**Library Progress International** *Print version ISSN 0970 1052*
*Vol.44 No. 3, Jul-Dec 2024: P.13967-13993* *Online version ISSN 2320 317X*

*Original Article* *Available online at www.bpasjournals.com*

# Anomaly Detection in Fog Computing: State-of-the-Art Techniques, applications, Challenges, and Future Directions

**Girish Wali[1], Dr. Chetan Bulla [2]**

[1] Senior Vice President CitiBank, waligirish@gmail.com
ORCID ID: https://orcid.org/0009-0006-3213-3168

[2] Senior Engineer, Empower, bulla.chetan@gmail.com
ORCID iD: https://orcid.org/0000-0002-6463-0509

.**How to cite this article:** Girish Wali1, Dr. Chetan Bulla (2024) Anomaly Detection in Fog Computing: State-of-the-Art Techniques, applications, Challenges, and Future Directions. *Library Progress International,* 44(3), 13967-13993.

**ABSTRACT**
The fog computing provides a platform for various time critical real-time applications. It reduces the communication latency by placing computational resources to near to IoT devices. The anomaly detection in fog computing is very essential aspect to improve the quality of service. An anomaly is a value, a status of resources or outcome that deviates from expected or normal values which affects the Quality of services. Anomaly detection is an important data analysis task which is useful for identifying various issues in fog computing environment. This paper presents detailed study of anomaly detection key factors, general architecture and anomaly detection techniques. It describes the state of the art anomaly detection techniques with comparative analysis. Finally, various applications and challenges/issues with future directions are discussed. This paper helps researchers, engineers and scientists to know the state-of-the-art works for anomaly detection techniques in fog computing environment.

**Keywords:** Fog computing, Anomaly detection, Key factors, Applications and Challenges

## I. Introduction

The new developing technology called Internet of Things (IoT) links many mechanical devices, digital machines, electrical and electronics and other kinds of computing equipment's [1]. Every one of these devices has various kinds of sensors that sense the data and distribute it over the cloud /fog infrastructure for processing free from human-to--computer contact. Originally presented for industry and business use, IoT today finds workplaces, residences, and hospitals. This shift results mostly from the use of smart gadgets in daily life. These smart devices may operate in any environment, are less expensive, simple to use, and have low maintenance expenses. Shorter reaction times and mobility support—such as smart cities and hospitals—allow time-sensitive IoT applications to exist. To fix the bandwidth and latency problem, Cisco has developed the fog computing whereby computational resources are transported to the network edge.
Mostly of the data processing and calculations in the distributed computing system known as fog computing are carried out in the virtualized and non-virtualized edge devices [2]. It is also connected with the cloud to handle heavy-weight processing and save big amounts of data. Modern paradigm known as fog computing stretches the traditional cloud computing architecture toward the edge of the network. While other applications are handled at cloud servers, fog nodes handles time-sensitive and crucial applications. minimal latency, geographically scattered applications, and minimal network bandwidth usage make fog computing well fit for IoT applications. Monitoring has become more important in enhancing the Quality of Service (QoS) [3] of the fog computing environment given the exponential expansion of the Internet of Things.

Various kinds of anomaly appear at several stages of fog computing architecture. Figure 1 presents the architecture of fog computing with various flaws. Three levels comprise it: terminal, fog, and cloud layers. Different sensors and actuators found in the terminal layer sense the surroundings and forward the data to the fog layer. Comprising little processing capability, the fog layer is an intermediary layer. The layer of fog analyzes data processing needs. Should a fog layer fail to process, it moves to the layer of clouds. Workload contentions, CPU hog, and memory leak plague the infrastructure supporting cloud and fog computing [10]. The CPU hog is the process wherein the CPU runs all of its cycles via endless loops or infections. Memory leaks suffer from memory exhausts by

introducing anomalies to utilize all memory capacity; workload congestion results in saturation in resource components by imitating user demands.
Most fog-based applications make use of IoT devices or sensors highly dependent on the precise operation, which is somewhat challenging to ensure. Cheap and prone to failure or inaccuracy are these sensors.

Furthermore, sensors are placed in possibly hostile environments, so their damage is significantly more likely. The abnormalities in the sensor have Spike, Drift, Noise and constant as their causes. The noise is a rise in the change of a series of consecutive data samples; the sensor is recorded as a constant value in a given situation; the drift is deviated in the computation of the sensor values. The Spike is a brief duration in the sensed values that differs from other observations. Data corruption or false readings most likely results from data transfer between sensors and fog nodes. A sensor anomaly is the great differences in sensed values and read data in fog nodes. The irregularities in fog computing lower performance, availability and dependability, increase shutdown time, maintenance cost, energy consumption, computational overhead, and change behavior. Thus, it is vital to find the abnormalities in order to raise the fog computing services' performance. Collecting different data and analyzing it to get the state of computing resources depends critically on a monitoring system [4].

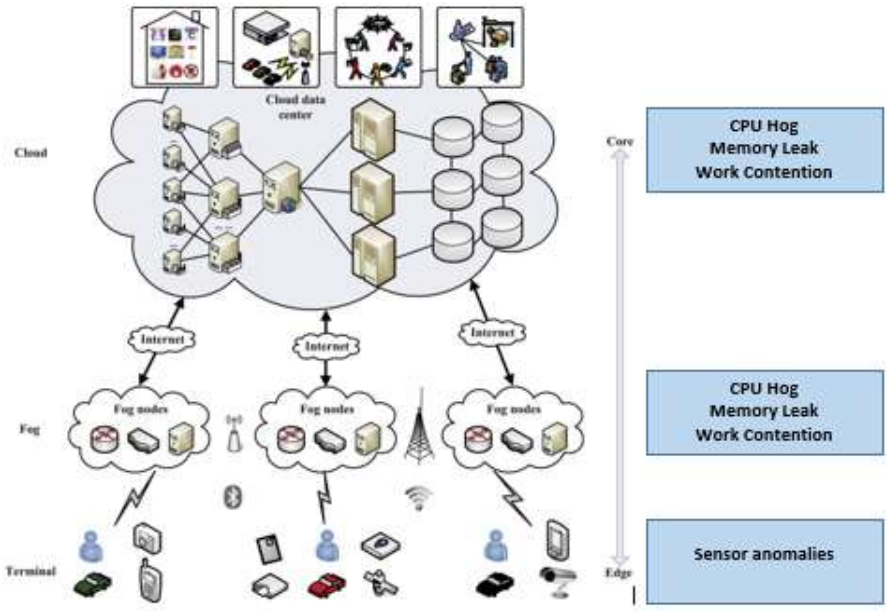


Fig 1 : Fog Computing Architecture with possible anomalies

Fig 2 simple example of anomalies. N1 and N1 are the normal regions where certain data belongs to these regions. The same data point that is far away from these regions are O1,O2, and O3 are anomalies

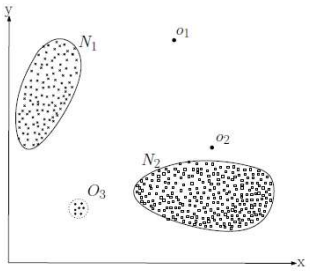


Fig 2: Simple example of anomalies

There are various reasons for anomalies in fog-enabled cloud computing environments [6]. The significant

reasons for anomalies are sensor malfunction, resource competition, resource bottleneck, misconfiguration, software defects, hardware failures, and external attacks in fog computing infrastructure. The Sensors are integrated with inexpensive hardware with limited Computational, communicational storage. They are located in harsh environments, so defective sensors may generate anomalous data on a continuous. Further, improper planning resource management may lead to deadlocks, resource bottleneck, or resource competition and generates anomalous data. The anomaly detection model should have the following characteristics to improve the performance of the machine

1. Online Predictions: The anomaly detection algorithm must identify normal or abnormal state before receiving the subsequent data. The online anomaly detection works in two stages i) Prediction and ii) Detection. The algorithm predict next n timesteps values and check for the difference between actual and predicted data. If the difference is more than it is considered as anomaly. These two action must perform parallely.
2. Live detection: The algorithm must learn continuously without the need to store the entire stream. Detection of live anomaly in a complex environment can boost unregulated breakdowns. The input data for anomaly detection model continuously and model detect the anomaly if there is certain difference between actual and predicted data.
3. Unsupervised algorithm: The training and testing model for anomaly detection algorithm should not contain label on input data. The unsupervised algorithms are more efficient compared to supervised algorithms but these are more complex and consume more time.
4. Dynamic: The model must be adapted to complex and dynamic conditions, so that data streaming and analytics produce proper results.
5. Robust: The model should detect anomalies within short period of time and minimize false positives and false negatives.

The rest of the paper is organized as follows: Basics of anomaly detection is introduced in section 2. In Section 3, the state-of-the-art of anomaly detection approaches are discussed. Section 4 presents the various application of anomaly detection algorithms. Section 5 discusses challenges in anomaly detection approaches and its future research directions. The Section 6 concludes.

## II. Factors of the anomaly detection problem

This section discusses the various aspects that are related to anomaly detection problems. The various factors are required to consider in designing the solutions to the anomaly detection problems.

### Affects of Anomalies

The anomalies deviate the normal operation of the system. It creates many problems in operation and maintenance. The various problem arises due to anomalies are as follows

1. Anomalous nodes in fog computing would cause abnormal latencies, data loss, and task failure etc on transmission paths. It leads to additional retries, and retransmissions of data can be taken, resulting in increased energy consumption and traffic in the communication link
2. It disrupts the regular operation of the cloud, increases resource computation cost and memory consumption. The anomalous value confuses the system operation and take into a different direction. For example an anomalous value: CPU utilization value 10% even actual value is 90%. So more resources are allocated when the system is overload state. After resource allocation, CPU could not handle all the request. At this time Performance is degraded.

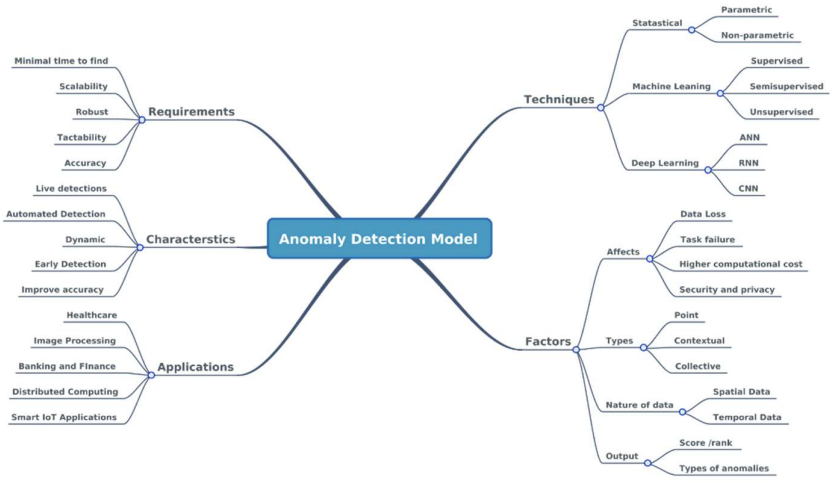


Fig 3 : Different aspects of Anomaly detection classifications

**Causes of anomalies**

a. **Data from different classes**: An item could belong to several kinds of class based on data from several backgrounds. Good instances of anomalies happening and found using class labels include cases like credit card fraud, Intrusion detection, outcome of sickness, erroneous test result.
b. **Natural Variation**: A Gaussian distribution rapidly lowers the probability of a data object. These are regarded as abnormalities. Others refer to these as outliers as well. Outliers and the center of the distribution have a distance far larger than that of any other data object. Every other item is close to the average object, or central of distribution.
a. ***Data Measurement and Collection Errors***: These sorts of errors happen if we gather false data or if data measurement deviates from the norm. Eliminating such mistakes would be the ultimate aim as they compromise the quality of the data. The issue now is, how can one lower such mistakes? One can do this by means of data pre-processing or data cleaning.

