IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

A Systematic Review on Anomaly Detection

Jaiprakash Prajapati[1] and Prof. Nilesh Choudhary[2]
Student, Department of Computer Engineering[1]
Professor, Department of Computer Engineering[2]
Godavari College of Engineering, Jalgaon, Maharashtra, India

Abstract: *Anomaly detection has been used for many years to perceive and extract anomalous points from data. This is an important question that has been explored in various research areas and application domains. Many anomaly detection techniques are specifically designed for specific application domains, while others are more general. Many data science strategies had been used to come across anomalies. One widely used technique is deep machine learning, which play an important role in this field. This research paper provides a systematic literature review analysing ML models for detecting anomalies. Our review analyses the models from four perspectives: the Problem nature and challenges, Classification and formulation, Review of past work, and the future opportunities. When applying a given technique to a particular domain, these assumptions can be used as guidelines to assess the effectiveness of the technique in that domain. We also discuss the computational complexity of the technique, as this is an important issue in real application domains. We hope that this paper will provide a better understanding of the different directions in which research has been done on this topic, and how techniques developed in one area can be applied in domains for which they were not intended to begin with.*

Keywords: Anomaly Detection, Outlier Detection, Machine Learning, Variational Autoencoders

I. INTRODUCTION

Anomaly detection refers to the problem of finding patterns in data that do not match expected behaviour. These nonconforming patterns are often called anomalies, outliers, contradictory observations, exceptions, deviations, surprises, peculiarities, or impurities in various domains of application. Of these, anomaly and outlier are the two most commonly used terms associated with anomaly detection. sometimes interchangeable. Anomaly detection is widely used in various applications i.e., Fraud detection in credit cards, insurance, and healthcare; intrusion detection in cyber security; failure detection in safety-critical systems, airplanes& ATMs; military surveillance for hostile activity.

Data outlier or anomaly detection was studied in the statistical community as early as the 19th century [Edgeworth 1887]. Over time, various anomaly detection techniques have been developed in several research communities. Many of these techniques are specifically designed for specific application domains, but there are also more general techniques. Anomalies fall into three main categories [1], [2], [3].

- **Point Anomalies**: It is considered the simplest anomaly shape. It occurs when anomalies deviate significantly from expected patterns. To detect this type of anomaly, we need to observe all the points that can be detected as deviating from other data flows. Figure 1 shows a point anomaly.

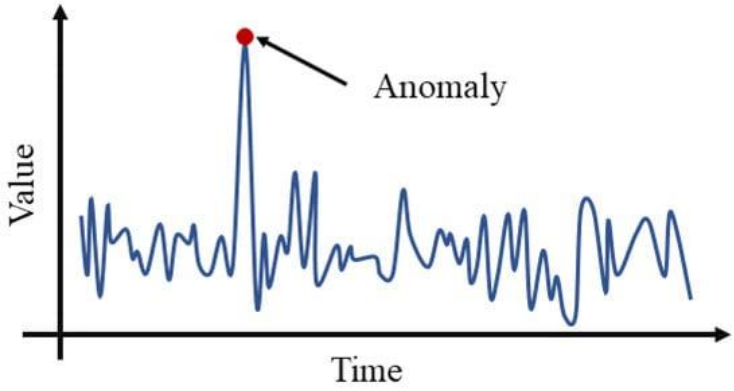

Fig. 1: Point Anomaly Detection

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

- **Contextual Anomalies**: A contextual anomaly is a deviation from an expected behavioural pattern that can be explained by the specific circumstances in which it occurs. For example, even a student who usually gets good grades may get a bad grade on a test in a subject that he/she is not good at. An example context anomaly is shown in Figure 2.

