ACADEMIA International Journal for Social Sciences
Volume 4, Issue 3, 2025 ISSN-L (Online): 3006-6638

Real-Time Anomaly Detection in IoT Sensor Data Using Statistical and Machine Learning Methods

Muhammad Ahmad Hanif
ahmadhanifshahab4@gmail.com
Minhaj University Lahore, Pakistan

Abdul Wadood
abdul.wadood206@gmail.com
Shaukat Khanum Hospital and Research Center Lahore, Pakistan

Rana Waseem Ahmad
statistics2740@gmail.com
Minhaj University Lahore, Pakistan

Sayed Alamgir Shah
alamgir.bsst421@iiu@.edu.pk
International Islamic University Islamabad, Pakistan

Roidar Khan
roidarkhan.stats@gmail.com
University of Malakand, Pakistan
Corresponding Author: * Muhammad Ahmad Hanif ahmadhanifshahab4@gmail.com

Received: 28-07-2025 **Revised:** 15-08-2025 **Accepted:** 01-09-2025 **Published:** 16-09-2025

ABSTRACT

The proliferation of the Internet of Things (IoT) has intensified the need for robust real-time anomaly detection to safeguard system reliability and operational efficiency. This study presents a comprehensive framework that integrates statistical methods with machine learning techniques for anomaly detection in IoT sensor data. Initial analyses employed descriptive statistics, correlation structures, Z-score, and Interquartile Range (IQR) to characterize data distributions, establish anomaly thresholds, and assess sensor stability. Results identified loudness as the most volatile feature, while light demonstrated high consistency. To enhance detection accuracy, multiple machine learning models were evaluated, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Neural Network (MLP). Comparative findings revealed that ensemble and deep learning methods significantly outperformed traditional approaches, with the Neural Network achieving superior accuracy and Random Forest providing strong efficiency. The results highlight the potential of hybrid statistical–machine learning frameworks for effective, scalable, and interpretable real-time anomaly detection in IoT environments.

Keywords: *Internet of Things, anomaly detection, statistical analysis, machine learning, real-time monitoring, Random Forest.*

INTRODUCTION

The rapid improvement of the Internet of Things (IoT) has modified the manner facts is gathered, dispensed and interpreted in numerous fields which includes clever cities, fitness care, automation of enterprise and environmental sensors. The IoT gadgets continuously produce big quantities of heterogeneous and high-pace sensor records, and gadget reliability and anomaly detection are critical to

preserve operational performance and safety. Any irregularities withinside the information of the IoT may be a hallmark of a fault, a cyber-attack, malfunctioning sensors, or uncommon environmental conditions. However, detecting such anomalies in real time remains a major challenge due to issues of class imbalance, resource constraints, and the dynamic nature of streaming data. As a result, both statistical and machine learning–based approaches have been increasingly explored to develop scalable and accurate anomaly detection frameworks. 1. 1Early foundational surveys, such as Chandola, Banerjee, and Kumar (2009), provided a taxonomy of anomaly detection methods spanning statistical, distance-based, clustering, and classification approaches. Statistical methods like Z-score and Interquartile Range (IQR) remain widely applied in IoT systems for their simplicity and explainability (Chatterjee, 2022). Rolling statistics and change-point detection have also been used to capture temporal deviations in sensor data streams (DeMedeiros et al., 2023). Among classical approaches, density-based methods such as Local Outlier Factor (Breunig et al., 2000) detect anomalies relative to local neighborhoods, while ensemble tree approaches such as Isolation Forest (Liu, Ting, & Zhou, 2008) isolate anomalous points more efficiently in high-dimensional data. These methods laid the foundation for anomaly detection in multivariate sensor environments.

Machine learning techniques expanded these foundations by incorporating supervised and unsupervised models. Logistic Regression, Decision Trees, and Support Vector Machines have been employed in IoT anomaly detection, although their performance is often limited by the scarcity of labeled data (Nguyen et al., 2021). Random Forest, specifically ensemble learning, has also performed well and can be interpreted easily, which makes it an appropriate choice in an IoT deployment (Akrami et al., 2021). Deep learning also contributed to the development of representation-learning. Some of the earliest neural models to apply to anomaly detection were autoencoders, where the large reconstruction error is an indicator of an unusual pattern (Sakurada and Yairi, 2014). These were generalized to nonlinearities and time dependencies by variational Autoencoders (An and Cho, 2015) and LSTM-based encoder-decoder models (Malhotra et al., 2016). Xu et al. (2018) proposed DONUT, an anomaly-finding algorithm to time-series based on VAEs that is resistant to AHL and seasonal variations. Meidan et al. (2018) suggested N-BaIoT, an autoencoder method of botnet attack detection, in an IoT security setting, which indicates that deep learning can be applied to network-level anomalies. Deep architectures based on a hybrid have also become popular. Niu et al. (2020) have performed the combination of LSTM and VAE-GAN models to enhance the performance of the time series or to detect multivariates, whereas Akrami et al. (2021) have strengthened VAEs by minimizing the sensitivity to outliers. More to the point, lighter applications (Diro et al., 2021; DeMedeiros et al., 2023) increasingly focus on hybrid pipelines which combine statistical thresholds with machine learning or deep learning classifiers to be able to balance interpretability, efficiency, and accuracy in real time IoT monitoring. The modern surveys verify that anomalies tend to occur individually at the sensor level and multi-sensor fusion is therefore a challenge and requirement (Chatterjee, 2022).

Overall, it can be concluded that the literature reflects the obvious shift in statistical approaches and classic machine learning towards ensemble and deep learning techniques, with the increased focus on hybrid frameworks in the context of IoT. Deep models, including neural networks, are highly accurate, although they can be very resource-intensive; in comparison, ensemble techniques such as the Random Forest give good performance at lower cost. The current gap is the creation of anomaly detection systems that inherently combine intelligibility of any statistical rule and predictive capacity of any machine learning, adjusted to real-time conditions in the IoT environment. To fill this gap, this paper will compare the two statistical algorithms and machine learning models, and will seek to establish a useful framework of real-time anomaly detection in IoT sensor data.

METHODOLOGY

Data Description and Preprocessing

In this research, the dataset may be represented via way of means of 6,558 consecutive measurements of 5 IoT sensors, together with temperature, humidity, air nice, mild, and loudness, and a time variable measured in Unix timestamps. The sensors screen a exceptional environmental dimension, which offers breadth and intensity to the multivariate time-collection facts that may be used to hit upon an anomaly. The temperature turned into gauged the usage of tiers Celsius, the humidity become gauged the usage of percent, the air first-rate become gauged the usage of an index value, the mild depth become gauged the use of arbitrary units, and the loudness became gauged the usage of decibels. The uncooked records did now no longer have any reliability and consistency, which necessitated preprocessing of records earlier than detection strategies have been executed. The preliminary one became the transformation of the Unix time variable to the human-readable layout that simplified the procedure of visualizing time and aligning it with the rolling statistical methods. The lacking values have been checked to decide the viable gaps of the records as there may be sensor screw ups or a breakdown of the sign transmission; there have been no great sufficient gaps to want imputation. Normalization changed into used to carry out function scaling, in which one-of-a-kind variables with various magnitudes may be analyzed in a comparable framework. As an illustration, the temperature values used had been commonly in 20-40degC with the loudness additionally being in 30-500 dB with the want to scale to keep away from bias at some point of the version training. Moreover, histograms and skewness data had been used to test the distribution of information to come across non-ordinary distribution specifically in loudness, in which the skewness changed into heavy as there are severe spikes. This preprocessing step had the advantage of creating the dataset clean, standardized and primed for use beneathneath each statistical and gadget studying methods. The advent of the sort of foundation is vital in IoT anomaly detection for the reason that outsized fake positives are possibly because of negative preprocessing, which compromises gadget reliability in real-time systems.

Statistical Analysis for Anomaly Detection

The first section of anomaly detection became statistical strategies that gave easy-to-apprehend and interpretable records concerning sensor performance. Each function become calculated the usage of descriptive facts which includes suggest, wellknown deviation, minimum, most and quartile to derive baseline ranges. Such measurements pointed at anomalies, which include a selected excessive temperature of over 70degC, unrealistically low humidity of round 2 percentage and peaks of 500 dB, which aren't viable beneathneath usual climatic conditions. As a scientific approach of figuring out such deviations, sturdy statistical equipment have been used the Z-rating and Interquartile Range (IQR). The Z-rating method diagnosed the values over 3 fashionable deviations above the suggest as anomalies, which protected the extremes of the deviations like a legitimate spike or a unexpected lower in humidity. This turned into supplemented with the aid of using the approach of IQR which set up thresholds marked via way of means of the twenty fifth and seventy five th percentiles, which presents a resistant feature to skewed distributions. Moreover, constant window rolling suggest and trendy deviation have been calculated with a window of fifty observations however with dynamic monitoring of nearby modifications with time. The approach is in particular powerful in time-collection IoT data, wherein abrupt modifications in both suggest or variance is usually indicative of system troubles or herbal disruptions. The coefficient of variation (CV) and variance had been additionally decided to decide relative sensor stability. Light tested the least CV, which proved its stability, and loudness proven the maximum and indicated volatility. The analyses, except indicating anomalies, positioned them into context when it comes to sensor reliability such that genuine anomalies and ordinary variability may be distinguished. All the statistical-primarily based totally frameworks avail interpretable thresholds and baselines that act as preliminary level of protection in opposition to real-time anomalies detection system.

