International Journal of Scientific Research in Science, Engineering and Technology

Print ISSN - 2395-1990
Online ISSN : 2394-4099

OPEN ACCESS

Available Online at : www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET19115110

Anomaly Detection in ERP Systems Using AI and Machine Learning

Naveen Kumar

Wipro Limited, East Brunswick, New Jersey, USA

ARTICLE INFO

Article History:

Accepted: 25 May 2019
Published: 08 June 2019

Publication Issue :
Volume 6, Issue 3
May-June-2019

Page Number :
522-530

ABSTRACT

ERP systems are critical in the administration of corporate activities and address huge volumes of transactional and operation information. However, since ERP systems combine many operations of an organization into one system it is prone to a possibility of developing an anomaly that can come from an erroneous data input or even instances of hacking hence causing operational insecurity and loss. This work seeks to understand how AI (Artificial Intelligence) and Machine Leaning (ML) can be used to determine abnormalities in ERP systems. Conventional methods of anomaly detection do not allow for detailed recognition and handling of complex patterns; thus, AI and ML are suitable for dynamic systems. The paper discusses different forms of anomalous situations in ERP systems and examines the potential of different learning techniques in increasing the effectiveness of anomaly identification. The framework that is presented provides for the incorporation of ML relied anomaly recognition in ERP systems to optimize operational efficiency as well as error identification in real time.

Keywords - ERP, ML, AI, Anomalies and Application Software.

I. INTRODUCTION

ERP systems are complex application software that provides the technological support to manage complex business operations in the twenty first century organization. To serve as the backbone of most organizational activities, ERP systems control and coordinate the primary organizational functions, including the financial, human resource, supply chain, sales and marketing, customer relationship management, etc. As these procedures are integrated in one solution, ERP systems let organisations function effectively, reduce the impact of mistakes, and make sound business decisions based on accurate data [1]. The amount and types of data that are processed and made available through ERP systems is enormous and diverse, and includes almost all transactional data such as purchase orders, sales receipts, personal data in the form of employees' and customers' information. At the same time, this amount and variety of data can bring improvements in operations' transparency, but it brings problems of data quality, protection, and

information systems' stability. ERP systems, when fail, experience issues resulting from data corruption problems with system configuration mistakes, or potential cyber threats; this results in financial loss, reputation impacts, regulatory non-compliance or operation interruption. Consequently, the preservation of the correctness and dependability of ERP systems is a concern for any organisation that uses them [2].

The use of anomaly detection is generally a significant approach in ERP system reliability. The aim of anomaly detection is to detect outliers or 'abnormal' behavior/anomalies in a system, data etc, which may flag a problematic situation in the system/data, incorrect data etc. ERP and its anomalies: There are several reasons for these anomalies, which may be originated from a human being by entering incorrect data, there could be system mis-configurations, attempts to breach the security system or probability behavior arising from a change in the business environment. The early diagnosis of such abnormalities is also important to proactively solve the possible issues in advance of becoming critical failures [3].

Consequently, most conventional approaches to anomaly detection in ERP systems have been rule-based systems. Such approaches include one setting up a number of fixed criteria or parameters that are not directly related to the size or frequency of transactions, for example, all transactions exceeding a certain dollar limit are considered suspicious while all orders placed during night time are potential outliers. Rule based systems are very basic and very useful in detecting known that is easily predictable anomalies but they have definite limitations in the case of complex anomalies or anomalies that are as yet unknown. Of these, semi-automatic and automatic working systems have inherent problems, notably they are rigid, and perhaps they may fail on complex patterns or anomalies not envisaged when defining the rules into the system[4]. The concepts of AI and ML provide a new approach to anomaly detection in ERP systems as flexible and more adaptive than other conventional methods. AI/ML functions can learn the behavior patterns of a community of users, build the models of 'standard' activity and update their perspectives, drawing on additions to the existing data set. Such an approach helps AI and ML models to detect variances and patterns that are hard or impossible to detect through the use of rules [5]. For instance, machine learning algorithms can detect changes in normal frequency, time of day and user access and treat any of these as an anomaly.

AI and ML-based anomaly detection systems are best-suited to changing ERP environments since the data itself, users and transaction types might differ constantly. This flexibility is necessary in organizations where ERP application is usually adjusted to meet the company's requirements or changed in line with changed legislation [6].