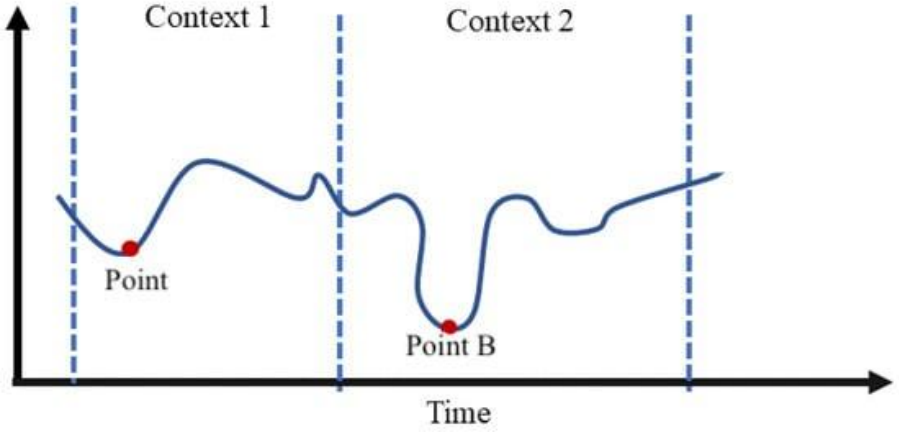

Fig. 2: Contextual Anomaly Detection

- **Collective Anomalies**: Collective anomalies occur when individual data points appear normal when viewed in isolation. However, looking at groups of these data points reveals unexpected patterns, behaviours, or results. Irregular heartbeat is an example of a collective anomaly, these unexpected events could be events that occur in an unexpected order or combination. As example, take again the possibility of credit card fraud. This happens when multiple purchases viewed individually appear to fit your normal spending activity. However, looking at these purchases collectively can reveal unusual patterns and behaviours. An example for collective anomaly is shown in Figure 3.

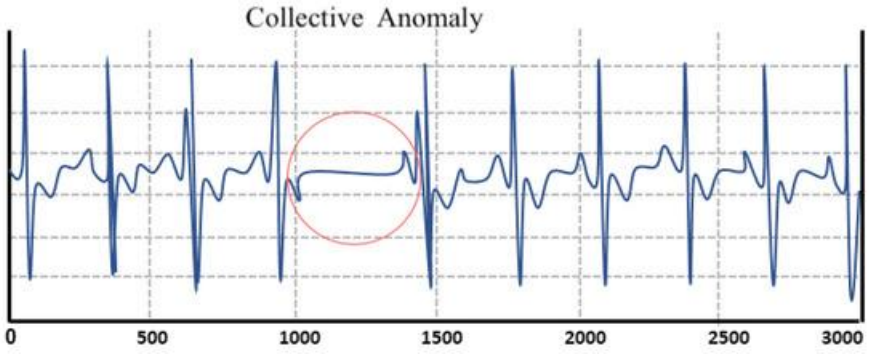

Fig. 3: Collective Anomaly

II. REVIEW OBJECTIVE

This paper aims to review the past work in the area of anomaly detection using machine learning techniques by the categorizing the past work in four major parts:

- **Problem nature and challenges**: This section explains the complexity of some of the inherent problems underlying anomaly detection, computational and algorithmic complexity and the resulting largely unsolved challenges.
- **Classification and formulation**: In this section we formulate the current methods for anomaly detection into three main frameworks: General feature extraction, learning from representations of regularity, and end-to-end Anomaly score detection.
- **Comprehensive Literature Review:** Reviewing relevant studies at leading conferences and journals from multiple relevant communities, including machine learning, data mining, IOT, computer vision and artificial intelligence, to create a comprehensive literature review of research advances. To provide a complete

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

introduction, we outline basic assumptions, objective functions, key intuitions, and their capabilities to overcome some of the above challenges through all categories of methods.

- **Future Opportunities:** In this section we have described the number of possible future opportunities and their implications for addressing the associated challenges.

