International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 10-No.4 (2024) pp. 898-910
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Overview of Anomaly Detection Techniques across Different Domains: A Systematic Review

Venkatraman Umbalacheri Ramasamy*

Distinguished Software Architect at Walmart Global Tech, Sunnyvale, CA, 94086, United States
* **Corresponding Author Email:** venkatrsrch@gmail.com - **ORCID:** 0009-0005-5451-962X

Article Info:

DOI: 10.22399/ijcesen.522
Received : 17 October 2024
Accepted : 29 October 2024

Keywords :

Anomaly Detection,
Machine Learning,
Deep Learning,
Artificial Intelligence,
Healthcare and Cyber security.

Abstract:

An anomaly, defined as something that deviates from what is normal, expected, or usual. It signifies abnormality or an irregularity that stands out from typical behaviours or patterns. Detecting anomalies is significant among numerous sectors due to the reasons of signal potential difficulties or opportunities. For an instance, in retail, detecting anomalies in sales data might prompt for further analysis into operational issues or customer behaviour to reduce losses and capitalize on its trends. Hence, different techniques are used for Anomaly Detection. However, anomaly detection using manual method are measured for time consuming, prone to error and can be tedious process. Therefore, different approaches have been considered for anomaly detection as AI (Artificial Intelligence) methods are efficient, faster, provides high level accuracy by effectively detecting the abnormalities. Owing to these aspects, this paper focuses on compiling different techniques and emphasizes on reviewing all anomaly detection using numerous techniques like ML (Machine Learning) and DL (Deep Learning) classifiers, statistical methods, one-class classification, clustering and density-based models which helps with identifying and comprehending the diversity of detection techniques that are applied in various domains like finance, retail, healthcare and cyber security. Various existing researches on anomaly detection are reviewed in the study. In addition to an overview, certain studies also deals with applications of detection models and future trends are reviewed in precise. Finally, the challenges are identified through the analysis of existing researchers and future recommendations are provided for overcoming the gaps that are intended to create promising work in this area.

1. Introduction

Main Security is considered as the significant need of present computer systems, and in recent time, it absorbs huge importance as the severity and number of malicious attacks growth [1]. One of the major branches of intrusion detection systems is anomaly detection that assess to identify the various variants of the attacks [2]. In today's technologies landscape, cyber-attacks are increasingly common and more sophisticated. Existing intrusion detection models are providing inadequate in addressing severe threats. This rise in malicious activity is driven by the growing demands on network systems, prompting attackers to intensify their efforts in these areas [3]. Several network or organizations based environments consume undertaking hundreds of attacks on network habitually. Recently, researches uses Artificial Intelligence (AI) based technology for anomaly classification. Advantages with Deep Learning (DL) and Machine Learning (ML) offers several applications in network security. Anomaly classification involves the use ML and DL methods to analyse network traffic then identify patterns that deviate from the norm. The support of ML and DL techniques, several methods are tried and tested to find the anomalies in network architecture. Correspondingly, weak authentication can allow the attackers to analyze and eavesdrop the traffic [4]. Despite of all functions and features, the network security is still considered to be a significant concern. The ongoing prevalence of cyber-attacks can lead to the theft of sensitive and valuable information, as well as issues such as ransomware, phishing targeting banks, insider threats, and hacking. This situation raises important questions and challenges about the partnerships between

Financial Technology (FinTech) companies and traditional banks. The economic impact of Cybersecurity risks on FinTechs is crucial for the sustainability of traditional banks [5]. Concurrently, there are specific concerns related to innovations in information technology and systematic operations.

1.1 Background

Anomaly classification is a process of identifying and categorizing unusual network activities. These anomalies can be indicative of possible security threats, like unauthorized access attempts, malware infections, or data breaches [6]. In the current technological atmosphere, cyber-attacks are increasingly prevalent and becoming more sophisticated, while existing intrusion detection models are proving inadequate in addressing severe threats. The increasing demands on network systems, are malicious attackers are growing their contribution in these areas. Many organizations and network-based atmospheres have engaged in hundreds of attacks on network regular basis. Anomaly classification is a critical aspect of network security which involves identifying and categorizing unusual network activities that deviate from normal behaviour [7]. Correspondingly, anomaly detection is also known as outlier detection which is key ML commission with several applications, including rare disease detection [8], social media analysis [9], and anti-money laundering [10] and intrusion detection [11]. Anomaly detection processes aimed to classify data examples that deviate considerably from the mainstream of data substances and various methods have been established in the last few decades.

Anomalies detection is vital in assisting decision-makers in concerning the root causes and creating efficient decisions based on real-time scenario of analysing data and information through providing insights into departures from certain activity [12]. Therefore, both strategic planning and operational effectiveness could benefited through this. Significantly, financial organizations utilise anomalies detection systems to monitor on transactions and consider for odd trends which seems to be fraudulent services [13]. Likely, banks could stop financial harms for both clients and themselves through reporting abnormalities. In scientific studies anomalies prediction could reveal the novel phenomena or patterns which demands precise examination and result in inventive ideas or findings. Likewise, in manufacturing, it could find drawbacks or irregularities in procedures or goods, thereby guaranteeing consistent quality and developed customer satisfaction [14].

1.2 Objectives of the review

- To summarize the numerous anomaly detection techniques used in different domains, including statistical methods, ML and DL approaches, and hybrid models.
- To categorize and analyse how different anomaly detection techniques are applied in specific fields such as finance, retail, healthcare, cyber security.
- To assess the performance metrics commonly utilized for evaluating anomaly detection techniques, including exactness, value of precision, Probability of detection, and F1 measures and how these metrics differ across various applications.

1.3 Paper Organisation

The paper organized in following progressive method. Section 1 illustrates a brief introduction regarding the anomalies detection and its significance. It also depicts the significance of the present research. Section 2 describes the methodology of prevailing scholarly research works related to the proposed research. Section 3 provides a detailed analysis of various detection techniques. Section 4 explains the application of anomalies detection in various domains. In following, section 5 presents the challenges and section 6 and 7 describes the future recommendations and conclusion of the present study.

2. Methodology

D. Cabrera and D. Antons have presented that the SR (systematic review) method which the current study applied[15,16]. The approach enables leading the reviews structurally as well as objectively. Moreover, methods permits representation of wider picture of the topic of anomaly detection techniques across different domains by classifying results in primary and related topic under consideration.

