linked to subsequent power shifts. It's essential to have an in-depth understanding of both these outlier types if a meaningful analysis of the data is to be undertaken.

The detection of unusual patterns or events within data streams can provide valuable insights and help identify potentially harmful or fraudulent behavior. However, processing high-dimensional [12, 13] and complex data streams [14] in real-time [15] remains a challenging task. In recent years, many statistical [16], machine learning [17], and deep learning [18] techniques have been proposed to tackle these challenges.

This chapter provides an overview of the current state-of-the-art models for anomaly detection in time series, including their strengths and limitations. We also explore new areas of research as current focuses of researchers in this field, that are being explored by implementing recent techniques, models of machine and deep learning to solve unique challenges posed by high-dimensional and complex data, high-volume data streams, and the need for real-time processing.

One of the primary challenges of anomaly detection in time series is dealing with high-dimensional and complex data. Traditional statistical methods such as ARIMA [19] and exponential smoothing [20] have been used in the past but lack the flexibility to handle complex data with multiple attributes. Newer techniques, including machine learning-based approaches such as isolation forests [21], autoencoder-based methods [22], and deep learning techniques such as LSTM and CNNs [18], have shown promising results in detecting anomalies in high-dimensional and complex data. However, these approaches also have their limitations, including high computational costs and the need for extensive training data [23].

Another challenge in anomaly detection in time series is dealing with high-volume data streams in real-time [24]. Traditional batch processing methods are not suitable for real-time applications where timely detection is critical. As a result, new methods such as online anomaly detection algorithms [25] and sliding window-based [26] approaches have been proposed, which process data in a continuous and efficient manner.

In addition to the current state-of-the-art models, techniques, and areas of research, this chapter also identifies some of the current issues and suggests future directions for research concerning anomaly detection systems. These include the need for more interpretable models, the development of novel unsupervised methods, the integration of domain knowledge, and the need to address the issue of data imbalance.

Sections of this chapter discuss recent algorithms, models separately as well as within current wide areas of research i.e., Forecasting, Real time anomaly detection, Dealing with Big and high dimensional data processing problem, Anomaly detection using Artificial Intelligence (AI), Industrial Control Systems. There will be a literature review of many models and strategies implemented in these areas. All these areas as well as models will be solving challenges as discussed before. Overall, the purpose is to provide readers a literature review about current focuses of researchers and to facilitate further development and research in time series data anomaly detection.

## 2. Big data and high dimensionality

Modern data sets are increasingly high-volume, high-velocity, and high-variety, making it difficult to identify anomalies with accuracy. Researchers and developers have started to investigate new approaches for coping with this complexity. When the number of features grows exponentially, more data is needed to create accurate models - leading to sparse and isolated data points. Gartner [13] defines big data as a collection of attributes that require cost-effective analytics to generate insight. The key challenges associated with big data are described in the 5 Vs: value, veracity, variety, velocity and volume [27].

The next stated paragraphs will give an overview of how real-time big data processing can be used to detect anomalous events through machine learning algorithms, as well as its current limitations and challenges.

McNeil et al. [28] examined existing tools used to detect malware on mobile devices. It was noted that these methods lacked the capability to incorporate group user profiling, which is necessary to automate behavior-based dynamic analysis for focused malware identification. To overcome this, they demonstrated a scalable architecture called SCREDENT, which allows users to classify, identify, and forecast possible target malwares in real time. While initial evaluation indicated that the approach had promise, further testing failed to demonstrate desired results.

In reference [29], a new architecture was proposed to detect threats in real-time using stream processing and machine learning. This architecture promotes an environment with minimal human oversight, allowing for improved detection of both known and previously unseen cyberattacks in order to hone attack classification and anomaly detection capabilities. However, their results did not benefit from the open KDD dataset as much as expected.

In reference [30], complex network infrastructure with vast logs files became the focus of another approach which looks to assess security logs that contain various device information through data mining and machine learning techniques. This proposed approach is split into two phases: defining/configuring the detection mechanism and then executing it at runtime. Nevertheless, the practical implementation turned out to need more automation due to its high human intervention levels; similarly, its output accuracy was not precise enough.

Recent research has highlighted the use of machine learning models for anomaly detection. But, the inability of inherent system performance to keep up with increasing network traffic is an issue that needs to be addressed. To do so, a novel model utilizing Hadoop, HDFS, MapReduce, cloud and multiple machine learning algorithms were developed. Weka interface was used to assess accuracy and efficiency through naive bayes, decision tree and support vector methods. However, the implementation of cloud infrastructure and real-time data streaming has not been sufficiently discussed as yet in this project [31].

Research paper [32] focuses on anomaly detection in streaming data, and provides a new approach to evaluate online anomaly detection with entropy and Pearson correlation. Big data streaming components like Kafka queues and Spark Streaming are used as a means of ensuring scalability and generality, although some processes which were potentially complicated by long batch processing periods or data limitations were not resolved.

Researchers have proposed a method for anomaly detection in smart grids that uses real-time, minimal energy consumption [33]. The proposed in-memory distributed framework, comprised of Spark Streaming and Lambda System, is viable for

scalable live streaming. However, it did take longer to train the model and there were scheduling issues with real-time tasks.

In reference [34], another framework was presented - one which involves sensor data preprocessing: anomaly detection using principal statistical analysis and Bayesian networks; as well as sensor data redundancy elimination using static or dynamic Bayesian networks (SBNs/DBNs). Included were two algorithms: static sensor data redundancy detection algorithm (SSDRDA) and real-time sensor data redundancy detection algorithm (RSDRDA), both serving to reduce redundant data in either static datasets or real-time scenarios respectively.

Anomaly detection in time series requires modifications in the above framework of approaches for efficiency. Anomaly detection in real-time big data analytics is a promising area of study, particularly when machine learning techniques are incorporated. Advancements in this field are likely to yield high accuracy and efficiency. Thus, the potential benefits of such research cannot be underestimated.

Following points will lay out the foremost research challenges in this field, in order to promote progress.

**Redundancy:** Managing real-time big data from diverse sources is challenging. Current technologies like Hadoop and spark fail to address redundancy, data quality and reliability, cost [35], and storage schema [36]. A new framework is needed to tackle these complexities.

**Computational cost:** Anomaly detection requires multiple techniques, increasing computation cost. Large datasets and high dimensionality cause algorithmic instability and computational expense [23]. Big data and cloud technology enable parallel and distributed processing and reduce computing costs. Cheaper processors and high chips improve system power and data processing in real-time, minimizing computational expenses.

**Nature of Input data:** Input data has instances with binary/categorical or continuous attributes, and can be univariate or multivariate. Anomaly detection algorithm selection depends on data diversity and attribute type [37]. A hybrid framework using unsupervised machine learning algorithms can detect anomalies in different datasets.

**Noise and missing value:** Network sensor streaming data has various types and can produce false alarms due to noise and missing values from high speed [37]. Noise can hide true anomalies [38]. An auto noise cleansing module in the detection framework can remove unnecessary features and handle missing values.

