

Pinto, Sarah Oliveira; Sobreiro, Vinicius Amorim

Article

Literature review: Anomaly detection approaches on digital business financial systems

Digital Business

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Pinto, Sarah Oliveira; Sobreiro, Vinicius Amorim (2022) : Literature review: Anomaly detection approaches on digital business financial systems, Digital Business, ISSN 2666-9544, Elsevier, Amsterdam, Vol. 2, Iss. 2, pp. 1-22, <https://doi.org/10.1016/j.digbus.2022.100038>

This Version is available at:

<https://hdl.handle.net/10419/337914>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>



Literature review: Anomaly detection approaches on digital business financial systems

Sarah Oliveira Pinto^{*}, Vinicius Amorim Sobreiro

Department of Management, University of Brasília, Campus Darcy Ribeiro, Brasília, Federal District 70910-900, Brazil

ARTICLE INFO

Keywords:

Anomaly detection
 Fraud detection
 Outlier detection
 Financial systems
 Accounting
 Systematic literature review

ABSTRACT

Anomaly detection approaches have become critically important to enhance decision-making systems, especially regarding the process of risk reduction in the economic performance of an organisation and the consumer costs. Previous studies on anomaly detection have examined mainly abnormalities that translate into fraud, such as fraudulent credit card transactions or fraud in insurance systems. However, anomalies represent irregularities in system patterns data, which may arise from deviations, adulterations or inconsistencies. Further, its study encompasses not only fraud, but also any behavioural abnormalities that signal risks. This paper proposes a literature review of methods and techniques to detect anomalies on diverse financial systems using a five-step technique. In our proposed method, we created a classification framework using codes to systematize the main techniques and knowledge on the subject, in addition to identifying research opportunities. Furthermore, the statistical results show several research gaps, among which three main ones should be explored for developing this area: a common database, tests with different dimensional sizes of data and indicators of the detection models' effectiveness. Therefore, the proposed framework is pertinent to comprehending an existing scientific knowledge base and signals important gaps for a research agenda considering the topic of anomalies in financial systems.

1. Introduction

The large volume of online financial transactions and the high amount of corresponding records makes it impossible for human specialists to identify anomalies present in financial systems (Gianini et al., 2020, p. 549) and to distinguish these anomalies from a random/systematic error which may not be directly related to financial operations and products and, most of the times, it can only be identified through validation tests (Carneiro, Figueira, & Costa, 2017, p. 94; Climent, Mompalmer, & Carmona, 2019, p. 888). Moreover, this becomes increasingly complex with the digital transformation that constantly changes services, businesses and, consequently, people's lives (Hanafizadeh & Kim, 2020, p. 1). This was even more evident during the COVID-19 pandemic in 2020, which accelerated and transformed services, and incorporated digital innovations into various businesses' strategies (Shaikh, Sharma, & Karjaluo, 2020, p. 1).

This era of digitisation and transformation of organisations within the so-called fourth Industrial Revolution encourages the use of digital technologies to improve the way processes are done in order to monitor,

diagnose, predict, and optimize data and performance (Karki & Porras, 2021, p. 2). This establishes a new ecosystem in which digital business models become pillars of the Industrial 4.0 (Shaikh et al., 2020, p. 1), due to the increasing arrival and use of emerging technologies such as big data, cloud computing, Internet of Things, artificial intelligence, blockchain, and cryptocurrencies (Hanafizadeh & Kim, 2020, p. 1).

However, human needs and demands move advances towards technological development (Karki & Porras, 2021, p. 1). In the context of digital transactions, value creation lies in combining new structures and new forms of governance that are able to connect information with people (Marikyan, Papagiannidis, Rana, & Ranjan, 2022, p. 2). As much as digital technologies are important in the digital transformation process, a combination of technologies with a business logic that can actually insert a business and its form of management into the digitalisation is necessary (Hanafizadeh & Kim, 2020, p. 1). Based on this, the digitisation of financial business and processes using new technologies enables companies' and individuals' optimized and constant access to financial tools (Bollaert, de Silanes, & Schwiendbacher, 2021, p. 1), which can help the decision-making process in financial and operational

^{*} Corresponding author.

E-mail addresses: sarah.oliveira.pinto@gmail.com (S.O. Pinto), sobreiro@unb.br (V.A. Sobreiro).

terms, seeking to prevent and detect financial anomalies.

However, this proximity of contact with online financial systems also enables individuals to carry out transactional tasks with ease (Timmons, McGowan, & Lunn, 2019, p. 161), such as payments, bank transfers, purchasing insurance, buying and selling financial assets or withdrawing from social security. The increasing adoption and evolution of modern financial technologies and IT-applications, as well as the development of artificial intelligence and automated innovations (Cho & Chen, 2021, p. 2), in daily activities creates fast evolving vulnerabilities and exposes systems to cyberthreats and malicious attacks (Wessels, van den Brink, Verburgh, Cadet, & van Ruijven, 2021, p. 1). Thus, as much as financial technologies make financial services and processes more efficient through the application of new forms of technology (Thompson, 2017, p. 152), the flexibility arising from these online systems reveals a weakness that contrasts with the growing need for fraud and anomaly detection systems (Gianini et al., 2020, p. 549).

The development and spread of personal computers, tablets and mobile phones exert a great influence on consumers' lives (Albayati, Kim, & Rho, 2020, p. 1) and allows this financial vulnerability the digitisation of transactions causes, especially considering that, since the 2000s, the number of cell phone users has grown exorbitantly (Zaffar, Kumar, & Zhao, 2019, p. 62; Verbeke, Dejaeger, Martens, Hur, & Baesens, 2012, p. 211) and mobile applications are widely used around the world (Shaikh et al., 2020, p. 3; Sokolova, Perez, & Lemerrier, 2017, p. 62), thus lowering barriers (Shaikh et al., 2020, p. 1) and enabling users to verify balances, deposit checks, and carry out various transactions (Zaffar et al., 2019, p. 62). The so-called mobile money, a financial technology payment service, has changed the way cell phones are used not limited to communication, entertainment, and internet browsing because it is no longer necessary to have direct contact or cash to conduct financial transactions. In addition, services can be accessed anywhere and anytime (Daragmeh, Lentner, & Sági, 2021, p. 1). This has resulted in several financial platforms and services being launched worldwide, resulting in more than 556 million registered mobile money accounts (GSMA, 2017, p. 17).

In a June 2021 study of the digital platform BankMyCell, Turner (2021, p. 1) indicated that the number of smartphone users around the world comprises 48.20% of the entire population. As a phone format that has advanced features, an intelligent operating system and several applications, the smartphone allows connecting data networks with the Internet. In addition to that, Handley (2019, p. 1) produced a report, by the World Advertising Research Center, signalling that approximately 51% of all smartphone users access the Internet only through their own devices, and that by 2025, 72.6% of general Internet users will adhere to the Internet access via smartphones alone, totalling the equivalent of approximately 3.7 billion people. For this reason, the convenience of being able to use the internet easily enhances social interactions, online financial transactions and increased productivity. However, this significant use raises concerns about the security of data and personal information, mainly of financial content (Jones & Chin, 2015, p. 561).

Given this significant increase in interactions, both private organisations' and government agencies' use of e-commerce platforms had great growth in the last few decades (Misra, Thakur, Ghosh, & Saha, 2020, p. 254). As a result, these platforms made it even more possible to connect the ease of accessing the Internet through smartphones with the option of carrying out financial transactions virtually. Thus, the growing popularity of these new emerging technologies creates opportunities for fraudsters to cause financial loss (Misra et al., 2020, p. 254). Therefore, smartphones may have been facilitators of fraud and data misuse, and the understanding and identification of anomalies caused as a result may be of fundamental importance for creating a strategic decision-making system, considering that identification methods of anomalies enable the learning of data network characteristics with the strategic purpose of distinguishing abnormalities (Stripling, Baesens, Chizi, & van den Broucke, 2018, p. 15), and considering how these approaches become relevant theoretical foundations for estimating probabilities and

predicting fraud, validating a more structured decision support system (Serrano-Silva, Villuendas-Rey, & Yáñez-Márquez, 2018, p. 78).

The development of an efficient system for detecting fraud and various anomalies is also important for reducing losses in financial institutions (Höppner, Baesens, Verbeke, & Verdonck, 2022, p. 291), since well-defined business management models in the financial area enable organisations to benefit from their internal strengths, as well as monitor and seize opportunities for improvement (Hanafizadeh & Marjaie, 2021, p. 1). As much as financial companies in general, often encouraged by law, are more likely to employ processes and systems that ensure their data are authentic and high quality (Elsaify & Hasan, 2021, p. 2), with the increasing entry of the financial sector into digital business, it is necessary to establish stricter mechanisms for accessing and controlling this data, in order to review each point and each transaction to make certain fraud behaviours do not arise (Guo & Liang, 2016, p. 11). Through smarter approaches, the creation of regulatory systems that consider the access, analysis, and processing of data in real time is possible (Barroso & Laborda, 2022, p. 13); therefore, gathering information about existing techniques in the scientific knowledge base is of paramount importance in enabling researchers and managers to identify the best techniques according to the scenarios in which they find themselves and to explore their capabilities for obtaining value and mitigating risk (Konopik, Jahn, Schuster, Hoßbach, & Pflaum, 2022, p. 3).

Large deviations called anomalies or abnormalities, in detail, represent inconsistencies based on economic paradigms or behavioural expectations (Parnes, 2020, p. 292). Hence, these anomalies can be defined as a data set that behave according to patterns different from those identified as normal or typical, thus being considered exceptions or peculiarities (León, 2020, p. 1; Rubio, Barucca, Gage, Arroyo, & Morales-Resendiz, 2020, p. 2; Stripling et al., 2018, p. 15; Wu & Zhang, 2006, p. 375). These data outliers can be portrayed as frauds because they represent a criminal change in a data group's behaviour (Pourhabibi, Ong, Kam, & Boo, 2020, p. 7). Regarding this, anomaly detection methods are important data mining processes for identifying trends and patterns in diverse sets of financial data (Parnes, 2020, p. 308) and the development of research that can improve such methods is of paramount importance to reducing financial losses, since organisations increasingly rely on data to make operational and strategic decisions (Elsaify & Hasan, 2021, p. 2). Additionally, for optimal development of their business models, managers should be aware of the warning signals that may appear in order to take the necessary actions (Koyuncugil & Ozgulbas, 2012, pp. 6238–6239), implementing the techniques and tools that can further assist in this process (Barroso & Laborda, 2022, p. 11).

Given the current scenario, the importance of identifying anomalies and/or frauds in the most diverse areas of knowledge is highlighted, considering this process makes it possible to face situations of irregularities or abnormalities arising from deviations or the adulteration of data. Although this issue has been extremely important in recent years, it is not recent and has been discussed quite regularly in academic circles. For instance, in the area of medicine, Chen et al. (1999, p. 123) studied methods of detecting anomalies in the human body, which constituted pathologies; Duftschmid and Miksch (2001, p. 39) addressed the topic of clinical guideline errors, which could result in application errors; and Smith et al. (2004, p. 458) studied visual identification of neurological problems.

Also as an example, in the area of computer systems, Waite (1999, p. 221) aimed to distinguish abnormal data performance with probabilistic techniques, while Saleh et al. (2001, p. 979) analysed synchronization anomalies in software development, and the study of Perdisci et al. (2009, p. 864) focused on using anomaly detection methods to prevent attacks on computer networks. In the area of network systems, Muniyandi et al. (2012, p. 174) and Wu and Zhang (2006, p. 376) discussed anomalies to identify interferences and intrusions. Furthermore, Yasami and Safaei (2017, p. 52), Kaur and Singh (2016, p. 213) and

Savage et al. (2014, p. 62) researched about identifying anomalies on social networks in order to find abnormal activities. In a more recent perspective, Ardizzi et al. (2018, p. 105), Singh and Best (2019, p. 2) and Rtayli and Enneya (2020, p. 13) addressed techniques for detecting anomalies in financial systems, exploring issues such as outliers in payment systems, evidence of money laundering and fraud in financial transactions. Not surprisingly, this shows how this theme remains recurrent because it is important and characterizes a research trend.