III. LITERATURE SURVEY

Anomaly detection has been the focus of many researchers. For example, Yu [4] published a comprehensive study on anomalous intrusion detection techniques including statistics, machine learning, neural networks, and his data mining detection techniques. Also, Tsai et al. [5] reviewed intruder detection, but the author focused on machine learning techniques. They gave an overview of machine learning techniques written between 2000 and 2007 that were developed to solve the problem of intrusion detection. In addition, the authors compared relevant studies based on classifier design type, dataset, and other metrics. Similarly, Patcha and Park [6] published extensive research on anomaly detection and attack detection techniques, and Buczak and Buvan [7] examined machine learning and data mining methods for cyberattack detection. They provided a description of each method and addressed challenges in using machine learning and data mining in cybersecurity. Finally, Satpute et al. [8] presented a combination of various machine-learning techniques and particle swarm optimization to improve the efficiency of anomaly detection in network intrusion systems.

Anomaly detection is an important topic that has been explored in various research and implementation areas. Many anomaly detection methods are specifically designed for specific applications, while others are more general. Chandra et al. [1] provided a comprehensive overview of anomaly detection techniques and applications. Board reviews of various machine learning and non-machine learning techniques, such as statistical and spectral detection methods, were discussed in detail. In addition, this study shows several applications of anomaly detection. Examples include cyber intrusion detection, fraud detection, medical anomaly detection, industrial damage detection, image processing detection, text anomaly detection, and sensor networks. The same author introduced another of his studies [3] on anomaly detection for discrete sequences. The authors provided a comprehensive and structured overview of existing research on the problem of detecting anomalies in discrete/symbolic sequences. Furthermore, Hodge and Austin [9] published a comprehensive study on machine learning and statistical anomaly detection methods. The authors also compared and illustrated the advantages and disadvantages of each individual method. On the other hand, Agrawal and Agrawal [10] proposed an anomaly detection investigation using data mining techniques.

Since network anomaly detection is an important research area [11], [12], many studies have focused on this topic. Buyan et al. [13] published a comprehensive study on network anomaly detection. They identified the types of attacks commonly encountered by intrusion detection systems and described and compared the effectiveness of various anomaly detection methods. In addition, the author also described a network defender tool. Similarly, Gogoi et al. [14] reviewed extensive research on known distance-based, and density-based methods, and supervised and unsupervised learning in network anomaly detection. On the other hand, Kwon et al. [15] mainly focus on deep learning techniques such as Constrained deep belief networks based on Boltzmann machines, deep recurrent neural networks, and machine learning methods suitable for detecting network anomalies.

Some research focuses primarily on detecting anomalies in specific domains and applications. In [16], the authors presented a general overview of broad cluster-based fraud detection and compared these techniques from different perspectives. In addition, Sodemann et al. [17] presented anomaly detection in automated surveillance and provided various models and classification algorithms. The authors considered research studies by problem area, approach, and method. In addition, Zuo [18] gave an overview of his three most widely used anomaly detection techniques in the field of geochemical data processing. Fractal/Multifractal Models, Constitutive Data Analysis, and Machine Learning (ML), but the authors focus primarily on machine learning techniques. On the other hand, they [19] examined a framework for log-based anomaly detection. The authors considered six representative anomaly detection methods and evaluated each. The author also compared and contrasted the accuracy and validity of his two representative datasets of the production protocol. Furthermore, Ibidunmoye et al. [20] gave an overview of the detection of performance-related anomalies and bottlenecks in computer systems. The authors identified the basic elements of the problem and categorized existing solutions.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

IV. RESEARCH CHALLENGES AND FUTURE DIRECTIONS

Although various anomaly detection techniques have been proposed in the literature, there are still some issues that need to be resolved for anomaly detection. There is currently no single best approach to this problem. Rather, there are some techniques that are better suited to specific data types and specific application domains. Below is an brief overview of the main challenges identified in the state-of-the-art researched.