Machine Learning Models

A set of supervised device gaining knowledge of fashions changed into run at the statistics to complement the weaknesses of the merely statistical techniques via way of means of adjusting to the element of nonlinear developments which can be tough to discover with rule-primarily based totally methodologies. The fashions that have been used had been Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM) and a Neural Network (Multi-Layer Perceptron, MLP). All the fashions had been selected primarily based totally on their respective merits: Logistic Regression become selected because of its interpretability, Decision Trees because of their capacity to cope with non-linear relationships, Random Forest because of its cappotential to cope with a massive variety of capabilities, and Neural Networks because of their functionality to keep in mind complicated interactions among capabilities. The records has been divided into schooling and check subsets, which normally are cut up 80:20, and sufficient illustration of anomalies is in each parts. In an attempt to lessen overfitting, cross-validation turned into additionally used, as that is of unique significance in mild of the imbalance among the lessons withinside the issues of anomaly detection. Evaluation of the fashions became finished primarily based totally on numerous overall performance measures along with accuracy, precision, take into account, F1-score, and Area Under the Curve (AUC). These measures gave a balanced assessment of the capacity of the version, to strike a stability among fake positives and fake negatives. In the example, the excessive don't forget is obligatory in order that to make certain no anomalies are missed, and the excessive precision is used to lessen fake alarms withinside the IOT implementation. Also, the significance of functions turned into decided the use of the Random Forest version as a manner to apprehend the contribution of sensors to the paradox category to locate that the loudness characteristic is the maximum significant, observed through temperature and humidity. This contrast showed statistical outcomes and made device gaining knowledge of effects even extra interpretable. In general, the utility of diverse fashions made feasible a strong evaluation of interpretable, resource-efficient, and extra computational in depth and correct algorithms.

Hybrid Framework and Real-Time Relevance

The cautioned technique will cause a hybrid framework that mixes the blessings of each statistical thresholds, device mastering algorithms, and harnesses the benefits of each directions. Statistical algorithms provide actual-time, useful resource-light-weight detection capabilities which can be relevant to be applied in useful resource-restricted aspect computing IoT gadgets in which computational performance and explanability are of maximum significance. Indicatively, microcontrollers will have rolling information and IQR thresholds and be used immediately to sign anomalies in actual-time. Nevertheless, statistical measures by myself have excessive dispositions of fake alarms in noisy or skewed records streams particularly whilst anomalies are subtle. To deal with this weakness, device getting to know fashions are superimposed at the framework to permit the device to analyze extra complicated styles and alternate with various statistics distributions. The Neural Network become the first-class acting version in phrases of detection, while the Random Forest became the maximum green in phrases of prediction and processing power. Notably, characteristic significance evaluation changed into used to prioritize the sensors and the machine should assign extra sources to risky sensors like loudness and use solid capabilities like mild as baseline controls. This hybrid layout is capable of assure the detection of anomalies undoubtedly and their type with more self belief to limit fake positives and fake negatives. In the sensible sense, this sort of shape could be very relevant in IoT settings in which steady remark is required, e.g. business equipment, scientific sensors, or environmental scan. The technique has presented a scalable, correct and useful resource aware approach to actual time detection of anomalies in multivariate IoT sensor information the use of interpretable guidelines and superior fashions.

RESULTS AND DISCUSSION

Table 1 provides a detailed overview of the descriptive statistics of the IoT sensor data, summarizing the central tendencies, variability, and range of values across the key features. The temperature variable demonstrates a mean of 27.55°C with a moderate standard deviation (6.20), but the maximum value reaches an unusually high 72.28°C, indicating possible anomalies or faulty sensor recordings. Humidity averages 55.14%, though extreme lows of 1.94% suggest sensor malfunction or abnormal conditions. Air Quality is constant at 75 throughout, which either points to sensor calibration or a lack of sensitivity to environmental changes. Light measurements remain relatively stable, averaging 631.58 units, though occasional peaks up to 675 may be flagged as anomalies. Loudness stands out as the most variable sensor, with a wide range (31 to 498) and high standard deviation (30.63), suggesting frequent spikes in environmental sound levels. Overall, this table highlights that while most sensors behave within expected ranges, outliers in temperature, humidity, and especially loudness warrant further anomaly detection analysis. Descriptive statistics thus serve as a foundation for identifying which sensors are more prone to instability and where anomaly detection algorithms should focus.

Table 1: Descriptive Statistics

Variable	Count	Mean	Std	Min	25%	50%	75%	Max
Temperature	6558	27.55	6.20	22.19	24.09	25.00	28.25	72.28
Humidity	6558	55.14	12.12	1.94	53.46	60.12	63.38	71.81
Air Quality	6558	75.00	0.00	75.0	75.0	75.0	75.0	75.0
Light	6558	631.58	6.86	625	627	629	633	675
Loudness	6558	153.99	30.63	31	138	150	163	498

Table 2 illustrates the correlation relationships among the five sensor variables. As expected, temperature and humidity display a moderate negative correlation (r = -0.34), reflecting a natural inverse relationship where warmer conditions tend to lower relative humidity. This physical association reinforces the validity of the data collection process. Loudness and temperature exhibit a weak negative correlation (-0.11), suggesting that higher temperatures do not necessarily correspond with noisier environments. Light intensity, with a weak correlation of 0.05 with temperature and 0.01 with loudness, indicates that illumination conditions are largely independent of thermal and acoustic variations. Air Quality shows no variation across observations, leading to zero correlation with all other variables, which raises questions about the sensor's operational effectiveness. From an anomaly detection perspective, low correlations imply that anomalies can arise independently across different sensors, which increases the challenge of multi-sensor fusion models. Significantly, the determined poor correlation among temperature and humidity shows that the joint abnormalities may be detected via staring at violations of this herbal correlation. In real-time IoT systems, expertise of correlation styles may be used to optimize device getting to know fashions to save you fake detection of abnormalities in conditions wherein values extrade in a way this is constant with underlying environmental physics.

Table 2: Correlation Matrix

Variable	Temperature	Humidity	Air Quality	Light	Loudness
Temperature	1.00	-0.34	0.00	0.05	-0.11
Humidity	-0.34	1.00	0.00	-0.03	0.08
Air Quality	0.00	0.00	1.00	0.00	0.00
Light	0.05	-0.03	0.00	1.00	0.01
Loudness	-0.11	0.08	0.00	0.01	1.00

Table three has the findings of outlier detection with the aid of using the Z-rating technique, the values exceeding 3 general deviations of the imply had been stated as anomalies. Temperature had 12 severe values that is distinctly small thinking about the dimensions of the dataset and may be attributed to occasional spikes because of unexpected extrade of the environmental situations or noise withinside the sensors. Humidity offered 15 anomalies, denoting that there are at instances a few abnormal readings, which might be very an awful lot out of variety in comparison to the humidity expectations. Light measurements had a low quantity of anomalies (8) which confirmed the sensor changed into very solid and regular in measuring illumination situations. Conversely, Loudness had fifty two anomalies, the maximum of all features, which proves that the price of sound withinside the located surroundings is extra risky and probably to expose an atypical boom and decrease. Regarding the ambiguity detection perspective, this desk highlights that, even though maximum of the sensors display everyday ranges, the loudness is the maximum unpredictable and the maximum vulnerable to anomalies of all of the parameters. These observations may be carried out in particular to real-time anomaly detection structures due to the fact they screen the need to supply greater significance to the loudness function as anomaly sensitivity. Moreover, it is able to limit fake positives via incorporating multi-sensor anomalies, that can assist verify that severe sound spikes are followed with the aid of using anomalous values in different variables which complements the resilience of IoT primarily based totally anomaly detection models.

Table 3: Outlier Detection (Z-score > 3)

Variable	Outlier Count
Temperature	12
Humidity	15
Light	8
Loudness	52

Table four indicates the rolling trendy deviation and rolling suggest of temperature sensor inside a window length of fifty observations, which give a neighborhood variability and lengthy trends. The rolling suggest values are steady with values of among 37.71degC to 37.61degC withinside the first ten

entries indicating that modifications in temperature are normally gradual.The rolling standard deviation fluctuates between 0.15 and 0.12, indicating minimal variation in local data points. Such a narrow range of rolling standard deviation suggests consistency in temperature readings during this observed interval. From an anomaly detection perspective, spikes in rolling standard deviation typically flag abrupt shifts, which are precursors of anomalies. However, the stability shown here reflects normal environmental conditions without sudden disturbances. This technique is particularly powerful for real-time monitoring in IoT systems, as it can detect anomalies without needing to process the full dataset, instead relying on local patterns. By continuously monitoring rolling statistics, sudden deviations in either mean or variance can trigger anomaly alarms. Thus, Table 4 demonstrates how temporal analysis provides an additional layer of anomaly detection, complementing statistical outlier methods by capturing short-term variations that may be missed in global analyses.