**Types of Anomalies**

The anomalies are classified into three types [5]. These are point anomaly, collective anomaly and contextual anomaly. Figure 2 shows the classification of anomalies. These anomalies are defined as follows:

*1. Point anomalies*: If a certain time-step value deviates considerably from the other data, it is considered a point anomaly. For example, the normal CPU utilization rate is in between 30 to 80, if CPU utilization rate for certain time-step 99 is considered as point anomaly
*2. Collective anomalies*: a sequence of points are deviates from normal values are considered as collective anomaly. For example: the more number of timesteps values of CPU utilization rate deviates largely from other values are considered collective anomalies.
*3. Contextual anomalies*: The certain variation in timesteps values are considered as anomaly in some application but not in other application. These types of variation are called contextual anomalies. Having a daily humidity in summer in one country is normal, while the same humidity in winter is observed as an anomaly.

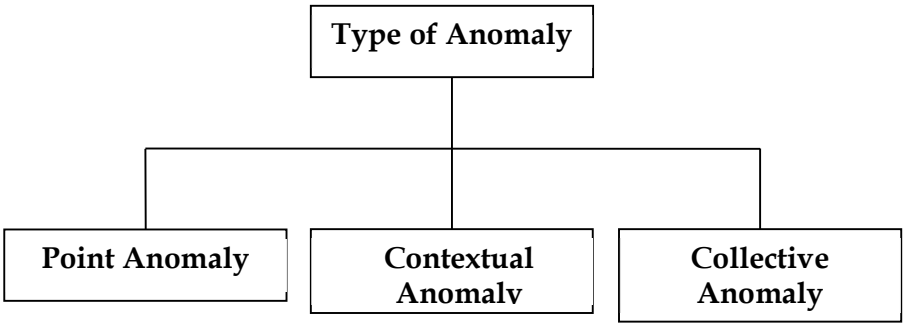


**Fig 2: Types of anomalies**

**Nature of Input Data**

The nature of input data is a significant factor in designed the anomaly detection models [5]. The input data is categorized into temporal and spatial data. The spatial data can be medical images, protein sequences, DNA structures etc. and the temporal data can be time-series data, like resource utilization, energy consumption, and sensors values at the particular timestamp. The data may be either equal or unequal interval with timestamps.

1. Data instances: a data instance can be described using a set of attributes or feature. The data instance can spatial, temporal or graph data.
2. The nature of attributes can identify the applicability of anomaly detection techniques. The attributes are categorized binary, categorical or continuous. Each data instance consist of either single attribute or multiple attributes. The single attributed value is called univariate and multiple attributed value is called multivariate data. The multivariate data either mixtures of different types of data or similar type of data. The data is also categorized based on its application like Evolving (Dynamic) data, streaming data, continuous data, Correlated data, contaminated data, and Big Data.

**Data Labels**

The labels associated with the data instance show whether that instance is normal or anomaly. It should be noticed that collecting labelled data that is reliable and representative of all forms of activity is always prohibitively costly [5]. Labeling is often performed manually by a human specialist and therefore takes significant work to produce a branded training data set. Usually, having a branded set of anomaly data instances to include all potential anomaly behavior is more difficult than getting labels for regular behavior. Moreover, anomaly behavior is also complex in nature, e.g. new forms of anomalies may occur, for which there are no labelled training data. In certain situations, such as air defense, anomalies would result in tragic accidents which would thus be extremely rare.

**Output of anomaly detection techniques**

The anomaly detection model either produce the percentage of anomaly or type of anomaly**.**

- **Score**: Anomaly score defines the level of outliers for each data point. The data instances may be ranked according to the anomalous score. Threshold based anomaly detection needs to be adopted and the threshold value is selected by subject matter expert to identify the anomalies.
- **Label/Binary**: The output of anomaly detection model is in terms of true or false / anomalous or normal data.
- **Type of anomaly:** The algorithm must classify and display its type; point anomaly or collective and There are three types of anomalies that discussed in previous topic.

**Dataset used for anomaly detection models**

When dealing with real-time anomaly detection problem, it is essential to validate the performance of proposed techniques. As anomalies occur unpredictably at any moment, it is a challenging task to generate them manually. To make avail the dataset, some academia and industry research domains capture anomalies on the field. In this section, the popular dataset that includes real-time and artificial data are discussed.

1. The **Numenta Anomaly Benchmark (NAB)** [56] provides time-series datasets for real-time anomaly detection in streaming AWS data produced by sensor systems. It includes different data such as system temperature, CPU usage, memory utilization, etc. It consists of over 50 real-time and artificial temporal data files and a novel scoring system optimized for real-time applications. NAB is a modular, open source programming archive. Numenta will work to create a group around NAB to incorporate data files and validate additional algorithms.

2. The **Bot-IoT dataset:** The BoT-IoT dataset [57] was created by modelling a real network system in the Cyber Range Lab of the UNSW Canberra Cyber Center. It contains a mixture of regular and botnet traffic. Dataset source files are supported in a variety of formats, including original pcap files, created argus files, and csv files. It is used specifically to verify an anomaly detection strategy aimed at separating normal from irregular flows within the network of sensors.

3. **Outlier Detection Datasets (ODDS)** It provides freely have access to a wide array of outlier exploration datasets of ground reality [58]. The main objective is to include datasets from various domains and to present them to the scientific community under a single umbrella. These dataset arranged based on their type.

4. **Yahoo Webscope:** The Yahoo Webscope Program [59] is a reference collection of fascinating and technically important datasets for non-commercial use by researchers and other scientists**.** It contains a labelled anomaly detection dataset consisting of temporal data and pair of timestamp and values. All databases have been checked to comply with Yahoo's data security requirements, including stringent privacy controls.

5. **UCI Machine Learning Repository** [60] The UCI Machine Learning Repository is a set of libraries, domain concepts, and data creators used by the machine learning applications for the evaluation of machine learning based anomaly detection algorithms. The collection was created by David Aha and fellow graduate students at UC Irvine as an ftp archive in 1987. Since then, teachers, educators and scholars around the world have been widely used as the main source of machine learning data sets. The 47 UCI data sets are available to publicly available to evaluate anomaly detection methods.

6. Unsupervised Anomaly Detection Benchmark (UADB) [61]: These datasets may be used to benchmark unsupervised anomaly detection algorithms. Datasets have been collected from various sources and are largely based on datasets primarily used for supervised machine learning. Through publishing these updates, a comparison of various algorithms is now feasible for unsupervised anomaly detection.

7. UCSD anomaly detection dataset [62]: The UCSD Anomaly Detection Dataset is collected of images captured with a stationary camera mounted on the top of traffic signals, overlooking the pedestrian walkways. The density of the crowd in the walkways was complex, varying from minimal to very pack. In standard situations, only pedestrians are included in the film. Abnormal incidents are triggered by either: Circulation by non-travelers in walkways or Anomaly in pedestrian orientation trends.
8. IEEE Dataset: It offers a sustainable framework for all data owners to support science and the broader goal of IEEE Advancing Technology for Humanity [51]. It provides various dataset for all the science and technology domains. The two datasets are presently available for use: Motion-Print Control Task Movements and Laser Heat Treatment Thermal Videos.

**Performance evaluation Parameters**

To evaluate the proposed anomaly detection model various performance parameters [63] used these are precision, recall, F1-score, accuracy and error rate etc. These parameters are defined in following sections:
(i) Precision, P, which is the ratio of true positives to the sum of true positives and false positives,

$$Precesion = \frac{True\ Positive}{True\ Positive + Fa \quad Positive} \quad (1)$$

(ii) Recall, R, which is the ratio of true positives to the sum of true positives and false negatives,

$$Recall = \frac{True\ Positive}{True\ Positive + Fals \quad Negative} \quad (2)$$

iii) R-precision [49]: It is the ratio of correctly predicted positive anomalies found within the first n possible anomalies, where n is the amount of true anomalies. As the number of true anomalies is very small due to the size of the dataset, the R-precision value would be very small.

iv) Average precision (AP) [50]: mean of precision scores in the ranks of all anomaly items. It mainly used to understand performance of anomaly detection models:

$$AP = \frac{1}{|a|} \sum_{t=1}^{|a|} P\,t \quad (3)$$

v) AUC [4]: The Receiver Operating Characteristic (ROC) curve is a graphical representation of the true positive rate and the false positive rate, where the true (false) positive rate indicates the percentage of anomalies rated between the highest possible anomalies.

(vi) F1-score, F1, it represents the accuracy of the model, which is calculated by taking mean of Precision and Recall. The F1-Score is defined as:

$$F1\ Score = \frac{2*(Precision* Recall)}{Precision \quad Recall} \quad (4)$$

The F1-score recapitulates both Precision and Recall, the highest F1 detection rule as the highest anomaly detection technique, true positives apply to the correctly predicted anomalies. False positives are the non-anomalies that wrongly classify as anomalies. False negatives refer to anomalies wrongly reported as non-anomalies.

The accuracy is another important key performance parameters to evaluate performance of the anomaly detection model. The accuracy is defined as follows

vii) $$Accuracy = \frac{TP+TN}{TP+TN+FP+F} \quad (5)$$

viii) $$Error\ Rate = \frac{FN+FP}{TP+FP+TN+FN} \quad (6)$$

In the optimization process, the error for the present state of the model must be continuously determined. This process involves the selection of the correct error function, commonly referred to as the loss function that can be used to calculate the loss of the model. Further, these weights can be changed in order to reduce the loss and increase output in the next evaluation.

ix) The RMSE is the right choice of time series prediction model, and it is defined as

$$Root\ Mean\ Squared\ Error = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y^i - y_p^i)^2} \quad (7)$$

The mean squared error is measured as the average squared difference between the expected and the actual values. The ideal value of RMSE is 0.0 which represents no errors in the prediction or highest accuracy. The result is always positive, regardless of the mark of the prediction. Squaring means that larger errors result in more error than smaller ones, which means that the model is punished for making larger mistakes.

x) Correlation coefficient: The similarity of rank of Spearman and the Pearson correlation, are also used as evaluation measurements. It puts more prominence on the possible anomalies ranked at the top by adding weights. It is very helpful to understand the relations between two objects.

xi) Rank power (RP): An anomaly ranking algorithm would be more effective if it places true anomalies in the top and usual findings in the bottom of the list of anomaly candidates. Ranking power is such a parameter and measures the overall ranking of true anomalies. The formal definition is as follows

$$Rank\ Power = \frac{n\,(n+1)}{2\sum_{i=1}^{n} Ri} \quad (8)$$

Where n is the number of anomalies in the highest possible objects, and Ri is the rank of the $i^{-th}$ true anomaly. For a fixed value of t, a higher value means better results. When the anomaly candidates are real exceptions, the power of rank is equal to one. Other output metrics are CPU usage, memory consumption, bandwidth consumption, detection time, etc.

**III. General Architecture of Anomaly Detection**

In this section, the general architecture and requirements of the anomaly detection model is discussed. Fig 3 shows

the architecture of anomaly detection models. Monitoring system plays a vital role in detecting anomalies in distributed computing. Various Roles of Monitoring in anomaly detection are Data Collection, System Observability, Profiling vs. Tracing, and Influence of sampling interval. The monitoring system collects the historical data and preprocesses it to improve the readability of data. Monitoring system collects the data by employing efficient data collection techniques. Then it preprocesses the data by adding missing values and removing unwanted data.