Among the articles analysed on this topic, there are unresolved or not considered problems in previous research, which may be relevant to this article. More precisely, considering recent literature review articles, the chosen analysis methods consisted of reviewing and examining data mining techniques for detecting financial fraud and classifying them according to the types of fraud (Al-Hashedi & Magalingam, 2021, p. 5), identifying and analysing researches using Graph-Based Anomaly Detection (GBAD) techniques in the context of financial fraud detection (Pourhabibi et al., 2020, p. 1) and exploring advantages and disadvantages of methods that may reveal fraudulent and dishonest behaviour in organisations (Houdek, 2020, p. 1), as shown below:

It reveals the most frequent financial fraud detection technique used in data mining that can result in disclosing fraud in financial domains with high-performance accuracy and provides a new classification framework for financial fraud as well as creating a roadmap for researchers and practitioners seeking to better understand this field.

Al-Hashedi and Magalingam (2021, p. 3).

We focus on research undertaken between 2007 and 2018 to provide a synthesized understanding of the state-of-the-art GBAD methods, identify key research issues raised against application contexts, and establish future directions to expand GBAD research in fraud detection.

Pourhabibi et al. (2020, p. 2).

This article supplies an interdisciplinary and multilevel overview of methods how to identify (not only) organizational misdeeds, the pros and cons of the use of each individual method, and the limitations and strengths of their results based on cherry-picked current studies.

Houdek (2020, p. 1).

Although there are such systematizations of literature, the approach used in all of them is mainly aimed at detecting financial fraud, not specifically considering the identification of anomalies present in the systems, which may or may not configure such frauds. For example, they may indicate illegal money laundering financial transactions (Ardizzi et al., 2018, p. 106) through abnormal behaviour patterns in the data. Observing this context, *the objective of this literature review is to systematize the knowledge about anomaly detection in the financial sector in order to consolidate the main techniques used in this area and identify existing gaps.* To carry out this research, we used the five-step technique by Jabbour (2013, p. 145) and Lage Junior and Godinho Filho (2010, p. 14), later successfully disseminated by other papers (Amui, Jabbour, de Sousa Jabbour, & Kannan, 2017, p. 311; Barboza, Kimura, Sobreiro, & Basso, 2016, pp. 328–329; Ferraz, Falguera, Mariano, & Hartmann, 2021, p. 7; Singh et al., 2019, p. 164; Stefanelli, Teixeira, Oliveira, Ferreira, & Sehnem, 2019, pp. 2050–2051; Talan & Sharma, 2019, p. 355). In sum, these authors recommended classifying and coding characteristics related to relevant articles; building brief summaries of the contributions discussed; analysing and describing strengths, weaknesses and methods addressed by each one; creating a framework of analysed articles; and providing an agenda for future research pertinent to the topic.

Thus, this systematization can contribute to identifying relevant techniques and results for this area of study from a decision maker's perspective because the management of decision support systems, which help in predicting risks and fraud through anomaly detection, has a direct effect on an organisation's economic performance (Severino & Peng, 2021, p. 2). Based on this, a manager needs to choose the techniques and the most relevant technological intervention parameters to

be applied to its systems, so that the financial management can be done in a more efficiently way through robust and coherent risk metrics (Byers, Popova, & Simkins, 2021, p. 1) that ensure the continuity of its processes, reinforcing cybersecurity (Wessels et al., 2021, p. 1) and preventing the organisation from being the target of illegal acts or errors and transactional deviations and, thus, enabling the corporate survival (Konopik et al., 2022, p. 1).

Hence, in broad terms, the five-step technique used in this paper can indicate financial anomaly detection approaches by correlating them to different factors according to the predefined classifications that will drive the descriptive analysis (Cabral & Dhar, 2019b, p. 2245), so future researchers will be able to find a database and statistics that inspire them to develop their own studies (Mallinguh & Zeman, 2020, p. 163) and to improve further the methods of detecting anomalies within the financial arena. In contrast, managers of organisations will be able to make more reasonable choices of techniques that enable greater interpretability according to the scenarios in which they find themselves, in order to optimize the decision-making process (Sá, Pereira, & Pappa, 2018, p. 26). This will benefit not only researchers and managers, but also financial systems in general, since a more precise allocation of methods in specific systems will be available to connect the dynamic capabilities of detection techniques with the process of digital transformation, defining opportunities and transforming available resources (Konopik et al., 2022, p. 2), which can result in business models that deliver financial services more efficiently (Barroso & Laborda, 2022, p. 1).

In addition to this introduction, the rest of the article is structured as follows: a detailed presentation of the research method used is presented in Section 2. Next, in Section 3 the classification/coding method used in the relevant articles is indicated. Section 4 is structured to present the results obtained and discussions relevant to the present literature review. Finally, the conclusions of this work as well as guidelines for future research are presented in Section 5.

2. Research method

This study presents a systematic literature review, which provides a comprehensive view of an area of research (Silva et al., 2020, p. 2) through identifying relevant articles based on selection criteria (Zhang & Benyoucef, 2016, p. 96) and synthesizing data, queries, and research techniques (Ain, Vaia, DeLone, & Waheed, 2019, p. 2). It allows for an in-depth analysis of scientific knowledge as it selects a dominant database in which research is performed, using a combination of keywords related to the topic of study, and reading relevant articles (Hosseini & Ivanov, 2020, p. 2). This is done to identify peer-reviewed articles that can help the scientific decision-making process (Rajaian, Cater-Steel, & Lane, 2017, p. 45) based on recognizing research questions and research agenda suggestions for future investigations within the relevant theme (Wu, Ngai, Wu, & Wu, 2020, p. 1).

Several authors (Ahmed, Deokar, & Lee, 2021, p. 3; Henrique, Sobreiro, & Kimura, 2019, p. 227; Amani & Fadlalla, 2017, p. 34; Seuring, 2013, p. 1514) have already excellently used this method and it has proven to be advantageous for the scientific community. Within the theme of anomaly detection, this research technique was recently used by authors such as Al-Hashedi and Magalingam (2021, p. 6), Pourhabibi et al. (2020, p. 1) and Houdek (2020, p. 1), considering mainly anomalies that signal the existence of fraud. Based on this, the present study aims to complement these previous researches. Furthermore, its focus on identifying techniques for detecting anomalies on digital business financial systems that also manifest themselves in other ways in addition to fraud distinguishes this work.

Thus, the literature review model to be developed was proposed by Jabbour (2013, p. 145) and Lage Junior and Godinho Filho (2010, p. 14) and follows the five steps briefly outlined below:

- First Step: Conduct a systematic search of available articles in a reliable database of variations of techniques used to detect anomalies.
- Second Step: Develop a classification and coding framework.
- Third Step: Apply the classification framework to synthesize and present study trends on anomaly detection.
- Fourth Step: Outline a framework portraying the composition of the literature review, its main results and what it represents according to the pre-established coding.
- Fifth Step: Analyse and discuss the results and, based on that, propose suggestions for future research.

This literature review technique has already been used and disseminated in several papers published in international journals (Henriques, Sobreiro, & Kimura, 2018, p. 145; Lee et al., 2021, p. 68415; Mariano, Sobreiro, & Do Nascimento Rebelatto, 2015, pp. 34–35; Masudin & Fernanda, 2019, p. 2; Nazário, Silva, Sobreiro, & Kimura, 2017, pp. 116–117; Salim, Rahman, & Wahab, 2019, p. 1449). As scientific studies seek to describe, explain, predict, or evaluate phenomena (Abreu, Kimura, & Sobreiro, 2019, p. 197), this review method makes tracking the evolution and main contributions within the field of study possible (Cruz, Kimura, & Sobreiro, 2019, p.80) by mapping and identifying the main characteristics and topics of articles (Pinheiro et al., 2019, p. 844) and by grouping them for deeper exploration of the analysis theme (Ferreira, Sobreiro, Kimura, & Barboza, 2016, p. 114), creating a complete view of the existing knowledge (Henriques, Sobreiro, Kimura, & Mariano, 2020, p. 6). This is done through a search for and selection of articles using a combination of keywords that allows for increasing the scope and reach of the research (Jabbour et al., 2017, p. 291) and by elaborating a classification framework that proposes guidelines for categorization of the selected articles (Cabral & Dhar, 2019a, p. 26).

From that, this technique permits the content and characteristics of the papers to be broken down and allocated according to the pre-established encodings, creating a new perspective of literary analysis through the statistical data generated (Silva, Kimura, & Sobreiro, 2017, p. 92). Based on such data, this literature review model will uncover research trends in not only an isolated way, but also by considering combinations of several factors that involve the results of studies about the detection of anomalies in the digital financial sector, proving adequate to answer research questions (Pham, Hoang, & Phan, 2019, p. 848) and reveal gaps for additional investigation (Bajaj, Kashiramka, & Singh, 2020, p. 176). Such research gaps will allow the critical consideration of challenges and guide future research directions within the study area (Abu et al., 2021, p. 755; Khudzari, Kurian, Tartakovsky, & Raghavan, 2018, p. 52).

The benefits that can come from this can be found both in the academic sphere and in the financial sphere, as this method allows the identification, evaluation and synthesis of existing works produced by researchers, scholars, and practitioners (Okoli, 2015, p. 880). The resulting graphical abstract can function as a visual tool for future researchers of this theme to use as a basis or starting point to develop more targeted or pertinent research in the area, as it represents a creative, concise and impactful way of communicating scientific information (Ojha, 2021, p. 218). Subsequently, these studies can have a positive influence on financial systems that use these more accurate models, because powerful fraud detection systems can not only detect but also prevent anomalies, reducing possible financial losses (Höppner et al., 2022, p. 291) and enabling a higher protection of the data (Elsaify & Hasan, 2021, p. 2), which enables the creation of healthy environments for the development and transformation of the digital business ecosystem (Hanafizadeh & Kim, 2020, p. 1).

Considering this, the first step of this research was conducted between June and July of 2021. A meticulous search was conducted using keywords referring to the topic of study, contained in the titles and abstracts of the articles to be chosen and analysed. To carry out such research, we chose the ScienceDirect® database, operated by Elsevier®.

As of now, this database has a portfolio of 2650 journals with more than 24,000 editors (Plc, 2021, p. 14), subdivided into the areas of Health Sciences, Life Sciences, Physical Sciences and Engineering and Social Sciences and Humanities. The database provides international access to scientific publications in different areas. ScienceDirect distributes and allows access to scientific contents and integrates researchers with peer-reviewed primary scientific and medical researches (Plc, 2021, p. 14). This is done by accessing scientific articles based on keywords, author name, journal/book title and other advanced types of research.

Therefore, in a more detailed way, to carry out this literature review, the keywords included “anomaly detection finance”, “outlier detection finance”, “anomaly detection economics”, “anomaly detection financial fraud”, “anomalous activity finance”, and other variations. Also, words were used in the advanced search section using the boolean marker “OR” to increase the chances of the search success, such as, “financial”, “finance”, “fraud”, “machine learning”, “anomaly detection”, and “detection”. To further narrow the search in the article database, filters related to years of publication were used, with articles published from between 2017 and 2021 being chosen only, aiming at a more current literature review. Moreover, because this literature review is about detecting anomalies in financial systems, the subject areas “Economics, Econometrics and Finance” and “Business, Management and Accounting” were chosen.

Considering all the searches, we found a large number of articles with the keywords and filters established. However, after recognizing the titles of each article, it was possible to notice that within the keyword variations, the searches returned many repeated articles, which reduced the number of articles considered new. For example, the keywords “anomaly detection finance” resulted in 71 articles of which only 22 were considered for the literature review. In addition, by consulting the abstracts of the articles found, it is possible to identify research that did not fit the topic to be studied, not involving, for example, methods or techniques for detecting anomalies, such as the paper of Bouzguenda (2018, p. 273) that addressed the effect of emotional intelligence in the financial decision-making process, or the article of Jung and Jeong (2020, p. 1), which presented a case study that explored techniques for predicting social media engagement level.

Thus, after selecting articles based on the listed criteria, we identified 66 articles relevant to the present study. However, the thorough reading of each of these remaining articles signalled some articles that were not compatible with the selection criteria, such as the study of Frömmel et al. (2021, p. 1) on low liquidity anomalies in China's stock beta, with a focus on exploring such anomalies and not on anomaly detection techniques within the context of the Chinese financial market. Another example of an unconsidered article was the research of Cumming et al. (2018, p. 130), which addressed market abuse, analysing differences in the intensity of law enforcement in European countries; therefore, not discussing issues related detecting anomalies or fraud that could cause such abuse. In summary, 48 articles fit the selection criteria considered for the literature review. This history of research and selection of articles is presented in Table 1.