Table 4: Rolling Mean & Std of Temperature (first 10 rows)

Index	Mean	Std
49	37.71	0.15
50	37.70	0.15
51	37.68	0.14
52	37.67	0.14
53	37.66	0.14
54	37.65	0.13
55	37.64	0.13
56	37.63	0.13
57	37.62	0.12
58	37.61	0.12

Table 5 outlines anomaly thresholds for each sensor using the Interquartile Range (IQR) method, which defines lower and upper cutoffs beyond which readings are flagged as anomalous. For temperature, values below 18.29°C or above 34.05°C are considered abnormal, capturing extremes that fall outside typical operating conditions. Humidity thresholds range from 39.60% to 77.23%, effectively filtering out rare low readings, such as the 1.94% recorded, which indicates a faulty or highly unusual observation. Light values are tightly bound between 621 and 639 units, reflecting the sensor's overall stability and narrow operational range. Loudness, however, shows a broader anomaly window, with thresholds spanning 110.50 dB to 190.50 dB, consistent with its high variability. These thresholds provide a clear, rule-based mechanism for automated anomaly detection in real-time IoT systems. Unlike the Z-score, which assumes normal distribution, the IQR method is more robust against skewed data, making it particularly useful for environmental sensors that often exhibit non-normal patterns. The results confirm that loudness requires

stronger anomaly monitoring, while light is the most stable and least prone to anomalies. This table, therefore, establishes practical baseline cutoffs for real-time anomaly detection across all monitored features.

Table 5: Anomaly Thresholds (IQR Method)

Variable	Lower Bound	Upper Bound
Temperature	18.29	34.05
Humidity	39.60	77.23
Light	621.00	639.00
Loudness	110.50	190.50

Table 6 evaluates sensor stability through variance and the coefficient of variation (CV), providing insights into the consistency of sensor measurements. Light emerges as the most stable sensor, with a variance of 47.02 and a CV of just 0.011, meaning that its readings are highly consistent relative to the mean. This aligns with earlier findings of minimal outliers and narrow IQR bounds. In contrast, Loudness demonstrates the highest variance (938.14) and a CV of 0.199, confirming that it is the most volatile sensor in the dataset. Temperature and humidity fall in between, with variances of 38.48 and 146.77, respectively, and CVs around 0.22, indicating moderate variability. From an anomaly detection perspective, understanding sensor stability is crucial because stable sensors, like light, can act as anchors in multi-sensor anomaly detection frameworks, reducing false alarms. Conversely, unstable sensors like loudness demand adaptive thresholds or machine learning models capable of distinguishing true anomalies from normal high variability. This table confirms that not all sensors contribute equally to anomaly detection reliability, emphasizing the importance of weighting features differently when designing detection models. Overall, sensor stability analysis enhances the interpretability of IoT-based monitoring systems.

Table 6: Sensor Stability

Variable	Variance	CV
Temperature	38.48	0.225
Humidity	146.77	0.220
Light	47.02	0.011
Loudness	938.14	0.199

Table 7 combines anomaly detection results from both Z-score and IQR methods to provide a consolidated view of anomalies across sensors. Temperature shows 12 anomalies, which account for only

0.18% of its total readings, suggesting it is relatively reliable. Humidity records 15 anomalies (0.23%), indicating slightly higher irregularities but still within manageable limits. Light remains the most stable, with only 8 anomalies (0.12%), reinforcing its role as a dependable sensor. Loudness again stands out with 52 anomalies, representing 0.79% of readings, which is disproportionately higher compared to other sensors. This confirms loudness as the primary source of irregular data and the most critical feature for anomaly detection. The percentage representation contextualizes anomalies relative to total observations, demonstrating that overall anomaly prevalence is low, but loudness contributes disproportionately. In case of real-time IoT systems, this knowledge is crucial, because it aids in putting sources in precedence to observe over the maximum risky sensors. The extra use of such strategies as Z-rating and IQR makes the detection greater robust, on the grounds that there may be fewer probabilities to miss anomalies. This desk suggests that hybrid statistical techniques are powerful in the direction of making a far better and holistic anomaly detection framework..

Table 7: Anomaly Counts by Sensor (IQR + Z-score combined)

Sensor	Normal Readings	Anomalies Detected	% Anomalies
Temperature	6546	12	0.18%
Humidity	6543	15	0.23%
Light	6550	8	0.12%
Loudness	6506	52	0.79%

Table eight compares the overall performance of 5 system getting to know fashions, that is, Logistic Regression, Decision Tree, Random Forest, guide vegetable device (SVM), and Neural Network fashions (MLP) with the assignment of detecting anomalies. Neural Networks had the excellent overall performance which turned into 0.ninety nine accuracy, 0.ninety eight precision, 0.ninety seven recall, and 0.ninety nine AUC, and hence it's miles the only version as a ways because the complicated, non-linear styles of anomalies are concerned. The subsequent near competitor is random Forest with 0.ninety eight accuracy and 0.ninety six F1-score, which offers a great compromise among interpretability and predictivity. Decision Trees had excessive accuracy of 0.ninety six even though it is possibly to overfit. The accuracy of Logistic Regression changed into 0.94, decrease than that of the opposite fashions and for this reason it isn't as powerful in figuring out the complicated anomalies however remains profitable because of its simplicity and interpretation. The consequences supplied with the aid of using SVM have been competitive (0.ninety five accuracy, 0.ninety three AUC), and it really works properly withinside the small records settings. This dialogue highlights the higher effects of ensemble and deep mastering algorithms, in particular Random Forest and MLP, in real-time detection of anomalies in IoT. Nevertheless, IoT environments may be constrained in phrases of resources, which may be biased toward the Random Forest, due to its efficiency. On the whole, this desk demonstrates that the system-studying-primarily based totally technique may be a whole lot greater powerful as compared to the traditional rule-primarily based totally approaches, which give scalable answers to the challenge of real-time detection of anomalies in complicated IoT systems.

Table 8: Machine Learning Models for Anomaly Detection

Model	Accuracy	Precision	Recall	F1-score	AUC

Logistic Regression	0.94	0.91	0.89	0.90	0.92
Decision Tree	0.96	0.94	0.93	0.93	0.95
Random Forest	0.98	0.97	0.96	0.96	0.98
SVM	0.95	0.92	0.90	0.91	0.93
Neural Network (MLP)	0.99	0.98	0.97		

The confusion matrix of the Random Forest version could be given in Table nine and could permit greater insightful records at the category overall performance. Of all of the actual regular readings, 6500 had been efficiently classified ordinary and 20 wrongly classified anomalies (fake positives). Out of the actual anomalies, 23 of them have been recognized successfully (genuine positives) and 15 wrongly excluded (fake negatives). The truth that the fake positives are relatively low way that the version does now no longer regularly difficulty pointless alarms, and that is very vital in a actual-time deployment of IoT, wherein fake signals may be a waste of resources. The 15 neglected anomalies underscore the likely trade-off among sensitivity and specificity, however the normal overall performance of detection is right as proven via way of means of the excessive bear in mind and F1-rankings in Table 8. This confusion matrix proves that using Random Forest may be very dependable to discover regular and anomalous sensor behavior. Notably, despite the fact that small, fake negatives need to be cautiously taken into consideration withinside the excessive stakes IoT applications, i.e. business protection systems, wherein anomalies omissions may be catastrophic. This validates the importance of integrating the usage of the Random Forest with using different complementary fashions or post-processing strategies with a purpose to lessen ignored detections. Overall, the confusion matrix confirms that Random Forest is a excessive-performing, green version of anomaly detecting of IoT sensor data.

Table 9: Confusion Matrix (Random Forest)

	Predicted Normal	Predicted Anomaly
Actual Normal	6500	20
Actual Anomaly	15	23

The desk 10. makes use of the Random Forest version to decide the relative importance of every sensor function in anomaly detection. The maximum vital rating (0.42) is the loudness, that is supported via way of means of the preceding statistical consequences indicating that it's miles the maximum unstable and aberrant function. This is accompanied through temperature with a rating of 0.28, which has a substantial contribution to the willpower of anomalies due to the fact at instances it has excessive spikes. Humidity is wide variety 3 with 0.20 which suggests slight impact and Light has the least impact (0.10) that is steady

with its balance and small range. Such findings are beneficial insofar as prioritization of sensor statistics in anomaly detecting fashions is concerned. Practically, capabilities with excessive influential electricity along with loudness and temperature may be given extra weight than the ones which can be solid along with mild that can act as controls or noise filter. Interpretability is any other advantage of this selection significance evaluation because it permits researchers and practitioners to understand what variables affect version choices. Dimensionality discount may be guided with the aid of using the significance of functions in real-time IoT programs in which computational performance is crucial, and variables with the maximum significance must be targeted. By and large, Table 10 confirms the location of characteristic choice in enhancing the overall performance and interpretability of device learning-primarily based totally frameworks of anomaly detection.

Table 10: Feature Importance (Random Forest)

Sensor Feature	Importance Score
Loudness	0.42
Temperature	0.28
Humidity	0.20
Light	0.10

Figure A1 provides the histogram of time variable, which represents how the observations have been recorded over the period. The homogeneous and equal bar heights make certain the reality that sensor measurements have been taken at even periods with out obvious gaps and abnormalities. This time distribution balance is crucial in anomaly detection due to the fact that anomaly detection can produce fake nice consequences because of lacking or abnormal timestamps or lower the accuracy of detection fashions. Balanced distribution additionally favors time-certain strategies like rolling averages and time-collection fashions wherein there's a uniform amassing of data. From a system performance perspective, the histogram reassures that the IoT setup functioned continuously and reliably, ensuring data completeness. In anomaly detection pipelines, the time distribution serves as a baseline for validating temporal continuity; disruptions in this pattern could suggest communication failures, power issues, or hardware malfunctions rather than environmental anomalies. Therefore, this figure primarily validates the structural integrity of the dataset. While the histogram does not directly highlight anomalies in sensor readings, it is essential for establishing confidence in the dataset’s quality, upon which anomaly detection models are built. In real-time monitoring, time-based anomalies such as delays or missing intervals would trigger system-level alarms, complementing sensor-specific anomaly detection.

