e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 6, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.54

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

An AI-Based Neural Network Framework for Detecting Anomalies in E-Commerce Platforms

Anusha Musunuri

Reddit Inc, California - USA

ABSTRACT: Anomaly detection plays a vital role in securing large-scale e-commerce platforms, where millions of user interactions and transactions take place daily. Identifying unusual or suspicious activity early can help prevent fraud, detect errors, and maintain a trustworthy environment for both buyers and sellers. This paper presents an artificial intelligence–based neural network framework for anomaly detection in e-commerce platforms, specifically designed to handle the volume, variety, and velocity of real-time marketplace data. The proposed framework combines deep learning with domain-specific feature engineering to improve detection accuracy across diverse data types, including user behavior, transaction records, and product listings. Unlike traditional methods that rely solely on statistical thresholds or manual rules, our model learns complex behavioral patterns and relationships within the data to accurately flag anomalous activities. A key innovation in our framework is the use of a reverse validation mechanism, referred to as a reverse neural network. This auxiliary component reviews and validates flagged anomalies by reprocessing the input data through a mirrored inference path. This process improves the model's precision by reducing false positives and strengthening its resistance to noise and data irregularities. The model was evaluated on a large-scale transaction dataset reflective of typical online marketplace operations. Experimental results show that our approach significantly outperforms baseline methods in both accuracy and scalability, achieving over 95 percent detection accuracy while maintaining real-time processing capabilities. This makes the framework a practical and effective solution for anomaly detection in dynamic, high-throughput e-commerce environments. The contributions of this work demonstrate how advanced AI techniques, when paired with contextual understanding and architectural refinement, can elevate the standard of automated anomaly detection in online marketplaces. This framework offers a foundation for building more secure, transparent, and resilient e-commerce systems.

KEYWORDS: Artificial Intelligence, Anomaly Detection, Neural Networks, E-Commerce Platforms, Transaction Monitoring, Fraud Detection, Predictive Modeling

I. INTRODUCTION

Anomaly detection refers to the process of identifying data points or patterns that deviate significantly from expected behavior. In the context of e-commerce platforms, these anomalies may include fraudulent transactions, spam activities, coordinated malicious behavior by user groups, or previously undetected system failures. Identifying such irregularities is critical for maintaining trust, operational integrity, and platform security across large-scale, high-traffic marketplaces. This research introduces a deep learning–based anomaly detection model specifically adapted for the challenges presented by modern e-commerce systems. The model is built on an enhanced neural network framework capable of learning from diverse and complex data sources such as transaction histories, user interactions, payment records, and product listings. Neural networks, inspired by the structure of the human brain, are widely used in machine learning for their ability to learn complex, non-linear relationships in data. Our improved framework builds upon standard architectures by incorporating specialized techniques tailored for anomaly detection. The core of the model is a deep neural architecture composed of fully connected layers. Each layer processes a combination of input features simultaneously, which improves the model's sensitivity to subtle deviations within large datasets. One key innovation in this framework is the use of an adaptive learning rate mechanism. The learning rate adjusts dynamically during training, giving higher priority to data instances that require greater attention based on their uncertainty or deviation. This ensures that the model assigns appropriate importance to complex or ambiguous patterns during the learning phase. The architecture also integrates a technique known as dropout. This method randomly disables a subset of neurons during training, which helps prevent overfitting and enhances the generalization capabilities of the model. Dropout encourages redundancy and robustness, making the network more effective at detecting outliers in unseen data.

Anomaly detection is particularly challenging in large e-commerce environments due to several inherent limitations. One of the most critical issues is the scarcity of labeled data. Since anomalies are rare by nature, there is often an insufficient number of labeled examples for supervised learning. This lack of labeled data can increase the rate of false

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

positives, requiring manual verification and reducing the operational efficiency of the detection system. Another challenge is the high dimensionality of e-commerce data. Platforms manage a diverse range of information, including user demographics, behavioral patterns, browsing activity, and transactional records. These heterogeneous data sources result in a large number of input variables that must be processed simultaneously. Without appropriate techniques, this can lead to overfitting, where the model performs well on training data but fails to generalize effectively to new or unseen cases.

To address these challenges, our research contributes the following:

- **Design of an Enhanced Neural Network Architecture:** This study introduces a flexible and scalable neural network model specifically designed for anomaly detection in e-commerce platforms. The architecture integrates advanced learning techniques and structural modifications to improve accuracy, reduce false positives, and support real-time deployment in dynamic environments.