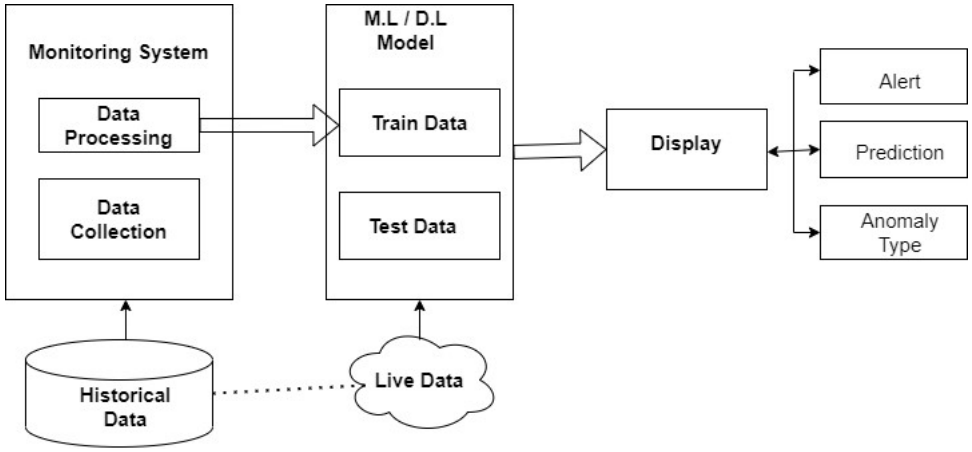


Fig 3: Anomaly detection general architecture

To improve the accuracy of anomaly detection techniques machine learning or deep learning algorithms are used. The statistical methods perform better but fail to provide accurate results. Machine learning algorithm trained by existing data that contains nominal and anomalous data. Training these data, anomaly detection model is developed. Once a model is developed, it is tested with test data to check the performance of developed model. If the performance is not up to the mark, then tune the model for better performance. Once the complete and accurate model is developed, real-time data is fed to identify anomalous data in real time data. Anomaly detection model will provide the results in terms of alert, prediction or information about type of alert. It alert user through monitoring system where it intimates user and service provider regarding anomalies in the system. To work towards proactive monitoring model, the fog computing predicts the future status certain resource health like CPU, memory, energy and bandwidth consumption in virtual machine. Further, the status or health of IoT devices needs to monitor to identify anomalies like temperature of the machine and its resources etc. So the predicting future system status and identifying anomalies in future system status improves the overall performance of system and avoid unconditional breakdowns. Figure 4 shows the overall steps in anomaly detection model.

**Requirement of Anomaly detection model**

1. ***Minimize false positives*:** A false positive is an error in the binary classification where the anomaly detection algorithm wrongly identify as anomaly. So it is essential to minimize these kind of error. The anomaly detection algorithm must identify anomalies accurately.
2. ***Alert within minutes*:** Once the anomalies are detection, it must notify or alert the users in the environment to avoid further consequences in short period of time. The anomaly detection algorithm should not take more time to alert.
3. ***Scale to millions of time series*:** The anomaly detection models should handle large amount of data. The IoT devices generates high volume of data so it is essential to handle these data effectively without comprising the performance.
4. ***Be robust to missing data***: Metrics may have missed data points at times. The most popular decomposition implementation (R's STL) does not support missing values, so we need to take care to allow these kinds of gaps in our decomposition process.
5. ***Take actionability into account*:** Many anomalies are far more significant and actionable than others. It is essential to frame rules to differentiate these anomalies and take actions or alerts based on application requirement.

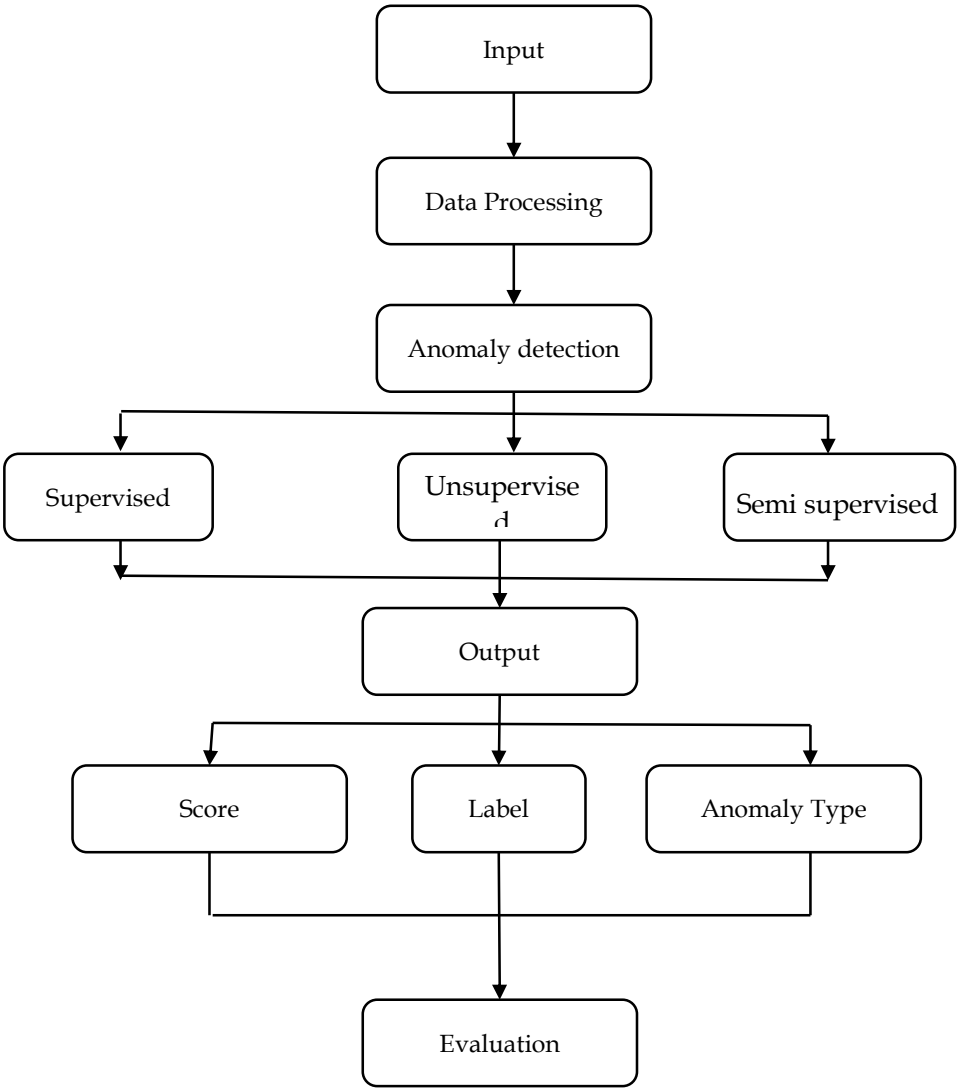


**Fig 4: Flow of anomaly detection model**

**Anomaly Detection Approaches**

The anomaly detection methods are classified into three categories based its revolution. Initially statistical methods are used to detect the anomalies. When machine learning and deep learning proved for highest accuracy in all the fields then machine learning and deep learning models are used. These three categories are Statistical approaches, ii) Machine Learning Approaches and iii) Deep learning approaches. Figure 5 shows the types of anomaly detection approaches. The statistical approaches used Regression based and distance based models. The machine approached uses cluster based models, density based models, distance based model and tree based model. Similarly, deep learning model used neural network based models. In this section, we discuss recent approaches for detecting anomalies.

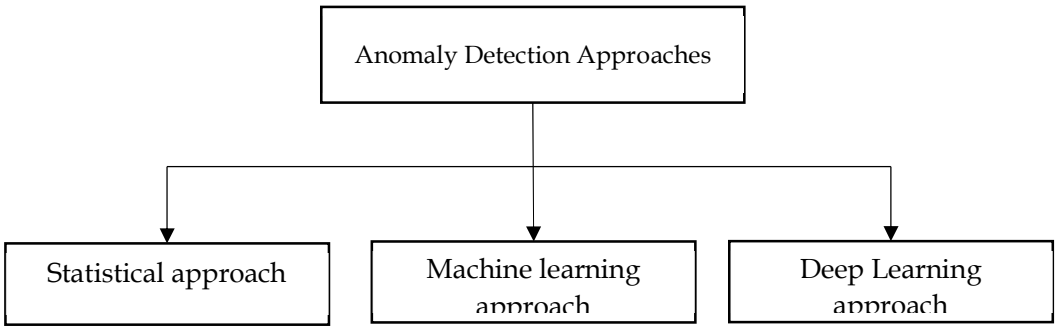


Fig 5: Anomaly detection approaches

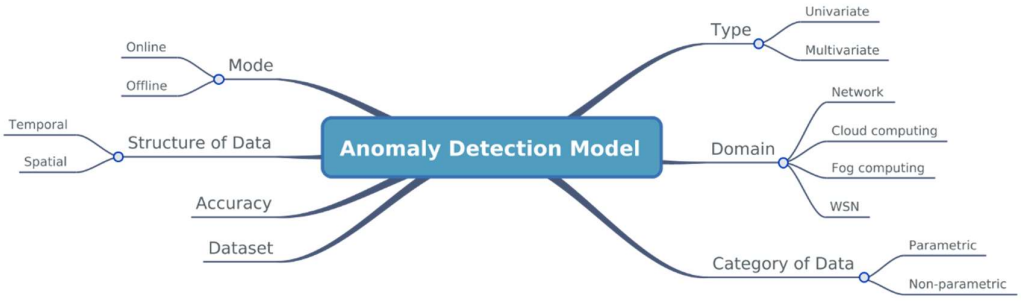


Fig 6: comparison parameters for anomaly detection algorithm

In this section, the various anomaly detection approaches are discusses with various parameters. Figure 6 shows the comparison parameters of various start-of-the-art anomaly detection algorithm. Any anomaly detection model have different working principle with various types of data depending on application and its requirement. The following section discusses three anomaly detection approaches.

1. **Statistical approaches:** The core theory of any statistical anomaly detection technique is: "An anomaly is an observation which is suspected of being partially or wholly irrelevant because it is not generate by the stochastic model assumed" Statistical anomaly detection strategies are based on the following fundamental assumption: Standard data instances locate a stochastic model with a high probability field, while anomalies have a low probability. There are two types of statistical approaches these are parametric and non-parametric approaches. Parametric and non-parametric approaches have been used to build predictive anomaly detection models. Parametric methods presume understanding of the underlying distribution and approximate the parameters from the given data while not parametric. The parametric models further categorized as Gaussian model based and regression model based. In the Gaussian based model, parameters are distributed using Gaussian distribution and estimated through Maximum likelihood estimated (MLE). The regression model based anomaly detection technique includes two steps. The first step is training the model to learn using the train data. In the second step is to evaluate learned model during training phase using test data to determine the anomaly Score. With confidence, the residual value can be used as an anomaly score for the test results.

   A combination of the graph spectral features and the restricted Vector Autoregressive (rVAR) anomaly detection model is presented [11]. The proposed model uses the non-randomness metric time series simulation approach derived from the spectral features of the graph to capture the irregular behaviours and experiences of individuals. The proposed model uses the non-randomness time series modelling method derived from the spectral data of the graph to identify the irregular behaviors and individual interaction. The Vector Autoregressive model used to evaluate multivariate results. It records the changes and interference of several variables over time, where each variable is described by the lagged values of itself and other variables. A time series sensor data monitor and anomaly detection in the thermal power plant is proposed [12]. The proposed model uses Regularized Vector Auto Regression, which is used to capture linear interdependencies between multiple time series. The benefit is that the RVAR model does not require too much previous knowledge of the forces that affect the component. The only prior information required is a list of variables

that can be expected to influence each other.

An automatic anomaly detection system [13] is given for high-frequency water quality data from in-situ sensors. Once user criteria have been established and abnormalities described, a suitable anomaly detection approach is chosen based on import rank. High priority deviations were eventually overcome by regression-based approaches such as self-regressive integrated moving average models. Incorporation of several water-quality variables as covariates decreased efficiency due to dynamic interactions between variables. Feature-based approaches also performed well on high-priority anomalies and were equally less proficient in identifying lower-priority anomalies, resulting in high false negative rates. Vector Autoregressive Model-Based Anomaly Detection scheme is proposed for Aviation Systems [14]. The proposed model is used for detecting anomalies over multivariate data using Vector autoregressive model. The autoregressive model calculate a distance matrix among the objects based on their respective vector autoregressive exogenous models. Next the dissimilarities are identified by distance matrix and consider these are anomalies. The author claim that proposed model is scalable and work well for online anomaly detection schemes.