Other areas where the AI and ML based systems benefit the ERP systems are; they give the ERP systems ability to closely monitor and quickly detect any anomalies than conventional systems. Real-time anomaly detection is especially important for real-time ERPs, which require immediate reaction to avoid errors or fraud in a company's financials or inventory deficits in their supply chain [7].

1.1 Types of Anomalies in ERP Systems

i) Abnormalities within ERP systems refer to any shakeup or deviation from expected norm in data, process or user behavior. They all come with their unique challenges ranging from data inaccuracy, poor operations and security amongst others. ERP system anomalies are of three types namely data anomaly, operation anomaly and security anomaly and each of these anomalies vary in the manner in which they affect the performance of the system and operations in the business organization [8].

ii) Data Anomalies: Data anomalies in ERP systems occur when data entered, integrated, or processed contains inconsistency, inaccuracy, or contains

error. These anomalies negatively a ect the quality of data and are the root of skewed analytical results and incorrect business decisions.

iii) Operational Anomalies: Process irregularities are fluctuations that influence the routine ERP systems within an organization. Such problems may emerge because of the processes' inefficiency or misuse, ill-conceived concepts and may be the sign of such issues as ineffective workflow, users' mistakes, or shifting business needs requiring process modifications.

iv) Security Anomalies: Security anomalies refer to deviations that may suggest Other security breaches, threats or risks in ERP system. Of these, these are particularly essential because they present immediate threats to the data and the ERP system safety [9].

II. Machine Learning Approaches to Anomaly Detection in ERP Systems

AI and ML are making changes to complex ERP systems as they integrate into ERP systems to streamline their normal operation. The specific ERP application of AI allows supplementing predictive and prescriptive analyses with ERP systems, thus automating data validation, error detection, and report generation.

However, Machine Learning (ML) brings about flexibility by training the system from the data and improving the ERP systems' capability to identify breakages and future patterns. ML is more used in anomaly detection, predictive maintenance, and workforce management to recommend on a workforce plan by looking at a historical data [11].

The application of AI and ML in ERP systems helps to automate common chores and find new patterns that mean better security and organization for business. Combined all these technologies will take the ERP systems from a normal standard to more reliable, flexible, and capable of facing every increasingly demand of the new generation business world.

The fact is that there is no single solution to adapt ML to the ERP system because the data is rather versatile and the same concerns the kinds of anomalous circumstances that are possible. This section gives a comprehensive analysis of the various ML techniques that are frequently applied to anomaly detection in relation to ERPs classified according to conventional learning methods as follows; [12][13].

i) Supervised Learning: Anomaly detection in supervised learning uses labeled training datasets where an example of anomalies when analyzed enables the detection process. It is especially applicable in ERP systems, where certain forms of the anomaly, for example fraud, are known in advance.

ii) Unsupervised Learning: Unsupervised learning proves very useful whenever labeled data is not available or wherever the anomalies are not described. This approach is used commonly with regard to ERP systems that are subjected to a process of sustain incremental change and where business processes are dynamic.

iii) Semi-Supervised Learning: It is a technique that lies however in between supervised and unsupervised learning, using some labeled data together with a huge amount of unlabeled data. This is especially more so in ERP systems, where labeling the set may cover a significant amount of history, while the need to detect anomalies is largely important.

III. Anomaly Detection Framework in ERP Systems

In this paper, the formulation of an effective framework for anomaly detection in ERP Systems is a holistic process that goes through several critical phases. The base starts with the accumulation of data that includes transactions, change of inventories, HR activities and audit trails for observance of business activities and identification of irregularities). These sources of data give a basis for marking out the

financial frauds, or other inefficiency indicators, or even rather suspicious activity. This makes the data structured and suitable for analysis, common issues like, missing value problem and interoperability of data across multiple ERP modules for overall operations viewpoint are solved [14].

After pre-processing, the moves to model selection training are next to be discussed. Criteria on selecting the most suitable machine learning algorithms for this study are determined by the nature of the anomalous situations to be identified, the characteristics of the data and the training processes that focus on normal and anomalous conditions 15]. The last step is to continually screening for anomalies and assessing their validity in the use of a real-time model. Measuring Malaynization involves quantitative methods like productivity of Malay, ratio of medium and percentage of population using the Malay language, as well as ethnographic measurements like Malay's modal number, Malay language proficiency, Malay educated people and Malay Job positions. Further, constant updating of the model and applying feedback and learning mechanism in the next passes guarantees that the refined model will fit the dynamic ERP settings. Last of all, the integration of the detection framework with business process guarantees response to detected anomalies when business risks are managed effectively [16].