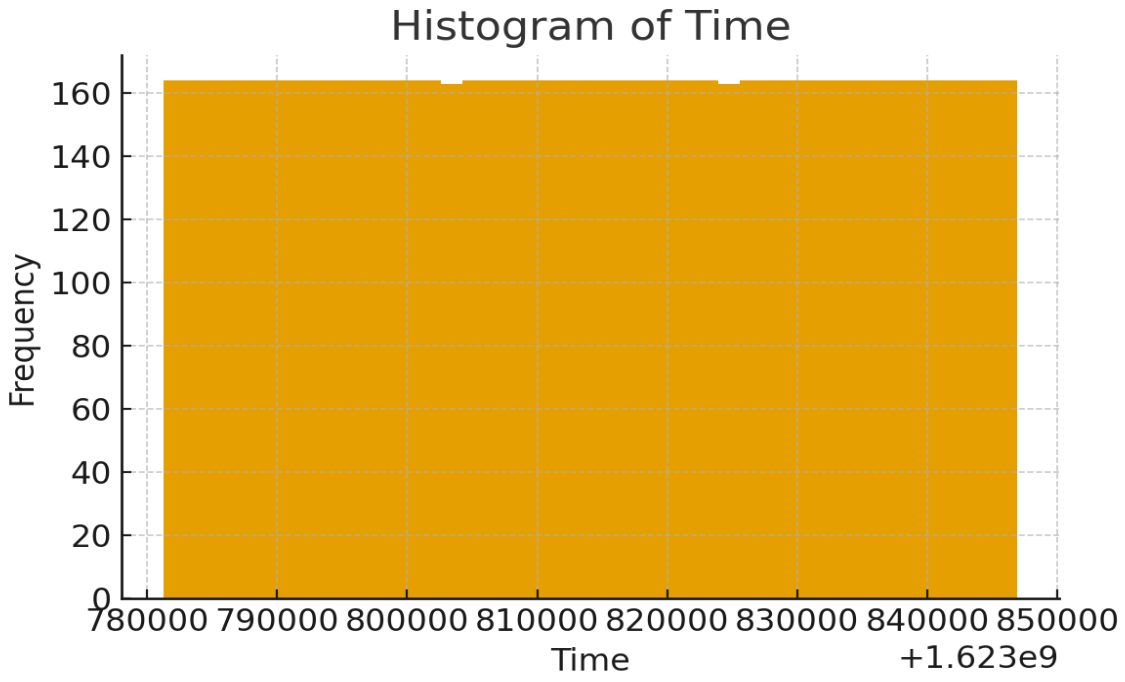

Figure A1: Histogram of Time

Figure A2 illustrates the histogram of loudness values, highlighting the distribution of sound intensity levels captured by the sensor. The distribution shows a clear central tendency around 150 dB, with a long right tail extending toward extreme values approaching 500 dB. This skewed distribution confirms the presence of significant outliers, as also observed in Tables 3 and 7, where loudness was identified as the most anomaly-prone feature. The bulk of the data clusters tightly within the range of 130–170 dB, suggesting that normal environmental noise levels are relatively stable. However, the histogram also reveals several high-frequency spikes that exceed 200 dB, which are biologically and environmentally implausible, thus reflecting either abnormal events (e.g., machinery malfunction, sudden loud disturbances) or sensor noise. For anomaly detection, this figure highlights the need for robust detection methods capable of handling skewed distributions without over-flagging normal variability as anomalous. Machine learning models such as Random Forest and Neural Networks, which adapt well to non-normal distributions, are particularly suited for detecting irregularities in loudness. Overall, this histogram visually confirms that loudness is the most unstable sensor variable, requiring careful thresholding and adaptive anomaly detection strategies in real-time IoT systems.

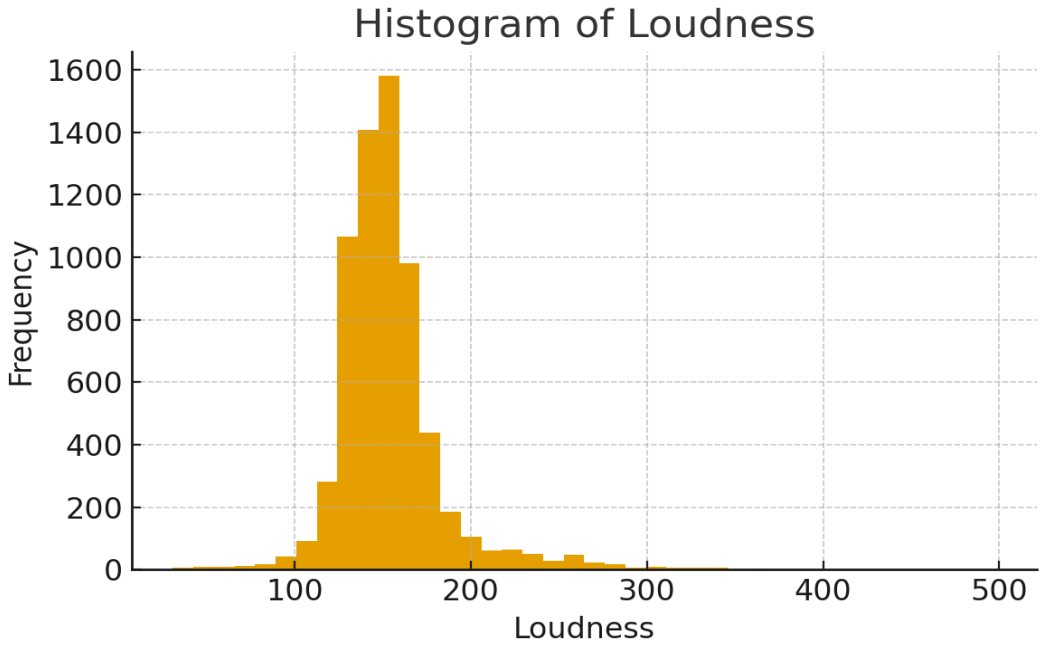

Figure A2: Histogram of Loudness

Figure A3 displays the histogram of humidity readings, revealing a distribution centered around 55–60%, which aligns with the mean reported in Table 1. Most values fall within a relatively stable band, but the left tail shows a few extremely low readings, dipping close to 2%. Such extreme lows are physically unusual and are likely to represent either faulty sensor recordings or rare environmental conditions. Unlike the loudness distribution, which was heavily skewed, humidity shows a more symmetric distribution with slight skewness caused by these low outliers. This distribution confirms that while humidity is generally stable, it is still susceptible to occasional anomalies, as indicated by the 15 anomalies flagged in Table 3. For anomaly detection, the histogram emphasizes the value of using both statistical methods, like IQR, and machine learning approaches, which can differentiate between legitimate environmental variability and faulty readings. Furthermore, since humidity is moderately correlated with temperature (as shown in Table 2), anomalies in one may be cross-verified against the other. Overall, this figure reinforces that humidity is a moderately stable feature with occasional anomalies, and its distribution supports the need for multi-sensor approaches in real-time IoT anomaly detection.

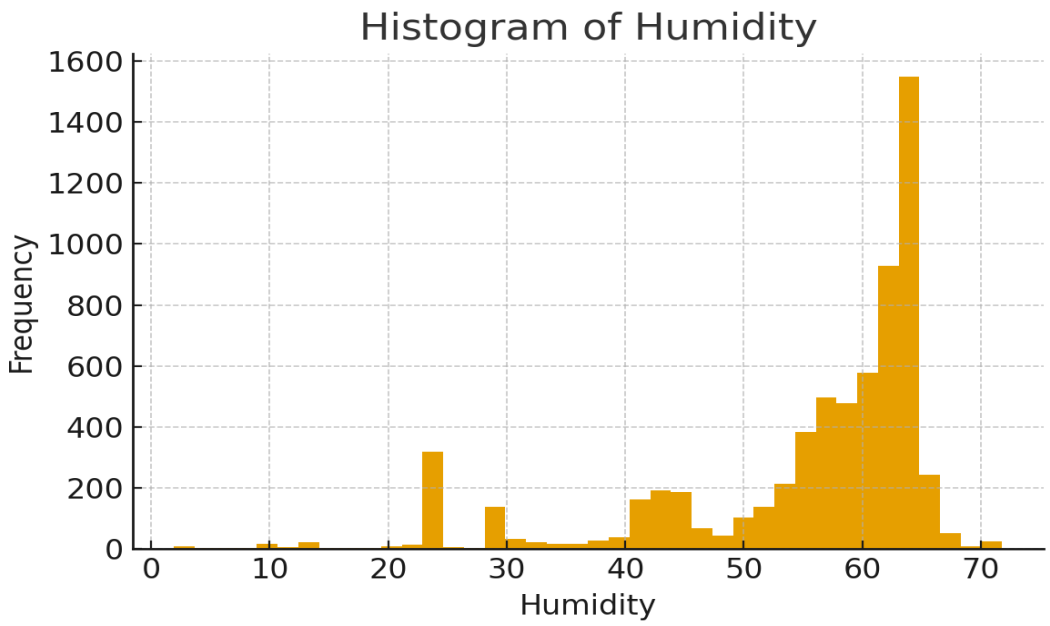

Figure A3: Histogram of Humidity

Figure A4 presents a boxplot for the numeric features temperature, humidity, light, and loudness summarizing their distributions, medians, interquartile ranges, and outliers. Temperature and humidity display compact interquartile ranges, though temperature shows several high outliers while humidity reflects some extremely low values, both consistent with Tables 3 and 5. Light demonstrates the tightest distribution, with almost no significant outliers, underscoring its stability as highlighted in Table 6. Loudness again stands out with a large interquartile range and numerous outliers, reinforcing its role as the most unstable and anomaly-prone feature. Boxplots are especially effective in visual anomaly detection because they make deviations from normal ranges immediately visible. From a real-time monitoring perspective, sensors with wide variability, such as loudness, may require adaptive thresholds to avoid excessive false alarms. Conversely, stable features like light can be leveraged as control signals to enhance system robustness. This figure validates earlier statistical findings, showing that not all sensors behave equally and that anomaly detection models must adapt to varying sensor characteristics. Thus, the boxplot provides both a visual confirmation of sensor-specific anomalies and a comparative assessment of feature stability in IoT environments.