**Parameters Selection:** Optimal parameters for machine learning algorithms are hard to select [39]. Real-time anomaly detector needs to consider multiple and single hyperparameters, which may change over time [40]. Parameter choice affects algorithm performance; eccentricity techniques can reduce selection processes [41].

**Inadequate Architecture:** Organizations need big data architecture for large real-time data. Existing architectures are insufficient. Real-time analytics and application components can create efficient environment [42]. Big data technologies and hybrid machine learning algorithms can solve architectural problems. Scalability for data in motion and at rest is achieved.

**Data visualizations:** Data or reports need effective, visual insights. Anomalies from connected devices can use heat maps, scatter plots, parallel coordinates and node-link graphs for 2D/3D views. 3D interaction needs data understanding and user rotation and zoom [43]. Opensource visualization techniques in frameworks can automatically select techniques for better user experience.

**Heterogeneity of data:** Unstructured data is varied and large, such as emails, faxes, form documents, social media posts, etc. Transcription is expensive. Hybrid



Machine Learning algorithms can identify data types quickly and accurately. Complex machine learning models can recognize heterogeneous information sources from unstructured text.

**Accuracy:** Anomaly detection with existing technologies is inaccurate. A hybrid machine learning algorithm can analyze large data from modern applications with low memory and power. Our team combines real time big data technologies with this algorithm for efficient and accurate results.

### 3. Transformer

The Transformer architecture, introduced by Vaswani et al. [44], has been widely used in various natural language processing tasks, such as machine translation and text classification. However, its effectiveness in anomaly detection is limited due to the lack of a specific mechanism to capture anomalous patterns. To address the shortcoming of Transformer architecture, researchers have created a variety of modifications to attempt to improve its performance, including the incorporation of an Anomaly-Attention mechanism. One such modification is the Anomaly Transformer, as illustrated in reference [45], which utilizes this mechanism to improve the detection of anomalous patterns in data. Thus, the Anomaly Transformer architecture represents an important advancement in the application of Transformers for anomaly detection.

Recent studies have focused on leveraging Transformer based architectures to benefit the time series anomaly detection task. These approaches are capable of modeling temporal dependencies, enabling better anomaly detection quality [45]. For instance, TranAD [46], MT-RVAE [47] and TransAnomaly [48] all fuse Transformers with VAEs i.e., neural generative models [49] or GANs [50], demonstrating improved performance in anomaly detection. These models have been explained further below.

**TranAD:** TranAD [46] is a robust adversarial training procedure designed to address small deviations in anomalies that the typical Transformer-based network may overlook. This GAN-style approach consists of two transformer encoders and two decoders in order to maintain stability. The results of an ablation study illustrate the performance of this architecture, with F1 scores dropping by nearly 11% when the Transformer-based encoder-decoder was replaced. This clearly demonstrates the efficacy of using Transformers for time series anomaly detection. It's valuable for modern industrial systems where instant detection of anomalies is compulsory.

**MT-RVAE:** There are two different approaches that combine VAE and Transformer to create novel models for time-series analysis. MT-RVAE, proposed by Wang et al. [47], uses a multiscale Transformer model to extract information from sequences of varying scales like complex satellite systems with several subsystems. Each subsystem's temporal features must be analyzed in correlation [47]. This addresses the limitations traditional Transformers had in being able to accurately analyze sequential data as these models were limited to local information extraction only. TransAnomaly, proposed by Zhang et al. [48], combines VAE with transformer for increased parallelization purposes. The combination of these two techniques is predicted to reduce training costs up to 80%.

**GTA:** GTA [51] leverages Transformers and graph-based learning to accurately detect anomalies in multivariate time series data, even when there are few dimensions or limited close relationships among sequences. This method features a multi-branch attention mechanism composed of global-learned attention, regular multi-head attention, and neighborhood convolution for increased accuracy, as well as a graph

convolution structure for modeling influence propagation processes. Thus, GTA seeks to provide an improved approach for analyzing and detecting anomalies in multivariate time series data than previous methods. It's valuable for internet connected sensory devices like smart power grids, water distribution networks as they remain under attack of cyber-attacks [51].

**AnomalyTrans:** AnomalyTrans [45] is a novel approach in distinguishing anomalies. Drawing inspiration from TranAD, AnomalyTrans makes it more difficult for anomalies to create strong connections with the entire time series, though retaining connectivity between adjacent time points. The model leverages Transformer and Gaussian prior-association to reach this objective. Through utilizing a minimax strategy to optimize the anomaly model, AnomalyTrans enforces restrictions on prior- and series-associations that result in a greater divergence between them.

**D3TN:** Disentangled Dynamic Deviation Transformer Network (D3TN) [52] is highly effective system for multivariate time series anomaly detection. It considers both short-term and long-term temporal dependencies as well as complex inter-sensor dependencies. To better model static topology, a new disentangled multi-scale aggregation scheme for graph convolutional neurons for fixed inter-sensor relationships was introduced. A self-attention mechanism was also employed to capture dynamic directed interaction in various subspaces that vary with time and unexpected events. Moreover, parallel processing of the time series helps simulate complex temporal correlations that span multiple time periods.

**DATN:** The Decompositional Auto-Transformer Network (DATN) [53] is a unique anomaly detection method for time series. This novel approach breaks complex time series into seasonal and trend components, before then renovating them with deep models. Additionally, the design integrates an auto-transformer block to detect important representations and dependencies based on seasonality and trends in the series. Furthermore, rather than using a traditional complex transformer decoder, we substitute it with a more efficient linear decoder.

Transformers have been applied in many real word scenarios for anomaly detection in time series like: SMD is a 5-week-long dataset acquired from one of the leading Internet companies with 38 characteristics. Pooled Server Metrics (PSM) was procured internally from multiple server nodes at eBay and consists of 26 variables. Besides these, both Mars Science Laboratory rover (MSL) and Soil Moisture Active Passive (SMAP) satellite datasets from NASA have been compiled as well, containing 55 and 25 features respectively with regard to the anomaly data derived from the Incident Surprise Anomaly (ISA) reports for spacecraft monitoring systems. Last but not least, the Secure Water Treatment (SWaT) dataset contains 51 indicators derived from continuous operation on a critical infrastructure system [45, 53].

The possible future challenges are indicated below in paragraphs.

**Inductive Biases for Time Series Transformers:** Transformers are powerful, general networks for modeling long-range dependencies. But they require quite a bit of data to train effectively and avoid falling prey to data overfitting. Time series data often follows seasonal or periodic patterns, as well as other trends, which suggests that incorporating this information into Transformers has potential to lead improvement in performance. For instance, recent studies have demonstrated the effectiveness of frequency processing [54] and capturing series periodicity [55]. Additionally, both explicitly allowing cross-channel dependency [56] and preventing it via channel-independent attention module [57] have yielded better models for certain tasks. The challenge then lies in finding a balance between designing inductive bias

to suppress noise while amplifying signal — a task whose solution is yet to come but promises exciting possibilities ahead.