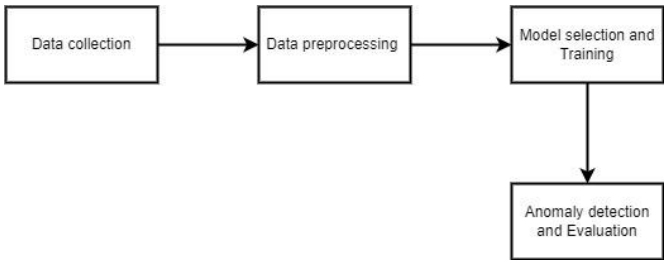

Fig. 1. ERP anomaly detection framework

IV. Related work

Parimi [17] investigated the use of deep learning of auto encoders, Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) to get better accuracy and efficiency of anomaly detection within SAP systems. Thus, this study's contribution is the systematic and extensive evaluation of deep learning architectures for financial anomaly detection, with solutions to issues, including interpretability and scalability and integrating real-time transaction monitoring. Thus, the findings of this paper offer a recital comparison of these techniques against conventional approaches to assess the efficiency of fraud risk reductions from SAP financial systems and overall enhanced operational OP3 resilience.

Recently, *Kohli* [18] developed a model which might be used for forecast of equipment failure relied on the data from SAP plant safeguarding module. Through the application of unsupervised learning technique of the clustering kind the author was able to record a class to cluster assessment accuracy of about 80%. Later that classifier model again developed with the help of different ML algorithms and was cross validated on two different datasets with the goal of predicting the equipment failure. In the case of the classifier model developed using ML algorithms like SVM and DT, the accuracy level and TPR to predict equipment failure was more than 95%. The proposed model forms the part of the AIMMS – Advanced Intelligent Control system needed in CPS for asset intensive organizations. *Parimi* [19] proposed the application of machine learning approaches for enhancing SAP systems in reporting and compliance. The study advances the knowledge in the area by describing the use of the techniques of supervised and unsupervised learning in fraud detection, the methods of anomaly detection for detecting any anomalies in financial data and the use of the techniques of predictive analytics for predicting financial results. Furthermore, it examines the combination of AI and Blockchain and looks at the implications as they affect the area of increased transparency and security in financial reporting. To the best of the authors' knowledge, *Župan et al.* [21] provided indications of applying semi-supervised learning opportunities in accountancy. One of the latest deep learning algorithms is applied for

reconstructing the journal entry key columns out of the above list. The model was trained and tested on a real-world dataset so it could be used as base for develop the great number of accounting and audit applications – as a module of detecting anomalies in the ERP software or as an independent application.

Mantere [23] developed a proof-of-concept (PoC). For testing of the selected approach and tools the PoC implementation is based on the Bro network security monitoring framework (Bro). In the PoC, the SOM algorithm is implemented in the Bro scripting language to showcase that Bro can serve as the base system. The implemented approach also corresponds to the minimal form of another concept developed during the research event-driven machine learning anomaly detection (EMLAD).

A RFM analysis and machine learning algorithms were used for combining the churn prediction based on mainly transactional data as it was suggested by *Aleksandrova* [24]. The data used in the present study is collected from the ERP system of a regional concrete production firm based in Bulgaria. That is why several machine learning algorithms have been applied to the given data set namely, Two-Class Boosted Decision Trees, Two-Class Neural Networks, Two-Class Decision Jungle, Two-Class SVM and Two-Class Logistic Regression. All of the experiments described were conducted in Azure Machine Learning Studio tools. The study found out that while RFM scores and metrics have limitations companies can sufficiently forecast the churn of their customers when machine learning algorithms are deployed.

The simulation tests for the evaluation of the proposed framework are carried out by *Wang* [26] as well as he deploying a prototype. The results of the tests prove high efficiency and high effectiveness of the in-memory columnar database in comparison with a traditional ERP system in terms of computational time and the amount of memory required. Moreover, it demonstrates the effective application of the in-memory columnar database to the cloud for use in continuous audit analytics. The third essay proposes a design of a fraud detection system with a recurrent architecture using changes in the deep learning technology.