The selected papers focused on methods and techniques for detecting anomalies, fraud and outliers in different financial systems, and a brief summary of each of these studies can be found in Table 2. After selecting these articles, as well indicated in the articles by Jabbour (2013, p. 145) and Lage Junior and Godinho Filho (2010, p. 14), we performed the second step of the research method in which the articles were classified into categories of analysis based on codes of distinction. The third step synthesized the existing scientific knowledge to enable a more comprehensive view of the study area. In line with the third step, the fourth step enabled a visual understanding of the results and trends in the current literature on detecting financial anomalies. Finally, the fifth step concluded the review by discussing the results and proposing a research agenda for future investigations. All the necessary steps for this research can be visually understood through the graphical abstract presented in Fig. 1, inspired by the studies Al-Hashedi and Magalingam

Table 1
Number of papers selected from *ScienceDirect*®.

| Keywords | Articles | | |
|--|---------------------------------|-------------------------|-----------------------|
| | Found at <i>ScienceDirect</i> ® | Added to the Study Base | Repeated or Withdrawn |
| Anomaly detection finance. | 71 | 22 | 49 |
| Outlier detection finance. | 169 | 7 | 162 |
| Literature review anomaly detection finance. | 118 | 5 | 113 |
| Anomaly detection economics. | 69 | 3 | 66 |
| Anomaly detection financial fraud. | 304 | 23 | 281 |
| Credit card fraud detection. | 70 | 1 | 69 |
| Anomalous activity finance. | 85 | 5 | 80 |
| Exclusion after reading. | – | – | 18 |
| Total of Papers. | – | 48 | – |

(2021, p. 5) and Pourhabibi et al. (2020, p. 3) developed.

In the next section, the second step of the research method will be presented: Classification and coding. From this, it is possible to categorize the selected articles and carry out a review of the existing literature in the last five years within the theme of detecting anomalies in financial systems.

3. Classification and coding

After the selection and reading of the articles in the present study base, a classification framework was built with codes to categorize the papers in a relevant way to the research theme of the review. Its structure is composed of 10 classification categories, numbered from 1 to 10, which are combined with a coding presented in the form of letters (A, B, C, D, and so on) juxtaposed to the numbers, which represent sub-categories of classification to promote further analysis of research trends in the context of methods and techniques for detecting anomalies in financial systems. In Table 3, the classification framework is presented. Considering this system of classification, the articles were analysed and incorporated into the structure, with codes referring to each category of analysis being assigned to each of them. It is worth noting that an article can receive more than one code in a given category, as it can fit into more than one classification subtopic.

Classification 1 involves the Economy Context in which the articles are inserted, considering the data's countries that are used to develop the research. The coding was done using the letters A-D. Codings "1A - Mature economy" and "1B - Non-mature economy" were chosen sub-categories based on the work of Jabbour (2013, pp. 145–146), with the aim of identifying whether the type of economy has any relationship with the development of fraudulent and criminal activities that may lead to the detection of anomalies in financial systems, considering that mature market economies have more stable economic structures while developing economies usually go through several structural changes (Borgersen & King, 2014, p. 102) that can make such actions viable. This category correlates the economic perspectives of the regions with the interference that triggers the anomalies in the financial systems, enabling the identification of any relationship between the development of more or fewer studies in countries with mature or immature economies over the last few years (Henriques et al., 2018, p. 146) by flagging the existence of incidents where financial concerns were aligned with anomaly detection (Ferreira et al., 2016, p. 8). It is relevant, since self-regulatory strategies used in different scenarios can be endowed with different approaches, which translates into different behavioural responses and may or may not facilitate the emergence of anomalies that

Table 2
Brief summary of the selected articles.

| Articles | Brief Summary |
|-----------------------------------|---|
| Al-Hashedi and Magalingam (2021). | Proposed a literature review to complement the sample of articles reviewed by other authors and also to present a comprehensive summary of recent research on the use of Data Mining techniques to promote fraud detection in the financial area. |
| Andor and Bohák (2017). | Proposed an innovative approach that promotes the identification of significant events in financial price time series from statistical tests to explore the detection of abrupt changes in the sample used. |
| Arboleda et al. (2018). | Used operators based on slice and dice operations, Benford's Law and forensic accounting fraud detection techniques to identify signs of possible fraud in data warehouses. |
| Ardizzi et al. (2018). | Implemented an econometric analysis model on financial deposits made in Italy to define a series of risk indicators that can distinguish illegal and anomalous components from those considered legal and within the normality level. |
| Asha and Suresh Kumar (2021) | Used different machine learning algorithms to identify the best performance in detecting fraudulent transactions made with credit cards. |
| Ausloos et al. (2021). | Explored the conformity of the distributions of the daily closing values of the S&P 500 and the corresponding daily logarithmic returns with the significant digits of Benford's Law. |
| Azevedo et al. (2021). | Evaluated the application of Benford's Law as a technique to detect payment fraud promoted by the Brazilian Bolsa Família benefits program in order to select regions that provide a suspicious distribution of capital. |
| Badal-Valero et al. (2018). | Presented a tool to identify, based on accounting data, fraudulent companies, detect criminals who carry out money laundering actions, and reduce the likelihood of errors in distinguishing between companies with illegal operations and those operating correctly. |
| Baesens et al. (2021). | Presented the use of data engineering techniques, broken down into feature engineering and instance engineering, to improve the performance of an analytical fraud detection model based on real payment transaction data from a European bank. |
| Baker et al. (2017). | Examined flaws that made it possible to detect fraud in the trade information system, which caused a financial meltdown in a European bank, in order to identify suitable control opportunities. |
| Byers et al. (2021). | Analysed the presence of outliers in commodity futures price data through a machine learning algorithm, which is also used to investigate the impact of these outliers on commodity risk measures. |
| Canhoto (2021). | Carried out an investigation into the use of machine learning algorithms to detect suspicious financial transactions, in addition to predicting and preventing money laundering and terrorist financing actions. |
| Carmona et al. (2019). | Explored the use of Extreme Gradient Boosting, a machine learning method, to improve the accuracy of predicting bank failures in the US by identifying key variables that can anticipate and predict banking patterns and even prevent financial crises. |
| Carneiro et al. (2017). | Presented the development and implementation of a risk scoring system that promotes the detection of credit card fraud |

(continued on next page)

Table 2 (continued)

| Articles | Brief Summary |
|------------------------------|--|
| | based on machine learning methods in an online retail company. |
| Carta et al. (2019). | Presented an innovative technique that combines machine learning and statistics to maximize the ability to detect fraudulent credit card transactions in an e-commerce. |
| Chen et al. (2020). | Used machine learning algorithms combined with credit score models to define a hybrid credit card fraud detection model capable of signalling the occurrence of discrepancies between a person's current situation and their situation in the near future. |
| Climent et al. (2019). | Applied a machine learning method to build a classification model that identifies and monitors a set of relevant indicators and variables that can help predict and prevent bank failure in the Eurozone. |
| Druică et al. (2018). | Analysed the compliance of the first two digits of the annual bank balances of a bank in Romania with Benford's Law based on statistical tests. |
| Forough and Momtazi (2021). | Presented a model that combines sequential data modelling with neural network algorithms to obtain better performance in the detection of credit card fraud. |
| Gianini et al. (2020). | Developed a rules management model for credit card fraud detection using a power-index based on Game Theory that evaluates and updates the rules and performance of their applications. |
| González-Sánchez (2021). | Analysed the existence of outliers in financial markets and investigated whether such deviations cause the time scaling process and what are the behavioural effects resulting from this possible correlation. |
| Houdek (2020). | Proposed a literature review in which it is possible to explore the advantages and disadvantages of methods that may reveal fraudulent and dishonest behaviour in organisations, as well as the pros and cons of each of these methods. |
| Jurgovsky et al. (2018). | Presented a technique of long-term memory networks to understand the individual purchase behaviour of each credit card holder in order to improve the detection of fraud in future transactions. |
| León (2020). | Presented a method of investigating payment networks in order to detect unusual changes in connection resources, which deviate from what is considered a normal network structure, signalling an anomalous network. |
| León et al. (2020). | Discussed the use of a machine learning method to understand the behaviour of payment systems at financial institutions in Colombia that serves as a tool to monitor and detect anomalies. |
| Lin et al. (2018). | Investigated the accounting manipulations made by corporate directors called fat cats and their impacts in contrast to the Benford's Law's significant number distribution model. |
| Lucas et al. (2020). | Proposed a machine learning model that quantifies the similarity between observed transaction sequences and fraudulent credit card transaction sequences of a European company, in order to improve the fraud detection process. |
| Mamede and Malaquias (2017). | Analysed the difference in behaviour patterns in Brazilian hedge fund stocks, in order to identify and understand anomalies caused by the Monday Effect that cause investment returns to be abnormal. |
| Misra et al. (2020). | Proposed a two-stage machine learning model, in which resources are first extracted from the dataset to promote the learning of significant transactional attributes and then |

Table 2 (continued)

| Articles | Brief Summary |
|---|--|
| | identify if there is fraud in credit card transactions. |
| Nikkel (2020). | Presented a reflection on the recognition of Fintech Forensics as a digital forensics tool to detect criminal actions that abuse of financial technologies to commit fraud, money laundering and extortion. |
| Parnes (2020). | Promoted the investigation of anomalies in the S&P 500 market index in different economic aspects to detect patterns of behaviour, and also examines deviations between the S&P 500 market index and nine American financial indicators. |
| Patil et al. (2018). | Proposed an analytical and self-adaptive framework model in order to identify credit card fraud and also compare the performance of different fraud detection algorithms in order to improve the model. |
| Pourhabibi et al. (2020). | Presented a literature review to consolidate existing knowledge about the use of Graph-Based Anomaly Detection Techniques in the context of financial fraud detection based on a mapping of research trends, methods and challenges. |
| Rocha-Salazar, Segovia-Vargas, and del Mar Camacho-Miñano (2021). | Presented a model for detecting suspicious financial transactions related to money laundering and terrorist financing based on an abnormality indicator based on variables that enable behavioural analysis of the transactions. |
| Rtayli and Enneya (2020). | Developed a new hybrid model that combines machine learning methods and techniques to improve the efficiency of predicting fraudulent credit card transactions. |
| Rubio et al. (2020). | Developed an autoencoder for a bank payment system in Ecuador to understand payment patterns and detect possible anomalies from neural networks. |
| Sá et al. (2018). | Presented the use of a bayesian network classifier algorithm as a method of detecting fraud in credit card transactions in order to classify and distinguish legitimate transactions from fraudulent transactions. |
| Sabetti and Heijmans (2021). | Developed an autoencoder based on machine learning methods that recognize patterns in a retail payment system that do not conform to expected behaviour and are considered anomalies. |
| Severino and Peng (2021). | Presented a comparative analysis of how the use of different machine learning models can contribute to the detection of fraud in property insurance policies. |
| Shi et al. (2018). | Conducted a statistical test to assess whether the price indices of financial reporting countries with a non-mature economy were in compliance with the significant digits of Benford's Law or if they had anomalous values. |
| Singh and Best (2019). | Demonstrated the application of visualization techniques that promote the investigation of financial transactions by detecting patterns that represent potential money laundering activities. |
| Stripling et al. (2018). | Presented a machine learning approach that promotes isolation-based anomaly detection in data from a European insurance company from the computation of conditional anomaly scores considered useful for identifying fraud. |
| Subudhi and Panigrahi (2020). | Proposed a new hybrid model that applies genetic algorithms with clustering techniques to improve the auto insurance fraud identification process. |
| Tardelli, Avvenuti, Tesconi, and Cresci (2021). | Presented two methods of automatic detection of financial misinformation in a social network through classification and |

(continued on next page)

Table 2 (continued)

| Articles | Brief Summary |
|--|--|
| Wang and Xu (2018). | regression to estimate the amount of inorganic content in financial discussions. Presented an automobile insurance fraud identification model that combines the use of Data Mining methods with the experience of analysing experts involved with the data to increase detection accuracy. |
| Yan et al. (2020). | Proposed a method that combines machine learning, statistics and computer simulations to promote the detection of car insurance fraud. |
| Zhai et al. (2017). | Presented a hybrid model composed of Data Mining techniques capable of identifying anomalous patterns and behaviours in commercial exchanges made in an American stock market. |
| Zoričák, Gnip, Drotár, and Gazda (2020). | Presented a bank failure prediction tool that, based on the detection of relevant financial attributes, can identify which companies are healthy and which are not in accordance with financial abnormality data. |

may alter economic contexts (Florack, Keller, & Palcu, 2013, p. 127). In line with this categorization, Classification 2 was structured, which signals the Geographical Research Location when indicating the continents related to the countries of the research data.