**Transformers and GNN for Time Series:** As datasets with multi-dimensional and spatial-temporal characteristics become more widespread, it's essential to have tools which can effectively capture the complexities that these data represent. Graph Neural Networks (GNNs) is one method of modeling dependencies and relations with each other between dimensions. Recent studies have shown that combining GNNs with Transformers/Attentions leads to impressive performance improvement in areas such as traffic forecasting [58, 59] and multi-modal forecasting [60, 61], knowledge of latent causality and the underlying clarity of spatial-temporal performance can be increased with a greater comprehension. It is an important development that could result in more effective use of Transformer-GNN hybrid models for spatial-temporal modeling in time series going forward.

**Pre-trained Transformers:** As the advances of large-scale Transformers using pre-training have yielded observable improvements across a wide range of natural language processing tasks [62, 63] and CV [64], research conducted on their efficacy for time series applications has been limited. Works existing to this day primarily focus on classification activities [65, 66]. In order to develop effective pre-trained Transformer models that are equipped to address a range of use cases within time series analysis, further examination will be required in the future.

**Architecture Level Variants:** Considering the success of Transformer variants in NLP and CV, it may be beneficial to transfer this concept over to time series data and tasks. We can look into more architecture-level designs for Transformers which may optimize performance on time series specific models. Examples of these variants include lightweight [67, 68], cross-block connectivity [69], adaptive computation time [70, 71], and recurrence [72]. These architecture-level designs provide us with a whole new range of opportunities for improvement.

**Transformers with Neural Architecture Search:** Tuning Transformer hyper-parameters such as embedding dimension, heads, and layers can have a significant impact on performance. Thankfully, Neural Architecture Search (NAS) provides a means to automatically find architectures that optimize performance. Recently, NAS technologies in NLP and CV have been applied to transformers [73, 74]. For machine data which may be high dimensional yet long in length, this technique is especially important for designing memory- and computationally-efficient transformers. We anticipate further progress in this area as the industry gears up for more efficient time series Transformers.

## 4. Non-pattern anomaly detection

Non-Pattern Anomaly Detection is an underdiagnosed but powerful method of identifying anomalies in time series. Existing techniques use initial profiling to determine which behavior should be tagged as “normal” or “abnormal,” but this definition fails to capture the nuanced changes between situations in different conditions. Researchers recognized the importance of such a technique and emphasized its potential for detecting abnormalities even in the absence of statistical methods that often play a dominant role in machine learning processes. Team of researchers aimed to compare current machine learning algorithms relating to NP-AD approaches and assessed how various datasets demonstrated their capacity for anomalies on diverse situations [75].



## 5. Hybrid models

Multivariate time-series anomaly detection is a complex challenge due to the imbalance of anomalous data and its underlying intricacies. Combining different methods for detecting anomalies in time series has been well explored, resulting in improved accuracy. Notably, hybrid models combining statistical and deep learning approaches have been found to provide greater precision when determining uncertainty and quantifying forecasts associated with these models.

One example hybrid approach is called HAD-MDGAT – it's based on a GAT (graph attention network) combined with multi-channel temporal stacked Denoising Autoencoder (MDA), designed to learn temporal and spatial correlations among observations. Ablation study results show that MDA enhances anomaly detection accuracy dramatically; this model with an MDA layer scored 10.86% higher than one without the extra layer [76].

A research paper published in reference [77] outlined a novel Long Short-Term Memory (LSTM) network-based method for accurately forecasting multivariate time series data. In addition, the study featured an LSTM Autoencoder network-based approach coupled with a one-class Support Vector Machine (SVM) algorithm, which was employed for anomaly detection. Their findings demonstrated that the LSTM Autoencoder based method outperforms the previously proposed LSTM based method. Moreover, their proposed forecast approach surpassed several other methods by NASA. The LSTM based methodology is well suited to forecasting while the combination of an LSTM Autoencoder with the OCSVM is suitable for detecting anomalies [77].

MES-LSTM is a combination of a multivariate forecasting model and Long Short-Term Memory, a form of Recurrent Neural Network (RNN). Accurate attribution is an important part of any system as it reinforces confidence in the mechanics and makes sure learning processes are not based on spurious effects. While MES-LSTM does a great job of anomaly detection, overall performance could still benefit from improvement [78].

A hybrid deep-learning model that integrates long short-term memory (LSTM) and autoencoder (AE) networks was proposed for anomaly detection tasks in Indoor Air Quality (IAQ) time series data. The LSTM cells are stacked together to learn the long-term dependencies in time-series data, while the AE helps identify an optimal threshold based on reconstruction loss rates across all sequences. This powerful combination helps detect outliers with precision and efficiency [79].

A SeqVAE-CNN model to carry out unsupervised deep learning for anomaly detection. This model takes inspiration from Variational Autoencoders (VAEs) and Convolutional Neural Networks (CNNs), creating a Seq2Seq structure that can capture both temporal relationships and spatial features in multivariate time-series data. The experimental results of their model on 8 datasets from different domains suggest it has a higher performance for anomaly detection; indeed, the highest AUROC and F1 scores have been observed when using our model [80].

Researchers propose a hybrid model of VAE-LSTM for unsupervised anomaly detection in time series. This model combines the features extracted from the VAE module [81], which capture local patterns for short windows, with the Long Short-Term Memory (LSTM) module, which captures long-term correlations in the time series. Additionally, Electrical power grids are vulnerable to cyber-attacks, existing attack detection methods are limited so to tackle Graph Convolutional Long Short-Term Memory (GC-LSTM) with a deep convolution network has been proposed to

further improve time series classification and analysis with respect to anomaly detection and attack graph models [82].

A new hybrid anomaly detector that merges two detection approaches i.e., Key Performance Indicators (KPIs) that are used in physics and Unsupervised Variational Autoencoder (VAE), thereby improving accuracy and decreasing the possibility of overlooking defective elements in safety-critical scenarios. Performance is discussed in comparison to different VAE architectures like long short-term memory (LSTM-VAE) and bidirectional LSTM (BiLSTM-VAE). Additionally, the efficient choice of hyperparameters in these structures can be optimized with the help of a genetic algorithm as presented in reference [83].

Due to the many advantages of conventional anomaly detection in time series models, further innovations in this area have the potential to yield beneficial results. In fact, tackling the complexities associated with real-world time series requires advanced solutions, such as hybridization of hybrid classes. Research shows that this technique can provide great improvements in terms of forecasting accuracy and has been gaining much attention recently.

## 6. Forecasting and anomaly

Time Series Forecasting has always been a useful tool to detect trends, patterns of any data. It's about predicting the next time stamps using previous or existing trends. Anomaly Detection and Time series forecasting have been interlinked several times by researchers in this field. Several Machine Learning algorithms have been implemented, sometimes merged with each other to derive another novel strategy to predict whether the next time stamp is normal or abnormal.

The power of forecasting lies in its potential to revolutionize healthcare. The goal? To empower medical professionals to take proactive and timely action, reducing patient transfers and hospital stay lengths, ultimately leading to improved survival rates. But the accuracy of predictions relies heavily on expertly combining machine learning algorithms like autoencoders and extreme gradient boosting (XGBoost) [84].