An anomaly detection model for heterogeneous, multivariate, variable-length time series [15] is proposed. The proposed model is used for the aviation safety domain, where data objects are flights and temporal sensor readings, and pilot switches. The main objective of proposed model is to detect anomalous flight segments, due to various factors like mechanical, environmental, or human in order to identifying operationally significant events and highlight potential safety risks. The proposed framework uses a semi-Markov switching vector autoregressive (SMS-VAR) model [16] to detect anomalies based on dissimilarities in temporal values. TO find dissimilarities the proposed model predict multistep ahead and checks with actual values. The difference between these crosses the threshold then it is identified as similar values. The framework is scalable, due to the inherent parallel nature of most computations, and can be used to perform online anomaly detection. Online anomaly detection model is proposed for VMware based cloud data centers. The proposed model employed a Chi-square based statistical anomaly detection technique in Spark for online streaming data. VMware' performance data such as CPU and Memory is fit into variable length windows to perform Chi-squared based comparison for anomaly detection. The algorithm works well for univariate and homogeneous data.

| Ref. No | Model | Univariate/ Multivariate data | Online /Offline | domain | Temporal /Spatial | Dataset | accuracy | Parametric |
|---|---|---|---|---|---|---|---|---|
| 11 | Autoregressive | univariate | Offline | Network | Temporal | NA | 80% - ARIMA 89% - HWDS | Parametric |
| 12 | Regularized Vector Auto Regression | univariate | online | thermal plant | Temporal | thermal power | 70 | Parametric |
| 13 | Autoregressive | univariate | offline | situ sensors | Temporal | Water Quality | 92.4 | Parametric |
| 14 | Vector Autoregressive Model | Multivariate | Online | IoT | Temporal | FOQA | 91.6 | Nonparametric |
| 15 | Semi-Markov Switching Vector Autoregressive | Multivariate | offline | Aviation sensor | Spatial | FOQA | 92.3 | Non parametric |
| 16 | Chi-square Test | univariate | online | VMware Data | Temporal | Real-time Data | 90 | parametric |
| 17 | Hyper ellipsoidal Clustering | univariate | Offline | Fog Nodes | Temporal | ISSNIP | 92 | Nonparametric |
| 18 | Incremental Clustering | univariate | online | Network | Temporal | Real-time Data | 95 | nonparametric |
| 19 | Hidden Markov Models | Univariate | Online | Cloud Edge | Temporal | Real-time Data | 73.6 | Nonparametric |
| 20 | Feature Clustering | Univariate | Offline | Cyber system | Temporal | KDD | 99.6 | parametric |
| 21 | Contextual anomaly detection framework | univariate | offline | IoT | Temporal | HVAC, temperature and traffic | 90 | parametric |
| 22 | Wavelet transformation | Univariate | offline | WSN | Temporal | NA | 81 | Non-parametric |
| 23 | Spectral anomaly detection | Univariate | Offline | WSNN | Temporal | NA | 90 | Non-parametric |
| 24 | Linear Transformations | Multivariate | Offline | MSN | Spatial | Industrial Dataset | 90 | parametric |

Table 1: Anomaly detection scheme using statistical methods

A novel anomaly detection Fog-Empowered anomaly detection using an efficient hyper ellipsoidal clustering algorithm [17]. The end nodes stream the data to fog layers and anomaly detection process is carried out in fog layer and cloud layer. Initially fog node collect the sensor data and built clusters at sensor level and find sensor anomalies by calculating anomaly scores. All the data coming from sensors merged together to build fog level clusters and perform anomaly detection by calculating anomaly score. All the data aggregated and send the cloud for further process. The proposed work compared with existing centralized and distributed anomaly detection model and shows improvement in accuracy. An incremental clustering approach is proposed to classify anomalous and normal data based on log lines [18]. The proposed model incorporates a semi-supervised concept for incremental log data clustering that provides the basis for a new on-line anomaly detection approach based on log data streams. The anomaly detection scheme works in two phases, during

training phase cluster map is prepared which represents the normal behaviors of system. In the detection phase, each line is compared with cluster map lines and identify as anomaly if there is lot of deviation from new line and map lines. The proposed model is offline anomaly detection and not works on temporal data.

An anomaly detection and prediction algorithm is proposed based on a statistic model called the Hidden Markov model [19]. The suggested model refers to the aspect of mapping measurements of the underlying infrastructure issues. The model forecasts runtime data and tests for anomaly by contrasting real and expected data. During training phase historical execution data is collected to study resource consumption and asses overall performance. Once training data is collected, HMM built the container that are installed at fog nodes. Based on resource utilization measurement, HMM built normal behavioral containers and check new data to calculate the differences. if the difference is more than specified threshold then those data are considered as anomaly. It also senses variations in reaction time depending on their usage of resources.

An anomaly detection based on Feature extraction [20] using Evolutionary feature clustering by adopting fuzzy logic method. The proposed system addresses dimensionality reduction to improve scalability and availability. The main objective is to transform the initial connection representation so that its equivalent representation has reduced noise affect and achieves better classification or detection rates. For Experiment evaluation KDD dataset is used and proved highest accuracy compared to Cluster Center and Nearest Neighbors (CANN). The fuzzy membership function is used for better classification and detection rates.A contextual framework for the identification of anomalies for large sensor data is proposed [21]. The paradigm consists of two stages. The first stage is content detection where it is used to identify anomalies in real world applications. Second, context detection is used to replant the outcome of the content detector, to classify as content and contextually anomalous. The context detector uses the profile model, which are groups of similar data points created by a multivariate clustering algorithm.

An anomaly detection scheme for multi-sensor multivariate time series model is introduced [22]. The model applies multivariate orthogonal space transformations to enhance performance. The monitoring system is used to collect the data in time. Device Identification Step functions as a fusion operation by looking for associations and dependencies between sensor channels to measure the status of system variables. Three different techniques have been proposed. Firstly, Non-linear Finite Impulse Response Models (NFIR), Non-linear Output Error models (NOE), Non-linear Box-Jenkins models (NBJ). The use of multivariate orthogonal space transformations makes it possible to generate more compact and precise models due to an integrated dimensional (noise) reduction step. Fault detection is based on the discovery of gradual deviations (untypical occurrences) in the residual temporal signal. A graph-driven filter based on a novel spectral anomaly detection system is proposed [23]. The unsupervised anomaly detection paradigm is designed for the network of wireless sensors. Further, the sensor measurements are displayed as signals on the line. Graphs are chosen in the proposed method to derive valuable proximity information for the measured data. The related graph-based filters are then used to map graph signals to regular and anomaly subspaces, and the resulting projections are used to detect data anomalies. The benefit of the proposed d anomaly detection scheme over the traditional spectral method, main component analysis (PCA), is that the integration of structural information established a priori in addition to data provides localized transformations leading to efficient distributed anomaly detection.

2. **Machine learning approaches**: The machine learning algorithm are more popular due to its accuracy and atomicity. These algorithm are used in most of applications. Currently most of Artificial Intelligence application adopted the machine learning algorithm. It works on three components: 1) Task, 2) Experience and 3) Performance. The main objective of machine learning algorithm to enhance the performance using its expense.

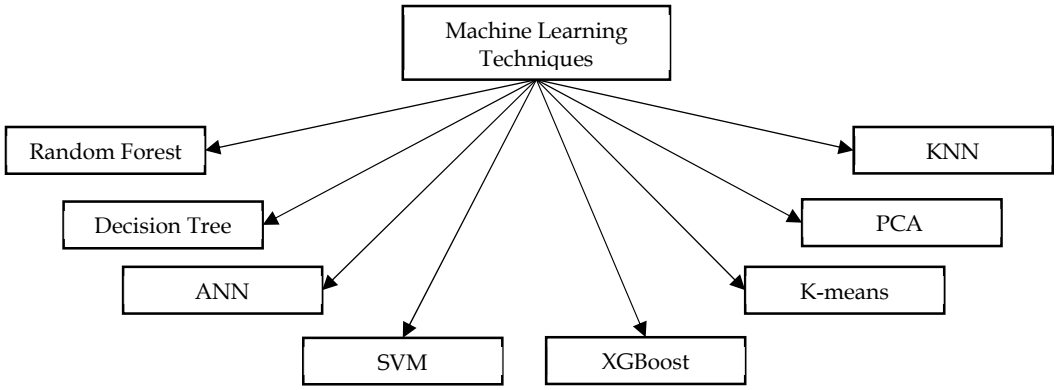


Figure 7: Machine learning algorithm for anomaly detection model

The industrial internet of things (IIoT), contains hundreds of hardware equipment's, *instruments*, Micro controllers and sensors. The anomaly detection model is proposed to identify sensor anomalies in IIoT [25]. Initially, to handle large amount of sensors data, data compression is adopted. Based on the different industrial process requirements, different threshold values are set to identify anomalies. The k-means clustering algorithm is used to analysis anomalous data values and the test values far the train values are considered as anomalous data points. The data compression and decompression consume more computational power. An anomaly detection model proposed for Internet of Things [27] using machine learning approach. Two robust machine learning algorithms Inverse weight clustering and C4.5 decision tree algorithm. IWC is extended version of k-means algorithm that is used to cluster the related data into group based on similarity measure. The decision tree algorithm is used classify the data by building the decision tree. The accuracy of anomaly detection model is 97%. The model evaluation done only temperature and humidity dataset. . The fog network model contain IoT devices in larger scale and delay is important aspect. To improve the availability, unsupervised machine learning based anomaly detection model is proposed [26]. The proposed work integrate the k-means and Principle component analysis to detection anomaly in shorter time. The proposed procedure and machine learning implementation not only encourages the collection of fog nodes, but also gives a new example of anomaly detection using fog nodes.

The IWC consume more computational resources like CPU and memory to execute the code repeatedly. A Real-time Anomaly Detection System (RADS) [29] is developed for Cloud data centers using one class classification algorithm. A window-based time series analysis to unknown behavior of attacks problem in cloud. The RADS can detect different types of virtual machine anomalies and crypto mining attacks [55]. The proposed model gains 90-95% accuracy with 0-3% false negative rates. The experiment evaluation consider a few VM only. The combination of linear regression and random for eash machine learning algorithm used to built anomaly detection model for multi cloud environment [30]. The popular open source data sets are used to evaluate the proposed model. The linear regression is used for time series prediction and Random forest is used categorization and anomaly detection. The experiment results shows that 93.6% for categorization and 99% anomaly detection accuracy. Scalping is more important and severe as increasing use of Internet and e-commerce. A new integrated method called SADM (Scalping Anomaly Detection Method) [31] to identify the special anomalies, and scalping. The model collect mobile traffic data from ISPs (Internet Service Providers) to process the NFP data (Network Fingerprint data), which records detailed event logs on mobile Internet. Then It extract features and select the important and powerful features. The propsed model uses Robust PCA model to construct powerful and robust feature set. The clustering method k-mean is used to find anomalies. 94.7 accuracy is recorded in experiment evaluation.

An unsupervised anomaly detection algorithm is proposed in [32], which combines Sub-Space Clustering (SSC) and One Class Support Vector Machine (OCSVM). The Subspace clustering is an improved version of classical clustering algorithm used to find clusters in dataset and one class SVM is used find the anomalies without any prior knowledges based on difference between cluster values and actual values. To evaluate the proposed model, a standard anomaly detection data called NSL-KDD dataset is used. the accuracy is 95% to 99% is recorded in experiment evaluation. The Fused-AI interpretabLe an anomaly detection system presented in [33] that combine knowledge-driven techniques and data-driven techniques. The data driven apply internal processing based on the raw input data and knowledge driven used to find the cause of problems/faults/anomalies. Machine Semantic information is built into a machine learning methodology to improve expressivity. At the same time, feedback on the defects and irregularities that have arisen is given as input to improve adaptiveness using semantic rule mining

methods. This current approach is tested on the basis of a statistical maintenance scenario for trains. it has been proved that proposed approach reduce downtime and gives further insight into typical problems. The various machine learning algorithm are used to predict and find the anomalies in the Internet of Things applications [34].These machine learning algorithm are Decision Tree (DT), Artificial Neural Network (ANN), Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF). Various performance metrics accuracy, precision, recall and F1 Score are evaluated. The highest accuracy is recorded is 99.4 for ANN, DT and RF algorithms.