2.1 Research Question

The research questions of the current study are follows:

- What are the primary categories of anomaly detection techniques, also how will they differ across various domains?
- What are the emerging trends in anomaly detection methodologies, particularly with the advent of deep learning and Machine learning?
- What future directions can be anticipated for research in anomaly detection across different domains?

Figure. 1 *Review Research Questions*

The above research questions are responded through exploring the anomalies detection, then, examining the conventional tools and techniques which are utilised. Certainly, difficulties and issues would be underlined and recognised. Hence, the key-words foremost the present exploration in eminent records have computed. Such keywords are Anomaly Detection - Machine Learning - Deep Learning - Artificial Intelligence – Finance - Healthcare - Cyber security is illustrated in Figure 1.

2.2 Sources selection

The determination of this review paper to study anomaly detection in various domains in the last five years, work completed between 2020 and 2024 is included in the exclusion criteria based on the year of publication. The language of publication served as an additional exclusion criterion. Non-English publications were filtered out of the exploration results through setting search engine for excluding them before showing the outcomes from digital libraries. Lastly, the presence criteria is based on a multifaceted examination of the results obtained. These elements consist of the abstract, conclusion, title, and keywords. Thus, the act of manually reviewing and selecting which publications to consider important or not is called analysis. Table. 1 illustrates the searching results. Table. 1 depicts the searching results. The present review uses various articles for acquiring optimal results. IEEE, MDPI, Springer, Taylor & Francis, Elsevier, Sage, Wiley and Others have 14, 10, 7, 1, 5, 2, 2 and 7 of publications respectively. Therefore, the total number of publications used in this work is around 48. Furthermore, figure. 2 signifies graphical representation of the publications used. From figure 3, it clears that the present review used IEEE more when comparing to other publications. However, Taylor & Francis have 1 publication in this review. Similarly, Table.2 describes year based counts for the publications.

Table 1. *Number of Publications*

Article Name	Number of Publication
IEEE	14
MDPI	10
Springer	7
Taylor & Francis	1
Elsevier	5
Sage	2
Wiley	2
Others	7
Total	**48**

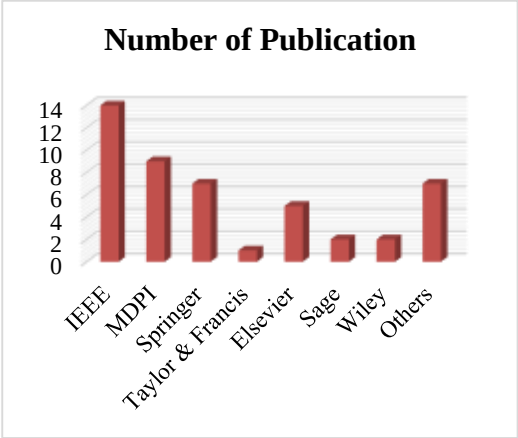

Figure. 2 *Usage of Publications*

Table 2*. Number of References per Year*

Year	References
2020	11
2021	11
2022	9
2023	11
2024	6

From Table 2, it clears that the present review used last five years of publications for effective review on anomalies detection techniques. The year 2020, 2021, 2022, 2023 and 2024 has 11, 11, 9, 11 and 6 of publications respectively.

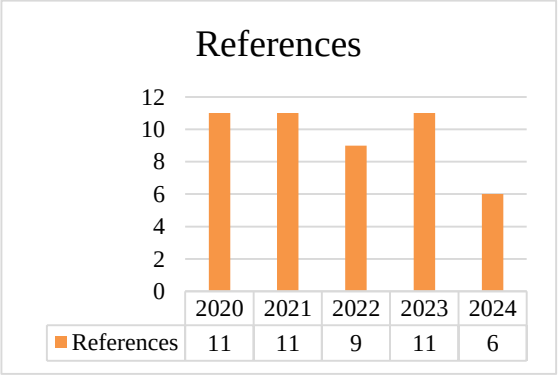

Figure. 3 *Graphical Representation of Used references*

3. Literature Search

This section provides about the analysis of various existing researches on Anomaly detection along with other techniques for the prediction on intrusion classification system.

Galeano-Brajones has evaluated experimentally on entropy-based resolution for mitigate and detect DDoS and DoS attacks in the scenarios of IoT utilising a SDN data plane [17]. It developed the application called proof-of-concept at Ryu SDN controller top which has been detected the DDoS and DoS attacks in terms of entropy values. To analyze the effects of metrics in various conditions, evaluated performance of the application in three state, the first state is general test bed where the bandwidth, entropy values are measured during attacks. Whereas the second and third state focus on IoT scenario. The results have occurred in three scenarios and attained better accuracy. Likewise, Polat [18] has detected DDoS attacks in SDN specific features were obtained for dataset in normal conition and under DDoS attack traffic, therefore, by using ML models and has generated a new dataset from feature selection methods, the data were trained and verified with Support Vector Machine (SVM), Naive Bayes (NB) ANN and KNN classification model of conventional dataset. Different number of features can change the accuracy. The test results have shown that utilized and wrapped feature selection with the KNN classifier has attained 98.3% accuracy in attack detection.

Correspondingly, Novaes [19] has deployed GAN (Generative Adversarial Network) outline better-quality the presentation metres evaluated to detect novel attacks, also functional confrontational training to make the system less sensitive to adversarial attack. Used a DL approach framework for the detection of DDoS attacks. It utilised public CICDDoS 2019 dataset which contains the most 12 up-to-date common types of DDoS attacks such as NTP, MSSQL, LDAP, DNS, SSDP,630 SNMP, NetBIOS, UDP, UDP-Lag, SYN, WebDDoS (ARME), and TFTP. The results have shown that the LSTM, GAN and MLP have attained 90.29%, 94% and 92.12% accuracy respectively. Contrastingly, Gadze [20] has focused on possibility and efficiency of DDoS attack detection and mitigation. The existing method has investigated the DL simulations such as, LSTM and Convolutional Neural Network (CNN), showcasing the method by converting raw networking traffic into image data for enhanced classification accuracy. It has generated an own dataset for testing and training, which significantly contributes to the robustness of the detection mechanism. The results have shown that the RNN-LSTM produced 89.63%, Naïve Bayes produced 82.61% and SVM produced 86.85% of accuracy.