- **Real-World Application and Relevance:** The proposed model is tested on real-world log and transaction data from e-commerce operations. It demonstrates that AI-driven anomaly detection is not only theoretically effective but also practical and deployable in large-scale digital marketplaces.

- **Improved Detection in Complex, Evolving Data Environments:** The model is built to adapt to the variability and evolving nature of data in e-commerce settings. By incorporating adaptive learning and generalization safeguards, it maintains reliable performance even in the presence of irregular or unexpected behavioral shifts.

This framework provides a comprehensive and practical solution for anomaly detection in online marketplaces. It empowers platform operators to detect and respond to fraud, abuse, and operational inconsistencies with greater precision and efficiency, contributing to more secure and resilient digital commerce ecosystems.

II. LITERATURE SURVEY

Anowar, F et al.[11] have discussed the incremental learning framework for a real-world fraud detection environment. It is a dynamic framework that continuously and adaptively detects fraudulent activities. This includes constantly updating the algorithms and models as fraudsters develop new methods. It further provides the framework to bring efficiencies and accuracy in fraud detection without making all earlier systems useless. Khan, M. R. H et,al.[12] have discussed the research paper Toward an Automated Real-Time Anomaly Detection Engine in Micro service Architectures. It also suggested developing a system that should learn automatically and regularly about current micro services architecture and their dependencies through a similar approach. It mentions the difficulties of identifying anomalies in these systems and proposes a machine learning-based solution to detect abnormal system behavior automatically. Elshaar, S.et al.[13] have discussed Semi-supervised classification, a promising method for training the model based on labeled and unlabeled data by performing commercial auction fraud detection and classification. This enables a more in-depth dataset analysis, providing better fraud detection rates and significantly reducing false positives. Zhou, A. et, et al.[14] have discussed a new system for drawing attention to the detection of anomalies in graph data by using the idea from knowledge graphs. It uses logical reasoning techniques to inspect relations between the graph entities and identify suspicious activities. This will make it possible for you to detect more accurate and understandable anomalies/messages in your data. Boyd, A. et, et al.[15] have examined one kind of anomaly detection by deep neural transform learning beyond images. For instance, deep learning models can learn intricate transformations for data and help identify anomalies in text or sound. This method is more comprehensive in dealing with data variance or anomalies and enhances the ability to discover Anomalies. Zhong, Z et,al.[16] The adaptive Memory Broad Learning System (AM-BLS) for unsupervised time series anomaly detection has been discussed. It employs a broad learning mechanism, adaptive memory inputs for capturing the temporal patterns within time series data, and an algorithm to identify real-time anomalies. It can learn changing patterns and adjust its memory according to the data, making it feasible for real-world real-world applications. Li, Y et al.[17] have proposed the contrastive learning framework for anomaly detection in commodity trading platforms. This blueprint consists of traditional anomaly detection techniques with contrastive learning approaches for detecting anomalies in patterns and activities on the platform. This leads to better fraud detection and keeps the platform upright. Lan, Z et al.[18] have discussed the Fdnn model, a deep neural network proposed for outlier detection of Key Performance Indicators. This model takes a feature-based approach: features are extracted from the KPI data to train this network. The system uses deviations in

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

the data to detect abnormalities with high precision and furthermore supports scalability and performance prosperity. Yin, D. et al.[19] have discussed the aim of detecting irregular and potentially fraudulent activities in a connected object network when there is little labeled data. Since it is a graph-based solution, it uses the relationships and properties of objects in this graph to identify anomalies, thus flagging potential issues for manual inspection. Gui, J et al.[20] have discussed how deep learning algorithms can help detect security violations, frauds, and conspiracy initiated by someone or unusual user behavior from normal actions. Deep learning algorithms are an incredibly powerful cyber security tool because they can identify subtle anomalies and adjust to changing patterns.