The combined approach of Self-Organizing Maps and Particle Swarm optimization algorithm is used to design and developed anomaly detection model [35]. The self-organized maps are most common unsupervised neural network approach used to cluster the data. The Particle swarm algorithm is used for optimization and improve the accuracy. It also handles non-linear problem smoothly to get higher accuracy. The proposed model used for crisis management like forest fires and their detection. The forest fires dataset is used to evaluate the proposed model and recorded 83.2% accuracy. A new improved anomaly detection model is design using probabilistic neural network (PSO-PNN)[28] for the detection and recognition anomaly detection. The proposed model works in two stages. First in the recognition process, the cloud user behavior is converted to understandable format. In the second stage, Multi-layer neural network is used to classify and identify anomalies. The particle swarm optimization is used to improve the accuracy of the model by tuning hyperparameters. The UNSW-NB15 dataset is used for evaluating anomaly detection model. The results shows that highest accuracy 98% is achieved. To improve the scalability and accuracy of anomaly detection model, K-means and XGBoost system is used in modelling anomaly detection scheme. K-means algorithm is used to classification and clustering the data based on their distances. XGBoost uses these classification details to produce a interpretable rules. These rules are the basis for generalizing its implementation over a large range of unseen occurrences in a distributed computing system. The

| Ref. No | Model | Univariate/ Multivariate data | Online /Offline | domain | Temporal /Spatial | Dataset | accuracy | Remarks |
|---|---|---|---|---|---|---|---|---|
| 25 | K-means Clustering | Multivariate | Offline | IIoT | Temporal | RealTime | 70% | Compression consume more computation power |
| 26 | K-mean and PCA | Multivariate | Offine | Cloud | Temporal &Spatial | NA | 84% | Real-time dataset is used |
| 27 | IWC and K-means | Univariate | Offline | IoT | Temporal | Intel IoT | 97% | IWC consume more CPU and Memory |
| 28 | PNN and PSO | Multivariate | Offline | Cloud | Temporal | UNSW-NB15 | 98% | Require more computation resources |
| 29 | One class classification | Univariate | Offline | Cloud | Temporal | Cloud | 93% | Smaller number of VM are considered |
| 30 | Linear regression & Random Forest | Univariate | Offline | Cloud | Temporal | Cloud | 93.6% | More computational resource required |
| 31 | K-mean and PCA Local Outlier Factor | Univariate | Online | Network | Spatial | NFP | 94.7% | Few parameters are considered |
| 32 | Sub-Space Clustering &One Class Support Vector Machine. | Multivariate | Offline | General | Temporal | NSL-KDD | 95%-99% | More computational resource required |
| 33 | AI Fusion | Multivariate | offline | Train | Spatial | Train | 95.4% | Less scalable |
| 34 | ANN, DT and RF | Multivariate | Offline | General | Temporal | DS2OS | 99.4% | Five ML algorithm were tested |
| 35 | SOM and PSO | Univariate | Offline | Crisis managemen | Spatial | Forest Fires | 83.2% | Less scalable |
| 36 | k-means and XGBoost | multivariate | offline | Network Managemen | spatial | NASA Hypertext 18 Transfer Protocol | NA | Accuracy is not evaluated |

3. Anomaly detection using Deep Learning Models: To improve the accuracy of machine learning by adopting various optimized neural network models are used. These models considered as deep learning model. The deep learning model train the data and model in depth by introducing new concepts to reduce the complexity and improve the accuracy. Further, the anomaly detection in real time critical application requires highest accuracy[7][8]. So various deep learning model such as Convolution neural network and Recurrent neural network,

| Ref. No | Model | Univariate/ Multivariate data | Online /Offline | domain | Temporal /Spatial | Dataset | accuracy | Remarks |
|---|---|---|---|---|---|---|---|---|
| 37 | Hierarchical; Edge model | multivariate | online | Edge | Temporal | power consumption | 93-99 | More computation cost |
| 38 | Auto encoder | Univariate | Online | WSN | Temporal | WSN Testbed | 95 | Latency between cloud and sensors is more |
| 39 | vector convolutional deep learning | Multivariate | Offline | IoT | Temporal | BoT-IoT | 98 | class imbalance problem |
| 40 | RNN | Multivariate | online | Aircraft | Temporal | FOQA | 99 | Limited features considered |
| 41 | Deviation Networks | Multivariate | Offline | security | Spatial | Security | 80-100 | Scalable |
| 42 | dLSTM | Multivariate | Offline | General | Temporal | ECG | 98 | Window based |
| 43 | LSTM & ACO | Multivariate | Online | IoT | Temporal | FOQA | 99 | initial cost to build multiple model is more |
| 44 | LSTM and SVDD | Multivariate | Online | General | Temporal | realtime | 95 | More computation cost |
| 45 | RNN-ABC | Multivariate | Online | IoT | Temporal | NSL-KDD | 95.02 % | Consume more time & CPU for optimization |
| 46 | CNN | univariate | online | IoT | Spatial | | 83% | More computation cost |
| 48 | 3D Resnet | Univariate | Online | IoT | Temporal | UCF-Crime | 89% | Training time is more |

An anomaly detection model is proposed using Deep learning technique called Deep Neural Network (DNN) model [37]. The main objective of proposed work is to reduce the detection delay and increase the accuracy of anomaly detection. An adaptive anomaly detection approach for hierarchical edge computing (HEC) systems is used to detect anomalies in edge, fog and cloud layer simultaneously. Three anomaly detection DNN models are proposed for IoT devices, edge servers, and cloud. Based on contextual information extracted from input data, to anomaly detection scheme is applied. To select contextual data a single-step Markov decision process is used. The result shows that 93-99% of accuracy and reduce 40% detection delay. The internet of Things and Wireless sensor network are used to bridge the gap between the physical and cyber worlds and they explore the lot of attacks. It is essential identify anomaly in various events of interest. An anomaly detection model is proposed using deep learning model called auto encoder neural network. The proposed model have two design parts, first, anomaly detection at sensor level, and cloud level. Second, the most processing part of anomaly detection is processed at cloud with less frequency. The evaluation of model is done using WSN testbed. The highest accuracy is recorded in evaluation of proposed model.

To mitigate the security attacks and improve the scalability in Internet of Things applications, an anomaly detection model[39] is developed using vector convolutional deep learning (VCDL) approach. The anomaly detection process performed in fog nodes and IoT traffic training is distributed in all the fog nodes. The proposed model classify IoT traffic as normal or attack then this data passed to cloud to mitigate type attacks. The BoT-IoT dataset is used for experiment evaluation and proved that proposed model is more scalable and obtained highest accuracy. To identify anomalies and faults in Aircraft flights, A Recurrent Neural network with Long Term Short memory model and Gated Recurent Unit model is used to develop an anomaly detection model. The main

objective of proposed work is to detect anomalies in latent features in scalable manner. The experiment results shows improved in accuracy, recall and precision values. The long short term memory model is used for prediction purpose as it is remember previous n values to predict the next n values with highest accuracy. A novel anomaly detection model [41] is introduced to address various problems such as scalability, accuracy and indirect optimization in existing anomaly detection models. The proposed model uses nural deviation leanring to leverage end-to-end anomaly score learning to improve the accuracy. Five different security dataset are used to perform experiment evaluation.

A delayed Long Short-Term Memory (dLSTM) model [42] is used to develop an anomaly detection scheme for temporal data. The proposed model works in two stages. First, a multistep prediction strategy is developed using dLSTM model. Second, an anomaly detection model is deployed where the prediction error is considered as anomaly based on certain threshold. The various prediction are proposed and best one is selected that gives highest accuracy. To evaluate proposed model, real and artificial data is used and observed highest accuracy. The recurrent neural network (RNN) is used to predict the vibration of aircraft engine using long short-term memory (LSTM)[43] model. the main advantage of LSTM model is that it is more robust and generalized model for prediction. The multiple LSTM RNN model were proposed with different parameters. The Ant Colony Optimization (ACO) is used to develop and improve the LSTM cell. The best LSTM model is selected which gives highest accuracy and least prediction error. The initial cost to build multiple model is more. Similarly, the LSTM model with joint optimization technique [44] is used to find anomalies in distributed environment. The LSTM model is used to extract temporal features and identify more relevant features vectors to detect anomalies. Later, labelling process is carried out using Support Vector Data Descriptor (SVDD) model. To evaluate the proposed model, a loss function is introduces and with real time dataset to demonstrate the evaluation. The training algorithm with different parameters are given for optimizing or tuning accuracy. A swarm optimization technique with Recurrent Neural Network is used to develop an anomaly detection model [45]. The Long Short Term Memory model is used to predict the next n timesteps and An artificial bee colony used to optimize the hyperparameters to improve accuracy and reduce the loss. The experiment evaluation is carried out using python and NSL-KDD dataset. The overall accuracy of model is 95.02% is recorded.

Over the past few years Convolutional Neural Networks (CNN) [46] have become popular choice for image processing application to enhance the performance. AN anomaly detection scheme is developed using CNN to quality inspection in automated industries. CNN model works well even larger image dataset to classify and find the anomalies. The proposed model adopted window based solution and optimize the loss function Mean Squared Error to improve the accuracy. similarly an anomaly detection scheme is introduced in [47] for automated monitoring solutions using CNN. The model defines a transfer learning function that pretraind the learning model on large scale univariate temporal dataset and then optimize the weight of network on small scale, univariate to identify anomaly. The anomaly detection in video surveillance is major challenges in most of the application. A background subtraction approach and deep learning model is adopted to detect online anomalies. model detect multiple objects of different shape and sizes by pixel-wise foreground segmentation.

A deep multiple instance learning based anomaly detecting framework [48] for surveillance videos is proposed. The supervised learning algorithm called Pseudo-3d ResNet network with batch normalization layer is added after each convolution layer to extract useful features. The three layer convolution network with 50% dropout regularization is used to detect anomalies to improved detection accuracy. Experimental results show improves the accuracy of abnormal behavior detection compared to existing models. To identify anomalous sound event in public places, a model uses WaveNet method [49]. The wavenet is generative model based on CNN used to model acoustic pattern detection in various temporal data. In the proposed model, wavenet is used as a predictor to detect waveforms in unknown acoustic patterns as wavenet is capable of modelling temporal structures. The proposed model detect anomalous sound events more accurately than conventional methods based on reconstruction errors of acoustic features. A real-world dataset called subway station is used to evaluate the model. The proposed method reduces prediction error compared to state of art model. Performance anomaly detection[50] is most important aspect to improve the QoS and overall performance of distributed service in service oriented architecture. It is process of detecting nonconformist occurrence and events that leads to system falldown. Based on dynamic features of QoS, anomaly detection model is proposed using RNN model. In the proposed model, three different anomaly detection models are developed based SimpleRNN, LSTM and Gated Recurrent Unit (GRU). The experiment results shows that 85% to 89% of accuracy is observed.

**Applications of Deep Anomaly Detection**

The distributed computing environment become more popular and adopted in multiple applications. In this section, various applications [9][64] of anomaly detection are discussed.