In the similar vein, Al-Dunainawi [21] has employed the hybrid CNN-LSTM model to detect for distinguishing the DDoS attack in SDN based networks, leveraging a custom dataset tailor for this specific application. The integration of both convolutional and recurrent neural network architectures, allowing for enhanced features extraction and temporal analysis of network traffic. The innovative combination significantly improves detection capabilities, leading to an impressive accuracy in identifying DDoS attacks, thereby demonstrating the models strength. Correspondingly, Kyaw [22] has effectively detected the DDoS attack and classified the normal or traffic in SDN environment using polynomial SVM method which offers a comparison against linear SVMs. The prevailing model has applied polynomial SVM to compare the linear SVM through using scapy (a packet generation tool and RYU controller). Hence, the dataset has collected by creating volume-based normal traffic and DDOS attack traffic. According to the results, the polynomial SVM has attained 3% accuracy more and 34% low false alarm rate than linear SVM compared to its linear counterpart, demonstrating significant improvements in detection performance.

Similarly, Wong [23] has presented a mechanism of anomaly detection utilising CNN and LSTM models, marking the significant advancement in field. This approach combines DL approach to effectively analyz network traffic patterns, enabling more accurate detection of anomalies. These models have been trained on network data that are extracted from the packet capture files. It has utilised a huge scale real network traffic dataset and benchmark intrusion detection dataset which provide robust foundation in training. The prevailing model has attained 97.2%, 88.9% and 92.3% of accuracy on CTU-13, ISCX-IDS and NSL-KDD datasets respectively. In parallel, Nadeem [24] has evaluated several significant feature selection method for ML on DDoS detection, introducing an approach by employing RFE (Recursive Feature Elimination) to optimize feature subset. It has used NSL-KDD dataset which contains 41 features and 52,800 records, the dataset is well regarded for its comprehensive representation of normal and attack traffic, providing a solid foundation for model evaluation. The results have shown that the RF classifier has

attained better accuracy using the features subset through the RFE method. The study also compared the performance of other ML algorithm, such as SVM and KNN, showcasing the superiority of RF in term of exactness ad computational efficiency.

Contrastingly, conventional model from Mukherjee who has predicted various anomalies on combination of various features in the dataset through applying ML models [25]. It has utilised a comprehensive dataset which is from Kaggle, consisting that 357,952 samples along with 13 features, the dataset includes a diverse range of IoT device traffic, encompassing both normal and anomalous behavior, enabling the model to learn complex patterns effectively. The advanced features selection technique to identify the most informative attributes, optimizing the performance of ML algorithm. By leveraging ensemble methods and DNN, the model has attain the state of art results in accurately detecting various type of anomalies. The approach lies in its ability to adapt to the dynamic nature of IoT networks, ensuring reliable anomaly detection even in the face of evolving treats. It has been attained better results. Similarly, Fotiadou [26] has presented a variant DL architectures to the issue of anomaly prediction on network logs which are obtained by pf Sense firewall along with chief target the classification of event type in the prevailing system. It has utilised network intrusion dataset. The results of the existing model have attained 96.34%, 87.14% and 83.13% of accuracy in LSTM, MLP model and RF respectively.

Hwang [27] has presented an effective anomaly detection mechanism which has been named as D-PACK. It comprises of unsupervised DL model like Auto encoder and CNN for auto profiling traffic patterns and extracting the abnormal traffic. It has used Mirai-CCU dataset, which provided a comprehensive representation of both normal and anomaly network traffic pattern, ensuring a robust training environment. The data include wide range of IoT device traffic, encompassing various types of attacks. By leveraging the strength of both auto encoder and CNN, D-PACK attains state of art and has attained satisfactory results. The approach lies in its ability to adapt to the dynamic nature of IoT networks, ensuring reliable anomaly detection even in the face of evolving threats. Correspondingly, a 5-layer auto encoder based model for a network anomaly classification tasks, showcasing an initial architecture that enhances the traditional auto encounter design. This model has evaluated on NSL-KDD dataset which is well-known for its comprehensive representation of network traffic, containing both normal and attack instances. The dataset comprises 41 features and 52,800 records, providing a robust training environment for the model. By implementing advanced techniques such as feature normalization and dropout layers, the model effectively reduces over fitting and improves generalization It has attained 90.61% of accuracy in detection performance [21].

Congruently, CSE-CIC-IDS2018 data-set has been employed to evaluate the traditional model, reflecting contemporary network traffic conditions. It has applied RNN, CNN, DNN, CNN+RNN, LSTM and CNN+LSTM models have been constructed to detect the anomaly attacks in network to enhance anomaly detection capabilities. By systematically tuning hyper parameters and applying rigorous data pre-processing techniques, the models achieved an impressive accuracy rate exceeding 98%, significantly outperforming existing intrusion detection systems. The use of this comprehensive dataset allows the model to efficiently identify a diverse variety of anomaly attacks, demonstrating its robustness in real-world applications. This innovative approach [29] not only improves detection performance but also sets a benchmark for future research in network intrusion detection systems. Figure. 4 illustrates that basic mechanism of anomaly detection using DL models.

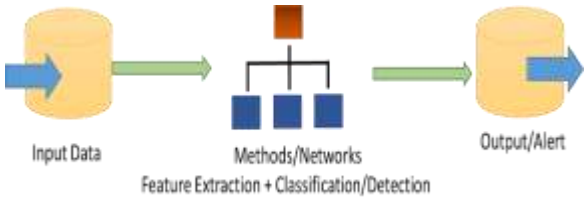

Figure. 4 *Basic Process for Anomaly Detection in DL[30]*

Figure. 4 demonstrates the fundamental workings of a DL-based algorithm, where data from IoT environments are the input and a binary alert system that tells the user if the data are abnormal or typical is the output.

3.1 Anomaly Detection Techniques

Korea O [30] has demonstrated that anomaly detection techniques are indispensable for recognising unusual outliers or patterns in data across various fields. It includes statistical methods, ML and DL approaches, Ensemble methods, Clustering techniques, Density-based methods and One-class classification. The selection of a technique is contingent upon various aspects, with but not limited to the type of information, that accessibility of labelled data, the wanted degree of

interpretable, and computing limits. Each technique possesses pros and downsides. In real-world applications, hybrid systems that integrate several techniques or adaptively choose the best methodology depending on data properties may provide better detection robustness and performance. In following, figure 5 deliberates the techniques of anomaly detection.