Table 1.Comparison of existing techniques used in literature

Study	Techniques Used	Purpose	Results/Performance	Key Contribution
Parimi [17]	Deep learning (Autoencoders, RNN, CNN)	Anomaly detection in SAP systems	Improved accuracy and efficiency in detecting anomalies	Systematic evaluation of deep learning for financial anomaly detection
Kohli [18]	Unsupervised learning (Clustering), ML algorithms (SVM, DT)	Predicting equipment failure	Class-to-cluster accuracy of 80%; SVM and DT achieved >95% accuracy and TPR	Developed a model for predictive maintenance in SAP Plant Maintenance
Parimi [19]	Supervised & unsupervised learning, AI, Blockchain	Fraud detection, predictive analytics	Improved transparency and security in financial reporting	Explores AI and Blockchain integration in ERP systems
Kroll et al. [20]	Timed hybrid automata	Predictive maintenance for industrial plants	Enhanced anomaly detection through real-time monitoring	Combines anomaly detection and data acquisition in one process
Župan et	Semi-supervised learning	Anomaly detecti	Applied deep learning to	Semi-supervised

al. [21]		on in ERP accounting	reconstruct journal entries	learning for ERP financial anomaly detection
Bhattacharyya [22]	Neural networks, pattern recognition	Cloud resource anomaly detection	Achieved 98.3% accuracy in anomaly detection	Real-time cloud resource monitoring with feedback integration
Mantere [23]	Self-Organizing Map (SOM)	Network security anomaly detection	Proof of concept for Bro framework using SOM	Event-driven machine learning anomaly detection concept
Aleksandrova [24]	RFM analysis, ML (SVM, Neural Networks, Decision Trees)	Churn prediction based on transactional data	Accurate customer churn prediction using ML algorithms	Combining RFM analysis with ML for customer churn prediction
Ahmed et al. [25]	Clustering-based techniques	Anomaly detection in ERP systems	Compared various clustering techniques	Survey of clustering-based anomaly detection methods
Shanthamallu et al. [27]	Machine Learning algorithms	General anomaly detection applications	Review of ML applications in various fields	Comprehensive survey of ML algorithms and software tools

A summary of the comparison made in this paper between the techniques that have been proposed by other authors with respect to the use of analyses for the detection of anomalies in ERP systems, is embodied in Table 1. It includes the name of the studies; methods used in them (for example, deep learning, clustering, hybrid automata); the goal of the studies (for example, fraud detection, predictive maintenance); measures or performance that were achieved during the study (for example, accuracy, true positive rate); and contributions of the studies.

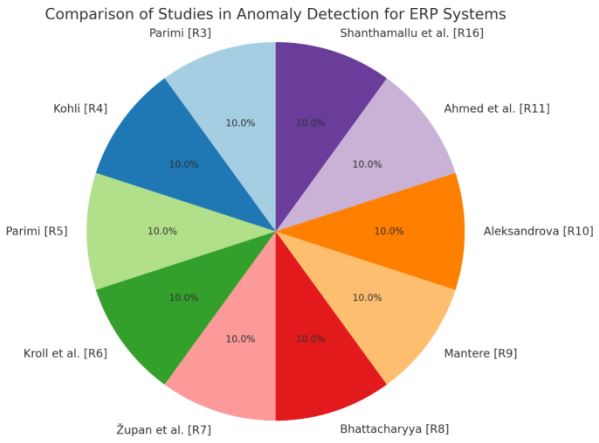

Fig. 2. Comparison of studies used in literature

Furthermore, figure 2 also reveals the overall comparison of the studies used in the literature review in the present study. Here, previous approaches and methods adopted across various studies for detecting anomaly in ERP systems are enumerated. In this process, the approaches adopted by each study are analyzed as per the kind of AI/ML employed, special area of interest of the study including financial anomaly detection or predictive maintenance, and the outcomes or principal contributions to the field.

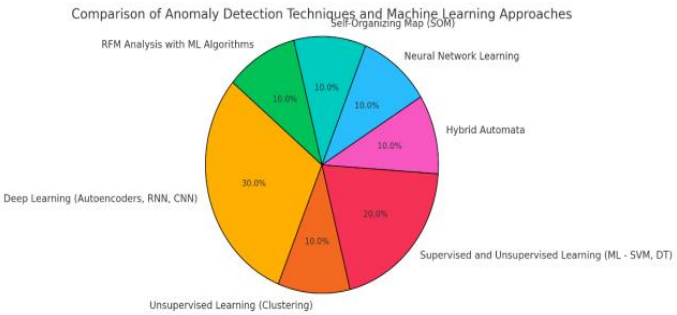

Fig. 3.Comparison of techniques with respect to ML techniques utilized by authors

Figure 3 presents the comparison of techniques used in the reviewed study with regards to the machine learning method used. It therefore divides them according to the types of ML algorithms used and these include; Supervised learning, unsupervised learning and the hybrid learning algorithms.