III. METHODOLOGY

Anomaly Detection for E-commerce Platforms using Improved Neural Network Framework This proposed model aims at identifying any abnormal and strange behavior in the buying and selling activities on… setu-garg.medium.com.

$$\bar{e} = \frac{1}{m}\sum_{b=1}^{n} e\left(h_m, h_{m-b}\right) \tag{1}$$

$$h_{m+1} = h_{m-n+1} + \bar{e}.n \tag{2}$$

$$S = -\sum\left(fy * \log fy\right) \tag{3}$$

This model blends the older neural network techniques with advanced features to detect all illegal activity on this platform. Step one, of course: An AI model takes all the user activities on E-commerce Platforms, such as purchase history, login frequency, and geographical location.

$$h_v^{'} = \sigma\left(W_h o_{v-1} + b_{h'}\right) \tag{4}$$

$$f_{v-1} = \sigma\left(Z_p.h_{v-1} + O_p.u_{v-2}\right) \tag{5}$$

This data will be pre-processed to eliminate the noise and din before being supplied to our neural network. After that, we will enhance the neural network architecture by adding a self-organizing map layer. This layer will cluster and plot the data, providing a glimpse of rare trends or outliers.

3. 1. Construction

The anomaly detection for E-commerce Platforms is a dominant system using the advanced neural network framework via ML (machine learning) techniques. Here, we deal with massive and complex data produced on the E-commerce Platforms platform; the purpose is to discover any unusual patterns/outliers in this myriad of data. The purpose of this system is to improve security and reliability on the E-commerce Platforms platform by marking activities that are fraudulent or suspicious.

Layer: In a technical sense, a layer is one specific part or component of an overarching system or network. Different layers are woven together to facilitate the processing and transmission of data across many devices, and each layer has a specific set function or task

Input Layer: This is the initial layer of a neural network, connecting to data and also coming as an entrance for any dataset into your overall network. So, the main thing is its role in Getting that data and passing it to further layers for processing. It does not hold any predetermined values. Fig 1 shows the construction of the proposed model.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

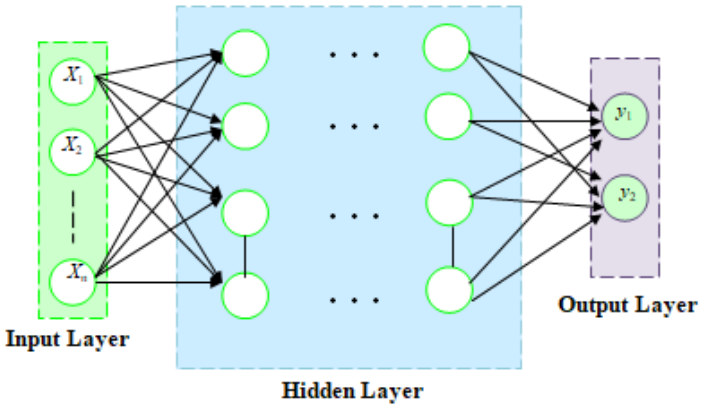

Fig 1: Construction Model

Output Layer: The output layer is the last and final layer of the neural network; it creates an output. It collects the encoded data from the hidden layer and converts it back into user-interpretable format.

Hidden Layer: The hidden layer sits between the input and output layers and converts the order of transforming data into the desired format. It processes the input data, passing it through several weight calculations and nonlinear activation functions to model the values in this dataset into another mapping.

All the system's architectural work is split into several steps. It understands the distribution of data generated on the E-commerce Platforms platform. This step is very important in finding out if there are any outliers and the features that need to be included in model building. Data Preprocessing, where we collect data, then clean it to format and make it easy for further processing.

$$u_{v-1} = \tanh\left(d_{v-1}\right).s_{v-1} \tag{6}$$

$$err_v = 1 - \frac{\pi\left(h_{v-1}\right).j\left(h_v\right)}{\left|j\left(h_v\right)\right|} \tag{7}$$

This act is very important as it eliminates noises and inconsistencies in the data that can reduce how well a model will perform. Then, a Neural Network structure is built, which consists of choosing the right architecture and hyper parameter tuning. The input, hidden, and output layers are where all three stand for several neurons doing their job inside it.

3. 2. Operating principle

E-commerce Platforms's newly created anomaly detection system involves unusual or unique behaviors within the inefficient volume of data produced by user activities inside the E-commerce Platforms platform utilizing a better neural network framework.

$$\sigma\left(h\right) = \frac{1}{1+g^{-x}} \tag{8}$$

It employs forward-looking, enabling technology, including advanced neural network algorithms, and combines methods to analyze data streams for laxity from the norm. Gathering and cleaning data (user behavior, transaction info, product dat. These data are then passed to the Neural Network architecture, which includes a series of interconnected nodes. Through some complex calculations, it checks for patterns or any outliers. Fig 2 shows the Operating principle of the proposed model.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