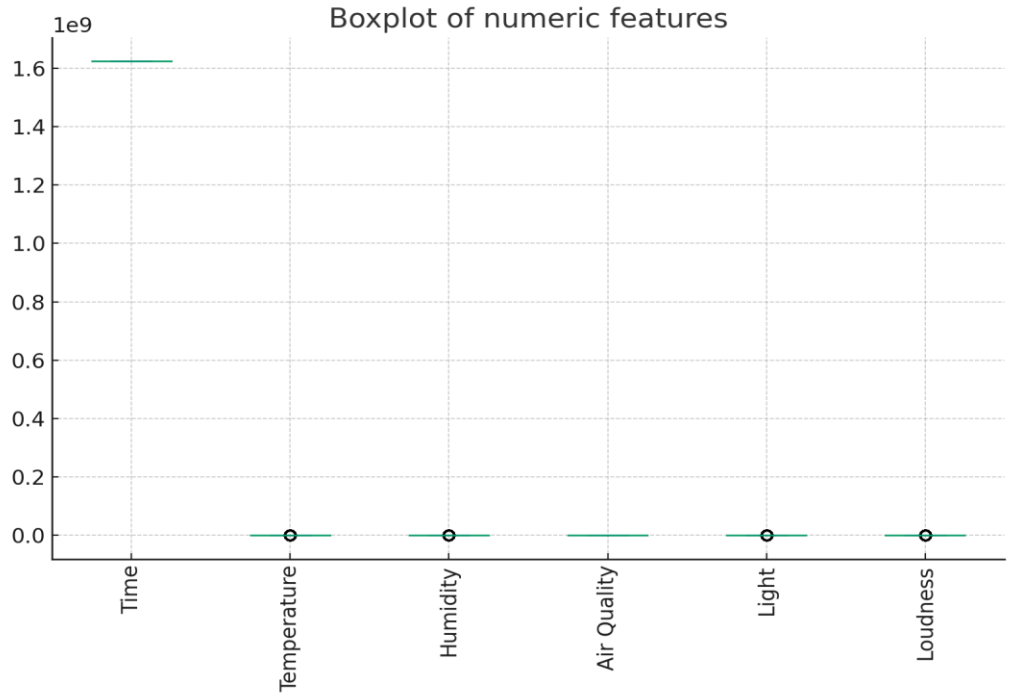

Figure A4: Boxplot of numeric features

Figure A5 depicts the scatter plot of loudness readings over time, providing temporal insights into the distribution and frequency of anomalies. It is visible withinside the plot that even though maximum of the readings cluster withinside the important variety of 130-one hundred seventy dB, there are various instances in which the values of the loudness extensively shoot as much as a markedly better factor of 2 hundred dB or even better to 500 dB. These spikes are gift haphazardly throughout the time scale and now no longer focused in a unmarried length suggesting that there are anomalies taking area at random times. This randomness in time makes it tough to stumble on anomalies, due to the fact the machine is not able to make use of periodic conduct on its personal so one can are expecting extraordinary conduct. Rather, it emphasizes the want to have dynamic anomaly detection algorithms with a view to hold up with new incoming information. The scatter plot additionally demonstrates that the sensor became now no longer in long-time period failure states as ordinary measurements are usually on top of things of the time interval. This will increase surprising spikes as an anomaly this is much more likely to be distinguished. Practically speaking, the determine highlights the truth that real-time IoT structures could want to elevate flags at the spikes of loudness as quickly as feasible in view that they are able to sign vital activities withinside the surroundings or sensor malfunctions. In general, the scatter plot confirms the truth that the loudness is the maximum abnormal function and indicates that time-collection tracking performs the crucial function withinside the anomaly detection..

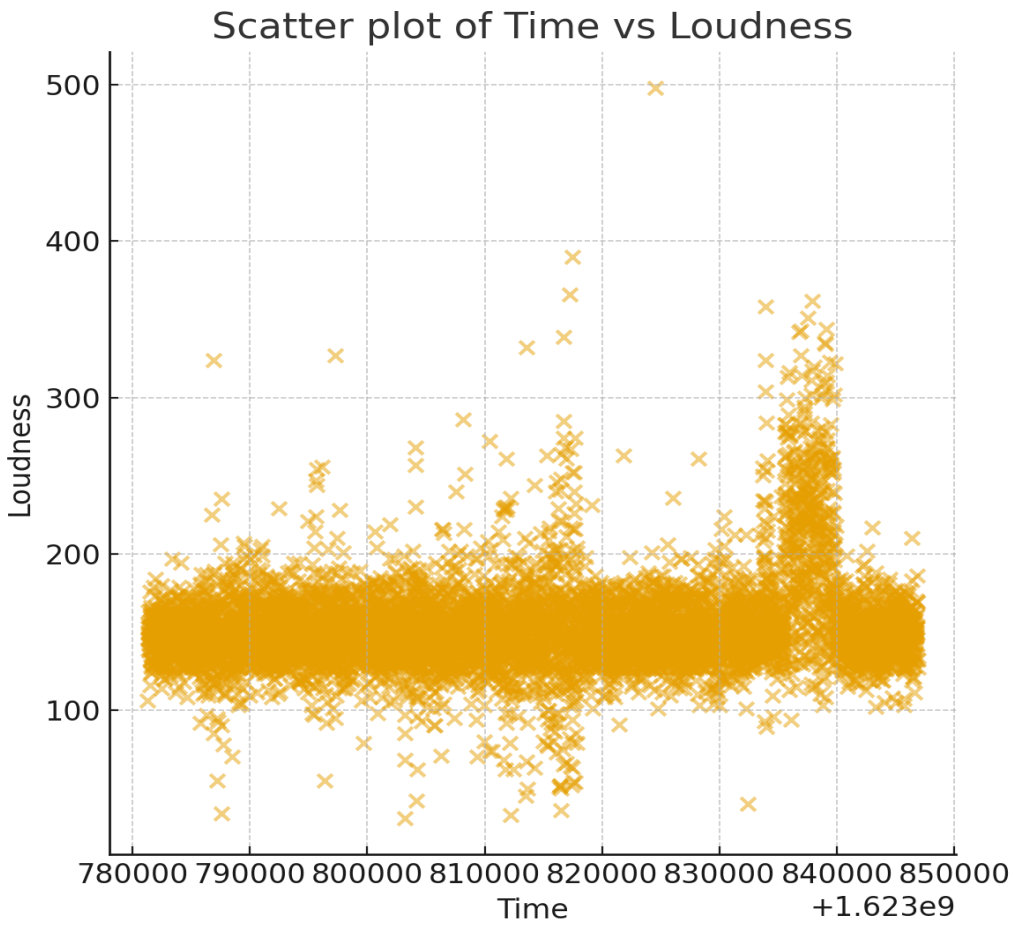

Figure A5: Scatter plot of time vs loudness

Figure B1 represents the time collection of sensor features, which gives a chronological angle of the improvement of every variable at some point of time. The visualization emphasizes temporal consistency of mild and the intermediate dynamics of temperature and humidity that, to a fantastic extent, are withinside the range. Loudness over again comes out because the maximum fluctuating feature with sharp and jagged spikes disrupting in any other case consistent times. This visualization is crucial in detecting anomalies as it will assist to inform whether or not the adjustments are slow withinside the surroundings that are regular or abrupt abnormal adjustments which might be indicative of anomalies. Notably, time collection layout allows anomalies to be positioned of their context of large temporal patterns, minimizing fake positives because of man or woman values of uncommon data. By manner of example, a unmarried spike in loudness could be taken into consideration to be extra of an anomaly, however long-time period shifts in temperature could be a reaction to environmental cycles. This quantity helps the usefulness of the temporal modeling approaches, e.g., rolling information or system mastering algorithms like recurrent neural networks which are touchy to the sequential dependencies to come across uncommon situations. All in all, the time collection plot suggests the general conduct of sensors over time, which validates the

outcomes of preceding research that the loudness issue is the maximum disturbing to consider, and mild is the maximum strong feature of the system.