Autoencoders excel in feature extraction [85]. They are uniquely adept at unsupervised anomaly detection when labeled data is scarce or nonexistent [86]. They're trained via reconstruction error, only triggering an alert if said error exceeds a pre-determined threshold - prompting a swift remedial response. As for XGBoost, this decision tree-based ensemble principle takes physiological variables from time  $t_i$  as input and outputs variables from the next temporal unit;  $t_{i+1}$  [84].

All told, tapping into modern technology's full potential could allow for massive improvements in healthcare outcomes - starting with careful utilization of solutions like autoencoders or XGBoost models.

Recent research has sought to compare the performance of supervised and unsupervised algorithms on physiological data. Heart rate data, due to its ubiquity and non-invasiveness, is ideal for predicting anomalies. Five algorithms were evaluated for detecting anomalies in heart rate -- two unsupervised techniques and three supervised methods. The models were tested on real heart rate data and findings demonstrated that both local outlier factor and random forests algorithms were effective in detecting abnormalities in this type of data. Additionally, results showed that simulated data can lead algorithms to a similar level as real labeled information when not available, enabling rapid initial deployment without prior knowledge [8].

DeepAnT is a deep learning-based anomaly detection approach for streaming and non-streaming time series data. It can detect a broad range of anomalies, from point anomalies to contextual and discords. Instead of learning about anomalies, DeepAnT uses unlabeled data to determine normal time series. The two key components of DeepAnT are its time series predictor (which uses CNN and takes context into account) and its anomaly detector module, which identifies whether an upcoming time stamp is normal or anomalous.

DeepAnT stands out against the competition by only needing a relatively small data set to generate a model. It utilizes parameter sharing of a convolutional neural network (CNN) which allows for good generalization capabilities. Unsupervised anomaly detection in DeepAnT removes the need for labeling, making it directly applicable to real-world scenarios with large streams of complex data from heterogeneous sensors. Neural networks are popular as they enable automatic feature discovery without having any prior domain knowledge; this capability is what makes them such excellent candidates for time series anomaly detection. DeepAnT optimizes through leveraging a CNN and raw data, making it more robust to variations than many other neural networks and statistical models on the market [87]. Using a data-driven approach can be beneficial in many contexts, especially when there is access to an abundance of untagged data. However, the data quality has a great impact factor on its accuracy; if too much of the dataset is contaminated (5% or more), then it could potentially lead to wrong inferences upon deployment. Additionally, selecting the right network architecture and hyperparameters are often difficult tasks. Nevertheless, new automated techniques have been developed that may assist in optimizing these settings instead of opting for human expertise [88]. Last but not least, one major drawback is the susceptibility to adversarial examples [89] which could restrict its usage in safety-critical system models. Luckily though, research into understanding and defending against such cases has increased progressively over time with some successful results achieved.

Light curve prediction and anomaly detection using LSTM neural networks is an important research area for time domain astronomy. A series of processing was done on star images collected from the National Astronomical Observatories of China using GWAC's mini-GWAC system, resulting in light luminance data over a period of time. Researchers explored a model of LSTM neural network to accurately predict light curves, with an optimal structure obtained through model training and validation; meanwhile, an anomaly detection mechanism based on prediction error was implemented. Results showed that this method has great potential when tested on real light curve data [90]. More historical data and certain well-known astronomical principles are needed to further improve upon this method.

Motorsports have limited access to sensors during competitions, limiting predictive capabilities and providing an edge for competitors. The proposed variational autoencoder-based selective prediction (VASP) framework addresses this challenge by combining the tasks of anomaly detection and time series prediction in one powerful approach. VASP consists of a variational autoencoder (VAE), an anomaly detector, and LSTM predictors which can all work together to help produce more robust predictions. Even if anomalies occur in the input signals, VASPs accuracy is not significantly impacted like that of other deep learning approaches such as long short-term memory (LSTM) neural networks. Try out VASP today to take your predictive insights to the next level with more effective technique [91].



## 7. Anomaly detection using AI

Time series data bring their own set of challenges when model analysis is applied, like notions of time and uncertainty, and the presence of drift. Typically, the time series window is broken down into two pieces with either sliding endpoints or landmark endpoints. In this paper, they categorize anomalies and outliers as the same, as presented in reference [92]. Detecting these outliers has been and remains an area of exploration for researchers and practitioners alike. Time series data is one of the most useful modalities available for a variety of applications. Upon analyzing this type of data, it becomes clear that outlier detection plays a key role. Companies such as Microsoft [93] have even created outlier detection services to monitor business data with triggers to alert them when outliers are present.

As stated in reference [94], AI assurance is an important process that must be incorporated throughout the engineering lifecycle of an AI system. This process should ensure the system is dependable and its outcomes valid, trustworthy, and ethical. Moreover, it should also be data-driven, explainable to all users, unbiased in its learning processes, and fair for all involved.

One of the recent hot algorithm in AI for Anomaly detection in time series is GAN proposed by Goodfellow [50], have become some of the most discussed topics in deep learning. The use of a generator in GANs helps to generate expected normal behavior, while a discriminator can distinguish between “normal” and “abnormal” behaviors. GAN technology has led to exciting new developments in deep learning. Generative adversarial networks (GAN) are an innovative form of AI that offer a powerful solution to the generative modeling problem. GAN is composed of two models - a generator used to create normal behavior and a discriminator used to distinguish between normal and abnormal behaviors. When dealing with imbalanced industrial time series data, GAN can be applied to derive an anomaly detection architecture that outperforms classic algorithms and other deep learning models such as big-GAN, ANOGAN and DBN [95]. The attached article further elaborates on the inner workings of GANs and their core design considerations. Additionally, drawing from research conducted by Li et al. [96], this architecture can feature a dynamic threshold generated by the discriminator which serves as a predictive warning for system failures or anomalies. GAN based approach is used to diagnose faults by generating much higher anomaly scores when a fault sample is fed into the trained model [95].

Moving on to next one GTAD, researchers have developed a new anomaly detection algorithm for multivariate time series, called Graph Attention Network and Temporal Convolutional Network for Multivariate Time Series Anomaly Detection (GTAD). This algorithm takes into account the correlation and temporal dependencies that many other existing algorithms fail to address. GTAD promises to provide better results when it comes to spotting anomalies in complex data sets. GTAD is an unsupervised approach powered by graph attention networks and temporal convolutional networks [97].

TadGAN is a breakthrough in unsupervised anomaly detection that makes use of Generative Adversarial Networks (GANs). At the core of the system are Long Short-Term Memory (LSTM) Recurrent Neural Networks, which provide an excellent base model for creating Generators and Critics. TadGAN is unique in its ability to capture temporal correlations with cycle consistency loss for more accurate time-series data reconstruction [98].

**Future concerns:** Combining information between different dimensions of multivariate time series is a key focus of future work [95] in AI algorithms. When it comes to GAN-based anomaly detection models, there can be difficulties in determining



the right sliding window length and maintaining stability during training. Further research is needed in order to more effectively train GANs [50].

## 7.1 AI based toolkits for automated anomaly detectors

**TODS:** TODS is a comprehensive automated Time Series Outlier Detection System with a modular design that enables easy construction of pipelines. It includes a range of primitives for data processing, time series analysis, feature analysis, detection algorithms and reinforcement methods. This makes TODS suitable for both research and industrial applications [99, 100].