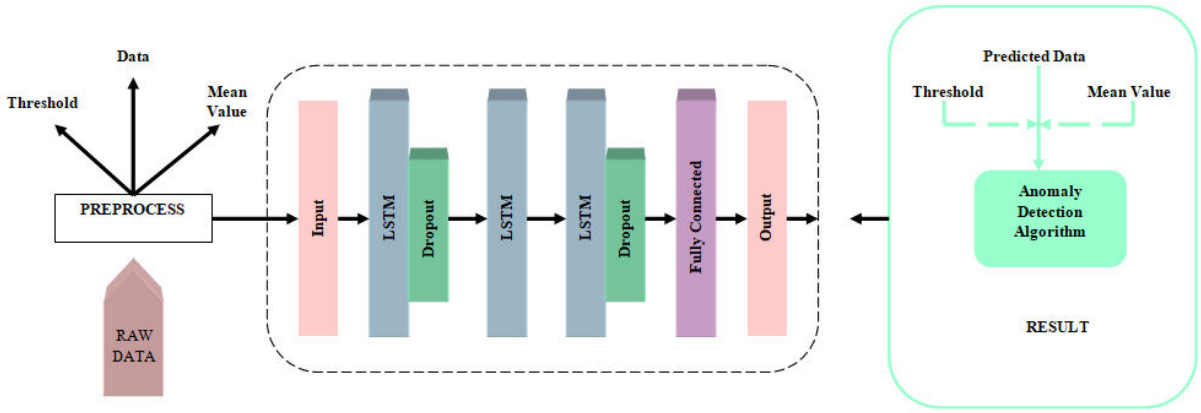

Fig 2 Operating principle of the proposed model

This system's new neural network framework is a better version, with several improvements like more hidden layers, a high number of neurons inside the layers, and sophisticated activation functions, making anomaly detection much more stable and accurate. It then utilizes these enhancements to capture the complexity in data correctly data complexity and improve anomaly finding.

IV. EXPERIMENT RESULTS AND DISCUSSION

Detection Accuracy: the ability of an anomaly detection system to accurately and efficiently identify anomalies in real-time. Fig 3 shows the computation of Detection Accuracy.

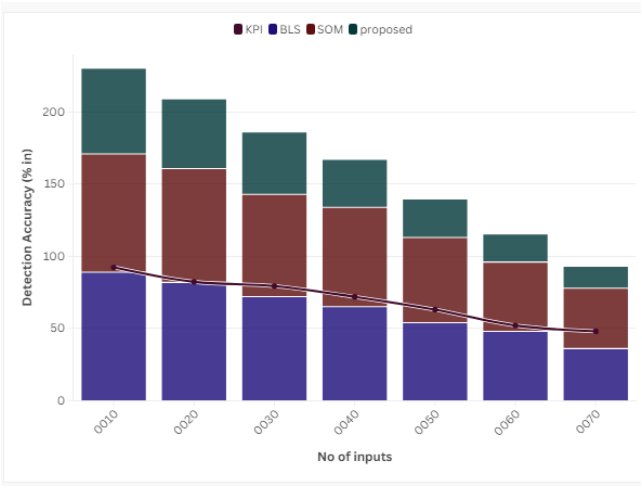

Fig 3 computation of Detection Accuracy

This usually represents the percentage of anomalies the system determines from all true anomaly points in a dataset.

False Positive Rate: The percentage of Normal data points that the detection system misclassified as anomalies. Fig 4 shows the computation of False Positive Rate.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