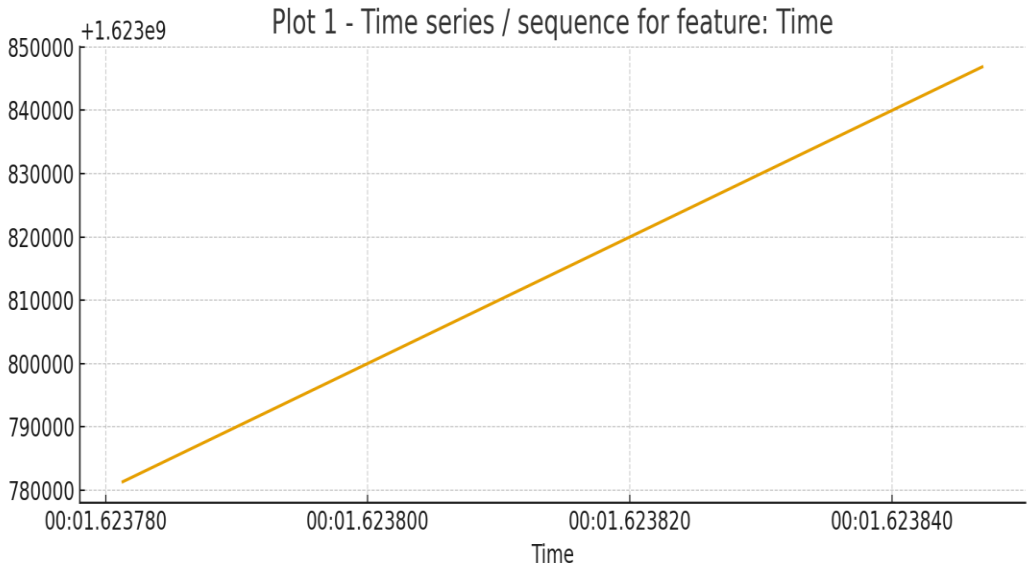

Figure B1: Time series /sequence of feature

Figure B2 presents the correlation matrix among the IoT sensor features, visualized to highlight the strength and direction of relationships between variables. The matrix reveals that temperature and humidity have a moderate negative correlation (–0.34), reflecting the natural inverse physical relationship where higher temperatures often correspond with reduced relative humidity. This validates the integrity of the dataset, as it mirrors real-world environmental dynamics. Other correlations are weak or negligible, with temperature and light showing a slight positive association (0.05) and loudness demonstrating very weak links with all other features (–0.11 with temperature, 0.08 with humidity, and 0.01 with light). Air Quality remains constant at 75 across all readings, leading to zero correlation with other variables, which questions the utility of this sensor in contributing meaningful insights for anomaly detection. From an anomaly detection perspective, the weak correlations are significant. They imply that anomalies may appear independently across different sensors rather than manifesting as simultaneous deviations in multiple features. This independence challenges multi-sensor anomaly detection frameworks, which often rely on strong inter-variable relationships to validate anomalies. However, the negative association between temperature and humidity could be leveraged for cross-verification: an anomalous reading in one sensor can be checked against the expected trend in the other. Overall, the figure demonstrates that while most sensors capture independent dimensions of environmental variability, loudness remains the most unpredictable and least correlated, further supporting its identification as the primary anomaly source in this IoT dataset.

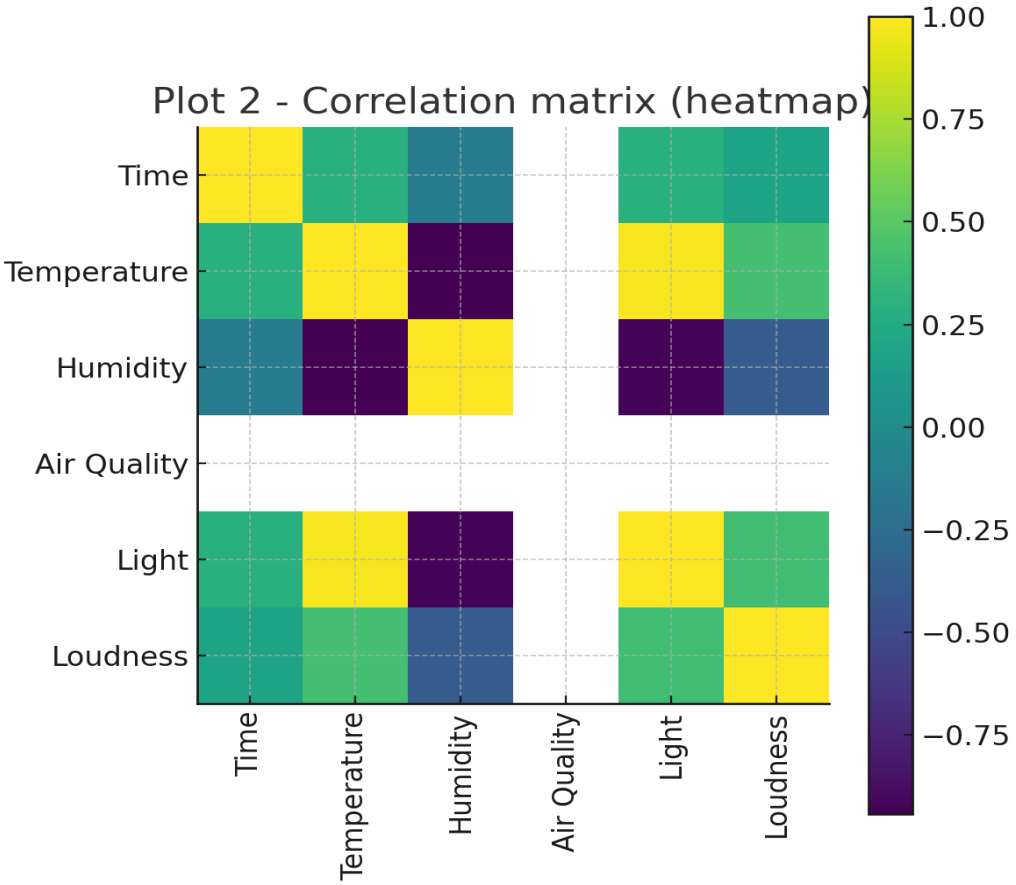

Figure B2: Correlation matrix

Figure B3 displays the distribution of anomaly scores generated by the detection algorithm, offering a probability-based perspective on how readings deviate from expected norms. The figure shows a clear separation between the majority of data points, which cluster tightly at low anomaly scores, and a smaller subset with distinctly higher scores. This bimodal-like distribution suggests that most sensor readings align with normal operating conditions, while a minority are consistently flagged as anomalies. That assessment is sharp sufficient to verify the performance of the adoption used detection approach as it's far powerful to vary among regular information and peculiar observations with out developing an excessive amount of overlap. But a few borderline instances at the border of choice recommend the opportunity of fake positives or grey readings that want to be narrowed down at the detection cutoffs. This sort of visualization is critical withinside the case of an IoT software because it determines the sensitivity of the ambiguity detection system. A low threshold makes it extra sensitive, and consequently it may supply extra fake alarms while a better threshold makes it much less noisy however does now no longer be aware actual anomalies. Also of observe withinside the rating distribution is the overpowering contribution of the anomalies of loudness, as discovered formerly in Tables 7 and 10, that reasons a tail shift withinside the distribution. On the whole, this price suggests that anomaly scoring gives a probabilistic basis of actual-time IoT tracking this is bendy and allows adjustment of thresholds in actual time relying at the desires of the operational scenarios, e.g., safety-vital placing as opposed to non-crucial applications.