1. **Evolving Data Stream**: As vast amounts of knowledge are categorized in the form of data streams characterized by such anomalies, the difficult task of detecting anomalies in evolving data streams must be tackled efficiently [21]. The data stream poses external detection problems such as Managing data by changing the way you detect, update incoming data, and capture underlying detected changes in limited memory and time [22]. Data evolution involves algorithms that use new insights over time to adjust configurations and parameters. Unlike static datasets, detection algorithms cannot adapt to complex conditions such as Ever-changing business domains [23]. Furthermore, most existing ones are inefficient at detecting anomalies in data streams and have poor performance requirements [24]. Anomaly detection in a real-time data stream environment, known for its evolving characteristics, suffers from low detection accuracy and high false positive rate [25]. Evolving data streams are a challenge that needs to be addressed in anomaly detection environments [26,27].
2. **Unknowingness**: Anomalies are associated with many unknowns. e.g., Instances of sudden behaviour, data structure, or unknown distribution. New terrorist attacks, scams, network intrusions, etc. remain unknown until actually occurs.
3. **Feature-Evolving**: Data sources that evolve with functionality also make it difficult to solve the anomaly detection problem. The problem is that the data changes and so do the properties of the data. By comparison, new and old dimensions of data appear and disappear over time. The field is fascinating and has many potentials uses such as Outlier detection in application systems where sensors are regularly switched on and off (representing the number of dimensions) [28].
4. **Heterogeneous anomaly classes**: Because anomalies are random, one class of anomaly can have very different anomaly properties than another class. For example, in video surveillance, unusual events such as robberies, traffic accidents, and robberies are visually very different.
5. **Rarity and class imbalance**: Anomalies are usually rare instances of data, as opposed to regular instances, which often dominate the data. That makes it difficult, if not impossible, to collect a large number of flagged anomalous cases. This makes large data labelled unusable for most applications. The class imbalance is also due to the fact that anomalies are usually much more costly to misclassify than normal instances.
6. **Windowing**: Accuracy is limited (windowing) due to short data processing used based on fixed interval timing [29]. Another major challenge is determining the optimal frequency for model retraining. This is because most current approaches use predefined interval times [29,30].
7. **Ensemble Approaches**: Another growth area is the ensemble approach. Ensemble methods are well established to increase the efficiency of anomaly detection by detecting and enforcing time accuracy [31]. Another valuable and possible research approach is ensemble detection of deviations, which is expected to improve the detection accuracy of algorithms. More specific models may be recommended to address unexplored areas. We recommend exploring the ensemble first to detect anomalies in the data stream environment. However, this research area is still largely unexplored, and a more comprehensive model is needed.
8. **Nature of Input Data**: In the context of IoT anomaly detection, many existing challenges need to be solved. As reported by Azimi et al. [32], the availability of labelled data is a key issue in his IoT anomaly detection, as anomaly occurrences may not be regular. Furthermore, retrieving the actual system data is complex and requires a lengthy process to retrieve the operating system data [33]. There are significant gaps in capturing knowledge logs and formalizing sensory data flows, developing models, and validating them in real-world environments. Many experiments were reported during the research, mainly related to the normal operation of the system [33]. The most advanced methods are based on training for normal behaviour, and anything different from data labelled as normal is considered abnormal. More accurate and reliable techniques are

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

needed to handle complex datasets in real-world scenarios. Moreover, the availability of datasets suitable for general anomaly detection is a key issue for training and validation of real-time anomaly detection techniques [34]. Such records should reflect a variety of new normal and anomalous behaviours, should be clearly identified and constantly updated to prevent the threat of new types of anomalous behaviour. Most existing anomaly detection datasets often lack labelling, low attack diversity, and compatibility with real-time detection [35]. A new anomaly detection dataset requires a realistic environment with a variety of normal and anomaly scenarios. Furthermore, when testing new anomaly detection systems, we need to create a basic truth containing anomalies to increase confidence in the dataset.