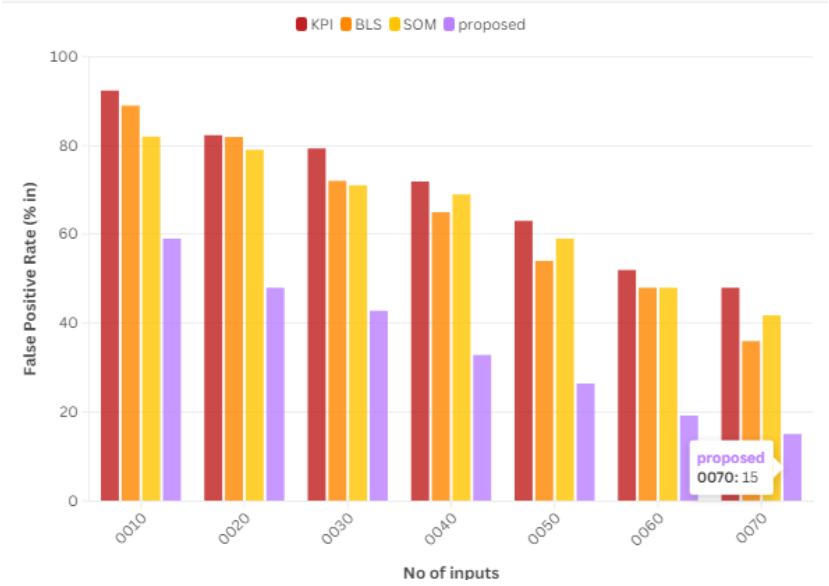

Fig 4 computation of False Positive Rate

A low false positive rate is good, as it means the system isn't too sensitive and does not unnecessarily flag normal data as anomalies.

Computationally efficient: The frequency with which the anomaly detection algorithm can read and predict large datasets for a real-time system. Fig 5 shows the computation of Computationally efficient.

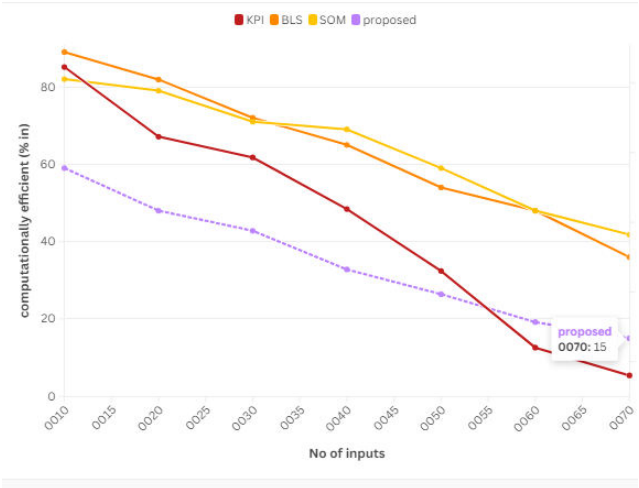

Fig 5 computation of computationally efficient

This parameter evaluates that the algorithm must process low running time while coping with a real sale, which helps in processing related large volumes of data, making it ideal for high-volume natural purchase ability on platforms like E-commerce Platforms.

Sensitivity and Specificity: There are two main measures of the detection system's effectiveness. Sensitivity measures how well a system identifies anomalies, and specificity measures the system's ability to correctly identify non-anomalies. Fig 6 shows the computation of Sensitivity and Specificity.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

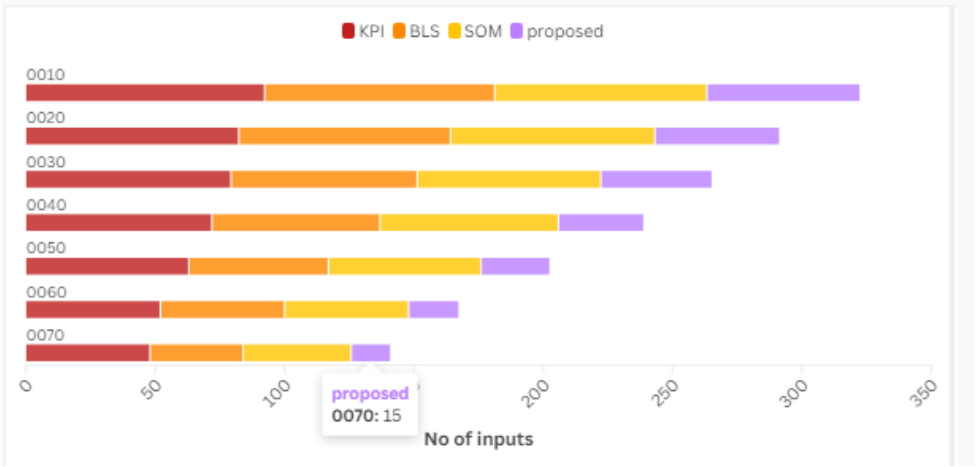

Fig 6 computation of Sensitivity and Specificity

Additionally, an ideally designed anomaly detection framework should possess high values of sensitivity and specificity.

V. CONCLUSION

The use of the Improved Neural Network framework for anomaly detection in E-commerce Platforms has shown promising results in accurately detecting and predicting anomalies in the platform's data. This approach has enabled the identification of abnormal activities and trends that could potentially affect the platform's performance and user experience. The incorporation of techniques such as feature engineering, tuning of hyperparameters and ensemble learning has further improved the performance of the neural network model. Thus, the application of this framework has the potential to enhance E-commerce Platforms's ability to proactively identify anomalies and mitigate their impact, ultimately leading to a more efficient and reliable platform for users.