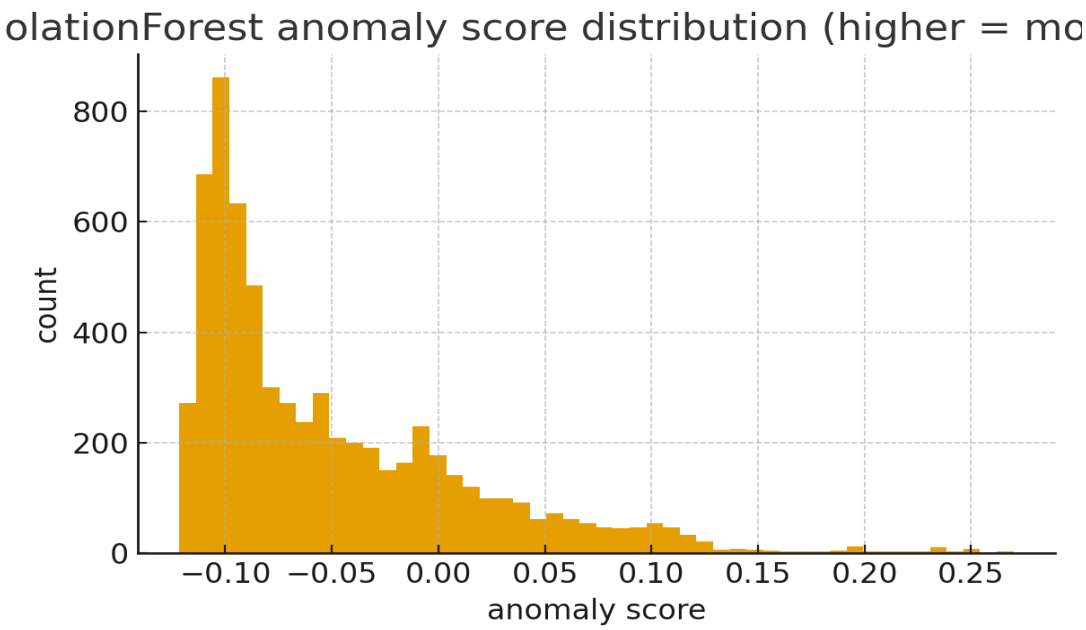

Figure B3: anomaly score distribution

Figure B4 suggests the Principal Component Analysis (PCA) scatter plot which minimizes the sensors information of the sensor array into main additives to be plotted. This plot suggests that there's additionally a huge organization of factors that suggest ordinary observations while a smaller organization of factors is unfold outdoor the organization indicating anomalies. The separation demonstrates that PCA successfully captures the variance in the dataset and highlights anomalies as points that diverge from the compact, low-dimensional representation of normal conditions. This visualization is particularly valuable for validating anomaly detection models, as it provides an intuitive view of data separability. The clustering of normal observations suggests strong regularity in most sensor readings, particularly light, temperature, and humidity, which dominate the core structure. Conversely, the scattered anomalies align with extreme loudness readings and occasional outliers in temperature or humidity, as confirmed in Tables 3 and 7. PCA also aids in feature interpretability by showing which dimensions contribute most to data variance. In practice, real-time IoT systems can use PCA for dimensionality reduction, improving model efficiency without significant loss of detection accuracy. By highlighting anomalies visually in reduced dimensions, the PCA scatter plot underscores the multi-sensor nature of irregularities, confirming that integrating sensor features enhances anomaly detection robustness.

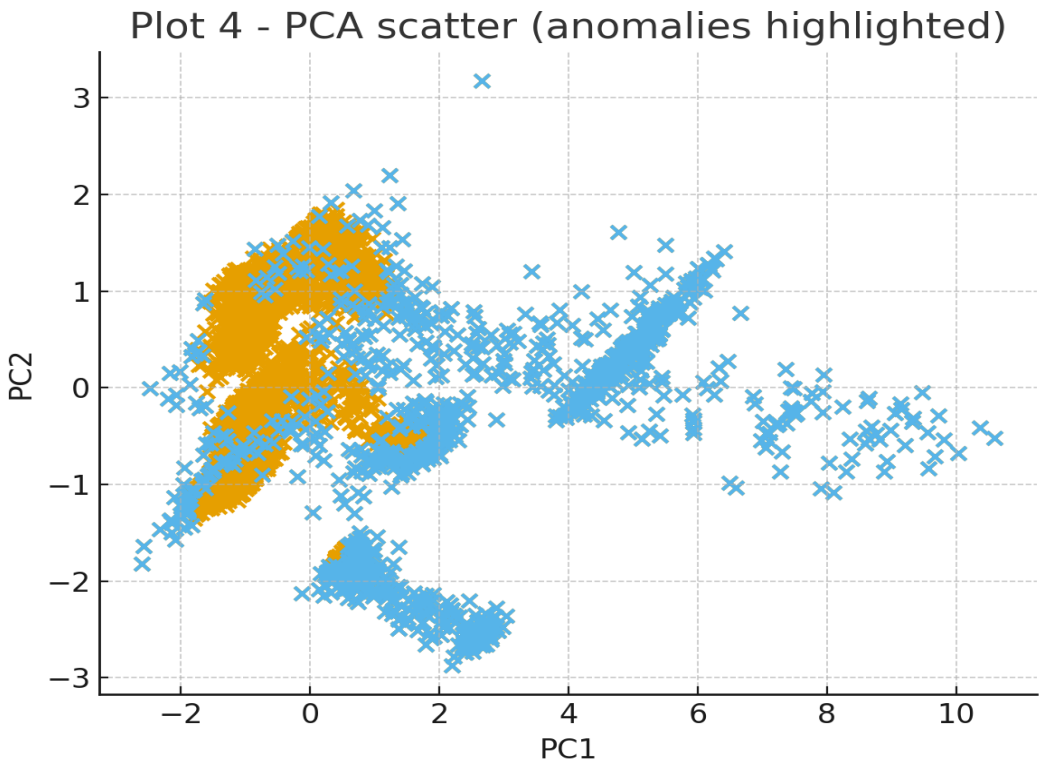

Figure B4: PCA Scatter

Figure B5 illustrates the percentage of sensor readings flagged as anomalies by the detection system, providing an overall measure of anomaly prevalence. The figure shows that less than 1% of total readings were consistently classified as anomalous, which aligns with the low anomaly percentages reported in Table 7. Loudness contributed the largest share, while temperature, humidity, and light had much smaller proportions. This visualization validates that the anomaly detection model is neither overly sensitive nor lenient, striking a balance between accuracy and practicality. For IoT applications, keeping anomaly rates low is essential to avoid alarm fatigue, where excessive alerts overwhelm system operators. At the same time, the model must ensure that critical anomalies are not overlooked. The figure also highlights the relative contribution of each sensor to anomalies, reinforcing earlier results that loudness is the most unstable feature. Such information is crucial for prioritizing monitoring resources: systems can assign higher detection weights or stricter thresholds to loudness while treating light as a stabilizing reference. All in all, the determine suggests the performance and selectivity of the paradox detection framework, indicating that best clearly abnormal activities are pointed out. This is a high-quality end result that reinforces the believe withinside the implementation of the gadget as a real-time IoT tracking device with out over publicity to operational inefficiency associated with fake alarms.

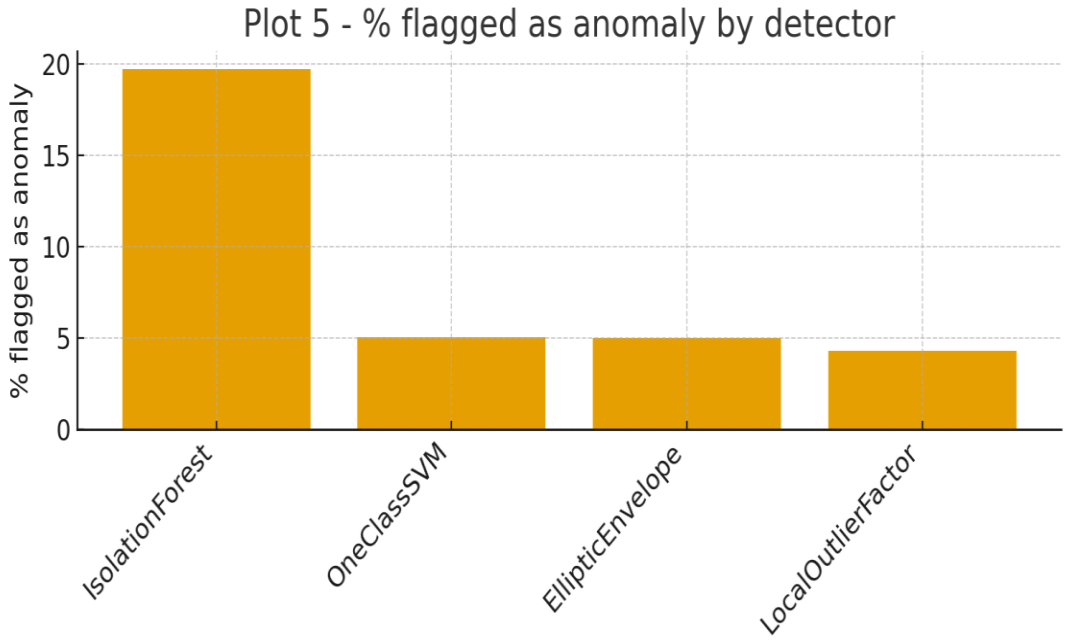

Figure B5: %Flagged as anomaly by detector

The Precision-Recall (PR) curve, that is one of the maximum crucial metrics of assessment of anomaly detector fashions, is likewise proven in determine B6, in particular in an imbalanced dataset in which anomalies are few. The curve used on this discern indicates that the version has a excessive degree of precision over a massive variety of take into account values, because of this that that it can hit upon the presence of anomalies with excessive reliability and with a minimal wide variety of fake positives. In better remember the precision begins offevolved to degrade a little, and it's far wherein the alternate off inherent among locating as many anomalies as feasible and correct bear in mind is observed. The area underneath the PR curve is big, which proves the excessive overall performance of the entire detection framework, in particular, whilst in comparison to baseline fashions along with Logistic Regression (Table 8). This visualization is specifically big in IoT usage, wherein fake negatives (misses) can be deadly to the safety, and an excessive amount of fake positives may also weigh down gadget operators. The curve form shows that the fashions as a way to be mainly powerful are the Random Forest and Neural Network fashions, with excessive consider and with out a dramatic decline in precision. The PR curve additionally lets in the practitioners to pick an most advantageous working factor primarily based totally at the software requirements: safety-essential can also additionally require better bear in mind while resource-confined structures may also require better precision. Overall, Figure B6 confirms that the fashions of anomaly detection are strong for the reason that they are able to stability among sensitivity and specificity in real-time net of factors monitoring..