Listing the geographic location where the papers were developed allows not only an analysis correlated with economic factors from the first classification, but also a more detailed examination of the combined effects of geographic knowledge and geographic research activity (Hohberger & Wilden, 2022, p. 706). This is because studies carried out in different localities exhibit particular characteristics (Boubakri, Guedhami, & Saffar, 2016, p. 364; Henriques et al., 2018, p. 146) that reflect at the available resources, the level of communication and the degree of prosperity and financial activity of organisations (Chen, 2016, p. 323) and affect control skills and data monitoring (Boubaker, Derouiche, & Lasfer, 2015, p. 24).

Therefore, the geographic analysis complements the first classification in a more specific way (Ferreira et al., 2016, p. 9), seeking to identify whether financial anomalies may be related to specific regions or countries, and may be considered endemic. This classification is relevant, since countries around the globe have different levels of maturity in their economic, financial and legal systems (Nazário et al., 2017, p. 118), which may appear in the fight against anomalies when they are drawn in order to create financial gains or losses and may reflect in the increase or decrease of academic discussion in face of the existence of financial anomalies. The coding was done considering the letters A-H, aiming to identify niches or specific research groups relevant to the theme that can provide insights into how anomalies are seen and treated according to the place where they are identified and determine whether this factor influences the academic approaches developed.

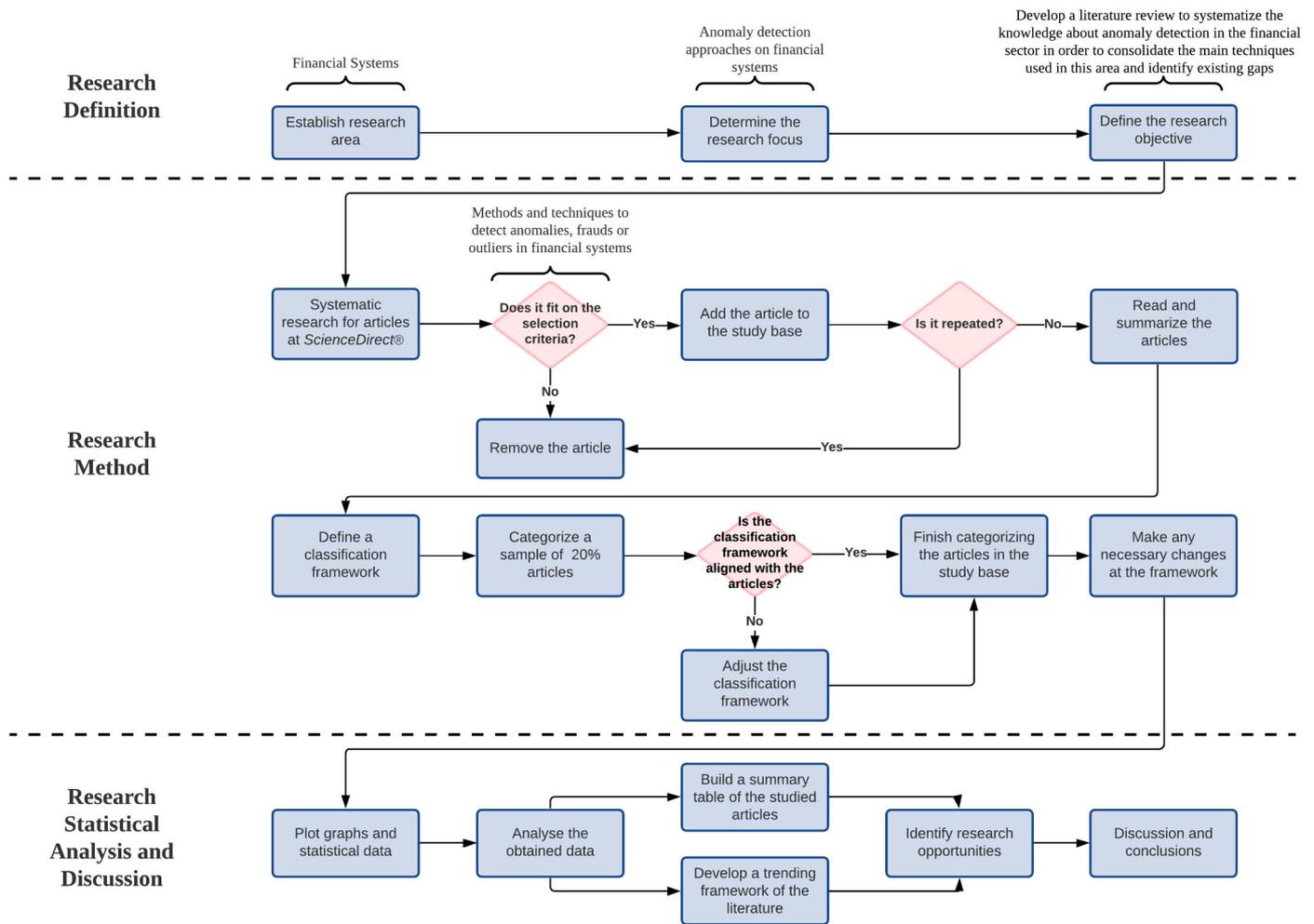


Fig. 1. Systematic literature review graphical abstract.

Notes: Based on the studies of Pourhabibi et al. (2020, p. 3) and Al-Hashedi and Magalingam (2021, p. 5).

Notes: 1A - Mature economy; 1B - Non-mature economy; 1C - Not informed; and 1D - Not applicable.

Table 3
Classification and coding framework.

| Classification | Meaning | Codes |
|----------------|---------------------------------|--|
| 1 | Economic Context. | 1A - Mature economy. 1B - Non-mature economy. 1C - Not informed. 1D - Not applicable. |
| 2 | Geographical Research Location. | 2A - America. 2B - Europe. 2C - Asia. 2D - Africa. 2E - Oceania. 2F - All continents. 2G - Others. 2H - Not applicable. |
| 3 | Research Method. | 3A - Qualitative. 3B - Quantitative. 3C - Conceptual. 3D - Modelling. 3E - Survey. 3F - Others quantitative approaches. 3G - Single case study. 3H - Multiple case study. |
| 4 | Application Area. | 4A - Banking. 4B - Actuarial. 4C - Financial market. 4D - Social security. 4E - Financial accounting. 4F - Others. 4G - Not applicable. |
| 5 | Focus. | 5A - Credit card. 5B - Insurance. 5C - Financial welfare benefits. 5D - Financial transactions. 5E - Financial statements. 5F - Stocks and commodities. 5G - Others. 5H - Not applicable. |
| 6 | Dataset Used. | 6A - Real. 6B - Synthetic/Artificial. 6C - Others. 6D - Not applicable. |
| 7 | Availability. | 7A - Restricted access. 7B - Unrestricted access. 7C - Others. 7D - Not applicable. |
| 8 | Technique Used. | 8A - Machine learning. 8B - Statistics. 8C - Bayesian statistics. 8D - Benford's law. 8E - Graph methods (visual). 8F - Computer simulations. 8G - Others. 8H - Not applicable. |
| 9 | Number of Strategies. | 9A - One. 9B - Two. 9C - Three. 9D - Four or more. 9E - Others. 9F - Not applicable. |
| 10 | Journal's Subject Category. | 10A - Health sciences. 10B - Life sciences. 10C - Physical sciences and engineering. 10D - Social sciences and humanities. |

The articles in these categories classified as “1C - Not informed” and “2G - Others” were those that used some type of data to construct the research, but did not specify their geographic origin, such as, for

example, the study developed by [Chen et al. \(2020, p. 3\)](#) which adopted a dataset with selected credit card transactions from the Kaggle¹ and UCI² machine learning repositories to present its research method, but did not make it explicit the countries of the chosen data. On the other hand, articles classified as “1D - Not applicable” in the economic context category and “2H - Not applicable” in the research geographic location category did not fit the previously established classification forms. Articles classified with these codes did not present data that refer to geographic sources of study, such as the literature review developed by [Pourhabibi et al. \(2020, pp. 2–3\)](#) which used articles within the theme of fraud detection, published in international journals to promote an in-depth analysis of graph-based anomaly detection techniques in this area.

Classification 3, structured on the A-H coding scale, signals the Research Method used by the authors on the development of their research on anomalies. The idea behind this classification is to recognize which methods are most used within the anomaly detection field in order to identify research trends. Considering that pattern identification is often related to various computer simulations through algorithms, statistics and distributions, especially using machine learning, bayesian statistics and Benford's Law, it is expected that the majority of articles are classified as “3B - Quantitative”, as the adoption of quantitative models promotes promising and rapid analyses that can easily map the object of study and provide relevant insights to the research ([Ciaccio et al., 2019, p. 120](#)). It is worth noting that the classification “3C - Conceptual” was used for articles that promote a literature review, since such articles are in line with the objectives of a conceptual research, based on the observation, investigation and analysis of methods and concepts to identify research gaps ([Soulé, Michonneau, Michel, & Bockstaller, 2021, p. 2](#)). However, this classification was also created to identify the existence of conceptual research that establishes common practices to be developed in the detection of anomalous patterns in order to better understand if there is a structural pattern of research performance that allows for more significant comparisons between them.

Classification 4 refers to the Application Area in which the researchers develop anomaly detection methods, with a coding scale ranging from A-G. These areas were chosen based on the trends of studies developed in the articles and relate the various sectors of the financial industry where fraud ([Azevedo, Gonçalves, Gava, & Spinola, 2021, pp. 1–2](#); [Yan et al., 2020, pp. 115–116](#); [Sá et al., 2018, pp. 21–22](#)), data tampering ([Ausloos, Ficcadenti, Dhesi, & Shakeel, 2021, pp. 2–3](#); [Druică, Oancea, & Vălsan, 2018, pp. 77–79](#); [Shi, Ausloos, & Zhu, 2018, pp. 879–880](#)), capital misuse ([Singh & Best, 2019, pp. 1–2](#); [Ardizzi et al., 2018, pp. 105–106](#); [Badal-Valero, Alvarez-Jareño, & Pavía, 2018, pp. 24–25](#)), among other actions that may indicate abnormalities in the systems, can occur.

The main objective of this classification is to identify the various tools or techniques used to identify the anomalies related to a specific area so the expertise arising from them and the considered approach can be applied in other areas of knowledge. This is because with the advancement of communication and technologies in the financial sector, a highly connected world that offers a wide range of information easily accessed through anonymous accounts has been created, which became an incentive for fraudsters to manipulate the various financial systems, hiding their malicious activities among the large amounts of data ([Hooi et al., 2017, p. 1](#); [Pourhabibi et al., 2020, p. 1](#)). In this way, identifying the approaches used according to specific areas can serve as basis and inspiration for their applications in new systems and subsequently prevent the spread of illegal acts, either in financial systems already in place or in those entering the market through digital transformation ([Nikkel, 2020, p. 1](#)).

In conjunction with this, Classification 5 indicates the approach

¹ Kaggle website: <https://www.kaggle.com/https://www.kaggle.com/>.

² UCI Machine Learning website: <https://archive.ics.uci.edu/ml/index.phphttps://archive.ics.uci.edu/ml/index.php>.

Focus, within the application areas, chosen by the researchers to develop and study anomaly identification techniques. This classification aims to further narrow the understanding of the anomaly detection approaches chosen for specific research focuses and application areas, in order to act as a support for future researchers to be able to more easily identify how the different anomaly identification techniques are chosen and how they correlate with different financial areas based on their characteristics and needs. The coding scale covers the letters A-H. The articles that received the encoding “5G - Others” cover study sectors that do not fit into the proposed sub categorizations, such as the study proposed by [Arboleda et al. \(2018, pp. 14–17\)](#) which, in addition to financial statements, addressed sales operations from a data warehouse.