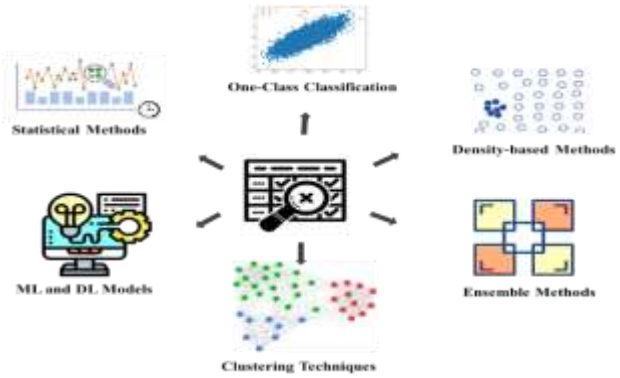

Figure. 5 *Various Techniques for Anomaly Detection*

- **Statistical Methods**: To find anomalous data points, statistical approaches for anomaly identification apply mathematical techniques. Z-scores are frequently used to quantify deviations from the mean and to identify outliers that fall outside of the normal range. The Interquartile Range (IQR) is another common method. These techniques are effective for preliminary anomaly detection attempts since they are simple and don't require any training data. These methods play a dynamic role in anomaly detection. It often involved in comparing experiential data against expected patterns or distributions to detect anomalies. Abisoye [31] has presented the process for detecting algorithm in MTS (Multivariate Time Series) data. It is based on fusion of Z-score, SED (Standard Euclidean Distance) and K-medoid. The experiments are made on 2 other distance metrics like Euclidean distance and City Block. It has attained 0.9978 of F1 measure and 0.8571 F1 measure on inter-quartile.

- **ML and DL approaches:** Algorithms are used in machine learning (ML) and deep learning (DL) techniques for anomaly detection in order to find odd patterns in data. Unsupervised learning is used by machine learning techniques like Isolation Forest and Local Outlier Factor to find anomalies in the absence of labeled data. On the other hand, deep learning methods, like autoencoders, need large datasets for training in order to recognize minor anomalies and learn intricate patterns. Ullah [32] has developed anomaly on the basis of IDS (intrusion prediction system) for the IoT networks. Later, it has implemented utilising CNN in 1D, 2D and 3 D. It has validated on four datasets namely IoT Network Intrusion, BoT-IoT, IoT-23 ID (intrusion detection) and MQTT-IoT-IDS2020 dataset which have been attained better values in terms of accuracy. In the similar vein, Alqahtani [33] has employed different ML classification algorithms such as NB (Naïve Bayes), DT Decision Tree (),RF (Random Forest), DTb (Decision Table), RDT (Random Decision Forest), ANN (Artificial Neural Network), to predict intrusions due to offer services in cyber-security domain. It has used KDD'99 cup data which comprises of 4898431 instances with 41Attributes. 93%, 94%, 90%, 92%, 91%, 91% and 90% of accuracies have attained by DTb, RF, RT, DT, NB, ANN and BN respectively.

- **Ensemble Methods:** Several algorithms are combined in ensemble approaches for anomaly detection to increase the resilience and accuracy of the detection. Through the use of different base detectors, these techniques overcome the drawbacks of single-method approaches and improve generalization over a range of datasets. Methods such as clustering-based picks and weighted averages balance accuracy and variety to maximize performance. Doreswamy [34] has presented an ensemble method utilising quorum and union combination techniques with 5 individual primary choosing techniques that are analysis of recursive feature elimination, sequential backward search, variance, variance threshold and shrinkage operator and least absolute selection. Experimental results have shown on UNSW-NB15 dataset. Minimum 12 and 6 features, RF with quorum and union feature sets have yielded 99.02% and 99% of F1 score respectively.

- **Clustering Techniques:** Similar data points are grouped together using clustering algorithms for anomaly identification, while outliers are defined as those that do not fit into any cluster. Well-liked techniques include K-means, which divides data into predetermined clusters, and DBSCAN, which handles noise and finds clusters of different sizes and shapes. These techniques analyze distances from cluster centroids or core locations, which successfully reveals anomalies**.** The clustering algorithms are aimed to separate the provided

unlabelled data into several clusters which could attain high outer dissimilarity and inner similarity, without relied on signs, labelled data for model training and explicit detection technique. G. Pu, L. Wang and J. Shen [35] have developed an unsupervised method for anomaly detection that integrates OCSVM (One Class Support Vector Machine) and SSC (Sub-Space Clustering) to predict attacks without any previous knowledge. It has evaluated on NSL-KDD dataset. It attained 238.88 s, 8.15 s, 0.69s and 1060.76 s of computation time on SSC-OCSVM, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), K-means and SSC-EA respectively.

- **Density-Based Methods:** Anomaly detection techniques based on density distinguish normal data points from those in low-density areas and mark them as anomalies. Methods such as DBSCAN and Local Outlier Factor (LOF) may identify clusters of any shape and handle noise well. These techniques are adaptable to a range of datasets because they don't require prior knowledge of the number of clusters. The conventional paper has presented a density-based unsupervised method to detect the anomalies in heart patients [36]. In this conventional method, basic features of data from the dataset is selected first on method of filter-based feature selection. DBSCAN clustering technique with adaptive parameters has been utilised to develop the clustering accuracy in healthy cases and to demonstrate anomaly instances in heart patients. It has used Heart Disease Prediction dataset. Experimental results have shown that the prevailing model has achieved 95% of accuracy.

- **One-Class Classification (OCC):** It is a specialized machine learning approach focused on identifying anomalies or outliers by training solely on "normal" data. Unlike traditional classification, which requires multiple classes, OCC models the characteristics of a single class to distinguish it from all others. This technique is particularly useful in scenarios where negative class examples are scarce or absent, such as fraud detection or fault monitoring. OCC methods typically involve estimating the probability density of the training set and flagging new instances that fall below a certain threshold as anomalies. Common algorithms include One-Class Support Vector Machines and Isolation Forests, which effectively capture the distribution of normal instances to identify deviations. OCC is a powerful method for anomaly detection that targets to identify annotations deviate considerably from the distribution in normal data. Xu and Hongzuo [37] have employed calibrated OCC for detecting anomalies, understanding the data normality and contamination tolerant through native anomaly based calibration and uncertainty modelling based calibration. The conventional paper has developed Calibrated One-class classification-based Unsupervised Time series Anomaly detection method (COUTA) which has been evaluated on ten datasets. Among 10 datasets, 6 datasets are multivariate and 4 are univariate datasets. It has achieved better performance in anomaly detection. Moreover, Table. 3 describes about the comparison analysis of various conventional researches based on anomalies detection techniques with its method and results.