1. **Intrusion Detection:** Intrusion Detection: Intrusion detection refers to the detection of malicious behavior (break-ins, penetrations, and other types of cyber abuse) in a computer-related device. These malicious attacks or intrusions are interesting from a computer security perspective. Intrusion is distinct from standard device behavior and thus anomaly detection methods are applicable in the intrusion detection domain. The key challenge for anomaly detection in this domain is the large amount data and input size [8]. The anomaly detection techniques need to be scalable to handle large input data. The online data analysis is required for the streaming fashion data. The intrusion detection system, is classified as Host-Based Intrusion Detection Systems (HIDS) and Network Intrusion Detection Systems (NIDS). In host based, the intruder attach on end devices OS system call , where as in network based intruder attach the information that is in transit. In order to identify or detect these attacks and traits on early stages, anomaly detection system is required. The anomalies in this intrusion detection in very difficult as the anomalies keeps changing over the time as the intruder changes their method/pattern to extract useful information [52].

2. ***Fraud Detection:*** Fraud detection refers to the prevention of fraudulent fraud of commercial institutions such as banks, credit card firms, insurance companies, mobile phone providers, the stock exchange, etc. The fraudulent customers may be the real clients of the company or may be acting as a customer. Fraud happens as these consumers consume the services offered by the company in an illegal manner [9]. Organizations are involved in the rapid discovery of such fraud in order to avoid economic damages. Anomaly detection plays important role in identify these fraud techniques are used to identify the frauds in an organization, these are Statistical Profiling using Histograms, Parametric Statistical Modeling, Non-parametric Statistical Modeling, Bayesian Networks, Neural Networks, Support Vector Machines, Rule-based Systems, Clustering Based, Nearest Neighbor based, Spectral and Information Theoretic [54].

3. ***Medical Anomaly Detection***: Detection of anomalies in the medical and public health domains generally deals with patient information [8]. Data may have anomalies due to a variety of causes, such as irregular medical condition or instrumentation failures or logging errors. Several methods have also been used to identify outbreaks of disease in a particular region. The identification of anomaly is therefore a very important topic in this domain and requires a high degree of precision. Anomaly detection is very crucial requirements in the medical fields.

4. ***Detection manufacture Defects in Industries***: Industrial units suffer damage due to repeated use and regular wear and tear. Such damage must be identified early in order to avoid further escalation and damages. Data in this domain is commonly referred to as sensor data because it is recorded using various sensors and processed for analysis. Anomaly detection techniques have been widely used in this area to detect such injury. Industrial damage identification can be further divided into two domains, one concerned with defects in electronic parts such as motors, engines, etc., and the other dealing with defects in physical structures.
5. **Image Processing**: The strategies for identifying irregularities in relation to images are either involved in any changes in the image over time (motion detection) or in areas that tend to be irregular on the static image [9]. Steganography is technique that hide the information in the image. Intruder may access and extract the hidden information. Once the intruder extract the image, the image may changes. To identify these kind of changes anomaly detection techniques are used. Due to the size of image or video frames, the anomaly detection become key challenge for researchers.
6. **Anomaly detection in text documents**: Anomaly detection techniques in this area are mainly used to detect novel subjects or incidents or news items in the compilation of documents or newspapers. Anomalies are triggered by a new interesting occurrence or a phenomenon [9]. The data in this domain is usually large in size and very small. It also has a temporal component as the records are gathered over time. The problem for anomaly detection techniques in this domain is to treat anomaly wide differences in documents belonging to one group or subject.
7. **Anomaly detection in Sensor detection**: Sensor networks have recently become an important research topic; more from a data analysis point of view, as data from sensor data obtained from different wireless sensors has many specific characteristics [64]. Anomalies in data obtained from a sensor network can either indicate that one or more sensors are dysfunctional or detect events (such as intrusions) that are of concern to researchers. Thus, anomaly detection in sensor networks may catch sensor failure detection or intrusion detection or both. A single sensor network can consist of sensors that capture various types of data, such as binary, discrete, continuous, audio, video, etc. The data are created in streaming mode. Sometimes the atmosphere in which the various sensors are deployed, as well as the contact channel, causes noise and missed values I n the data collected. Detection of anomaly in sensor networks faces a

range of special problems. Anomaly detecting techniques are necessary in order to work in an online approach. Owing to extreme resource limitations, anomaly detection strategies need to be lightweight. Another problem is that the data is obtained in a distributed manner, and therefore a distributed data mining approach is needed to interpret the data. Furthermore the inclusion of noise in the data obtained from the sensor makes the identification of anomaly more difficult, as it is now important to differentiate between interesting irregularities and unnecessary noise/missing values.

8. **Smart IoT Application**
   - i. **Smart Grid:** The Smart Grid can be seen as an energy network that benefits from two-way cyber-secure connectivity technology and computational intelligence for the generation, transmission, convergence of substations, delivery and usage of electricity in order to achieve the objectives of renewable, safe and secure facilities. In the electrical grid, suspected manipulation of the Smart Meter housing, under/overvoltage at particular sites, or remote switching breakdown. Regulated appliances are called anomalies. These anomalies have certain trends for different kinds of unusual acts. The techniques of anomaly detection may be used to classify this type of action.
   - ii. **Smart city**: Various smart city technologies attach big loT machines to real-world artefacts, and do also have some substantial benefits for urban life. The vast majority of loT devices over heterogeneous types of services, technologies, devices and protocols raise the sophistication of the IoT network control mechanism. The integration of different types of protocols with the Internet causes significant cyber security risks and vulnerabilities to attack detailsialy activity information in the citizen lives. The information may be medical records, financial reports, and property details. So it essential to stop this kind of leakages through anomaly detection models

   - iii. **Smart Hospitals**: Healthcare data is very important as it provide the details related to human health and survival. It also provides the insight of disease and its treatment to save human life. The smart hospital application collect, maintain and analysis patience data so provide proper treatment to patients. The survival rate of patient is increases through these kinds of applications. The most important medical data are: Electronic Health Record maintenance and retrieval, Bio-medical signal analysis, Medical image analysis and remove treatment in pandemic times etc.

**KEY ISSUES AND CHALLENGES IN ANOMALY DETECTION (In IoT)**

In this section, various challenges of anomaly detection techniques are discussed. These challenges are categorized as Input data challenges, data preprocessing challenges, Big data processing challenges [9], computational challenges and visualization challenges.

**1. Input Data challenges**

Various IoT devices are linked to fog node and these IoT devices are created same data frequently. In certain cases, collected data is redundant as two or more domain produce the same data. In addition, there are a variety of elements of the time-series structure as well as the context in which the data is generated and processed that may have an effect on the performance of the anomaly detection algorithm. It is also essential to collect the fog node virtual machine data to understand the possibility of anomalies in fog computing.

1) Contextual information: The fog computing environment surrounded by large number of IoT devices. it is essential to monitor the behavior of these devices to improve the QoS[65]. Further, the contextual knowledge required to identify the anomalies in the IoT applications to avoid uncertainties. A wide range of challenges needs to address in order to incorporate contextual information. .

1. Temporal Context: As the majority of IoT data is produced in the form of time-series data [36], it is essential to understand the relationship between these data to evaluate and enhance the performance of anomaly detection. For example in smart farming, it is essential to understand the relation between humidity and temperature sensor values at certain timestamp to find the anomalies.
2. Spatial Context: The several sensors are deployed in fog computing environment to monitor the same device and it is difficult to manage same context from multiple heterogeneous sensors. This becomes more difficult to address when the spatial context grows in size .
3. Hybrid context – In some application both temporal and spatial data needs to are monitor for anomaly detection. It is difficult to extract the relationship between spatial and temporal data.

2) Nature of input data: The input data has two distinct attribute types, i.e. binary, categorical, or continuous.

Further, each data instance is classified into two type's univariate or multivariate [5]. The complex nature of input data makes it impossible for anomaly detection techniques to pick the right algorithm to manage the particular data. The diverse nature of input data makes it very difficult for anomaly detection systems to pick the right algorithm to manage the particular data. Essentially, anomaly detection strategies will differ depending on the design of the attributes of the application. This problem will be solved by creating a hybrid unsupervised machine learning algorithm.

3) Heterogeneity of data: The IoT devices generate two types of data structured data and unstructured data. It is easy to handle structures data because it organized with proper mechanism and manageable way. The unstructured data is highly dynamic and completely raw data such as interactions meeting recording and fax transfers etc. Unstructured data is difficult to store in databases and difficult to interpret. Working with unstructured data is tedious and, of course, expensive. It is also not possible to turn all these unstructured data into structured one. The use of unsupervised hybrid machine learning algorithms would resolve the heterogeneous data problem.

2. **Data pre-processing: It** is an important step in Anomaly Detection. Before the data is sent to particulat anomaly detection model, it is essential to prepare the data in appropriate format. The data preparation involves adding missing values, data encryption, remove noise and normalize the data [53]. During data preparation, the following issues may arrive:

1. Missing data points: The sensor deployed in various environmental conditions. In some conditions sensors may not work and unable to produce/store the data. The missing value in dataset may not get proper results. So it essential to add these missing data by taking mean or median values.
2. Data corruption: External aspects or device malfunctioning may contaminates the data and makes it difficult to distinguish anomaly and corrupted data. It is difficult to understand which data is correct or which one is corruption. The identification of corrupt data and rectify that data itself is difficult and it must
3. Encrypted data: in some case the encrypted data is used to detect anomalies. So it is difficult for anomaly detection model to understand these encrypted data as its dimensions, types and ranges are different. So it difficult to find the anomalies with encrypted.
4. Sensor fusion: In IoT, different kind of sensor were installed for various purposes. These data is different in dimensions, types and ranges. So it is aggregate the data for further data processing to produce the results.
5. Real-time detection: Real-time anomaly detection includes high-speed streaming data and thus require a quicker response. Anomaly detection for live streaming data required large computational resources to store and process.
6. Noisy data: The IoT contains large number of electronic components like Sensors, microcontrollers, and actuator. These electronic component transfer the data from one device to another electronic devices. These transmission produces noisy data/ contaminated data & difficult to remove from edge computing devices in IoT before it forwarding it to the cloud.

**3. Big Data Challenges**: The performance of anomaly detection techniques is becoming complex as the data size increases[9]. The "volume" attribute of the big data emphasizes the power, storage and processing ability of the device to accommodate the increasing size of the data. Performance is a major challenge for anomaly detection approaches when dealing with massive data sets. When the data size is high, anomaly detection techniques may not give accurate results due to limited computational ability and associated factors. Few major challenges in anomaly detection techniques when dealing Big Data [9]:

1. *Uncertainty*: Data may be ambiguous due to external occurrences from vulnerable channels, such as inability to quantify the measure of characteristics, imprecision, vagueness, uncertainty, and misunderstanding. Information that cannot be depended on In full certainty, they are defined as uncertain.

2. Performance: The more computational power such as CPU and Memory is required to get the accurate results while detecting anomalies in High dimensional data.

3. Scalability: It is the ability of the technique to quantify the increasing scale and size of the data. The performance of anomaly detection will be degraded when the size of data increase. The previous instance or existing data is very important to identify the anomalies. The more computational capacity is required to handle large volume of data.

4. Asynchronous instances: Multiple asynchronous data points coming from different sources at different times and these data points are independent of one another. Some data points may arrive at any point in time and some

are missing or delay in arrival. Existing anomaly detection method stores the data as it static in nature. It is difficult to find Anomalies on streaming data as the incoming data is a highly volatile because the volume of data is not limited and cannot be stored indefinitely. So it is difficult to identify anomaly in such dynamic and heterogeneous context and it is essential to determine the specific time series context on these data instances of multiple streams. Another important issues is concept drift. The data distributions are changes over the time, so predicting certain values may not valid for next time step (A comprehensive survey of anomaly detection techniques for high dimensional big data)

5. Data Redundancy: The IoT devices generate large amount of redundant data frequently. So it difficult to manage this data in terms of processing and storing the data[9]. More processing and storage power may waste due to redundant data. It also reduces the performance of any distributed network. The redundant data leads to many performance issues. There are many state of the art methods are introduced but still it is essential to reduce the large amount of redundant data. Hence, designing a system that is capable of resolving and mitigating redundancy problems in time critical application becomes important.