Classification 6 is related to the Dataset Used for the construction of anomaly detection methods, considering a scale ranging from A-D. It is relevant to identify which articles used a database with real data to validate their study model and which used artificial or simulated data, since the use of real data, even when combined with synthetic data, produces a better version of the detection model ([Shtanko & Kulik, 2021, p. 712](#)). In addition, financial markets generate large volumes of data and, therefore, analysing the data used can reveal useful and valuable information for the decision-making process ([Wang & Weigend, 2004, p. 457](#)). Classification 7, Availability, also considers the data used for searches, indicating whether they have restricted access or unrestricted access from the A-D scale. The codes “6C - Others” and “7C - Others” are used, respectively, when the authors do not specify whether the database is real or artificial, and do not inform the origin of the data, and it is not possible to identify whether such data is public or private. On the other hand, the codes “6D - Not applicable” and “7D - Not applicable”, when used, signal that the article does not fit these forms of classification. For example, this occurs with articles defined as literature reviews, a method adopted by the authors [Al-Hashedi and Magalingam \(2021, pp. 5–6\)](#), [Pourhabibi et al. \(2020, pp. 2–3\)](#) and [Houdek \(2020, p. 1\)](#), which do not have a database with validation data for analysing techniques and methods.

Classification 8 referred to the Technique Used to detect the most diverse forms of anomalies in financial systems, based on a coding scale ranging from A-H. The encoding “8G - Others” is used when the articles to be classified do not correspond to the proposed classification options, but still present techniques capable of identifying anomalies, and the encoding “8H - Not applicable” is chosen when the article does not support this form of classification. Considering the variety of existing techniques for analysis in this literature review, Classification 9 was structured following the A-F scale in order to verify the Number of Strategies used in the papers to detect anomalies, that is, the number of techniques used in a single model of studies. This classification was chosen in order to enable analysis on the trend of using combined techniques, since normally hybrid models are more accurate than those with single processes ([Chen et al., 2020, p. 3](#)) and the combination of methods can improve the prediction procedure to be developed ([San-karrao, Ghose, & Rathinsamy, 2021, p. 2](#)).

Finally, Classification 10, structured in the A-D coding scale, was designed to identify the areas of greatest incidence of studies within the theme of detecting anomalies in financial systems, considering the Journal's Subject Category from ScienceDirect®. The categories are: Health sciences (medicine and dentistry; nursing and health professions; pharmacology, toxicology and pharmaceutical science; and veterinary science and veterinary medicine), Life sciences (agricultural and biological sciences; biochemistry, genetics and molecular biology; environmental science; immunology and microbiology; and neuroscience), Physical sciences and engineering (chemical engineering; chemistry; computer science; earth and planetary sciences; energy; engineering; materials science; mathematics; and physics and astronomy), Social sciences and humanities (arts and humanities; business, management and accounting; decision sciences; economics, econometrics and finance psychology; and social sciences). It is worth noting that as the area of finance is a field of study of the social sciences, it is expected that the

selected articles mostly fit into the “10D - Social sciences and humanities” coding.

4. Results and discussions

Based on the proposed classification framework, the 48 selected articles were reviewed, analysed and categorized according to their specifications, as shown in [Table 4](#). This classification seeks to enable an intelligible understanding and analysis of the knowledge and techniques developed in the selected papers. Based on this, this literature review on the topic of detecting anomalies in financial systems can outline study trends and identify future research opportunities according to a classification system that allows for the recognition of parity and difference points between surveys.

After classification, the categorized data obtained were statistically compiled using a *script* developed in the programming language *Python*, version 3.7, through *Anaconda Distribution*, by *Anaconda Inc*®. More precisely, this *script* automated the calculation of the descriptive statistics for each category. In addition, graphs were constructed from such data, which showed analysis and discussion of the results also in a visual way. The graphs were divided into two for each category, indicating the statistical data regarding the count of codes alone and the cluster of codes that appeared at the time of classification of each article, with the exception of classifications 9 and 10, in which there was no more than one indicated code for a single article. Therefore the statistical data only shows the count of isolated codes.

4.1. Economic context

This classification category was structured in order to list the selected articles according to their economic characteristics. Countries that have mature economies usually have more structured systems, with stricter laws and greater coercive pressure to promote regulation and inspection ([Raj, Agrahari, & Srivastava, 2020, pp. 2–3](#)). However, countries with a non-mature economy, even though they have different physical, financial and legal infrastructures, are increasingly investing in information technology ([Arnott, Jirachiefattana, & O'Donnell, 2007, p. 2078](#)) and creating decision support systems that can help both the public and private sectors in economic terms ([Chaudhry, Li, Xu, & Zhang, 2007, p. 1987](#)). Thus, identifying the economic context in which the anomaly detection models are applied in the articles is of paramount importance not only to verify the structural differences of the countries in question, but also to identify similarities and possible technological advances with regard to identifying fraud, misappropriation of capital or tampering with financial systems.

The statistical data regarding the classifications of the articles can be seen through [Fig. 2](#). The most used coding was “1A - Mature Economy”, covering 38% of the articles, which is in line with the research by [Raj et al. \(2020, p. 2\)](#) about the advances of more mature economies and, as a result, the number of surveys in these countries ends up being higher. However, it is necessary to observe that the classification percentage of the encoding “1B - Non-Mature Economy” is also significant, covering 28% of the articles. Thus, it is possible to verify that the search for efficient anomaly detection models and techniques is not limited only to countries with mature economies but rather configures a similarity between the two forms of economy. Nonetheless, in order to have comparative analysis at this level, it would be necessary for future research to combine a set of test data involving both mature and non-mature economy countries, as done in only one article among the selected ones ([González-Sánchez, 2021, p. 7](#)). There is, then, the first gap identified in this research, namely:

G₁: Develop papers that consider countries with mature and non-mature economies to identify similarities and divergences in their approaches to the detection of anomalies and managing decision systems.

In addition, a significant analysis that can be deepened is whether there is a correlation between the economic context of the research and

Table 4
Categorized articles based on the classification and coding framework.

| Selected Articles | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------------------------------|-------|----------|-------|-------|-------|-------|-------|----------|----|-----|
| Al-Hashedi and Magalingam (2021). | 1D | 2H | 3C | 4G | 5H | 6D | 7D | 8H | 9F | 10C |
| Andor and Bohák (2017). | 1A | 2A | 3B | 4C | 5F | 6A | 7B | 8B | 9A | 10D |
| Arboleda et al. (2018). | 1B | 2A | 3G | 4E | 5E/5G | 6A | 7A | 8D | 9A | 10C |
| Ardizzi et al. (2018). | 1A | 2B | 3B/3D | 4A | 5D | 6A | 7A/7B | 8B | 9A | 10D |
| Asha and Suresh Kumar (2021) | 1C | 2B | 3A | 4A | 5A | 6A | 7B | 8A | 9A | 10A |
| Ausloos et al. (2021). | 1A | 2A | 3B | 4C | 5F | 6A | 7B | 8B/8G | 9B | 10D |
| Azevedo et al. (2021). | 1B | 2A | 3B/3G | 4D | 5C/5E | 6A | 7B | 8B/8D | 9B | 10C |
| Badal-Valero et al. (2018). | 1A | 2B | 3B | 4A | 5D | 6A | 7C | 8A/8D | 9B | 10C |
| Baesens et al. (2021). | 1C | 2B | 3B | 4A | 5A | 6A | 7A | 8A | 9A | 10C |
| Baker et al. (2017). | 1A | 2B | 3G | 4A/4E | 5F | 6A | 7B | 8H | 9A | 10D |
| Byers et al. (2021). | 1C | 2G | 3B | 4C | 5F | 6A | 7B | 8A | 9A | 10C |
| Canhoto (2021). | 1A | 2B | 3G | 4A | 5D | 6A | 7A | 8A/8G | 9B | 10D |
| Carmona et al. (2019). | 1A | 2A | 3A | 4A | 5E | 6A | 7B | 8A | 9A | 10D |
| Carneiro et al. (2017). | 1C | 2G | 3B/3G | 4A | 5A | 6A | 7A | 8A | 9C | 10C |
| Carta et al. (2019). | 1C | 2B | 3B/3D | 4A | 5A | 6A | 7B | 8A/8B | 9B | 10C |
| Chen et al. (2020). | 1C | 2G | 3B/3D | 4A | 5A | 6C | 7B | 8A/8B | 9B | 10C |
| Climent et al. (2019). | 1A | 2B | 3B | 4A/4E | 5E | 6A | 7B | 8A | 9A | 10D |
| Druicá et al. (2018). | 1A | 2B | 3B | 4A | 5D/5E | 6A | 7B | 8B/8D | 9B | 10D |
| Forough and Momtazi (2021). | 1B/1C | 2A/2B | 3B/3D | 4A | 5A | 6A | 7B | 8A | 9B | 10C |
| Gianini et al. (2020). | 1C | 2G | 3B | 4A | 5A | 6A | 7A | 8A/8G | 9B | 10C |
| González-Sánchez (2021). | 1A/1B | 2F | 3B | 4C | 5F | 6A | 7B | 8B | 9A | 10D |
| Houdek (2020). | 1D | 2H | 3C/3H | 4G | 5H | 6D | 7D | 8H | 9F | 10D |
| Jurgovsky et al. (2018). | 1C | 2G | 3B | 4A | 5A | 6A | 7C | 8A | 9A | 10C |
| León (2020). | 1B | 2A | 3B | 4A | 5D | 6A/6B | 7A | 8A/8G | 9B | 10D |
| León et al. (2020). | 1B | 2A | 3B | 4A | 5D | 6A | 7A | 8A | 9A | 10D |
| Lin et al. (2018). | 1B | 2C | 3B | 4E | 5E | 6A | 7B | 8B/8D | 9B | 10D |
| Lucas et al. (2020). | 1A | 2B | 3B | 4A | 5A | 6A | 7A | 8A | 9A | 10C |
| Mamede and Malaquias (2017). | 1B | 2A | 3B | 4C | 5F | 6A | 7B | 8B | 9A | 10D |
| Misra et al. (2020). | 1C | 2B | 3B/3D | 4A | 5A | 6A | 7B | 8A | 9A | 10C |
| Nikkel (2020). | 1A | 2B | 3A | 4A | 5A/5D | 6A | 7B | 8G | 9A | 10C |
| Parnes (2020). | 1A | 2A | 3A/3B | 4C | 5F | 6A | 7B | 8A/8B/8E | 9A | 10D |
| Patil et al. (2018). | 1A | 2B | 3B | 4A | 5A | 6A | 7C | 8A/8F | 9B | 10C |
| Pourhabibi et al. (2020). | 1D | 2H | 3C | 4G | 5H | 6D | 7D | 8H | 9F | 10C |
| Rocha-Salazar et al. (2021). | 1B | 2A | 3B | 4A | 5D | 6A | 7A | 8A | 9A | 10C |
| Rtayli and Enneya (2020). | 1C | 2B/2G | 3B/3D | 4A | 5A | 6A/6B | 7A/7B | 8A/8G | 9C | 10C |
| Rubio et al. (2020). | 1B | 2A | 3B | 4A | 5D | 6A | 7C | 8A | 9A | 10D |
| Sá et al. (2018). | 1B | 2A | 3B/3G | 4A | 5A | 6A | 7A | 8A/8B | 9B | 10C |
| Sabetti and Heijmans (2021). | 1A | 2A | 3B | 4E | 5D | 6A/6B | 7D | 8A | 9A | 10D |
| Severino and Peng (2021). | 1B | 2A | 3B | 4B | 5B | 6A | 7A | 8A | 9A | 10C |
| Shi et al. (2018). | 1B | 2A/2B/2C | 3B | 4E | 5E | 6A | 7B | 8B/8D | 9B | 10C |
| Singh and Best (2019). | 1C | 2G | 3A/3G | 4A | 5D | 6A | 7A | 8E/8F | 9B | 10D |
| Stripling et al. (2018). | 1C | 2B | 3B | 4B | 5B | 6A | 7A | 8A | 9A | 10C |
| Subudhi and Panigrahi (2020). | 1A | 2A | 3B | 4B | 5B | 6A | 7B | 8A | 9B | 10C |
| Tardelli et al. (2021). | 1A | 2A | 3B/3G | 4C | 5F | 6A | 7B | 8A | 9B | 10C |
| Wang and Xu (2018). | 1B | 2C | 3B | 4B | 5B | 6A | 7A | 8A/8G | 9B | 10C |
| Yan et al. (2020). | 1C | 2G | 3B/3D | 4B | 5B | 6A/6B | 7A/7B | 8A/8B/8F | 9D | 10B |
| Zhai et al. (2017). | 1A | 2A | 3B | 4C | 5F | 6A | 7B | 8A/8B | 9B | 10C |
| Zoričák et al. (2020). | 1A | 2B | 3B | 4A/4E | 5E | 6A | 7B | 8A/8B | 9C | 10D |

the behaviour of individuals, that is, whether there are economic factors that favour the increase in criminal actions that signal financial anomalies and, consequently, they lead to an increase in the search for models that can identify these outliers and, thus, enable greater security for financial systems and for the individuals who benefit from them.