4. Application Domains

This section describes about the anomalies detection models in various domains like healthcare, finance, retail and cyber security. The table 4 signifies the domains and used techniques for detecting anomalies.

4.1 Finance

In finance industry, anomaly detection is utilized to detecting fraud, risk management and AML (Anti money laundering) compliance. Such algorithms analyze account activities, financial transactions and spending patterns to recognize the suspicious behavior like identity theft, unauthenticated transactions and fraudulent account practices. Similarly, Suseendran [38] has analyzed the overall operations of financial technologies, challenges of FinTechs and Banking industries, and cyber security when associating IoT with the internet. Additionally, the usage of IoT on FinTechs in the digital environment has been mentioned precisely. Contrarily, Al-Alawi [44] has demonstrated the advantages of applying cyber security and its significant effects on banking institutions. It has used an online questionnaire method to identify the risk types in Bahrain. The existing study has found that banks have been tackling cyber-attacks frequently, where 26% of financial industries have encountered online theft cases, and 23% have experienced damage to computer systems. The remaining 11% have faced hacking attempts. Correspondingly, the conventional study has examined how FinTechs are influenced by changes in banking and challenges, along with a specific

***Table 3**. Comparative Analysis*

Reference	Objective	Method	Dataset	Result
[23]	Mechanism of anomaly detection	DL method (CNN and LSTM)	Real network traffic dataset and benchmark intrusion detection dataset	97.2%, 88.9% and 92.3% of accuracy on CTU-13, ISCX-IDS and NSL-KDD datasets respectively.
[32]	Anomaly based IDS for IoT networks	DL (CNN in 1D, 2D and 3D)	IoT-23 intrusion detection, BoT-IoT and MQTT-IoT-IDS2020 dataset	Better Values
[34]	Mechanism of anomaly detection	Ensemble method (quorum and union combination techniques)	UNSW-NB15 dataset	99.02% and 99% of F1 score
[33]	To predict intrusions due to offer services in cyber-security domain	ML (NB, DT, RF, DTb, RDT, ANN)	KDD'99 cup data with 4898431 instances with 41 Attributes.	93%, 94%, 90%, 92%, 91% and 91% of accuracies have attained by DT, RF, RT, DTb, ANN and NB respectively.
[25]	To predict various anomalies	ML method	A comprehensive dataset which is from Kaggle, consisting that 357,952 samples along with 13 features.	Better Results
[29]	To detect the anomaly attacks in network	DL (RNN, CNN, LSTM, DNN, RNN + CNN and LSTM + CNN models)	CSE-CIC-IDS2018 dataset	An impressive accuracy rate exceeding 98%.
[35]	Mechanism of anomaly detection	Clustering Method (unsupervised method SSC-OCSVM)	NSL-KDD dataset	Computation time of SSC-OCSVM, SSC-EA, K-means, and DBSCAN are 238.88 s, 1060.76s, 8.15s, and 0.69s respectively.
[36]	To detect the anomalies in heart patients	Density-based unsupervised method (DBSCAN)	Heart Disease Prediction dataset	Achieved 95% of accuracy

***Table 4** Domains and utilised Techniques*

References	Techniques	Domains
[38]	Embracing IoT	Financial technology
[39]	Based on CNN	Retail
[40]	Lightweight auto encoder on MIoT (Medical Internet of Things).	Healthcare
[41]	MI (Mutual Information) and a DNN	Cyber Security
[42]	Co-AD (Concept based Anomaly Detection) model	Retail
[37]	Calibrated One-Class Classification (OCC)	Data science
[36]	Density-based unsupervised method	Healthcare
[43]	Combination of ensemble methods	Cyber Security
[31]	Statistical Methods	Data Science, ML, and Statistics

emphasis on blockchain technology [45]. It has performed a thematic analysis of FinTech banking industry. The prevailing study has found that FinTechs have the enormous potential to grow and influence banking institutions. In addition, it has revealed that blockchain applications' potential is not limited to FinTechs and payment transactions. However, there were growing interest in blockchain technology.

4.1 Retail

In the retail industry, anomaly detection is a vital tool that helps companies find odd patterns or behaviors in data. This can improve operational effectiveness, improve consumer satisfaction, and boost income. By using a variety of methods for data monitoring and analysis, this technology enables merchants to react quickly to unforeseen shifts in customer behaviour or operational metrics. Z. Da and Y. Dun [39] have presented a retail anomaly detection technique based on CNN known as the framework of complexity-classification anomaly detection (ClassAD). It embraces 2 modules namely classification module and complexity rate module. These modules combine input images of various feature complexity into various capacity networks that made omplete usage of feature extraction capability in CNNs. The conventional ClassAD has included with decomposition feature development technique to improve the inference speed. It has established UVM (Unmanned Vending Machines)-anomaly dataset for an evaluation which has achieved 96.28% of accuracy. Similarly, Kapoor [42] has developed a Co-AD (Concept based Anomaly Detection) model utilising ViT (Vision Transformer) which is able to ensign the misplaced ones without utilising prior idea bases like planogram. It has used an auto encoder model and outlier detection in latent space. Co-AD has achieved 89.90% of accuracy in images of retail objects from RP2K dataset. To state its utility, it has described the robotic mobile manipulation to correct the anomalies automatically by Co-AD. It has aimed in developing autonomous mobile robot solutions which decrease the requirements for human interference in the management of trade stores.

4.2 Healthcare

In healthcare, anomaly detection is utilized for sensor data and medical imaging anomaly detection, as well as for illness diagnosis and patient monitoring. Anomaly detection systems can recognise unusual condition of health, medical occurrences, or errors in diagnostic by examining sensor readings, patient health records, and medical imaging. This allows for early intervention and better patient outcomes. N. Shvetsova [46] has established a technique for image anomaly detection which relied on prevailing auto encoder method with re-designed training mechanism to manage high-resolution and complex images. It evaluated by 2 medical datasets comprising digital pathology and radiology images. The existing model has suggested a baseline for medical images anomaly detection tasks. Similarly, Abououf [40] has proposed online EAD (Event and Anomaly Detection) model utilising lightweight auto encoder on MIoT (Medical Internet of Things). The predicted abnormality has explained utilising XAI (explainable AI), KernelSHAP method, where anomaly explanation is utilised through ANN and to classify into abnormal or an event. It has conducted an intensive simulations which used the dataset of Medical Information Mart for Intensive Care (MIMIC) for numerous physiological data. Experimental results have shown that robustness in classification and detection of events.