6. Dimensionality: Dimensionality describes the number of distinct data properties collected in each observation [6], the data dimensionality defines the choice of method used, since certain methods are not suitable for higher-dimensional data. In contrast, the cost of processing higher-dimensional data can be higher than that of lower-dimensional data. IoT data is generated in two broad categories: Univariate data and multivariate data. The univariate data has single parameters and multivariate data contains multiple parameters. Anomaly detection over univariate streams is based on a comparison of the present observation with the local or global context of the time-series being analysed. This is contrasted with multivariate streams, where not only is the context of the stream relevant for the detection mission, but also the correlation between each of the measurements that combine to form the observation at a given time-step. It is therefore necessary to implement a dimensional reduction methodology for anomaly detection in order to improve accuracy and other QoS parameters.

7) Parameters selection: Finding appropriate parameters for any machine learning model can be difficult. These are called as hyperparameters [9]. To improve the accuracy of any machine learning model, it is essential to tune the parameters and its values. Parameters selection is very important key aspect in real time anomaly detection models. Further, a combination of parameters may perform well in some stages and may not perform well in other stages and vice versa. Parameters and its values were one of the major contributors to the machine learning algorithms' results. There are good number of state of the art works based hyperparameter optimization to improve the performance but these solution may not fit others. So it is essential to design and develop appropriate hyperparameter tuning mechanism for anomaly detection models.

**4. Computational challenges**

A variety of experiments have focused on combining or integrating multiple methods to improve the efficiency of anomaly detection, leading to an increase in the cost of computing. In addition, high dimensionality combined with large sample size produces problems such as high computing expense and algorithmic uncertainty. Using big data technology along with cloud, thus, can solve the issue of computational costs by integrating parallel and distributed processing, which allows to create several clusters that minimize the cost of computing.

**Time and Resource Constraints**: In IoT framework, the most device have low power with minimal processing resources, since the existing paradigm is commonly used to capture and process data at a centralized servers, normally using cloud or data center, and storage technology. This paradigm allows for greater resources to be leveraged for the computational process, but it also imposes some machine lag due to round trip delays as well as resource scheduling [5]. In some cases, this is permissible as it is not important to act immediately on the information gathered from the data, however when looking at the automation of the connected tools, it might be necessary for the data to be analyzed quickly and thus the reports to be produced as soon as possible after the data has been generated [9]. The use of Edge/Fog devices provides an ability for this to occur closer to the position of the data generated, but these devices are generally less driven than cloud providers and thus it is necessary to consider the computing cost of any analytical tasks conducted on that data. Most of devices in IoT framework are wireless. These wireless devices may not reliable. Due to packet loss and high traffic, retransmission becomes very essential. Repeated long-range wireless communications often incur considerable costs in terms of the use of batteries where the devices are self-powered, as such, where minimal processing may be done closer to the device and aggregated information transmitted at a lower frequency can improve the lifespan of each such device [3].

**Inadequate architecture**

The state-of-art architecture is manage anomaly detection in batch processing and small scale data, but is unable to handle big-data in real time [9]. Researchers are struggling to make the big data structure work better, but the architecture is fundamentally different from the big data architecture when it comes to real-time data. Real-time architecture components must incorporate an application and Big-data analytics to propose a new way of working environment that meets the requirement of online, offline, static and dynamic data. Big data technology is inefficient because it is not combined with current enterprise data; the same cannot be accomplished until big data correlates the convergence of different big data systems with hybrid machine learning algorithms to solve architectural problems. .

**5. Data visualizations**

The interpreted and evaluated data or analysis must be visualized by the user as well as produced by the report. However, the difficulty lies in the collection of suitable visualization tools for the identification of anomalies from the different linked devices [9]. Various visualization methods are used in existing system, these are categorized as text based or GUI based. The text methods shows the score and types of anomaly. The GUI based visualized are GUI oriented. Two dimensional and three dimensional graphs are used for visualization. Now a day's python and R language has interactive graph API where user can apply various operations like zoom-in, zoom-out and rotate at angle. it is difficult for anomaly detection approach to use right visualization model automatically. So it essential to develop anomaly detection approach that provide customization in visualization method.

**Reporting Method**

There are two primary ways in in which anomalous data may be reported [6], [48]:

1) Anomaly Score: An anomaly score is a value that represent percentage of deviation from the predicted value as described by the anomaly detection model used. There are a number of approaches for generating anomaly scores that are unique to each algorithm. This method is often used when carrying out a later analysis of the data obtained, as the researcher would wish to examine only the top-n anomalies within a given period of time. This score can also be useful in identifying and managing outliers when doing related statistical tasks, such as predictive analytics.

2) Labels: It is the observation of anomaly detection and outputs normal or anomalous. Most of the anomaly detection algorithm produce direct classification with certain threshold values considered. This method may be more commonly used when immediate monitoring is needed, such as the detection of faults in the device being tracked where the operator or owner of the system needs near-real-time notice of irregularities. More sophisticated anomaly detection systems, in particular those using a supervised learning method, which have several anomaly types, each with its own label, allowing separate alerts to be activated on the basis of the label assigned to it.

**Conclusion**

Anomaly detection model improve the quality of services and avoid unexpected breakdowns in fog computing environments. In this paper, basics of anomaly detection approaches, working principle and state-of-art anomaly detection approaches are discussed. This review work considered only few important and latest work to gain insight of currently anomaly detection approaches. The state of art anomaly detection approach comparison has been done by considering important matrices. Further, various applications of anomaly detection model in Internet of Things are presented. The challenges and possible future directions are highlighted to carry out further research to improve the Quality of Service.

**References**

[1] Arun Kumbi, Pavankumar Naik, Kirthishree C. Katti, Kiran Kotin, "A Survey Paper on Internet of Things Based Healthcare System", Vol5, Iss5, pp 1-4 Journal of Internet of Things and Cloud Computing, Science PG, 2017.

[2] P. Hu, S. Dhelim, H. Ning, and T. Qiu, "Survey on fog computing: architecture, key technologies, applications and open issues," Journal of Network and Computer Applications, vol. 98, pp. 27–42, Nov. 2017, doi: 10.1016/j.jnca.2017.09.002.

[3] Dr. Mahantesh N Birje Chetan Bulla, "Cloud Monitoring System: A Review", vol 1, issue 1, IJESM,2019.

[4] Dr. Mahantesh N Birje, Chetan Bulla, "Commercial and Open Source Cloud Monitoring Tools: A Review", International Conference on Emerging Trends in Engineering, Springer, 2019.

[5] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: 10.1145/1541880.1541882.

[6] G. Fernandes, J. J. P. C. Rodrigues, L. F. Carvalho, J. F. Al-Muhtadi, and M. L. Proença, "A comprehensive survey on network anomaly detection," Telecommun Syst, vol. 70, no. 3, pp. 447–489, Mar. 2019, doi: 10.1007/s11235-018-0475-8.

[7] B. Sharma, L. Sharma, and C. Lal, "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, Dec. 2019, pp. 146–149, doi: 10.1109/ICCIKE47802.2019.9004362.

[8] R. Wang, K. Nie, T. Wang, Y. Yang, and B. Long, "Deep Learning for Anomaly Detection," in Proceedings of the 13th International Conference on Web Search and Data Mining, Houston TX USA, Jan. 2020, pp. 894–896, doi: 10.1145/3336191.3371876.

[9] R. A. Ariyaluran Habeeb, F. Nasaruddin, A. Gani, I. A. Targio Hashem, E. Ahmed, and M. Imran, "Real-time big data processing for anomaly detection: A Survey," International Journal of Information Management, vol. 45, pp. 289–307, Apr. 2019, doi: 10.1016/j.ijinfomgt.2018.08.006.

[10] Ahmad, S., Lavin, A., Purdy, S., & Agha, Z. (2017). Unsupervised real-time anomaly detection for streaming data. Neurocomputing, Available online 2 June 2017, ISSN 0925-2312,

[11] E. H. M. Pena, M. V. O. de Assis, and M. L. Proenca, "Anomaly Detection Using Forecasting Methods ARIMA and HWDS," in 2013 32nd International Conference of the Chilean Computer Science Society (SCCC), Temuco, Cautin, Chile, Nov. 2013, pp. 63–66, doi: 10.1109/SCCC.2013.18.

[12] Y.-M. Wei, H. Yu, W.-J. Wang, Y.-H. Sun, J. Wang, and Z.-J. Song, "Anomaly Detection Based on Regularized Vector Auto Regression in Thermal Power Plant," MATEC Web of Conferences, vol. 35, p. 06002, 2015, doi: 10.1051/matecconf/20153506002.

[13] C. Leigh et al., "A framework for automated anomaly detection in high frequency water-quality data from in situ sensors," Science of The Total Environment, vol. 664, pp. 885–898, May 2019, doi: 10.1016/j.scitotenv.2019.02.085.

[14] I. Melnyk, B. Matthews, H. Valizadegan, A. Banerjee, and N. Oza, "Vector Autoregressive Model-Based Anomaly Detection in Aviation Systems," Journal of Aerospace Information Systems, vol. 13, no. 4, pp. 161–173, Apr. 2016, doi: 10.2514/1.I010394.

[15] I. Melnyk, A. Banerjee, B. Matthews, and N. Oza, "Semi-Markov Switching Vector Autoregressive Model-Based Anomaly Detection in Aviation Systems," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA, Aug. 2016, pp. 1065–1074, doi: 10.1145/2939672.2939789.

[16] M. Solaimani, M. Iftekhar, L. Khan, and B. Thuraisingham, "Statistical technique for online anomaly detection using Spark over heterogeneous data from multi-source VMware performance data," in 2014 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, Oct. 2014, pp. 1086–1094, doi: 10.1109/BigData.2014.7004343.

[17] L. Lyu, J. Jin, S. Rajasegarar, X. He, and M. Palaniswami, "Fog-Empowered Anomaly Detection in IoT Using Hyperellipsoidal Clustering," IEEE Internet Things J., vol. 4, no. 5, pp. 1174–1184, Oct. 2017, doi: 10.1109/JIOT.2017.2709942.

[18] M. Wurzenberger, F. Skopik, M. Landauer, P. Greitbauer, R. Fiedler, and W. Kastner, "Incremental Clustering for Semi-Supervised Anomaly Detection applied on Log Data," in Proceedings of the 12th International Conference on Availability, Reliability and Security, Reggio Calabria Italy, Aug. 2017, pp. 1–6, doi: 10.1145/3098954.3098973.

[19] A. Samir and C. Pahl, "Detecting and Predicting Anomalies for Edge Cluster Environments using Hidden Markov Models," in 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC), Rome, Italy, Jun. 2019, pp. 21–28, doi: 10.1109/FMEC.2019.8795337.

[20] G. R. Kumar, N. Mangathayaru, G. Narsimha, and A. Cheruvu, "Feature Clustering for Anomaly Detection Using Improved Fuzzy Membership Function," in Proceedings of the Fourth International Conference on Engineering & MIS 2018 - ICEMIS '18, Istanbul, Turkey, 2018, pp. 1–9, doi: 10.1145/3234698.3234733.

[21] M. A. Hayes and M. A. Capretz, "Contextual anomaly detection framework for big sensor data," Journal of Big Data, vol. 2, no. 1, p. 2, Dec. 2015, doi: 10.1186/s40537-014-0011-y.

[22] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger, M. Pichler, and H. Efendic, "Fault detection in multi-sensor networks based on multivariate time-series models and orthogonal transformations," Information Fusion, vol. 20, pp. 272–291, Nov. 2014, doi: 10.1016/j.inffus.2014.03.006.

[23] H. E. Egilmez and A. Ortega, "Spectral anomaly detection using graph-based filtering for wireless sensor networks," in 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, May 2014, pp. 1085–1089, doi: 10.1109/ICASSP.2014.6853764.