It is worth noting that a gap observed is the significant number of articles that do not indicate the origin of the data used to apply and test the functionality of their detection models, with 28% of articles classified in the category “1C - Not informed”, which covers studies which do not explain the location of the data and, therefore, do not allow an analysis of the economic context in which the research development is inserted. This is, therefore, the second opportune gap for future research, namely:

G₂: Indicate the origin of the data used for research, so that it is possible to relate them to their respective economic contexts.

4.2. Geographical research location

When producing a study on a topic, the analysis of how the geographic location factors in which the research is inserted is fundamental for the acquisition of information that can be useful for different

forms of investigation (Lozano, Schreiber, & Brynielsson, 2017, p. 18) and, consequently, such deeper contact with the information regarding the geographic position where the study is being conducted can be extremely important for understanding the different forms of expression of a research topic around the world. This occurs, since analysis about the location enable the understanding of possible behavioural patterns, spatial arrangements, groupings, in addition to probable relationships that may result from geographic phenomena (Pick, Turetken, Deokar, & Sarkar, 2017, p. 1).

Checking and pondering the questions about the geographic location of the researches developed, in this context, can enable the attribution of information about the possible objects of interest, which can be shown as important characteristics to be considered by a decision-maker (Hess, Rubin, & West, 2004, p. 198), given that such particularities can influence the interpretation of the situations under study. Based on this, the selected articles involving methods of detecting anomalies in different financial systems and frauds and criminal actions resulting from such abnormalities can be analysed according to the perspectives of the geographic position in which they were produced. It is worth noting that the perspective used for the analysis can be indicative of different forms of standard deviations, as, for example, in the study developed by

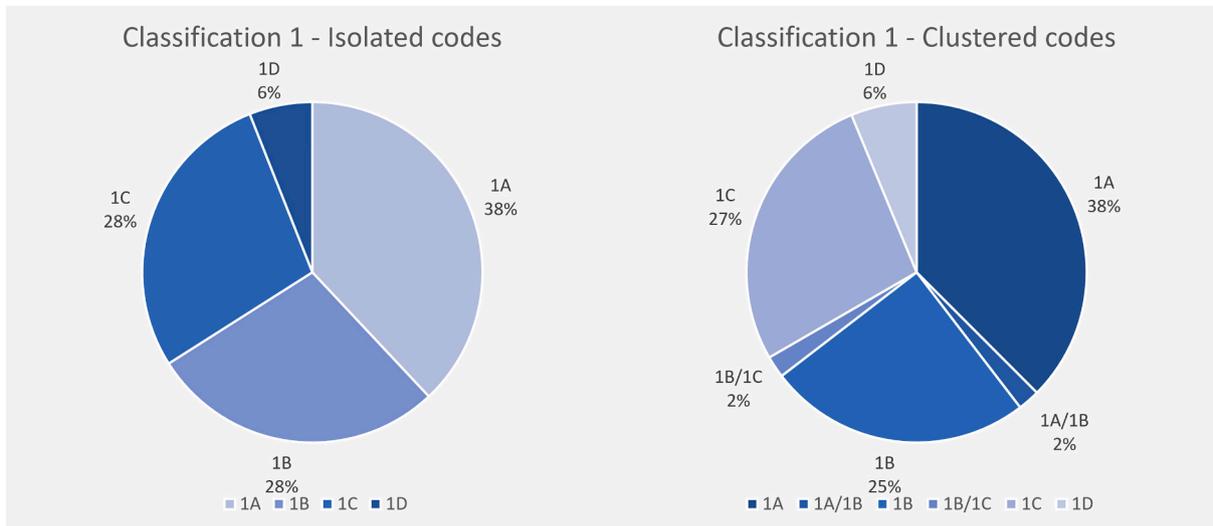


Fig. 2. 1 - Economic Context.
Notes: 2A - America; 2B - Europe; 2C - Asia; 2F - All continents; 2G - Others; and 2H - Not applicable.

Jurgovsky et al. (2018, p. 236) where fraudulent transactions can be analysed as anomalies in consumer behaviour or as a set of transactions that diverge from the class of genuine transactions. In this way, the point of view, when combined with the analysis of the geographic location of research produced on anomalies, signals the need to identify correlations between locations and the rates of criminal actions, or, between cultures and the way to position oneself about issues such as system security or punishment and justice.

Statistical data from the classification of selected articles were presented in Fig. 3. Most studies have been classified as “2A - America” and “2B - Europe”, so these two continents develop more research on how to identify anomalies in their financial systems, comprising 71% of the articles reviewed. Even among these articles, there are two that develop research considering countries on both continents (Forough & Monttazi, 2021, pp. 6–7; Shi et al., 2018, p. 880). Considering this, it is worth discussing here the possible causes of this expressive number of papers on financial anomalies in American and European countries, in order to provide an illustration of the circumstances that possibly provide or justify such statistical data, such as the previously discussed cultural patterns and, subsequently, positions about the situation.

Cultural dimensions are strongly used to find the reasons for the occurrence of variations in results between countries, as individuals who have divergent values, beliefs and preferences tend to act differently (Tung & Quaddus, 2002, p. 190). In addition, different cultures arouse dissimilar perspectives that, based on values, beliefs and norms, provide different human behaviours, and therefore generate attitudes and practices specific to each culture (Kankanhalli, Tan, Wei, & Holmes, 2004, pp. 183–184). Likewise, divergent cultures can explain significant differences in terms of financial sector development and related issues Khan et al. (2021, p. 8). These cultural effects can be observed, for example, in the different discourses used by the authors to address the existence of anomalies in financial systems considering America and Europe.

When it comes to America, Azevedo et al. (2021, p. 2) addressed corruption and embezzlement as a traditional and significant problem in the country of the survey data, such that not even social programs are free from criminal actions, and Parnes (2020, pp. 292–293) stated that the study of anomalies in conjunction with analysis of macroeconomic indicators can help not only in understanding the research topic, but also in understanding past events. With regard to Europe, Baker et al. (2017,

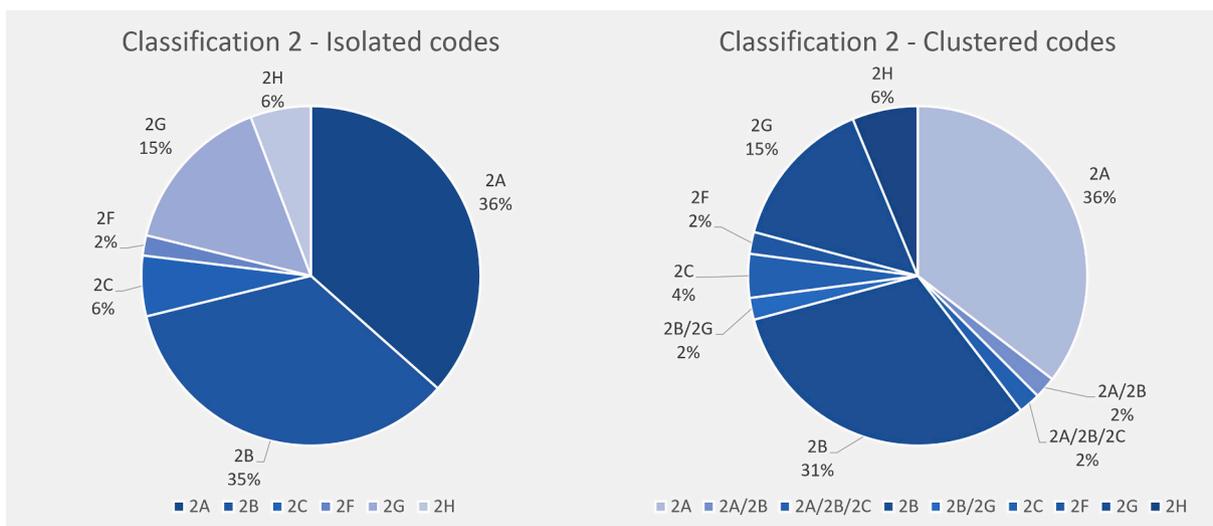


Fig. 3. 2 - Geographical Research Location.
Notes: 3A - Qualitative; 3B - Quantitative; 3C - Conceptual; 3D - Modelling; 3G - Single case study; and 3H - Multiple case study.

p. 20) discussed the need to examine the failures and deviations that caused the failure of a bank internationally known for its expertise and high profitability in order to generate directions for future research, and Ardizzi et al. (2018, p. 106) addressed the importance of using anomaly detection tools in conjunction with the premises of an Italian strategic intelligence unit. Thus, the discourse of research carried out in American countries reinforces that such problems are established in countries and are constant, while the discourse of European countries builds arguments that reinforce local strategies and expertise, being less destructive in relation to the existence of anomalies in the financial systems. Such approaches can be derived from cultural effects as well as the frequency of surveys across continents.

Still with regard to cultural effects, it can be seen that the lowest classification category, in statistical terms, was “2F - All continents”, with only one article classified (González-Sánchez, 2021, p. 7). This means that there are not many articles that develop comprehensive studies, considering a greater diversity of countries and regions around the world. The data set used in the articles usually includes isolated national contexts in which analysis are carried out on the effectiveness of methods for detecting anomalies in specific financial systems in a location, without any form of geographic comparison, as in the studies developed by Lin et al. (2018, p. 159), Druică et al. (2018, p. 77) and Mamede and Malaquias (2017, pp. 47–48). Furthermore, there were no articles classified as “2D - Africa” and “2E - Oceania”, which only highlights the scarcity of research on anomalies not only of comparative mode, but also isolated surveys on all continents. Countries in Africa, for example, have a vast population that does not have access to financial services, but with technological advances in the field of cell phones this situation could change (Zaffar et al., 2019, p. 62) and, therefore, studying correlations between population behaviour and the detection of anomalies in financial systems may indicate different insights on this topic. In this way, it is possible to identify the third and fourth research gaps, namely:

G₃: The lack of articles using data from Africa and Oceania, or even considering all continents in a single survey.

G₄: Characterize human behaviour in different cultures in line with the existence of anomalies in financial systems.

It is worth noting that understanding the relationship between geographic location and the cultural effects arising from it can help in a deeper understanding of why America and Europe have obtained more studies on the subject in question, as well as identifying the causes for not having so many studies in other continents. These gaps become an important opportunity for future research, in order to answer whether there is a correlation of cause and effect between different cultures and the adoption of criminal practices that result in anomalies, or, between

cultures and ways to prevent and address such occurrences.

4.3. Research method

Each context and research question relates to different and appropriate methods of analysis (Lo, Rey-Martí, & Botella-Carrubi, 2020, p. 221). An analysis of the research method of articles developed in the field of detecting financial anomalies makes it possible to identify common methods used in this area, as well as those that are not much explored and, thus, configure a research gap that could be studied and developed in future research. In Fig. 4 is presented the statistical classification and coding data of the selected articles, so that it is possible to visualize the different research methods chosen for these researches.

The most adopted coding in the classification was “3B - Quantitative”, so that the articles develop research methods that involve algorithms and numerical approaches to explore quantities of data. This result is in line with the studies by Bose et al. (2011, p. 557), which addressed how the development of information technology and decision sciences enable the use of quantitative research models, as they enable the automation and reduction of manual operations during the fraud detection process, for example. Thus, the quantitative research method is shown to be the most relevant in this theme, whether used in isolation or in combination with other methods, such as the studies of Forough and Momtazi (2021, p. 4), Chen et al. (2020, p. 3), Misra et al. (2020, pp. 258–259), Rtayli and Enneya (2020, p. 3), Yan et al. (2020, p. 118), Carta et al. (2019, p. 16) and Ardizzi et al. (2018, pp. 108–109) which, in addition to receiving 3B coding, were also coded as “3D - Modelling”, when developing a new detection model based on quantitative techniques.