4.3 Cyber security

In cyber security, anomaly detection is frequently used to find and eliminate a range of threats, including as malware infections, insider assaults, and network breaches. Anomaly detection systems can identify anomalous activity suggestive of malicious activity and initiate prompt remedial actions to reduce security risks by examining network traffic, system records, and user behavior. Cyber-security risk refers to the likelihood of exposure resulting from a data breach caused by cybercriminals or malicious insiders. From an industrial perspective, the cyber security risk is potential harm or loss caused by cyber-attacks related to a firm's critical structure. There are specific concerns in innovations and information technology and systematic operations. In the same way, Yaseen [43] has developed a theoretical model for network anomalies prediction to improve cyber security thereby, emphasizing the combination of ensemble methods, advanced feature engineering and ML models. Particularly, ensemble methods of RF, improve the robustness of the conventional framework through amalgamating strengths from vast models mitigating entire false negatives and false positives. The traditional model has employed an actual detection anomaly system utilizing MI (Mutual Information) and considered a DNN for IoT network. It has used IoT-Botnet 2020 dataset for an evaluation which has attained better accuracy [41].

5. Challenges in Anomaly Detection

Several existing researches have been limited by predicting the anomaly in the networks. It has several lacks and it is provided below.

- **Distribution of Imbalanced Data:** The datasets of Anomaly detection often exhibits distribution of imbalanced class, with usual instances expressively outnumbering the irregular instances. Imbalanced data distributions could lead to biased model training and decreased detection task, requiring techniques like cost sensitive learning, data resampling, or collective methods to state the class imbalance effectually [12].
- **Subtle Anomalies Detection:** Identifying subtle anomalies which deviate somewhat from regular activity amidst normal or noise variations in data values is a crucial challenge [47]. The subtle anomalies might exhibit distinct patterns, building them difficult to distinguish from regular activity utilising conventional detection models.
- **Mitigation of False Negatives and False Positives**: Reconciliation in trade-off among false negatives (deteriorating to predict actual anomalies) and false positives (incorrectly weakening normal cases as anomalies) is critical to ensure the effectiveness and reliability of systems on anomaly detection [4]. Reducing false alarms however, exploiting sensitivity of detection of anomaly needs watchful alteration of model parameters and edges [48].

6. Future Directions and Trends

The detection of anomaly, future directions are formed through technology developments, data landscapes evolving, and evolving challenges.

- **Advancements in ML:** Upcoming works in detecting anomalies would likely concentrated on advanced ml models to enhance accuracy, robustness and scalability. Methods like DL, reinforcement learning, meta-learning would continue to be discovered to consider complex patterns, evolving anomalies and high-dimensional data in dynamic areas.
- **Domain Knowledge Incorporation**: Integrating contextual information and knowledge in domain that anomaly prediction systems will change as necessary for enhancing in detecting exactness and decrease false alarms. Methods like knowledge graphs, ontologies and hybrid models will assess the combination of constraints, semantics and field of specific rules into anomaly detection mechanism.
- **Self-Learning and Adaptive Systems:** in future anomaly detection systems will become more adaptive and self-learning, allowing them to autonomously adjust to shifting data distributions, evolving anomalies and dynamic environments. The method like online learning, concepts drift detection, and self-tuning algorithms will empower these models to continuously learn and enhance their performance over time.
- **Multi-Modal Detection**: Multi-Modal Detection with propagation of heterogeneous modalities and sources of data, upcoming detection systems will required to manage vast data types, including unstructured and structured data, images, time-series and text data. The techniques like transfer learning, cross-modal learning, multimodal fusion, be able to comprehensive valuation among multiple modalities and data sources.
- **Interpretable and Explainable Models:** As anomaly detection models are becoming increasingly black-box and complex, it will be a rising requirement for interpretable and explainable methods which could offer understandings into model decisions and basic patterns of data. Methods like model agnostic details, rule extraction and feature significance analysis would enhance the trustworthiness and interpretability anomaly detection methods.
- **Anomaly Detection for Privacy-Preserving:** With rising concerns about security and data privacy, future detection systems would require to add privacy preserving methods for considering integrity and confidentiality of sensitive information. Methods like differential privacy, federated learning and secure computation would enable the detection models to work on anonymised or encrypted data without co-operating in discretion.
- **Edge Computing and Real-Time:** The rapidity and volume of information last to upsurge, future systems needs for work in real-time and network edge. Methods like edge computing, distributed processing and streaming analytics will enable such models to work proficiently in handling high data throughput and resource constrained areas.

All things considered, the requirement for further precise, privacy-preserving anomaly detection systems, adaptable, and scalable that can manage a variety of data sources, changing

settings, and new threats will shape the field's future directions. Anomaly detection systems may maintain their critical role in boosting security, increasing working effectiveness, and facilitating data driven decision-making through a range of disciplines by embracing these future directions.

7. Conclusion

In this paper, a systematic review of the topic anomaly detection techniques across different domains have been conducted. Detection of anomalies using manual techniques are considered time-consuming, slow, and prone to errors due to human intervention. Therefore, different techniques are used for Anomaly detection. Moreover, the paper focused on reviewing techniques on various domains, as it possess various beneficial properties. However, from the papers that have been reviewed, it can be examined that very few studies have focused on clustering based methods. However, most of the papers have focused on ML and DL based methods for detecting anomalies in various sectors like retail, health and cyber security. Apart from detection techniques, the paper has also focused on literature search on anomalies detection researches in various needs. Therefore, from the papers that have been reviewed, challenges like Distribution of Imbalanced Data, subtle anomalies detection and mitigation of false negatives and false positives are identified. The paper discussed about various applications of domains which predominantly includes medical, finance and retail sectors and analysis of different studies demonstrated that DL models are considered to be used widely for Anomaly detection. The research questions have been answered through this review, it may start as an initiative point for researchers, researching about anomalies detection in more advanced ways and the advancement of the future research work, can assist organization or an individuals to monitor their domain sector, detect the abnormalities affected by the attacks or malicious activities and aids in optimizing detecting practices and may also assist in various sectors for identifying the irregularities.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1]Riggs, H., Tufail, S., Parvez, I., Tariq, M., Khan, M. A., Amir, A., . . . Sarwat, A. I. J. S. (2023). Impact, vulnerabilities, and mitigation strategies for cyber-secure critical infrastructure. *Sensors (Basel)*. 17;23(8):4060. doi: 10.3390/s23084060.