**ANOVIZ:** ANOVIZ is an innovative anomaly detection solution for multivariate time series. It provides you with accurate detections, as well as easy-to-use visualizations and user interfaces to promote better explanation and assessment of the quality of those detections [101].

**AnomalyKiTS:** AnomalyKiTS is a system that allows end users to detect anomalies in time series data. It provides a range of algorithms, as well as an enrichment module to label identified anomalies. AnomalyKiTS offers four categories of model building capabilities, enabling users to select the best option for their needs [102].

**TranAD:** TranAD is an advanced model designed to provide superior recognition and diagnosis results. Our proprietary focus score-based self-conditioning and adversarial training technology extract multi-modal features, while MAML ensures quick and efficient on-the-fly training with minimal data. With TranAD, you get the best of both worlds: powerful detection capabilities and superior performance. TranAD has been proven to outperform existing baseline methods [46]. There is a range of data sizes, formats, and anomalies to consider when deciding which anomaly detection toolkit to use.

**Future Scrutinizes:** To maintain the quality of pipeline discovery system, researchers are planning on adding more primitives in the future as well as improved integral searchers to ensure optimal performance. To incorporate predefined rules efficiently into pipelines, researchers should also aim to develop learning-based active learning techniques for our reinforcement module. Existing solutions may not be comprehensive enough for certain applications, such as scenarios where semi-supervised or prediction-based unsupervised anomaly methods are needed.

The current research focused on using machine learning techniques to detect anomalies before they arise in future forecasting, leveraging stacked and bidirectional LSTM. The analysis produced promising results as noted in reference [103], validating the use of such models for anomaly detection. The review of AI-based energy monitoring and anomaly detection commercial solutions for buildings [104] provides an overview of the available systems. Efficient predictive maintenance of equipment in various industries requires the detection of anomalies in time-varying multivariate data. Researchers presented MTV (Multivariate Time Series Visualization), a visual analytics system that helps to streamline collaboration between humans and AI, facilitating the most ideal workflow [105].

## 8. Conclusion

In conclusion, this chapter has provided an insight into the cutting-edge models for anomaly detection in time series, discussed their merits and pitfalls, and highlighted new areas of research that are being explored to solve unique problems posed

by high-dimensional and complex data, high-volume data streams, and a need for real-time processing. These research areas have provided concrete examples of the applications of discussed models. Moreover, citations will help readers about how these models can be used in real world scenarios. We have also identified some of the current issues and suggested future directions for research concerning anomaly detection systems.

As the field of anomaly detection in time series continues to evolve and new challenges arise, it is crucial that researchers remain focused on developing innovative solutions that can effectively process high-dimensional and complex data in real-time. By better understanding the existing state-of-the-art models and the challenges that still need to be addressed, researchers can identify new opportunities for developing more effective anomaly detection systems.

We have not explored all the current algorithms, models, and new research areas. This chapter has just provided readers an overview of some current techniques and areas so that they can identify what is going on exactly in this field. Interested readers will definitely go out and do some more research about it and prepare themselves for further development within this growing field.

## **Acknowledgements**

The authors acknowledge the editor of the book for his support throughout the writing process.

## **Conflict of interest**

The authors declare no conflict of interest.

## **Notes/thanks/other declarations.**

The authors would like to thank the Editor of the book and the publisher for giving them a valuable opportunity to prepare a book chapter.

IntechOpen

IntechOpen


### **Author details**

Farrukh Arslan\*, Aqib Javaid, Muhammad Danish Zaheer Awan and Ebad-ur-Rehman  
Department of Electrical Engineering, University of Engineering and Technology,  
Lahore, Pakistan

\*Address all correspondence to: farrukh\_arslan@uet.edu.pk

### **IntechOpen**

---

© 2023 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Sen PC, Hajra M, Ghosh M. Supervised classification algorithms in machine learning: A survey and review. In: *Advances in Intelligent Systems and Computing*. Singapore: Springer Singapore; 2020. pp. 99-111
- [2] Ezugwu AE, Ikotun AM, Oyelade OO, Abualigah L, Agushaka JO, Eke CI, et al. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*. 2022;**110**(104743):104743
- [3] Petropoulos F, Apiletti D, Assimakopoulos V, Babai MZ, Barrow DK, Ben Taieb S, et al. Forecasting: Theory and practice. *International Journal of Forecasting*. 2022;**38**(3):705-871
- [4] Ratanamahatana CA, Lin J, Gunopulos D, Keogh E, Vlachos M, Das G. Mining time series data. In: *Data Mining and Knowledge Discovery Handbook*. Boston: Springer; 2009. pp. 1049-1077
- [5] Fu T-C. A review on time series data mining. *Engineering Applications of Artificial Intelligence*. 2011;**24**(1):164-181
- [6] Esling P, Agon C. Time-series data mining. *ACM Computing Surveys*. 2012;**45**(1):1-34
- [7] Hilal W, Gadsden SA, Yawney J. Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*. 2022;**193**(116429):116429
- [8] Šabić E, Keeley D, Henderson B, Nannemann S. Healthcare and anomaly detection: Using machine learning to predict anomalies in heart rate data. *AI & Society*. 2021;**36**(1):149-158
- [9] Sharma B, Sharma L, Lal C. Anomaly detection techniques using deep learning in IoT: A survey. In: *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*. Dubai, United Arab Emirates: IEEE; 2019. pp. 146-149
- [10] Gupta M, Gao J, Aggarwal CC, Han J. Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and Data Engineering*. 2014;**26**(9):2250-2267
- [11] Fox AJ. Outliers in time series. *Journal of the Royal Statistical Society*. 1972;**34**(3):350-363
- [12] Bommert A, Sun X, Bischl B, Rahnenführer J, Lang M. Benchmark for filter methods for feature selection in high-dimensional classification data. *Computational Statistics and Data Analysis*. 2020;**143**(106839):106839
- [13] Big data basics for digital marketers [Internet]. Gartner. [cited 28 April 2023]. Available from: <https://www.gartner.com/en/marketing/insights/articles/big-data-basics-for-digital-marketers>
- [14] Blázquez-García A, Conde A, Mori U, Lozano JA. A review on outlier/anomaly detection in time series data. *ACM Computing Surveys*. 2022;**54**(3):1-33
- [15] Ahmad S, Purdy S. Real-time anomaly detection for streaming analytics. *arXiv [cs.AI]*. 2016
- [16] Barbariol T, Chiara FD, Marcato D, Susto GA. A review of tree-based approaches for anomaly



detection. In: Springer Series in Reliability Engineering. Cham: Springer International Publishing; 2022. pp. 149-185

[17] Nassif AB, Talib MA, Nasir Q, Dakalbab FM. Machine learning for anomaly detection: A systematic review. IEEE Access. 2021;**9**:78658-78700

[18] Schmidl S, Wenig P, Papenbrock T. Anomaly detection in time series: A comprehensive evaluation. Proceedings VLDB Endowment. 2022;**15**(9):1779-1797