V. Future Directions in AI-Driven Anomaly Detection for ERP Systems

As AI and machine learning technologies continue to evolve, the future of anomaly detection in ERP systems is promising and multifaceted. Several emerging trends and methodologies can significantly enhance the capabilities of these systems:

i) Deep Learning: Models like CNNs and LSTMs are highly effective at identifying complex patterns in large datasets, particularly useful for detecting fraud and operational anomalies. Transfer learning can accelerate the deployment of these solutions across ERP modules.
ii) Reinforcement Learning (RL): RL dynamically adapts to changing data patterns, continuously refining detection criteria in real-time. This approach minimizes false positives and allows proactive adjustments to evolving threats and operational shifts.
iii) Hybrid Approaches: Combining machine learning techniques—such as unsupervised clustering for initial detection followed by supervised classifiers for detailed analysis—improves accuracy and reduces false positives. Ensemble methods can further enhance robustness.
iv) Explainable AI (XAI): XAI ensures transparency in anomaly detection, helping stakeholders understand why certain actions were flagged, which is crucial for compliance, user trust, and refining detection algorithms.
v) IoT Data Integration: As ERP systems incorporate IoT devices, real-time IoT data can be analyzed alongside traditional data to improve anomaly detection, identifying issues from sensor malfunctions or unusual operational conditions.
vi) Continuous Learning: Continuous learning algorithms adapt to evolving data streams without full retraining; ensuring anomaly detection remains relevant over time, particularly through techniques like online learning.
vii) Enhanced Data Governance and Privacy: Privacy-preserving techniques like federated learning allow models to be trained across decentralized data sources, ensuring privacy while maintaining effective anomaly detection in compliance with regulations.
viii) Real-Time Analytics and Decision Support: The integration of real-time analytics and decision support systems enables immediate insights into anomalies, allowing faster decision-making and risk mitigation, supported by dashboards and visualization tools for clarity.

VI. Conclusion

Therefore, the research establishes that the application of AI and ML in the context of anomaly detection improves resilience in ERP systems significantly, in relative to rule-based models that are rigid in their approach. The ability to separate, highly analyze, and make decisions based on past data, an ability to recognize multiple-interconnected components and find new and known deviations from norms make ML especially valuable in ERP settings with highly fluid data, users, and operations. Real-time anomaly detection coupled with automation means AI assists organizations in keeping the ERP system reliable, accurate, and secure. Sustaining mechanisms like continuous learning mechanisms, addressing techniques like deep learning, clustering, and usage of predictive analytics make sure that the ERP system stays nimble to business changes and offer strong support for the decision-making process and fraud detection.

VII. REFERENCES

[1]. Schreyer, Marco, Timur Sattarov, Damian Borth, Andreas Dengel, and Bernd Reimer. "Detection of anomalies in large scale accounting data using deep autoencoder networks." arXiv preprint arXiv:1709.05254 (2017).

[2]. Bergdahl, Jacob. "The AI Revolution: A study on the present and future application and value of AI in the context of ERP systems." (2018).

[3]. Alarifi, Suaad S., and Stephen D. Wolthusen. "Detecting anomalies in IaaS environments through virtual machine host system call analysis." In 2012 International Conference for Internet Technology and Secured Transactions, pp. 211-218. IEEE, 2012.

[4]. Kumar, Rohit. Machine learning and cognition in enterprises: business intelligence transformed. Apress, 2017.

[5]. Barta, Gergő. "The increasing role of IT auditors in financial audit: risks and intelligent answers." Business, Management and Education 16, no. 1 (2018): 81-93.

[6]. Arachchi, Samantha Mathara, Siong Choy Chong, and A. G. I. Madhushani. "Quality assurance and quality control in ERP systems implementation." American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS) 11, no. 1 (2015): 70-83.

[7]. Qiu, Judy, Bo Peng, Ravi Teja, Sahil Tyagi, Chathura Widanage, and Jon Koskey. "Real-Time Anomaly Detection from Edge to HPC-Cloud."

[8]. Mohamudally, Nawaz, and Mahejabeen Peermamode-Mohaboob. "Building an anomaly detection engine (ADE) for IoT smart applications." Procedia computer science 134 (2018): 10-17.

[9]. Tanaka, Hiroki, Hiroki Watanabe, Hayato Maki, Sakriani Sakti, and Satoshi Nakamura. "Single-trial detection of semantic anomalies from EEG during listening to spoken sentences." In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 977-980. IEEE, 2018.