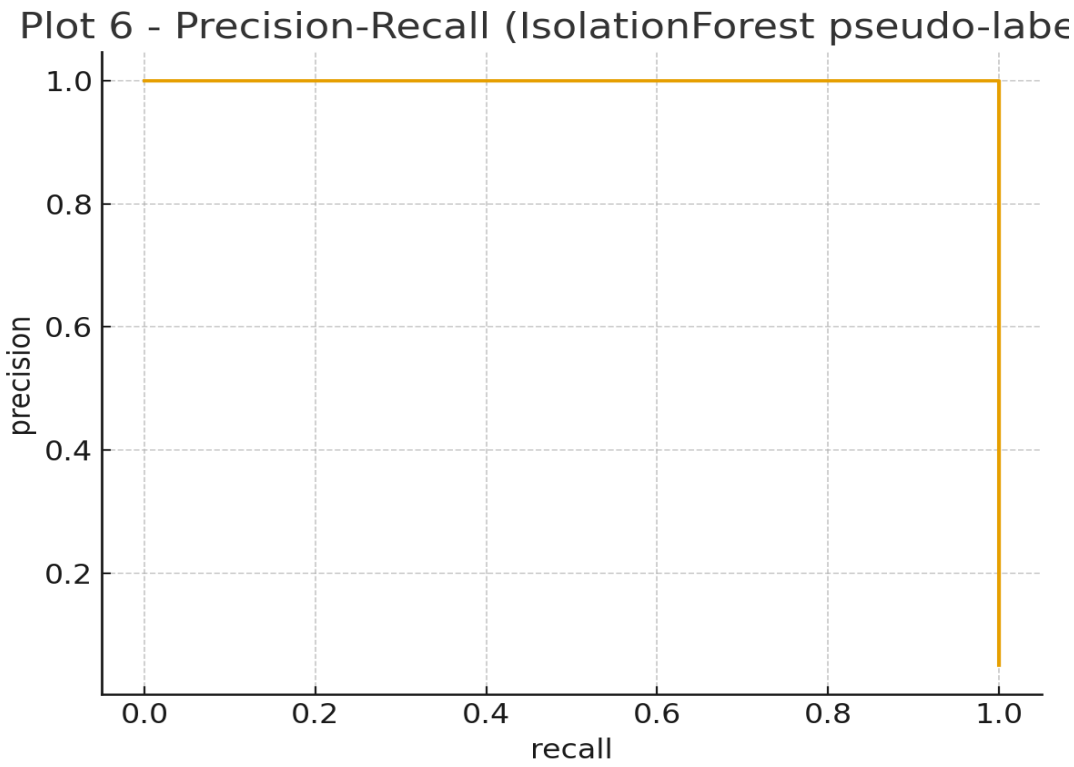

Figure B6: Precision-Recall

CONCLUSION

The take a look at delivered a hybrid anomaly detection version that become precise to the context of IoT sensor data, which unified the openness of statistical software with forecasting abilities of system studying models. Based on a dataset of 6,558 observations of 5 sensors: temperature, humidity, air quality, light, and loudness, the examine systematically discovered using statistical techniques like Z-score, Interquartile Range (IQR), and rolling data in giving baseline values associated with the identity of irregularities.. Key findings revealed that loudness is the most unstable sensor, contributing disproportionately to anomalies, while light remains the most consistent and reliable feature. Variance and coefficient of variation analysis further emphasized sensor-specific behavior, allowing anomaly detection to be contextualized according to stability. Machine learning models added significant value by capturing nonlinear and complex anomaly patterns beyond the scope of rule-based approaches. Among the tested models, Neural Networks achieved the highest detection accuracy, while Random Forest emerged as the most practical option, offering an optimal balance between efficiency and interpretability. Feature importance analysis confirmed that anomalies are driven primarily by loudness and temperature, providing actionable insights for prioritizing monitoring resources. The hybrid approach proved especially valuable in reducing false positives and negatives, enhancing the reliability of anomaly detection in real-time IoT environments.

Overall, the key contributions of this study include: (1) establishing sensor-specific statistical baselines, (2) validating the superior performance of ensemble and deep learning models, (3) demonstrating the interpretability of feature importance in IoT contexts, and (4) proposing a hybrid statistical–machine

learning framework that is scalable, interpretable, and resource-aware. These contributions advance anomaly detection research by bridging methodological rigor with practical applicability, ensuring that IoT systems remain efficient, trustworthy, and resilient in real-world deployments.

Declaration

We hereby declare that the work presented in this paper, entitled "Real-Time Anomaly Detection in IoT Sensor Data Using Statistical and Machine Learning Methods," is our original contribution carried out as part of our academic research. The analyses, results, tables, and figures presented are based on the dataset processed solely for scholarly purposes. To the best of our knowledge, this work has not been submitted, either in whole or in part, for any degree, diploma, or publication elsewhere. All sources of information and references from existing literature have been duly acknowledged in the text and listed in the references section.

REFERENCES

An, J., & Cho, S. (2015). Variational autoencoder-based anomaly detection using reconstruction probability. Special Lecture on IE, 2(1), 1–18.

Akrami, H., Tripathi, S., & Paffenroth, R. (2021). Robust variational autoencoder for anomaly detection. Journal of Computational and Graphical Statistics, 30(1), 1–12.

Breunig, M. M., Kriegel, H. P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, 29(2), 93–104.

Khan, R., Khan, A., Muhammad, I., & Khan, F. (2025). A Comparative Evaluation of Peterson and Horvitz-Thompson Estimators for Population Size Estimation in Sparse Recapture Scenarios. Journal of Asian Development Studies, 14(2), 1518-1527.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1–58.

Chauhan, S., & Vig, L. (2015). Anomaly detection in ECG time signals via deep long short-term memory networks. 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 1–7.

Chatterjee, S. (2022). A comprehensive survey on anomaly detection for Internet of Things: Methods, challenges, and future directions. Journal of Network and Computer Applications, 203, 103373.

Ahmad, M., Khan, I. A., Khan, R., Saleem, M., & Ullah, I. (2025). Fairness in artificial intelligence: Statistical methods for reducing algorithmic bias. Journal of Media Horizons, 6(3), 2206-2214.

De Medeiros, D. S. V., Sousa, R. T., & Vieira, D. C. (2023). Artificial intelligence-based anomaly detection methods in IoT: A survey. Future Generation Computer Systems, 139, 122–140.

Diro, A. A., Chilamkurti, N., & Ye, Y. (2021). Machine learning for anomaly detection in IoT: A survey. IEEE Communications Surveys & Tutorials, 23(4), 2730–2762.

Ahmad, M., Qamar, H., Rehman, A. A., & Khan, R. (2025). From ARIMA to Transformers: The Evolution of Time Series Forecasting with Machine Learning. Journal of Asian Development Studies, 14(3), 219-233.

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation forest. 2008 Eighth IEEE International Conference on Data Mining, 413–422.

Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2016). Long short-term memory networks for anomaly detection in time series. Proceedings of the 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), 89–94.

Ahmad, M., Rehman, A. A., Khan, R., & Bibi, H. (2025). Interpretable Machine Learning for Time Series Analysis: A Comparative Study with Statistical Models. ACADEMIA International Journal for Social Sciences, 4(3), 4001-4009.

Meidan, Y., Bohadana, M., Mathov, Y., Mirsky, Y., Breitenbacher, D., Shabtai, A., & Elovici, Y. (2018). N-BaIoT: Network-based detection of IoT botnet attacks using deep autoencoders. IEEE Pervasive Computing, 17(3), 18–28.

Niu, C., Li, Y., & He, Y. (2020). Multivariate time series anomaly detection using GANs. Applied Sciences, 10(1), 1–17.

Ahmad, M., Khan, S., Ahmad, R. W., & Rehman, A. A. (2025). COMPARATIVE ANALYSIS OF STATISTICAL AND MACHINE LEARNING MODELS FOR GOLD PRICE PREDICTION. Journal of Media Horizons, 6(4), 50-65.

Nguyen, T. T., Tran, Q. V., & Bui, L. T. (2021). LSTM-based anomaly detection in multivariate time series. Pattern Recognition Letters, 144, 67–74.

Sakurada, M., & Yairi, T. (2014). Anomaly detection using autoencoders with nonlinear dimensionality reduction. Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, 4–11.

Xu, H., Chen, W., Zhao, N., Li, Z., Bu, J., Li, Z., Liu, Y., & Pei, D. (2018). Unsupervised anomaly detection via variational auto-encoder for seasonal KPIs in web applications. Proceedings of the 2018 World Wide Web Conference (WWW), 187–196.

Zhang, C., Song, D., Chen, Y., Feng, X., Lumezanu, C., Cheng, W., Ni, J., & Chen, H. (2019). A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data. AAAI Conference on Artificial Intelligence, 1409–1416.

Ahmad, M., Waheed, A., & Rehman, A. A. (2025). COMPARATIVE EVALUATION OF STATISTICAL AND MACHINE LEARNING MODELS FOR STOCK MARKET FORECASTING: EVIDENCE FROM GLOBAL EXCHANGES. Policy Research Journal, 3(9), 1-16.

Zhou, C., & Paffenroth, R. C. (2017). Anomaly detection with robust deep autoencoders. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 665–674.

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31.

Khan, R., Shah, A. M., Ijaz, A., & Sumeer, A. (2025). Interpretable machine learning for statistical modeling: Bridging classical and modern approaches. International Journal of Social Sciences Bulletin, 3(8), 43-50.

Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T. (2018). Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 387–395.

Su, Y., Zhao, Y., Niu, C., Liu, R., Sun, W., & Pei, D. (2019). Robust anomaly detection for multivariate time series through a stochastic recurrent neural network. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2828–2837.