[2]Hosseinzadeh, M., Rahmani, A. M., Vo, B., Bidaki, M., Masdari, M., & Zangakani, M. J. S. C. (2021). Improving security using SVM-based anomaly detection: issues and challenges. *Soft Computing* 25(4), 3195-3223. DOI:10.1007/s00500-020-05373-x

[3]Duo, W., Zhou, M., & Abusorrah, A. J. I. C. J. o. A. S. (2022). A survey of cyber attacks on cyber physical systems: *Recent advances and challenges.* 9(5), 784-800.

[4]Hussein, A., Chadad, L., Adalian, N., Chehab, A., Elhajj, I. H., & Kayssi, A. J. J. o. C. S. T. (2020). Software-Defined Networking (SDN): the security review. *Journal of Cyber Security* 4(1), 1-66.

[5]Khan, M. A., & Malaika, M. (2021). *Central Bank Risk Management, Fintech, and Cybersecurity*. International Monetary Fund.

[6]Mothukuri, V., Khare, P., Parizi, R. M., Pouriyeh, S., Dehghantanha, A., & Srivastava, G. J. I. I. o. T. J. (2021). Federated-learning-based anomaly detection for IoT security attacks. *IEEE Internet of Things Journal* 9(4), 2545-2554. DOI: 10.1109/JIOT.2021.3077803

[7]Inuwa, M. M., & Das, R. J. I. o. T. (2024). A comparative analysis of various machine learning methods for anomaly detection in cyber attacks on IoT networks. *Internet of Things* 26, 101162. DOI: 10.1016/j.iot.2024.101162

[8]Zhao, H., Li, Y., He, N., Ma, K., Fang, L., Li, H., & Zheng, Y. J. I. T. o. M. I. (2021). Anomaly detection for medical images using self-supervised and translation-consistent features. *IEEE Transactions on Medical Imaging* 40 (12), 3641-3651

[9]Rahman, M.S., Halder, S., Uddin, M.A. et al. (2021). An efficient hybrid system for anomaly detection in social networks. *Cybersecur* 4, 10. https://doi.org/10.1186/s42400-021-00074-w

[10]Dumitrescu, B., Băltoiu, A., & Budulan, Ș. J. I. A. (2022). Anomaly detection in graphs of bank transactions for anti money laundering applications. *IEEE Access* 10, 47699-47714. DOI: 10.1109/ACCESS.2022.3170467

[11]Kale, R., Lu, Z., Fok, K. W., & Thing, V. L. J. a. e.-p. (2022). A Hybrid Deep Learning Anomaly Detection Framework for Intrusion Detection. arXiv: 2212.00966.

[12]Bammidi, T. R., Gutta, L. M., Kotagiri, A., Samayamantri, L. S., & krishna Vaddy, R. J. I. J. o. M. E. f. S. D. (2024). The Crucial Role of Data Quality in Automated Decision-Making Systems. *International Journal of Management Education for Sustainable Development* 7(7), 1-22.

[13]Bakumenko, A., & Elragal, A. J. S. (2022). Detecting anomalies in financial data using machine learning algorithms. *Systems* 10(5), 130. DOI: 10.3390/systems10050130

[14]Wang, Y., Perry, M., Whitlock, D., & Sutherland, J. W. J. J. o. M. S. (2022). Detecting anomalies in time series data from a manufacturing system using recurrent neural networks. *Journal of Manufacturing Systems* 62, 823-834. DOI: 10.1016/j.jmsy.2020.12.007

[15]Cabrera, D., & Cabrera, L. L. J. J. o. S. T. P. (2023). The Steps to Doing a Systems Literature Review (SLR).

[16]Antons, D., Breidbach, C. F., Joshi, A. M., & Salge, T. O. J. O. R. M. (2023). Computational literature reviews: Method, algorithms, and roadmap. *Organizational Research Methods* 26(1), 107-138 DOI:10.1177/1094428121991230.

[17]Galeano-Brajones, J., Carmona-Murillo, J., Valenzuela-Valdés, J. F., & Luna-Valero, F. J. S. (2020). Detection and mitigation of DoS and DDoS attacks in IoT-based stateful SDN: An experimental approach. *Sensors (Basel)* 20(3), 816. doi: 10.3390/s20030816.

[18]Polat, H., Polat, O., & Cetin, A. J. S. (2020). Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models. *Sustainability* 12(3), 1035. DOI:10.3390/su12031035

[19]Novaes, M. P., Carvalho, L. F., Lloret, J., & Proença Jr, M. L. J. F. G. C. S. (2021). Adversarial Deep Learning approach detection and defense against DDoS attacks in SDN environments. *Future Generation Computer Systems* 125, 156-167. DOI: 10.1016/j.future.2021.06.047

[20]Gadze, J. D., Bamfo-Asante, A. A., Agyemang, J. O., Nunoo-Mensah, H., & Opare, K. A.-B. J. T. (2021). An investigation into the application of deep learning in the detection and mitigation of DDOS attack on SDN controllers. *Technologies* 9(1), 14. DOI:10.3390/TECHNOLOGIES9010014

[21]Al-Dunainawi, Y., Al-Kaseem, B.R., & Al-Raweshidy, H.S. (2023). Optimized Artificial Intelligence Model for DDoS Detection in SDN Environment. *IEEE Access, 11*, 106733-106748. DOI:10.1109/ACCESS.2023.3319214

[22]Kyaw, A. T., Oo, M. Z., & Khin, C. S. (2020). Machine-learning based DDOS attack classifier in software defined network. *17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*,

[23]Wong, M. L., Arjunan, T. J. E. T. i. M. I., & Data, B. (2024). Real-Time Detection of Network Traffic Anomalies in Big Data Environments Using Deep Learning Models. *International Journal for Research in Applied Science and Engineering Technology* 16(1), 1-11. DOI: 10.22214/ijraset.2024.58946.