9. **Data Complexity and Noise**: Data complexity, Data set imbalance, unexpected noise, and redundancies in the data are among the major challenges in developing anomaly detection models [36]. Gathering useful information and insights requires a well-developed approach to curating datasets.
10. **Parameters Selection**: IoT data streams are often generated from non-stationary environments without sophisticated data distribution information. This influences the selection of a suitable model parameter set for anomaly detection [21].
11. **Data Visualization**: Visualization of anomaly analysis highlighted the presence of gaps. Visual system analysis requires the implementation of new techniques and solutions. Therefore, these gaps should be examined in relation to the area of anomaly detection processes [37].
12. **Time Complexity**: A key feature of data streams is the huge amount of continuously arriving data that requires algorithms to operate in real time. However, there is always a trade-off between accuracy and time complexity, making time complexity a major challenge in detecting anomalies [38,39,40].
13. **11. Accuracy**: Learning algorithms can detect and identify anomalous behaviour in real-time, these algorithms can improve accuracy, such as lowering the detection rate of false positives, especially in large sensor networksare still optimized for [26,27,34,41].
14. **Scalability**: Scalability is another important requirement for anomaly detection algorithms, as many algorithms become inefficient when processing large amounts of data [31].
15. **High Dimensional Data**: Most current data stream anomaly detection algorithms lose their effectiveness in the presence of high dimensional data [21]. Therefore, outlier detection requires accurate and efficient redesign of existing models. More specifically, when many features are observed, a set of outliers may only occur in a subset of dimensions at a given time. This group of outliers looks natural given the different dimensions and timeframes of the subsets. Many features is another challenge for anomaly detection algorithms in choosing the most important data features [37]. Therefore, feature reduction is important in choosing the most important ones that display the whole data.
16. **Diverse types of anomalies**: Three absolutely exceptional varieties of the anomaly were explored [3]. Point anomalies are man or woman times which might be anomalous w.r.t. the bulk of different man or woman times, e.g., atypical fitness signs of a patient. Conditional anomalies, a.k.a. contextual anomalies, additionally seek advice from men or women about anomalous times however in a particular context, i.e., facts times are anomalous withinside the particular context, in any other case every day. The contexts may be fairly exceptional in actual-international applications, e.g., unexpected temperature drop/boom in a specific temporal context, or fast credit score card transactions in uncommon spatial contexts. Group anomalies, a.k.a. collective anomalies, are a subset of facts times anomalous as an entire w.r.t. the opposite facts times; the man or woman individuals of the collective anomaly won't be anomalies, e.g., pretty dense subgraphs fashioned via way of means of faux bills withinside the social community are anomalies as a collection, however, the man or woman nodes in the one's subgraphs may be as every day as actual.

V. RESULTS AND DISCUSSIONS

Our review shows that despite advances in anomaly detection research, there are still many open challenges and problems to be solved. Moreover, the future direction is most urgently needed for anomaly detection approaches in data streams. Therefore, the focus of this study is limited to highlighting open challenges related to anomaly detection in data streams. Therefore, the difficult problem of efficiently detecting anomalies in evolving data streams must be

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

addressed. Most existing data stream anomaly detection algorithms lose their effectiveness in the presence of high-dimensional data. Therefore, current models need to be redesigned to detect anomalies accurately and efficiently. More specifically, if there are many features, there may be a set of anomalies that occur in only a subset of the dimensions at any given time. This set of anomalies looks natural compared to another subset of dimensions and timeframes.

Addressing the problem of anomaly detection in evolving data streams is also an important concern. This challenge arises as data evolves over time and data characteristics change. Additionally, new/old data dimensions will show/hide over time. In addition, another challenge is processing data in real-time when data points are constantly retrieved from the data source. In the past, data streams could be large and had to be processed all at once. Additionally, the algorithm must be able to process data in a given memory so that large amounts of data do not impact the processing power of the data stream. Therefore, data stream algorithms should not require unlimited storage for unlimited data points arriving in the system. Instead, it should be able to process data within the available memory. Additionally, the need for large-scale IoT implementations is growing rapidly, leading to significant security issues. However, there are concerns about the scalability of anomaly detection and how machine learning algorithms can handle large amounts of data. Scalability is a critical issue facing most existing anomaly detection technologies, and some of these technologies become inefficient when deployed at scale.

The techniques reviewed were evaluated with respect to their ability to perform data projections, handle noisy data, and work in limited memory and in limited time. In addition, the ability to deal with evolving data, high-dimensional data, evolving capabilities, and ultimately scalability is addressed.