[10]. Massaro, Alessandro, Angelo Calicchio, Vincenzo Maritati, Angelo Galiano, Vitangelo Birardi, Leonardo Pellicani, M. Gutierrez Millan et al. "A case study of innovation of an information communication system and upgrade of the knowledge base in industry by ESB, artificial intelligence, and big data system integration." International Journal of Artificial Intelligence and Applications (IJAIA) 9, no. 5 (2018): 27-43.

[11]. Sumaiya, P. "Enhancing user experience using machine learning." Int. J. Eng. Res. Technol.(IJERT) 7, no. 2 (2018): 353-358.

[12]. Halder, Soma, and Sinan Ozdemir. Hands-On Machine Learning for Cybersecurity: Safeguard your system by making your machines intelligent using the Python ecosystem. Packt Publishing Ltd, 2018.

[13]. Shin, Hyun-Jun, Kyoung-Woo Cho, and Chang-Heon Oh. "SVM-based dynamic reconfiguration CPS for manufacturing system in industry 4.0." Wireless Communications and Mobile Computing 2018, no. 1 (2018): 5795037.

[14]. Lamba, Anil, Satinderjeet Singh, Singh Balvinder, Natasha Dutta, and Sivakumar Rela. "Deriving intelligent data analytics using anomaly detection framework for IoT network and smart environments." International Journal For Technological Research In Engineering 4, no. 6 (2017).

[15]. Katona, A., Peter Panfilov, and B. Katalinic. "Building predictive maintenance framework for smart environment application systems." In Proceedings of the 29th DAAAM international symposium, pp. 0460-0470. 2018.

[16]. Pavaloiu, Alice. "The impact of artificial intelligence on global trends." Journal of Multidisciplinary Developments 1, no. 1 (2016): 21-37.

[17]. Parimi, Surya Sai Ram. "Leveraging Deep Learning for Anomaly Detection in SAP Financial Transactions." TIJER-TIJERINTERNATIONAL RESEARCH JOURNAL (www. TIJER. org), ISSN (2017): 2349-9249.

[18]. Kohli, Manu. "Using machine learning algorithms on data residing in SAP ERP application to predict equipment failures." International Journal of Engineering & Technology 7, no. 2.28 (2017): 312-319.

[19]. Parimi, Surya Sai Ram. "Optimizing Financial Reporting and Compliance in SAP with Machine Learning Techniques." TIJER-TIJERINTERNATIONAL RESEARCH JOURNAL (www. TIJER. org), ISSN (2018): 2349-9249.

[20]. Kroll, Björn, David Schaffranek, Sebastian Schriegel, and Oliver Niggemann. "System modeling based on machine learning for anomaly detection and predictive maintenance in industrial plants." In Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), pp. 1-7. IEEE, 2014.

[21]. Župan, Mario, Svjetlana Letinić, and Verica Budimir. "Journal entries with deep learning model." International Journal of Advance Computational Engineering and Networking (IJACEN) 6, no. 10 (2018): 55-58.

[22]. Bhattacharyya, Arnamoy, Seyed Ali Jokar Jandaghi, Stelios Sotiriadis, and Cristiana Amza. "Semantic aware online detection of resource anomalies on the cloud." In 2016 IEEE international conference on cloud computing technology and science (CloudCom), pp. 134-143. IEEE, 2016.

[23]. Mantere, Matti. "Network security monitoring and anomaly detection in industrial control system networks." (2015).

[24]. Aleksandrova, Yanka. "Application of machine learning for churn prediction based on transactional data (RFM analysis)." In 18 International Multidisciplinary Scientific Geoconference SGEM 2018: Conference Proceedings, vol. 18, no. 2.1, pp. 125-132. 2018.

[25]. Ahmed, Mohiuddin, Abdun Naser Mahmood, and Md Rafiqul Islam. "A survey of anomaly detection techniques in financial domain." Future Generation Computer Systems 55 (2016): 278-288.

[26]. Wang, Yunsen. "Designing continuous audit analytics and fraud prevention systems using emerging technologies." PhD diss., Rutgers University-Graduate School-Newark, 2018.

[27]. Shanthamallu, Uday Shankar, Andreas Spanias, Cihan Tepedelenlioglu, and Mike Stanley. "A brief survey of machine learning methods and their sensor and IoT applications." In 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA), pp. 1-8. Ieee, 2017.