However, as much as the quantitative method is predominant, it can be seen from the results that there is a lack of conceptual articles for this area of knowledge, that could make it possible to standardize common practices. Among the selected articles, those categorized as “3C - Conceptual” were literature reviews (Al-Hashedi & Magalingam, 2021, pp. 5–6; Houdek, 2020, p. 1; Pourhabibi et al., 2020, pp. 2–3), as this research method enables the abstraction of relevant aspects of a domain, in addition to identifying requirements and research objectives (Maass & Storey, 2021, p. 2), as well as conceptual research. Nonetheless, the results indicated that, in addition to these three literature review articles, there are no articles that use conceptual models. This type of research is critically important for academia, in addition to showing itself as a need for practitioners, as it defines methods, tools and theories relevant to the construction and development of research by creating a system for implementing global practices that guarantee quality applications (Trujillo, Davis, Du, Damiani, & Storey, 2021, p. 1). Thus, to the



Fig. 4. 3 - Research Method.

Notes: 4A - Banking; 4B - Actuarial; 4C - Financial market; 4D - Social security; 4E - Financial accounting; and 4G - Not applicable.

best of our knowledge, this characterizes the fifth research gap, namely:

G₅: The lack of conceptual studies in the literature that facilitate the comparison between articles within the same theme and research approach.

This gap, however, is not isolated, and signals the existence of the sixth research gap related to this theme, namely:

G₆: The lack of a specific indicator to verify the effectiveness of models and techniques developed for detecting financial anomalies, as well as a public database for standardized testing and measurement.

If there were the construction of theoretical references, an indicator and a public database with standardized data sets of different sizes, it would be possible for the authors to apply tests on their detection models in order to enable a concrete comparison of efficiency and accuracy identification of anomalies. This would allow, therefore, the measurement and comparison of academic solutions within a more plausible and real research pattern.

4.4. Application area

This classification category was created in order to identify the financial areas in which fraud, capital embezzlement and data tampering are carried out, in order to better understand which sectors are most affected and which are least affected according to the amount of research carried out in these areas. The statistical data referring to this classification can be visualized through Fig. 5.

The most recurrent coding among the selected articles was “4A - Banking”, comprising 27 articles representing 52.94% of the total, thus representing more than half. To Patil et al. (2018, p. 385), the banking area is a very important sector nowadays, since almost all individuals need to deal with banking issues, whether physically or virtually, but this virtualization of systems has made banking very vulnerable to various attacks. Nikkel (2020, pp. 1–2) stated that these new financial systems facilitate the transfer of money between individuals, payments for services and replace the need for physical money, and because of this, criminals abuse these financial technologies to promote fraud, extortion and money laundering. This justifies the great need for research into methods and techniques to detect anomalies on this financial area, in order to identify and mitigate such criminal activities.

It is worth noting that Climent et al. (2019, p. 885), in their research, talked about how banks are an important part of the economy of countries, and crises in their systems can destabilize financial markets and harm global economic growth. To Canhoto (2021, p. 441), as they keep records of their customers' identity and behaviour, it is possible to use this data set to detect and prevent criminal activity. In this context,

Carmona et al. (2019, p. 304) stated that this prevention approach helps in the process of trusting these customers with the banks, maintaining their reputation by avoiding bank failure. Thus, the significant number of models found to detect anomalies in banking systems explains a great need for security in this sector, which, if threatened, could interfere with the economy globally.

Considering such trust issues, it is worth emphasizing here that the significant results of research in banking can be related to the cultural dimensions addressed by Tung and Quaddus (2002, p. 190), causing variations in results according to the behaviour of customers in different locations. Picoto and Pinto (2021, p. 620), in their studies, identified that countries that use mobile banking the most, for example, are those that normally have not yet adequately developed their banking systems, which can influence data protection and enable criminal actions such as financial fraud. Therefore, there is the seventh research gap, namely:

G₇: Identify whether cultural effects can also influence application areas, leading some countries to need more rigid prevention systems than others.

On the other hand, the lowest classification encoding used was “4D - Social Security”, representing only 1.96% of the articles. To Azevedo et al. (2021, p. 2), corruption and fraud in this area can harm the beneficiaries' receipt of money. Thus, for future research within this theme, there is the eighth gap, namely:

G₈: Identify why Social Security is a less explored area and whether this is related to cultural effects or whether this sector only represents a low priority when it comes to developing anomaly detection systems.

4.5. Focus

The Focus category goes deeper into application areas in order to more precisely identify where financial abnormalities are concentrated. The statistical results can be observed through Fig. 6. The expressive numbers of articles classified as “5A - Credit card” and “5D - Financial transactions” justify the high result found for the ratings of banking in the “Application area” subsection, further explaining the need for high security in the banking sector.

The encoding that most stands out in this classification is “5A - Credit card”, representing 26.92% of the total number of articles. This is in line with the search results of Rubio et al. (2020, p. 3) that when it comes to financial contexts, anomaly detection model applications are mainly focused on fraudulent credit card transactions. To Forough and Momtazi (2021, p. 1) and Jurgovsky et al. (2018, p. 234), this result is due to the growth of the electronic form of payments that enables credit card fraud by creating opportunities for fraudsters to carry out criminal actions

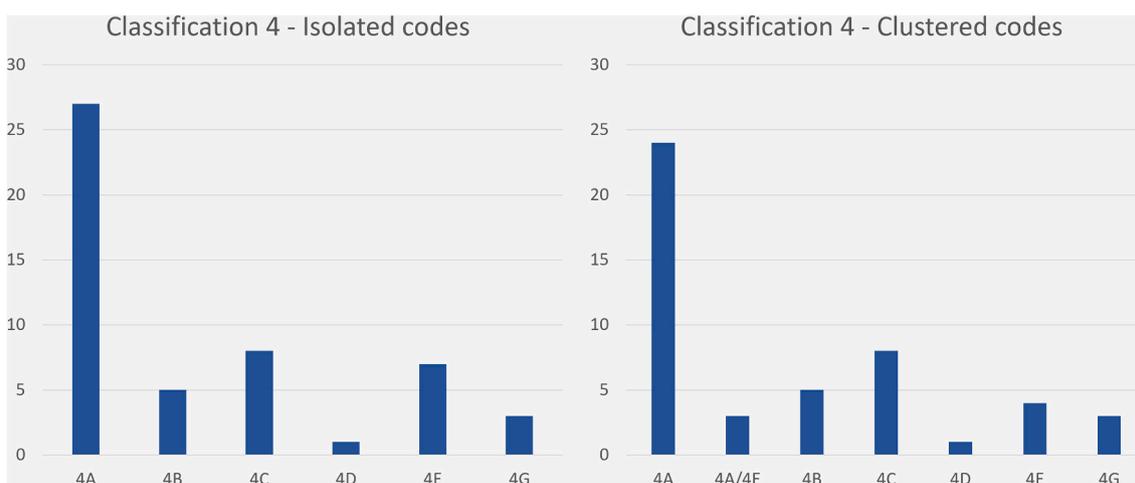


Fig. 5. 4 - Application Area.

Notes: 5A - Credit card; 5B - Insurance; 5C - Financial welfare benefits; 5D - Financial transactions; 5E - Financial statements; 5F - Stocks and commodities; 5G - Others; and 5H - Not applicable.

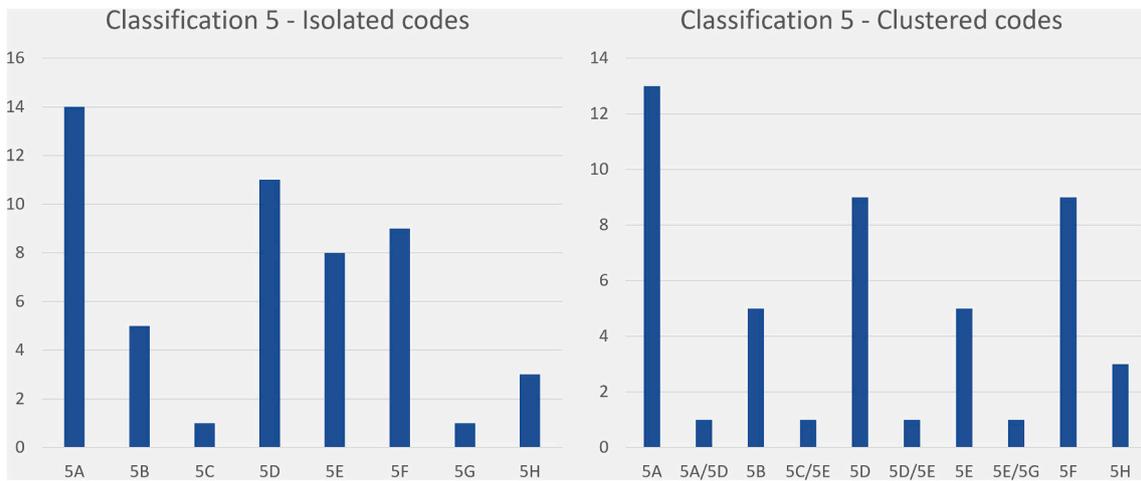


Fig. 6. 5 - Focus.
Notes: 6A - Real; 6B - Synthetic/Artificial; 6C - Others; and 6D - Not applicable.

(Misra et al., 2020, p. 254). In addition, the large volume of online financial transactions makes manual verification for anomalies unfeasible (Gianini et al., 2020, p. 549), which creates a series of setbacks when it comes to the development of methods for detecting abnormal data.

To Forough and Momtazi (2021, p. 9), one of the biggest challenges in identifying credit card fraud is the need for a model that is capable of working with real-time data. Carta et al. (2019, p. 13) presented the need to have this same real-time perspective, since, currently, the commercial dynamics allow the consumption of goods and services through any device connected to the internet. Lucas et al. (2020, p. 393), on the other hand, stated that the problem lies in the fact that many studies adopt credit card transactions as isolated events, rather than relating them as sequences. Carneiro et al. (2017, p. 91) signalled yet another challenge: more and more fraudsters are manipulating the data in order to appear to be in agreement with the data referring to the legitimate behaviour of the cardholders. Therefore, the divergent opinions, in fact, converge to the great need for research on this topic and explain the result obtained in this literature review.

In addition, a classification that is worth mentioning is “5B -

Insurance”, which was little used, covering only 9.62% of the total articles. According to Stripling et al. (2018, p. 13), fraudsters see insurance fraud as low risk and, at the same time, as a way to receive high monetary values, being considered by them as a more “safe” crime to commit. Also, Severino and Peng (2021, p. 1) stated that the insurance market is very profitable and moves a large amount of money per year. To Wang and Xu (2018, p. 87), the optimistic development of this market will allow even more capital to enter this form of industry. Thus, the lack of a significant number of researches with this approach signals the ninth gap, namely:

G₉: Why don't insurance fraud anomalies get as much attention when it comes to studying financial anomalies?

4.6. Dataset used

The analysis of the type of data set used to promote efficiency tests of the anomaly detection models developed in the articles aims to understand which would be the best data for measurement: real or artificially simulated. The statistical results of this category are indicated in Fig. 7.