VI. CONCLUSIONS

As the world becomes more and more data-driven and there is no standard approach for detecting anomalies in data, high-dimensional problems are inevitable in many application domains. Moreover, as the amount of data increases, the loss of precision becomes greater, and the computational cost becomes higher.

Identifying anomalous data points in large datasets with imbalanced dataset problems is a research challenge. This study provided a comprehensive overview of anomaly detection techniques related to data capabilities of volume and velocity. It is clear that further research and evaluation of strategies for detecting anomalies in datasets addressing high-dimensional, imbalance problems are needed.

To address this research problem, we propose future research directions for building new frameworks that can identify anomalous data points in datasets with high-dimensional, imbalance problems. The main contribution is to improve the balance between precision and recall for data anomaly detection problems.

REFERENCES

[1]. V. Chandola, A. Banerjee, and V. Kumar, “Anomaly Detection : A Survey,” ACM Comput. Surv., vol. 41, no. 3, pp. 71–97, 2009, doi: 10.1145/1541880.1541882.

[2]. R. A. Ariyaluran Habeeb, F. Nasaruddin, A. Gani, I. A. Targio Hashem, E. Ahmed, and M. Imran, “Real-time big data processing for anomaly detection: A Survey,” Int. J. Inf. Manage., vol. 45, no. February, pp. 289–307, 2019, doi: 10.1016/j.ijinfomgt.2018.08.006.

[3]. V. Chandola, A. Banerjee, and V. Kumar, “Anomaly Detection for Discrete Sequences: A Survey,” IEEE Trans. Knowl. Data Eng., vol. 24, no. 5, pp. 1–16, 2012.

[4]. Y. Yu, “A survey of anomaly intrusion detection techniques,” J. Comput. Sci. Coll., pp. 9–17, 2012, [Online]. Available: http://dl.acm.org/citation.cfm?id=2379707.

[5]. C. F. Tsai, Y. F. Hsu, C. Y. Lin, and W. Y. Lin, “Intrusion detection by machine learning: A review,” Expert Syst. Appl.,vol. 36, no. 10, pp. 11994–12000, 2009, doi: 10.1016/j.eswa.2009.05.029.

[6]. Patcha and J. M. Park, “An overview of anomaly detection techniques: Existing solutions and latest technological trends,” Comput. Networks, vol. 51, no. 12, pp. 3448–3470, 2007, doi: 10.1016/j.comnet.2007.02.001.

[7]. L. Buczak and E. Guven, “A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection,” vol. 18, no. October, pp. 1153–1176, 2016, doi: 10.1109/COMST.2015.2494502.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

[8]. K. Satpute, S. Agrawal, J. Agrawal, and S. Sharma, "A Survey on Anomaly Detection in Network Intrusion Detection System Using Swarm Optimization Based Machine Learning Techniques," in International Conference on Frontiers of Intelligent Computing, 2013, vol. 199, pp. 441–452, doi: 10.1007/978-3-642-35314-7.

[9]. V. Hodge and J. Austin, "A Survey of Outlier Detection Methodologies," Artif. Intell. Rev., no. 1969, pp. 85–126, 2004, doi: 10.4324/9781315744988.

[10]. S. Agrawal and J. Agrawal, "Survey on anomaly detection using data mining techniques," Procedia Comput. Sci., vol. 60, no. 1, pp. 708–713, 2015, doi: 10.1016/j.procs.2015.08.220.

[11]. V. Sharma, R. Kumar, W. H. Cheng, M. Atiquzzaman, K. Srinivasan, and A. Y. Zomaya, "NHAD: Neuro-Fuzzy Based Horizontal Anomaly Detection in Online Social Networks," IEEE Trans. Knowl. Data Eng., 2018, doi: 10.1109/TKDE.2018.2818163.

[12]. P. Zhao, Y. Zhang, M. Wu, S. C. H. Hoi, M. Tan, and J. Huang, "Adaptive Cost-Sensitive Online Classification," IEEE Trans. Knowl. Data Eng., 2019, doi: 10.1109/TKDE.2018.2826011.