[24] Y. Yang, J. Zhang, S. Song, C. Zhang, and D. Liu, "Low-Rank and Sparse Matrix Decomposition With Orthogonal Subspace Projection-Based Background Suppression for Hyperspectral Anomaly Detection," *IEEE Geosci. Remote Sensing Lett.*, vol. 17, no. 8, pp. 1378–1382, Aug. 2020, doi: 10.1109/LGRS.2019.2948675.

[25] DEQUAN KONG, DESHENG LIU, LEI ZHANG, LILI HE, QINGWU SHI, XIAOJUN MA, Sensor anomaly detection in the industrial internet of things based on edge computing, Turkish Journal of Electrical Engineering & Computer Sciences, 10.3906/elk-1906-55.

[26] Y. Li, Y. Zhang, Y. Liu, Q. Meng, and F. Tian, "Fog Node Selection for Low Latency Communication and Anomaly Detection in Fog Networks," in 2019 International Conference on Communications, Information System and Computer Engineering (CISCE), Haikou, China, Jul. 2019, pp. 276–279, doi: 10.1109/CISCE.2019.00069.

[27] Ahod Alghuried, A Model for Anomalies Detection in Internet of Things (IoT) Using Inverse Weight Clustering and Decision Tree, Dissertations, School of Computing, Technological University Dublin,2017.

[28] S. K. Peddoju, H. Upadhyay, and S. Bhansali, "Health Monitoring with Low Power IoT Devices using Anomaly Detection Algorithm," in *2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC)*, Rome, Italy, Jun. 2019, pp. 278–282, doi: 10.1109/FMEC.2019.8795327.

[29] S. Barbhuiya, Z. Papazachos, P. Kilpatrick, and D. S. Nikolopoulos, "RADS: Real-time Anomaly Detection System for Cloud Data Centres," *arXiv:1811.04481 [cs]*, Nov. 2018, Accessed: Dec. 28, 2020. [Online]. Available: http://arxiv.org/abs/1811.04481.

[30] T. Salman, D. Bhamare, A. Erbad, R. Jain, and M. Samaka, "Machine Learning for Anomaly Detection and Categorization in Multi-Cloud Environments," in 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), New York, NY, USA, Jun. 2017, pp. 97–103, doi: 10.1109/CSCloud.2017.15.

[31] C. Wu, K. Yu, and X. Wu, "Scalping Anomaly Detection Based on Mobile Internet Traffic Data," in Proceedings of the 2nd International Conference on Telecommunications and Communication Engineering - ICTCE 2018, Beijing, China, 2018, pp. 237–244, doi: 10.1145/3291842.3291905.

[32] G. Pu, L. Wang, J. Shen, and F. Dong, "A hybrid unsupervised clustering-based anomaly detection method," Tinshhua Sci. Technol., vol. 26, no. 2, pp. 146–153, Apr. 2021, doi: 10.26599/TST.2019.9010051.

[33] B. Steenwinckel et al., "FLAGS: A methodology for adaptive anomaly detection and root cause analysis on sensor data streams by fusing expert knowledge with machine learning," Future Generation Computer Systems, vol. 116, pp. 30–48, Mar. 2021, doi: 10.1016/j.future.2020.10.015.

[34] L. Wang, Y. Zhang, and R. Jin, "A Monitoring System for Anomaly Detection in Fog Manufacturing," in 2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS), Tampere, Jun. 2020, pp. 67–72, doi: 10.1109/ICPS48405.2020.9274741.

[35] M. Lotfi Shahreza, D. Moazzami, B. Moshiri, and M. R. Delavar, "Anomaly detection using a self-organizing map and particle swarm optimization," Scientia Iranica, vol. 18, no. 6, pp. 1460–1468, Dec. 2011, doi: 10.1016/j.scient.2011.08.025.

[36] M. V. Ngo, H. Chaouchi, T. Luo, and T. Q. S. Quek, "Adaptive Anomaly Detection for IoT Data in Hierarchical Edge Computing," arXiv:2001.03314 [cs, stat], Jan. 2020, Accessed: Dec. 27, 2020. [Online]. Available: http://arxiv.org/abs/2001.03314.

[37] M. V. Ngo, T. Luo, H. Chaouchi, and T. Q. S. Quek, "Contextual-Bandit Anomaly Detection for IoT Data in Distributed Hierarchical Edge Computing," arXiv:2004.06896 [cs, stat], Apr. 2020, Accessed: Dec. 27, 2020. [Online]. Available: http://arxiv.org/abs/2004.06896.

[38] T. Luo and S. G. Nagarajan, "Distributed Anomaly Detection using Autoencoder Neural Networks in WSN for IoT," 2018 IEEE International Conference on Communications (ICC), pp. 1–6, May 2018, doi: 10.1109/ICC.2018.8422402.

[39] B. A. N.G. and S. S., "Anomaly detection framework for Internet of things traffic using vector convolutional deep learning approach in fog environment," Future Generation Computer Systems, vol. 113, pp. 255–265, Dec. 2020, doi: 10.1016/j.future.2020.07.020.

[40] A. Nanduri and L. Sherry, "Anomaly detection in aircraft data using Recurrent Neural Networks (RNN)," in 2016 Integrated Communications Navigation and Surveillance (ICNS), Herndon, VA, USA, Apr. 2016, pp. 5C2-1-5C2-8, doi: 10.1109/ICNSURV.2016.7486356.

[41] G. Pang, C. Shen, and A. van den Hengel, “Deep Anomaly Detection with Deviation Networks,” in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage AK USA, Jul. 2019, pp. 353–362, doi: 10.1145/3292500.3330871.

[42] S. Maya, K. Ueno, and T. Nishikawa, “dLSTM: a new approach for anomaly detection using deep learning with delayed prediction,” Int J Data Sci Anal, vol. 8, no. 2, pp. 137–164, Sep. 2019, doi: 10.1007/s41060-019-00186-0.

[43] A. ElSaid, F. El Jamiy, J. Higgins, B. Wild, and T. Desell, “Optimizing long short-term memory recurrent neural networks using ant colony optimization to predict turbine engine vibration,” Applied Soft Computing, vol. 73, pp. 969–991, Dec. 2018, doi: 10.1016/j.asoc.2018.09.013.

[44] O. Karaahmetoglu, F. Ilhan, I. Balaban, and S. S. Kozat, “Unsupervised Online Anomaly Detection On Irregularly Sampled Or Missing Valued Time-Series Data Using LSTM Networks,” arXiv:2005.12005 [cs, stat], May 2020, Accessed: Dec. 30, 2020. [Online]. Available: http://arxiv.org/abs/2005.12005..

[45] A.-U.-H. Qureshi, H. Larijani, N. Mtetwa, A. Javed, and J. Ahmad, “RNN-ABC: A New Swarm Optimization Based Technique for Anomaly Detection,” Computers, vol. 8, no. 3, p. 59, Aug. 2019, doi: 10.3390/computers8030059.

[46] B. Staar, M. Lütjen, and M. Freitag, “Anomaly detection with convolutional neural networks for industrial surface inspection,” Procedia CIRP, vol. 79, pp. 484–489, 2019, doi: 10.1016/j.procir.2019.02.123.

[47] T. Wen and R. Keyes, “Time Series Anomaly Detection Using Convolutional Neural Networks and Transfer Learning,” arXiv:1905.13628 [cs, stat], May 2019, Accessed: Dec. 31, 2020. [Online]. Available: http://arxiv.org/abs/1905.13628.

[48] B. Lu, Z. Lv, and S. Zhu, “Pseudo-3D Residual Networks Based Anomaly Detection in Surveillance Videos,” in 2019 Chinese Automation Congress (CAC), Hangzhou, China, Nov. 2019, pp. 3769–3773, doi: 10.1109/CAC48633.2019.8996830

[49] T. Hayashi, T. Komatsu, R. Kondo, T. Toda, and K. Takeda, “Anomalous Sound Event Detection Based on WaveNet,” in 2018 26th European Signal Processing Conference (EUSIPCO), Rome, Sep. 2018, pp. 2494–2498, doi: 10.23919/EUSIPCO.2018.8553423.

[50] M. Hasnain, S. Ryul Jeong, M. Fermi Pasha, and I. Ghani, “Performance Anomaly Detection in Web Services: an RNN-based Approach Using Dynamic Quality of Service Features,” Computers, Materials & Continua, vol. 64, no. 2, pp. 729–752, 2020, doi: 10.32604/cmc.2020.010394.

[51] Prarthi Jain, Seemandhar Jain, June 3, 2020, "Anomaly detection dataset", IEEE Dataport, doi: https://dx.doi.org/10.21227/rt7n-2x60.

[52] M. Pirker, P. Kochberger, and S. Schwandter, “Behavioural Comparison of Systems for Anomaly Detection,” in Proceedings of the 13th International Conference on Availability, Reliability and Security, Hamburg Germany, Aug. 2018, pp. 1–10, doi: 10.1145/3230833.3230852.

[53] H. Fanaee-T and J. Gama, “Tensor-based anomaly detection: An interdisciplinary survey,” Knowledge-Based Systems, vol. 98, pp. 130–147, Apr. 2016, doi: 10.1016/j.knosys.2016.01.027.

[54] M. Ahmed, A. Naser Mahmood, and J. Hu, “A survey of network anomaly detection techniques,” Journal of Network and Computer Applications, vol. 60, pp. 19–31, Jan. 2016, doi: 10.1016/j.jnca.2015.11.016.

[55] S. Agrawal and J. Agrawal, “Survey on Anomaly Detection using Data Mining Techniques,” Procedia Computer Science, vol. 60, pp. 708–713, 2015, doi: 10.1016/j.procs.2015.08.220. https://doi.org/10.1016/j.neucom.2017.04.070

[56] Fernandes, G., Rodrigues, J.J.P.C., Carvalho, L.F. *et al.* A comprehensive survey on network anomaly detection. *Telecommun Syst* **70,** 447–489 (2019). https://doi.org/10.1007/s11235-018-0475-8., Springer, 2019.

[57] Nickolaos Koroniotis, Nour Moustafa, Elena Sitnikova, Benjamin Turnbull, “Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset”, https://arxiv.org/abs/1811.00701, 2018

[58] Shebuti Rayana, ODDS Library [http://odds.cs.stonybrook.edu], Stony Brook, NY: Stony Brook University, Department of Computer Science, 2016.

[59] Yahoo Webscope Dataset [https://webscope.sandbox.yahoo.com/], Yahoo Inc.

[60] Dua, D. and Graff, C., UCI Machine Learning Repository [http://archive.ics.uci.edu/ml], Irvine, CA: University of California, School of Information and Computer Science,2019.

[61] Goldstein, Markus, "Unsupervised Anomaly Detection Benchmark", https://doi.org/10.7910/DVN/OPQMVF, Harvard Dataverse, 2015.

[62] UCSD Anomaly Detection Dataset, [http://www.svcl.ucsd.edu/projects/anomaly/dataset.htm], 2010

[63] M. Braei and S. Wagner, “Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art,” arXiv:2004.00433 [cs, stat], Apr. 2020, Accessed: Jan. 09, 2021. [Online]. Available: http://arxiv.org/abs/2004.00433.

[64] M. Behniafar; A.R. Nowroozi; H.R. Shahriari. "A Survey of Anomaly Detection Approaches in Internet of Things". **The ISC International Journal of Information Security**, 10, 2, 2018, 79-92. doi: 10.22042/isecure.2018.116976.408

[65] G. Muruti, F. A. Rahim, and Z.-A. bin Ibrahim, “A Survey on Anomalies Detection Techniques and Measurement Methods,” in 2018 IEEE Conference on Application, Information and Network Security (AINS), Langkawi, Malaysia, Nov. 2018, pp. 81–86, doi: 10.1109/AINS.2018.8631436.

[66] Aleroud, A., Karabatis, G. Contextual information fusion for intrusion detection: a survey and taxonomy. Knowl Inf Syst 52, 563–619, Springer, 2017. https://doi.org/10.1007/s10115-017-1027-3.