[24]Nadeem, M. W., Goh, H. G., Ponnusamy, V., Aun, Y. J. C., Materials, & Continua. (2022). DDoS Detection in SDN using Machine Learning Techniques. *CMC* 71(1). DOI: 10.32604/cmc.2022.021669

[25]Mukherjee, I., Sahu, N.K. & Sahana, S.K. (2023). Simulation and Modeling for Anomaly Detection in IoT Network Using Machine Learning. *Int J Wireless Inf Networks* 30(2);173–189 (2023). https://doi.org/10.1007/s10776-021-00542-7

[26]Fotiadou, K., Velivasaki, T.N., Voulkidis, A.C., Skias, D., Tsekeridou, S., & Zahariadis, T.B. (2021). Network Traffic Anomaly Detection via Deep Learning. *Inf., 12(5)*, 215. DOI:10.3390/info12050215

[27]Hwang, R.-H., Peng, M.-C., Huang, C.-W., Lin, P.-C., & Nguyen, V.-L. J. I. A. (2020). An unsupervised deep learning model for early network traffic anomaly detection. *IEEE Access*, 8, 30387-30399. DOI: 10.1109/access.2020.2973023

[28]Xu, W., Jang-Jaccard, J., Singh, A., Wei, Y., & Sabrina, F. (2021). Improving Performance of Autoencoder-Based Network Anomaly Detection on NSL-KDD Dataset. *IEEE Access*, 9, 140136-140146. DOI:10.1109/ACCESS.2021.3116612

[29]Koren, O., Koren, M., & Peretz, O. J. E. A. o. A. I. (2023). A procedure for anomaly detection and analysis. *Engineering Applications of Artificial Intelligence* 117, 105503. DOI: 10.1016/j.engappai.2022.105503

[30]Rafique, S. H., Abdallah, A., Musa, N. S., & Murugan, T. J. S. (2024). Machine learning and deep learning techniques for internet of things network anomaly detection—current research trends. *Sensors (Basel)* 24(6):1968. doi: 10.3390/s24061968.

[31]Chikodili, N.B., Abdulmalik, M.D., Abisoye, O.A., Bashir, S.A. (2021). Outlier Detection in Multivariate Time Series Data Using a Fusion of K-Medoid, Standardized Euclidean Distance and Z-Score. In: Misra, S., Muhammad-Bello, B. (eds) Information and Communication Technology and Applications. *ICTA 2020. Communications in Computer and Information Science*, vol 1350. Springer, Cham. https://doi.org/10.1007/978-3-030-69143-1_21

[32]Ullah, I., & Mahmoud, Q. H. J. I. A. (2021). Design and development of a deep learning-based model for anomaly detection in IoT networks. *IEEE Access* 9, 103906-103926. doi: 10.1109/ACCESS.2021.3094024

[33]Alqahtani, H. (2020). Cyber Intrusion Detection Using Machine Learning Classification Techniques. *In book: Computing Science, Communication and Security* (pp.121-131) DOI:10.1007/978-981-15-6648-6_10

[34]Doreswamy, Hooshmand, M. K., & Gad, I. J. C. T. o. I. T. (2020). Feature selection approach using ensemble learning for network anomaly detection. *CAAI Transactions on Intelligence Technology 5*(4), 283-293. DOI: 10.1049/trit.2020.0073

[35]Pu, G., Wang, L., Shen, J., Dong, F. J. T. S., & Technology. (2020). A hybrid unsupervised clustering-based anomaly detection method. *Tsinghua Science and Technology* 26(2), 146-153. doi: 10.26599/TST.2019.9010051.

[36]Nanehkaran, Y., Licai, Z., Chen, J., Jamel, A. A., Shengnan, Z., Navaei, Y. D., . . . Computing, M. (2022). Anomaly Detection in Heart Disease Using a Density-Based Unsupervised Approach. *Wireless Communications and Mobile Computing*, Article ID 6913043, 14 pages DOI: 10.1155/2022/6913043

[37]Xu, H., Wang, Y., Jian, S., Liao, Q., Wang, Y., Pang, G. J. I. T. o. K., & Engineering, D. (2024). Calibrated one-class classification for unsupervised time series anomaly detection. arXiv:2207.12201v2

[38]Suseendran, G., Chandrasekaran, E., Akila, D., & Sasi Kumar, A. (2020). Banking and FinTech (financial technology) embraced with IoT device. Data Management, Analytics and Innovation: *Proceedings of ICDMAI 2019*, Volume 1,

[39]Da, Z., Dun, Y., Liu, C., Liang, Y., Xue, Y., & Qian, X. J. K.-B. S. (2023). Anomaly detection framework for unmanned vending machines. *Knowledge-Based Systems* 262, 110251. DOI: https://doi.org/10.1016/j.knosys.2023.110251

[40]Abououf, M., Singh, S., Mizouni, R., & Otrok, H. (2024). Explainable AI for Event and Anomaly Detection and Classification in Healthcare Monitoring Systems. *IEEE Internet of Things Journal*, 11, 3446-3457. DOI: 10.1109/JIOT.2023.3296809

[41]Ahmad, Z., Shahid Khan, A., Nisar, K., Haider, I., Hassan, R., Haque, M. R., . . . Rodrigues, J. J. J. A. S. (2021). Anomaly detection using deep neural network for IoT architecture. *Applied Sciences* 11(15), 7050. DOI: 10.3390/app11157050

[42]Kapoor, A., Sengar, V., George, N., Vatsal, V., Gubbi, J., P., B., & Pal, A. (2023). Concept-Based Anomaly Detection in Retail Stores for Automatic Correction Using Mobile Robots. *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 163-170. DOI: 10.1109/SMC53992.2023.10394209

[43]Yaseen, A. J. S. S. R. o. A. M. L. (2023). The role of machine learning in network anomaly detection for cybersecurity. *SSRAML SageScience*, 1(1), 1–15.

[44]Al-Alawi, A. I., & Al-Bassam, M. S. A. J. J. o. X. U. (2020). The significance of cybersecurity system in helping managing risk in banking and financial sector. *Journal of Xidian University 14*(7), 1523-1536. DOI: 10.37896/jxu14.7/174

[45]Varma, P., Nijjer, S., Sood, K., Grima, S., & Rupeika-Apoga, R. J. R. (2022). Thematic Analysis of Financial Technology (Fintech) Influence on the Banking Industry. *Risks* 10(10), 186. DOI: DOI:10.3390/risks10100186

[46]Shvetsova, N., Bakker, B., Fedulova, I., Schulz, H., & Dylov, D. V. J. I. A. (2021). Anomaly detection in medical imaging with deep perceptual autoencoders. *IEEE Access*, 9, 118571-118583. DOI: 10.1109/ACCESS.2021.3107163

[47]Guha, A., & Samanta, D. (2021). Hybrid approach to document anomaly detection: an application to facilitate RPA in title insurance. *International Journal of Automation and Computing*, *18*(1), 55-72.

[48]Palakurti, N. R. (2024). Challenges and future directions in anomaly detection. In *Practical Applications of Data Processing, Algorithms, and Modeling* (pp. 269-284). IGI Global.