[13]. M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network Anomaly Detection: Methods, Systems and Tools," IEEE Commun. Surv. TUTORIALS, Accept. Publ., pp. 1–34, 2013, [Online]. Available: http://ieeexplore.ieee.org/document/6524462/.

[14]. P. Gogoi, D. K. Bhattacharyya, B. Borah, and J. K. Kalita, "A survey of outlier detection methods in network anomaly identification," Comput. J., vol. 54, no. 4, pp. 570–588, 2011, doi: 10.1093/comjnl/bxr026.

[15]. D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," Cluster Comput., pp. 1–13, 2017, doi: 10.1007/s10586-017-1117-8.

[16]. M. Ahmed, A. N. Mahmood, and M. R. Islam, "A survey of anomaly detection techniques in financial domain," Futur. Gener. Comput. Syst., vol. 55, pp. 278–288, 2015, doi: 10.1016/j.future.2015.01.001.

[17]. A. Sodemann, M. P. Ross, and B. J. Borghetti, "A review of anomaly detection in automated surveillance," IEEE Trans. Syst. Man Cybern. Part C Appl. Rev., vol. 42, no. 6, pp. 1257–1272, 2012, doi: 10.1109/TSMCC.2012.2215319.

[18]. R. Zuo, "Machine Learning of Mineralization-Related Geochemical Anomalies: A Review of Potential Methods," Nat. Resour. Res., vol. 26, no. 4, pp. 457–464, 2017, doi: 10.1007/s11053-017-9345-4.

[19]. S. He, J. Zhu, P. He, and M. R. Lyu, "Experience Report: System Log Analysis for Anomaly Detection," Proc. - Int. Symp. Softw. Reliab. Eng. ISSRE, pp. 207–218, 2016, doi: 10.1109/ISSRE.2016.21.

[20]. O. Ibidunmoye, F. Hernández-Rodriguez, and E. Elmroth, "Performance Anomaly Detection and Bottleneck Identification," ACM Comput. Surv., vol. 48, no. 1, pp. 1–35, 2015, doi: 10.1145/2791120.

[21]. Maia, J.; Severiano, C.A.; Guimarães, F.G.; de Castro, C.L.; Lemos, A.P.; Galindo, J.C.F.; Cohen, M.W. Evolving clustering algorithm based on mixture of typicalities for stream data mining. Future Gener. Comput. Syst. 2020, 106, 672–684.

[22]. Peng, Y.; Tan, A.; Wu, J.; Bi, Y. Hierarchical Edge Computing: A Novel Multi-Source Multi-Dimensional Data Anomaly Detection Scheme for Industrial Internet of Things. IEEE Access 2019, 7, 111257–111270.

[23]. Gottwalt, F.; Chang, E.; Dillon, T. CorrCorr: A feature selection method for multivariate correlation network anomaly detection techniques. Comput. Secur. 2019, 83, 234–245.

[24]. Ding, N.; Ma, H.; Gao, H.; Ma, Y.; Tan, G. Real-time anomaly detection based on long short-Term memory and Gaussian Mixture Model. Comput. Electr. Eng. 2019, 79, 106458.

[25]. Xing, L.; Demertzis, K.; Yang, J. Identifying data streams anomalies by evolving spiking restricted Boltzmann machines. Neural Comput. Appl. 2020, 32, 6699–6713.

[26]. Bezerra, C.G.; Costa, B.S.J.; Guedes, L.A.; Angelov, P.P. An evolving approach to data streams clustering based on typicality and eccentricity data analytics. Inf. Sci. 2020, 518, 13–28.

[27]. Maciąg, P.S.; Kryszkiewicz, M.; Bembenik, R.; Lobo, J.L.; Del Ser, J. Unsupervised Anomaly Detection in Stream Data with Online Evolving Spiking Neural Networks. Neural Netw. 2021, 139, 118–139.