[19] Kozitsin V, Katser I, Lakontsev D. Online forecasting and anomaly detection based on the ARIMA model. Applied Sciences (Basel). 2021;**11**(7):3194

[20] Tang H, Wang Q, Jiang G. Time series anomaly detection model based on multi-features. Computational Intelligence and Neuroscience. 2022;**2022**:2371549

[21] Xu H, Pang G, Wang Y, Wang Y. Deep isolation forest for anomaly detection. arXiv [cs.LG]. 2022

[22] Thill M, Konen W, Wang H, Bäck T. Temporal convolutional autoencoder for unsupervised anomaly detection in time series. Applied Soft Computing. 2021;**112**(107751):107751

[23] Fan J, Han F, Liu H. Challenges of big data analysis. National Science Review. 2014;**1**(2):293-314

[24] Toledano M, Cohen I, Ben-Simhon Y, Tadeski I. Real-time anomaly detection system for time series at scale. In: Anandakrishnan A, Kumar S, Statnikov A, Faruquie T, Xu D, editors. Proceedings of the KDD 2017: Workshop on Anomaly Detection in Finance. PMLR; 2018. pp. 56-65

[25] Mason A, Zhao Y, He H, Gompelman R, Mandava S. Online

anomaly detection of time series at scale. In: 2019 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA). Oxford, UK: IEEE; 2019. pp. 1-8

[26] Ranjan KG, Tripathy DS, Prusty BR, Jena D. An improved sliding window prediction-based outlier detection and correction for volatile time-series. International Journal of Numerical Modelling. 2021;**34**(1):e2816

[27] Zhai Y, Ong Y-S, Tsang IW. The emerging “big dimensionality”. IEEE Computational Intelligence Magazine. 2014;**9**(3):14-26

[28] McNeil P, Shetty S, Guntu D, Barve G. SCREDDENT: Scalable real-time anomalies detection and notification of targeted malware in mobile devices. Procedia Computer Science. 2016;**83**:1219-1225

[29] Lopez MA, Gonzalez Pastana Lobato A, Duarte OCMB, Pujolle G. An evaluation of a virtual network function for real-time threat detection using stream processing. In: 2018 Fourth International Conference on Mobile and Secure Services (MobiSecServ), Miami Beach, FL, USA; 2018. pp. 1-5. DOI: 10.1109/MOBISECSERV.2018.8311440

[30] Goncalves D, Bota J, Correia M. Big data analytics for detecting host misbehavior in large logs. In: 2015 IEEE Trustcom/BigDataSE/ISPA. Helsinki, Finland: IEEE; 2015

[31] Cui B, He S. Anomaly detection model based on Hadoop platform and weka interface. In: 2016 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS). Fukuoka, Japan: IEEE; 2016. pp. 84-89

[32] Rettig L, Khayati M, Cudré-Mauroux P, Piorkowski M. Online Anomaly

- Detection over Big Data Streams. In: Braschler M, Stadelmann T, Stockinger K, editors. *Applied Data Science*. Cham: Springer; 2019. DOI: 10.1007/978-3-030-11821-1\_16
- [33] Liu X, Nielsen PH. Regression-Based Online Anomaly Detection for Smart Grid Data. arXiv (Cornell University); 2016
- [34] Xie S, Chen Z. Anomaly detection and redundancy elimination of big sensor data in Internet of things [Internet]. arXiv [cs.DC]. 2017
- [35] Bhadani AK, Jothimani D. Big data: Challenges, opportunities, and realities. In: *Effective Big Data Management and Opportunities for Implementation*. IGI Global; 2016. pp. 1-24
- [36] Hashem IAT, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Ullah KS. The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*. 2015;47:98-115
- [37] Chandola V, Banerjee A, Kumar V. Anomaly detection. *ACM Computing Surveys*. 2009;41(3):1-58
- [38] Erfani SM, Rajasegarar S, Karunasekera S, Leckie C. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. *Pattern Recognition*. 2016;58:121-134
- [39] Mirsky Y, Shabtai A, Shapira B, Elovici Y, Rokach L. Anomaly detection for smartphone data streams. *Pervasive and Mobile Computing*. 2017;35:83-107
- [40] Sarker RA, Elsayed SM, Ray T. Differential evolution with dynamic parameters selection for optimization problems. *IEEE Transactions on Evolutionary Computation*. 2014;18(5):689-707
- [41] Akoglu L, Tong H, Koutra D. Graph-based anomaly detection and description: A survey [Internet]. arXiv [cs.SI]. 2014
- [42] Katal A, Wazid M, Goudar RH. Big data: Issues, challenges, tools and good practices. In: *2013 Sixth International Conference on Contemporary Computing (IC3)*. Noida, India: IEEE; 2013. pp. 404-409
- [43] Shiravi H, Shiravi A, Ghorbani AA. A survey of visualization systems for network security. *IEEE Transactions on Visualization and Computer Graphics*. 2012;18(8):1313-1329
- [44] Vaswani A, Shazeer NM, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention Is all you Need. *NIPS* [Internet]; 2017 Available from: <https://www.semanticscholar.org/paper/Attention-is-All-you-Need-Vaswani-Shazeer/204e3073870fae3d05bcbc2f6a8e263d9b72e776>
- [45] Xu J, Wu H, Wang J, Long M. Anomaly Transformer: Time series anomaly detection with Association Discrepancy [Internet]. arXiv [cs.LG]. 2021 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/2110.02642>
- [46] Tuli S, Casale G, Jennings NR. TranAD: Deep transformer networks for anomaly detection in multivariate time series data [Internet]. arXiv [cs.LG]. 2022 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/2201.07284>
- [47] Wang X, Pi D, Zhang X, Liu H, Guo C. Variational transformer-based anomaly detection approach for multivariate time series. *Measurement (Lond)* [Internet]. 2022;191(110791):110791 Available from: <https://www.sciencedirect.com/science/article/pii/S0263224122000914>
- [48] Zhang H, Xia Y, Yan T, Liu G. Unsupervised anomaly detection