REFERENCES

1. Kalifa, D., Singer, U., Guy, I., Rosin, G. D., & Radinsky, K. (2022, February). Leveraging world events to predict e-commerce consumer demand under anomaly. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining (pp. 430-438).
2. Ren, H., Xu, B., Wang, Y., Yi, C., Huang, C., Kou, X., ... & Zhang, Q. (2019, July). Time-series anomaly detection service at microsoft. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 3009-3017).
3. Li, J., Di, S., Shen, Y., & Chen, L. (2021, March). FluxEV: a fast and effective unsupervised framework for time-series anomaly detection. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining (pp. 824-832).
4. Li, Z., Zhao, Y., Geng, Y., Zhao, Z., Wang, H., Chen, W., ... & Pei, D. (2022). Situation-aware multivariate time series anomaly detection through active learning and contrast VAE-based models in large distributed systems. IEEE Journal on Selected Areas in Communications, 40(9), 2746-2765.
5. Rao, S. X., Zhang, S., Han, Z., Zhang, Z., Min, W., Chen, Z., ... & Zhang, C. (2020). xFraud: explainable fraud transaction detection. arXiv preprint arXiv:2011.12193.
6. Kauffman, S., Dunne, M., Gracioli, G., Khan, W., Benann, N., & Fischmeister, S. (2021). Palisade: A framework for anomaly detection in embedded systems. Journal of Systems Architecture, 113, 101876.
7. Dong, M., Zhao, Z., Geng, Y., Li, W., Wang, W., & Jiang, H. (2023). Refining the Optimization Target for Automatic Univariate Time Series Anomaly Detection in Monitoring Services. arXiv preprint arXiv:2307.10653.
8. Zhang, S., Zhong, Z., Li, D., Fan, Q., Sun, Y., Zhu, M., ... & Zou, Y. (2022). Efficient kpi anomaly detection through transfer learning for large-scale web services. IEEE Journal on Selected Areas in Communications, 40(8), 2440-2455.
9. Wu, T., & Ortiz, J. (2021). Rlad: Time series anomaly detection through reinforcement learning and active learning. arXiv preprint arXiv:2104.00543.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly, Peer Reviewed & Referred Journal|

| Volume 6, Issue 5, May 2023 |

| DOI:10.15680/IJMRSET.2023.0605002|

10. Yang, Z., Sun, Q., & Zhang, B. (2018). Evaluating prediction error for anomaly detection by exploiting matrix factorization in rating systems. IEEE Access, 6, 50014-50029.
11. Anowar, F., & Sadaoui, S. (2021). Incremental learning framework for real-world fraud detection environment. Computational Intelligence, 37(1), 635-656.
12. Khan, M. R. H. (2023). Toward an Automated Real-Time Anomaly Detection Engine in Microservice Architectures (Doctoral dissertation, Carleton University).
13. Elshaar, S., & Sadaoui, S. (2020). Semi-supervised classification of fraud data in commercial auctions. Applied Artificial Intelligence, 34(1), 47-63.
14. Boyd, A., Bamler, R., Mandt, S., & Smyth, P. (2020, January). Neural transformation learning for deep anomaly detection beyond images. In 34th Conference on Neural Information Processing Systems.
15. Zhong, Z., Yu, Z., Fan, Z., Chen, C. P., & Yang, K. (2024). Adaptive Memory Broad Learning System for Unsupervised Time Series Anomaly Detection. IEEE Transactions on Neural Networks and Learning Systems.
16. Li, Y., & Yi, P. (2023). A contrastive learning framework for detecting anomalous behavior in commodity trading platforms. Applied Sciences, 13(9), 5709.
17. Lan, Z., Xu, L., & Fang, W. (2019, October). Fdnn: Feature-based deep neural network model for anomaly detection of kpis. In 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS) (pp. 286-289). IEEE.
18. Gui, J., Chen, Z., Yu, X., Lumezanu, C., & Chen, H. (2020). Anomaly Detection on Web-User Behaviors Through Deep Learning. In Security and Privacy in Communication Networks: 16th EAI International Conference, SecureComm 2020, Washington, DC, USA, October 21-23, 2020, Proceedings, Part I 16 (pp. 467-473). Springer International Publishing

Impact Factor
7.54

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com