[28]. Manzoor, E.; Lamba, H.; Akoglu, L. xStream: Outlier Detection in Feature-Evolving Data Streams. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 1963–1972.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

[29]. Vergeles, A.; Khaya, A.; Prokopenko, D.; Manakova, N. Unsupervised Real-Time Stream-Based Novelty Detection Technique an Approach in a Corporate Cloud. In Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, 21–25 August 2018; IEEE: New York, NY, USA, 2018; pp. 166–170.

[30]. Stiawan, D.; Idris, M.Y.; Malik, R.F.; Nurmaini, S.; Budiarto, R. Anomaly detection and monitoring in Internet of Things communication. In Proceedings of the 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 5–6 October 2016; Institute of Electrical and Electronics Engineers: New York, NY, USA, 2016; pp. 1–4.

[31]. Dong, Y.; Japkowicz, N. Threaded ensembles of autoencoders for stream learning. Comput. Intell. 2017, 34, 261–281.

[32]. Azimi, I.; Oti, O.; Labbaf, S.; Niela-Vilen, H.; Axelin, A.; Dutt, N.; Liljeberg, P.; Rahmani, A.M. Personalized Maternal Sleep Quality Assessment: An Objective IoT-based Longitudinal Study. IEEE Access 2019, 7, 93433–93447.

[33]. Fahim, M.; Sillitti, A. Anomaly Detection, Analysis and Prediction Techniques in IoT Environment: A Systematic Literature Review. IEEE Access 2019, 7, 81664–81681.

[34]. Moustafa, N.; Hu, J.; Slay, J. A holistic review of Network Anomaly Detection Systems: A comprehensive survey. J. Netw. Comput. Appl. 2019, 128, 33–55.

[35]. Wang, H.; Bah, M.J.; Hammad, M. Progress in Outlier Detection Techniques: A Survey. IEEE Access 2019, 7, 107964–108000.

[36]. Qiu, J.; Du, Q.; Qian, C. KPI-TSAD: A Time-Series Anomaly Detector for KPI Monitoring in Cloud Applications. Symmetry 2019, 11, 1350.

[37]. Vilenski, E.; Bak, P.; Rosenblatt, J.D. Multivariate anomaly detection for ensuring data quality of dendrometer sensor networks. Comput. Electron. Agric. 2019, 162, 412–421.

[38]. Yu, K.; Shi, W.; Santoro, N. Designing a Streaming Algorithm for Outlier Detection in Data Mining—An Incrementa Approach. Sensors 2020, 20, 1261.

[39]. Gibert, X.; Patel, V.M.; Chellappa, R. Deep Multitask Learning for Railway Track Inspection. IEEE Trans. Intell. Transp. Syst. 2017, 18, 153–164.

[40]. Santos, J.; Leroux, P.; Wauters, T.; Volckaert, B.; De Turck, F. Anomaly detection for Smart City applications over 5G low power wide area networks. In Proceeding of the NOMS 2018—2018 IEEE/IFIP Network Operations and Management Symposium, Taipei, Taiwan, 23–27 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–9.

[41]. Da Costa, K.A.; Papa, J.P.; Lisboa, C.O.; Munoz, R.; de Albuquerque, V.H.C. Internet of Things: A survey on machine learning-based intrusion detection approaches. Comput. Netw. 2019, 151, 147–157.

BIOGRAPHY

- Jaiprakash is an M. Tech candidate at the Dr. Babasaheb Ambedkar Technological University, researching the subjects of AI, NLP, and social networks. He is a seasoned Data Scientist with over a decade of professional modeling experience and a passion for developing business ideas into implementable solutions by harnessing the power of Machine Learning and Technology.

  A multidisciplinary data scientist with extensive experience in machine learning and data engineering. He works with technologies such as Docker, TensorFlow, Spark, cloud environments, relational and graph databases, APIs, Python, and R stacks. He has prototyped, developed, and deployed real-world data applications for many of the most prominent Indian and global brands. He can be reached at jaiprakash.prajapati@outlook.com.