- in multivariate time series through transformer-based variational autoencoder. In: 2021 33rd Chinese Control and Decision Conference (CCDC). Kunming, China: IEEE; 2021. pp. 281-286
- [49] Kingma DP, Welling M. Auto-Encoding Variational Bayes [Internet]. arXiv [stat.ML]. 2013 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/1312.6114>
- [50] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Networks [Internet]. arXiv [stat.ML]. 2014 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/1406.2661>
- [51] Chen Z, Chen D, Zhang X, Yuan Z, Cheng X. Learning graph structures with Transformer for multivariate time series anomaly detection in IoT [Internet]. arXiv [cs.LG]. 2021 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/2104.03466>
- [52] Wang C, Xing S, Gao R, Yan L, Xiong N, Wang R. Disentangled dynamic deviation transformer networks for multivariate time series anomaly detection. Sensors (Basel) [Internet]. 2023 [cited 28 April 2023];23(3):1104. Available from: <https://www.mdpi.com/1424-8220/23/3/1104>
- [53] Wu B, Fang C, Yao Z, Tu Y, Chen Y. Decompose auto-transformer time series anomaly detection for network management. Electronics [Internet]. 2023;12(2):354. DOI: 10.3390/electronics12020354
- [54] Zhou T, Ma Z, Wen Q, Wang X, Sun L, Jin R. FEDformer: Frequency Enhanced Decomposed Transformer for long-term series forecasting [Internet]. arXiv [cs.LG]. 2022 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/2201.12740>
- [55] Wu H, Xu J, Wang J, Long M. Autoformer: Decomposition Transformers with Auto-Correlation for long-term series forecasting [Internet]. arXiv [cs.LG]. 2021 [cited 28 April 2023]. pp. 22419-22430. Available from: <https://proceedings.neurips.cc/paper/2021/hash/bcc0d400288793e8bdcd7c19a8ac0c2b-Abstract.html>
- [56] Zhang Y, Yan J. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting [Internet]. 2023 [cited 28 April 2023]. Available from: <https://openreview.net/pdf?id=vSVLM2j9eie>
- [57] Nie Y, Nguyen NH, Sinthong P, Kalagnanam J. A time series is worth 64 words: Long-term forecasting with transformers [Internet]. arXiv [cs.LG]. 2022 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/2211.14730>
- [58] Cai L, Janowicz K, Mai G, Yan B, Zhu R. Traffic transformer: Capturing the continuity and periodicity of time series for traffic forecasting. Transactions in GIS [Internet]. 2020 [cited 28 April 2023];24(3):736-755 Available from: <https://research-information.bris.ac.uk/en/publications/traffic-transformer-capturing-the-continuity-and-periodicity-of-t>
- [59] Xu M, Dai W, Liu C, Gao X, Lin W, Qi G-J, et al. Spatial-Temporal Transformer Networks for traffic flow forecasting [Internet]. arXiv [eess.SP]. 2020 [cited 28 April 2023]. Available from: <https://paperswithcode.com/paper/spatial-temporal-transformer-networks-for>
- [60] Li L, Yao J, Wenliang L, He T, Xiao T, Yan J, et al. GRIN: Generative relation and intention network for multi-agent trajectory prediction. Advances in Neural Information Processing Systems [Internet]. 2021



- [cited 28 April 2023];34:27107-27118. Available from: <https://proceedings.neurips.cc/paper/2021/hash/e3670ce0c315396e4836d7024abcf3dd-Abstract.html>
- [61] Ding C, Sun S, Zhao J. MST-GAT: A multimodal spatial-temporal graph attention network for time series anomaly detection. *Information Fusion* [Internet]. 2023;89:527-536 Available from: <https://www.sciencedirect.com/science/article/pii/S156625352200104X>
- [62] Devlin J, Chang M-W, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North. Stroudsburg, PA, USA: Association for Computational Linguistics*; 2019. pp. 4171-4186
- [63] Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language Models are Few-Shot Learners [Internet]. *arXiv [cs.CL]*. 2020 [cited 28 April 2023]. p. 1877-901. Available from: <https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfc4967418bfb8ac142f64a-Abstract.html>
- [64] Chen H, Wang Y, Guo T, Xu C, Deng Y, Liu Z, et al. Pre-trained image processing transformer. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2021. pp. 12299-12310
- [65] Zerveas G, Jayaraman S, Patel D, Bhamidipaty A, Eickhoff C. A Transformer-based Framework for Multivariate Time Series Representation Learning. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2021
- [66] Yang C-HH, Tsai Y-Y, Chen P-Y. In: Meila M, Zhang T, editors. *arXiv [cs.LG] Voice2Series: Reprogramming Acoustic Models for Time Series Classification* [Internet]. 2021. pp. 11808-11819 Available from: <https://proceedings.mlr.press/v139/yang21j.html>
- [67] Wu Z, Liu Z, Lin J, Lin Y, Han S. Lite Transformer with Long-Short Range Attention [Internet]. *arXiv [cs.CL]*. 2020 [cited 28 April 2023]. Available from: [https://iclr.cc/virtual\\_2020/poster\\_ByeMPIHKPH.html](https://iclr.cc/virtual_2020/poster_ByeMPIHKPH.html)
- [68] Mehta S, Ghazvininejad M, Iyer S, Zettlemoyer L, Hajishirzi H. DeLight: Deep and Light-weight Transformer [Internet]. *openreview.net*. 2023 [cited 28 April 2023]. Available from: <https://openreview.net/forum?id=ujmgfuxSLrO>
- [69] Bapna A, Chen MX, Firat O, Cao Y, Wu Y. Training deeper neural machine translation models with transparent attention [Internet]. *arXiv [cs.CL]*. 2018 [cited 28 April 2023]. Available from: <https://aclanthology.org/D18-1338.pdf>
- [70] Dehghani M, Gouws S, Vinyals O, Łukasz JU, Google K, Google B. UNIVERSAL TRANSFORMERS [Internet]. *Arxiv.org*. [cited 28 April 2023]. Available from: <http://arxiv.org/abs/1807.03819v3>
- [71] Xin J, Tang R, Lee J, Yu Y, Lin J. DeeBERT: Dynamic early exiting for accelerating BERT inference. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Stroudsburg, PA, USA: Association for Computational Linguistics; 2020. pp. 2246-2251
- [72] Dai Z, Yang Z, Yang Y, Carbonell J, Le QV, Salakhutdinov R. Transformer-XL: Attentive language models beyond a fixed-length context [Internet]. *arXiv [cs.LG]*. 2019 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/1901.02860>
- [73] So DR, Liang C, Le QV. The Evolved Transformer [Internet]. *arXiv [cs.LG]*.



2019 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/1901.11117>

[74] Chen M, Peng H, Fu J, Ling H. AutoFormer: Searching transformers for visual recognition. In: In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE; 2021. pp. 12270-12280

[75] Tkach V, Kudin A, Kebande VR, Baranovskyi O, Kudin I. Non-pattern-based anomaly detection in time-series. Electronics (Basel) [Internet]. 2023 [cited 28 April 2023];12(3):721 Available from: <https://www.mdpi.com/2079-9292/12/3/721>

[76] Zhou L, Zeng Q, Li B. Hybrid anomaly detection via multihead dynamic graph attention networks for multivariate time series. IEEE Access [Internet]. 2022;10:40967-40978 Available from: <https://ieeexplore.ieee.org/abstract/document/9758699/>

[77] Nguyen HD, Tran KP, Thomassey S, Hamad M. Forecasting and anomaly detection approaches using LSTM and LSTM autoencoder techniques with the applications in supply chain management. International Journal of Information Management [Internet]. 2021;57(102282):102282 Available from: <https://www.sciencedirect.com/science/article/pii/S026840122031481X>

[78] Mathonsi T, van Zyl TL. Statistics and deep learning-based hybrid model for interpretable anomaly detection [Internet]. arXiv [cs.LG]. 2022 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/2202.12720>

[79] Nizam H, Zafar S, Lv Z, Wang F, Hu X. Real-time deep anomaly detection framework for multivariate time-series data in industrial IoT. IEEE Sensors

Journal [Internet]. 2022;22(23):22836-22849 Available from: <https://ieeexplore.ieee.org/abstract/document/9915308/>

[80] Choi T, Lee D, Jung Y, Choi H-J. Multivariate time-series anomaly detection using SeqVAE-CNN hybrid model. In: 2022 International Conference on Information Networking (ICOIN). Jeju-si, Korea: IEEE; 2022. pp. 250-253

[81] Lin S, Clark R, Birke R, Schonborn S, Trigoni N, Roberts S. Anomaly detection for time series using VAE-LSTM hybrid model. In: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain: IEEE; 2020. pp. 4322-4326

[82] Presekale A, Stefanov A, Rajkumar VS, Palensky P. Attack graph model for cyber-physical power systems using hybrid deep learning. IEEE Transactions on Smart Grid [Internet]. 2023;1-1 Available from: <https://ieeexplore.ieee.org/abstract/document/10017381/>

[83] Terbuch A, O'Leary P, Khalili-Motlagh-Kasmaei N, Auer P, Zohrer A, Winter V. Detecting anomalous multivariate time-series via hybrid machine learning. IEEE Transactions on Instrumentation and Measurement [Internet]. 2023;72:1-11 Available from: <https://ieeexplore.ieee.org/abstract/document/10015855/>

[84] Boloka T, Crafford G, Mokuwe W, Van Eden B. Anomaly detection monitoring system for healthcare. In: 2021 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA). Potchefstroom, South Africa: IEEE; 2021. pp. 1-6

- [85] Luo A, Yang F, Li X, Nie D, Jiao Z, Zhou S, et al. Hybrid graph neural networks for crowd counting. *Proceedings of the AAAI Conference on Artificial Intelligence* [Internet]. 2020 [cited 28 April 2023];**34**(07):11693-11700 Available from: <https://ojs.aaai.org/index.php/AAAI/article/view/6839>
- [86] Goldstein M, Uchida S. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLoS One* [Internet]. 2016;**11**(4):e0152173. DOI: 10.1371/journal.pone.0152173
- [87] Karadayi Y, Aydin MN, Ogrenci AS. Unsupervised anomaly detection in multivariate spatio-temporal data using deep learning: Early detection of COVID-19 outbreak in Italy. *IEEE Access* [Internet]. 2020;**8**:164155-164177 Available from: <https://ieeexplore.ieee.org/abstract/document/9187620/>
- [88] Zoph B, Le QV. Neural architecture search with reinforcement learning [Internet]. *arXiv [cs.LG]*. 2016 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/1611.01578>
- [89] Kurakin A, Goodfellow I, Bengio S. Adversarial machine learning at scale [Internet]. *arXiv [cs.CV]*. 2016 [cited 28 April 2023]. Available from: <http://arxiv.org/abs/1611.01236>
- [90] Zhang R, Zou Q. Time series prediction and anomaly detection of light curve using LSTM neural network. *Journal of Physics: Conference Series*. 2018;**1061**:012012
- [91] von Schleinitz J, Graf M, Trutschnig W, Schröder A. VASP: An autoencoder-based approach for multivariate anomaly detection and robust time series prediction with application in motorsport. *Engineering Applications of Artificial Intelligence* [Internet]. 2021;**104**(104354):104354 Available from: <https://www.sciencedirect.com/science/article/pii/S0952197621002025>
- [92] Haris M, Sharif U, Gupta K, Mohammed A, Jiwani N. Anomaly detection in time series using deep learning [Internet]. *Ijeast.com*. [cited 28 April 2023]. Available from: <https://www.ijeast.com/papers/296-305%20Tesda0706.pdf>
- [93] Ren H, Xu B, Wang Y, Yi C, Huang C, Kou X, et al. Time-series anomaly detection service at Microsoft. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York, NY, USA: ACM; 2019
- [94] Batarseh FA, Freeman L, Huang C-H. A survey on artificial intelligence assurance. *Journal of Big Data* [Internet]. 2021;**8**(1):60. DOI: 10.1186/s40537-021-00445-7
- [95] Jiang W, Hong Y, Zhou B, He X, Cheng C. A GAN-based anomaly detection approach for imbalanced industrial time series. *IEEE Access* [Internet]. 2019;**7**:143608-143619 Available from: <https://ieeexplore.ieee.org/abstract/document/8853246/>
- [96] Li D, Chen D, Jin B, Shi L, Goh J, Ng S-K. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. In: *Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series*. Cham: Springer International Publishing; 2019. pp. 703-716
- [97] He Y, Zhao J. Temporal convolutional networks for anomaly detection in time series. *Journal of Physics: Conference Series*. 2019;**1213**:042050
- [98] Geiger A, Liu D, Alnegheimish S, Cuesta-Infante A, Veeramachaneni K.

- TadGAN: Time series anomaly detection using generative adversarial networks. In: 2020 IEEE International Conference on Big Data (Big Data). Atlanta, GA, USA: IEEE; 2020. pp. 33-43
- [99] Lai K-H, Zha D, Wang G, Xu J, Zhao Y, Kumar D, et al. TODS: An automated time series outlier detection system. Proceedings of the AAAI Conference on Artificial Intelligence [Internet]. 2021 [cited 28 April 2023];**35**(18):16060-16062 Available from: <https://ojs.aaai.org/index.php/AAAI/article/view/18012>
- [100] Milutinovic M, Schoenfeld B, Martinez-Garcia D, Ray S, Shah S, Yan D. On Evaluation of AutoML Systems [Internet]. Automl.org. [cited 28 April 2023]. Available from: [https://www.automl.org/wp-content/uploads/2020/07/AutoML\\_2020\\_paper\\_59.pdf](https://www.automl.org/wp-content/uploads/2020/07/AutoML_2020_paper_59.pdf)
- [101] Trirat P, Nam Y, Kim T, Lee J-G. ANOVIZ: A visual inspection tool of anomalies in multivariate time series [Internet]. Github.io. [cited 28 April 2023]. Available from: [https://itouchz.github.io/files/AnoViz\\_AAAI23.pdf](https://itouchz.github.io/files/AnoViz_AAAI23.pdf)
- [102] Patel D, Ganapavarapu G, Jayaraman S, Lin S, Bhamidipaty A, Kalagnanam J. AnomalyKiTS: Anomaly detection toolkit for time series. Proceedings of the AAAI Conference on Artificial Intelligence [Internet]. 2022 [cited 28 April 2023];**36**(11):13209-13211 Available from: <https://ojs.aaai.org/index.php/AAAI/article/view/21730>
- [103] Girish L, Rao SKN. Anomaly detection in cloud environment using artificial intelligence techniques. Computing [Internet]. 2023;**105**(3):675-688. DOI: 10.1007/s00607-021-00941-x
- [104] Himeur Y, Ghanem K, Alsalemi A, Bensaali F, Amira A. Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. Applied Energy [Internet]. 2021;**287**(116601):116601 Available from: <https://www.sciencedirect.com/science/article/pii/S0306261921001409>
- [105] Liu D, Alnegheimish S, Zytek A, Veeramachaneni K. MTV: Visual analytics for detecting, investigating, and annotating anomalies in multivariate time series. Proceedings of the ACM on Human-Computer Interaction [Internet]. 2022;**6**(CSCW1):1-30. DOI: 10.1145/3512950