The most used encoding was “6A - Real”, covering 84% of the

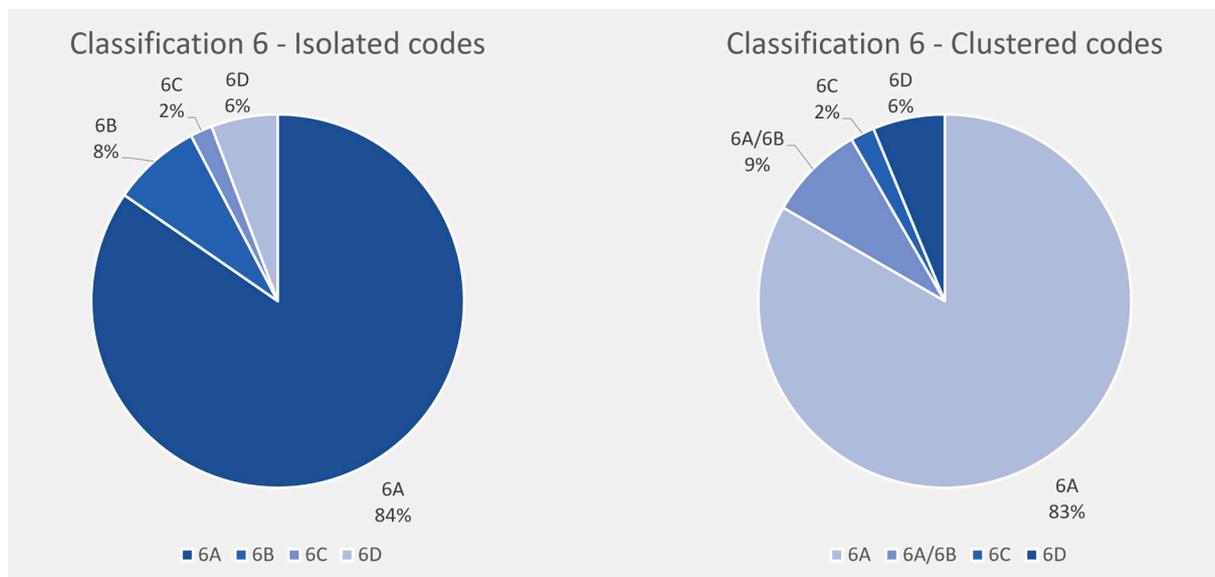


Fig. 7. 6 - Dataset Used.
Notes: 7A - Restricted access; 7B - Unrestricted access; 7C - Others; and 7D - Not applicable.

selected articles. This result reinforces the result obtained by Pourhabibi et al. (2020, p. 8) that most research studies on fraud detection models use real world data as their testing platform, while synthetic data only serve to simulate more specific scenarios and circumstances. To Severino and Peng (2021, p. 1), real data provide a significant advantage when compared to simulated data, both in terms of evaluation and in terms of practical applicability in the decision-making process. Events that take place in the real world cannot be considered monotonous (Andor & Bohák, 2017, p. 47), because they suffer variations that cause a high degree of data imbalance (Carta et al., 2019, p. 21) and, therefore, they usually need many adjustments and treatments (Badal-Valero et al., 2018, p. 28). This is necessary since the costs of not identifying an anomaly or fraud are very high when dealing with a real data set (Sá et al., 2018, p. 25). Nonetheless, despite the many challenges, the results indicate that the use of real data is still the most used to measure the effectiveness of models and techniques for detecting financial anomalies.

However, it is worth noting that the amount of data used in the tests varies from article to article, there is no standardized amount to serve as a basis for comparison and, therefore, there is no way of knowing whether a technique that is very effective in a small data set would be viable for the required processing of large databases. When it comes to processing and analysing a large volume of data that normally does not have a structured nature and size, such as, for example, the so-called Big Data, it is not possible to use traditional methods (Trujillo et al., 2021, p. 2). Existing algorithms for detecting anomalies are limited by efficiency and dimensionality problems, especially when processing high-dimensional sequence databases (Liu, Lin, He, & Cai, 2010, p. 270). This is even worse when it comes to real-time data processing, as it requires skill to interpret and manage the so-called 4 V's of Big Data (Volume, Velocity, Variety and Veracity) that are continuously generated, so that it is possible to provide adequate informational support for decision-makers (Fertier, Barthe-Delanoë, Montarnal, Trupitel, & Bénaben, 2020, p. 1). It is worth pointing out, then, the tenth research gap, namely:

G₁₀: Develop research models that take into account dimensionality problems and are thus able to work with voluminous and varied data in real-time.

As much as many systems that aid in the decision process have been developed, most of the activities were directed towards the construction of models that do not involve real-time data control (Chaturvedi, Hutchinson, & Nazareth, 1993, p. 213). Among the selected articles, only three developed techniques for detecting anomalies in real-time financial systems (Forough & Montazi, 2021, p. 10; Rubio et al., 2020, p. 15; Patil et al., 2018, p. 385) and were able to identify patterns in recurring transactions. In addition, many studies select data sets from the UCI Machine Learning Repository, which have low dimensionalities, with less than 100 features, while circumstances referring to the real world contain very high dimensional data, with thousands of features (Tsai & Chen, 2014, p. 80), such as the paper by Chen et al. (2020, p. 3). This analysis reveals the difficulties arising from the real-time decision process (Chaturvedi et al., 1993, p. 214), since the developed financial anomaly detection models are not necessarily capable of meeting the need to identify fraud, deviations or data tampering so that decision support systems can achieve the most appropriate choices, managing to detect and prevent criminal actions. Based on this, future research can address dimensionality problems in order to develop techniques that enable the processing of a large volume of data.

4.7. Availability

The availability of data used for research can enable a greater understanding of the study to be developed. Thus, this classification was made in order to understand whether the data used to test the anomaly detection models are restricted and how this can influence the interpretation of the results. Statistical data related to this classification can

be seen in Fig. 8.

The most frequent encoding was “7B - Unrestricted access”, used for more than half of the articles. However, the encoding “7A - Restricted access” also included a significant percentage of the articles, that is, 33%. Thus, even though most articles among the selected ones have used available data to check the consistency of the developed model, there is still a significant number of articles that do not provide their data, mainly due to the need to keep private data to prevent exposing companies, for example. Nonetheless, concerns about privacy intensify the challenges of data integration, sharing and analysis (Wimmer, Yoon, & Sugumaran, 2016, p. 51). To Canhoto (2021, p. 444) it's difficult to get access to companies' financial data because of the secret nature of this industry. Pourhabibi et al. (2020, p. 8) stated in their research that there are regulations that prohibit the transmission and distribution of confidential data, whether personal or financial, making data sharing difficult and promoting the use of mathematical assessment measures to detect fraud. An example of this is the search for Singh and Best (2019, pp. 3–4) that, despite using real data, they were obtained from banking transactions of a large financial entity and, thus, the information and specifics were removed from the research due to confidentiality.

This situation coincides with the sixth gap already discussed in the subsection “Research method”, that is, there is no standardized basis to verify the efficiency of the anomaly detection models developed. However, the non-availability of data is another factor that hinders the measurement, characterizing the eleventh research gap, namely:

G₁₁: The lack of data available for understanding the development of detection models in research.

This gap makes it not possible to fully interpret the studied technique, as well as making comparisons between research and models impossible, making it difficult to understand the effectiveness of identifying anomalies in financial systems, either due to lack of data or due to differences in size between the data. Thus, the need for a standardized database that enables such a comparison to be made is critical, especially considering that, in many studies, the data used for testing are not even available, which can make the evolution of this research topic imprecise and deficient. Future research, therefore, must consider structuring theories and tools that enable a more relevant and real research process.

4.8. Technique used

This classification category was structured in order to identify trends in techniques used in models for detecting anomalies in financial systems, as well as to indicate which techniques are being less used. This is done so that it is possible to identify not only research opportunities, but also the most pertinent and effective techniques for this field of research according to the frequency of use, considering that the investigation of common and unusual behavioural patterns on the financial area is extremely important for the efficient and safe control of the functioning of the systems (León, Barucca, Acero, Gage, & Ortega, 2020, p. 1). Statistical data regarding the classification and coding process of articles was presented in Fig. 9.

The most frequent coding among the articles was “8A - Machine learning”, covering 45.71% of the articles. The choice of this technique may be related to the fact that research methods involving machine learning were and are widely used when dealing with problems involving financial decisions (Kuzey, Uyar, & Delen, 2014, p. 128), and also with the large amount of open source frameworks and algorithms to develop the models (Khan & Al-Badi, 2020, p. 573), which enables the authors' access to the programming libraries of the techniques to be modelled. The research results of Delen (2010, pp. 501–502) confirmed the superiority of methods of machine learning to produce better ways to prevent situations. The learning capacity arising from the machine learning process enables adaptation and improves the performance of systems (Ali, Shamsuddin, & Ismail, 2012, p. 568), which justifies the authors' choice to build models for detecting deviations in data patterns.

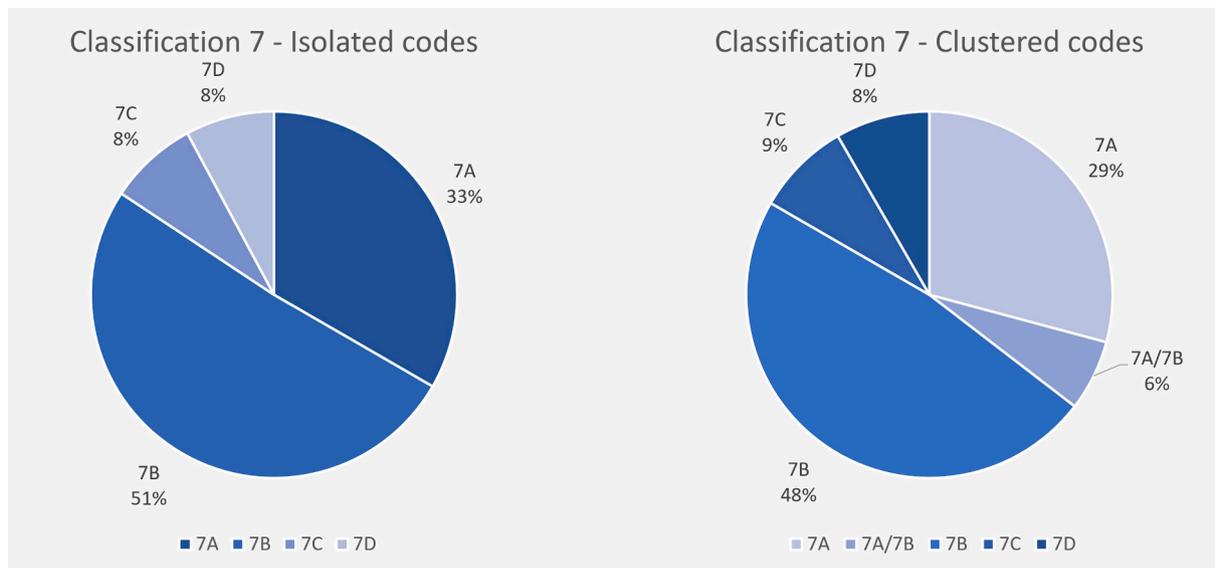


Fig. 8. 7 - Availability.

Notes: 8A - Machine learning; 8B - Statistics; 8C - Bayesian statistics; 8D - Benford's law; 8E - Graph methods (visual); 8F - Computer simulations; 8G - Others; and 8H - Not applicable.



Fig. 9. 8 - Technique Used.

Notes: 9A - One; 9B - Two; 9C - Three; 9D - Four or more; and 9F - Not applicable.

To León et al. (2020, p. 1), machine learning techniques lead the development of automated tools and techniques capable of controlling and supervising payment systems and the various infrastructures related to the financial market. Carmona et al. (2019, p. 308) discussed the fact that approaches that use machine learning techniques are important for the areas of fraud detection and anomalous events. Canhoto (2021, p. 442), in the same line of reasoning, stated that exploring the potential of machine learning techniques can be of great help in the global fight against crime, as they signal anomalies and outliers present within a data pattern (Baesens, Höppner, & Verdonck, 2021, p. 6). Carneiro et al. (2017, p. 91), on the other hand, talked about how, as much as machine learning algorithms improve the methods of detecting fraud, the use of techniques that are really practical and relevant to this research topic are rarely reported.

Furthermore, the authors disagree on which machine learning method is the most efficient for detecting anomalies. To Wang and Xu (2018, p. 87), the techniques of deep neural networks have a performance superior to other widely used machine learning models, such as random forests and support vector machine. Forough and Momtazi

(2021, p. 2) stated that, when it comes to real-life applications, the best method to be used is random forest. Badal-Valero et al. (2018, p. 33), however, talked about how random forest is actually the best approach only when considering the application of balancing techniques, and neural networks have superior performance only when they involve imbalanced data sets. In any case, the results indicate that machine learning techniques are likely to detect suspicious activities in real time, and this is an important factor in reducing economic losses in companies (Yan et al., 2020, p. 11).

It is worth noting that the results obtained from this classification present several combinations of codes that include the encoding “8A - Machine learning”, which signals the importance of using this combined technique even when the authors choose to use other detection models. However, it is worth emphasizing here the techniques developed based on Benford's Law for detecting deviations in data, symbolized by the encoding “8D - Benford's law”, which was used by 8.57% of the selected articles. It is typically developed on a hybrid model with statistical techniques that improve its results (Azevedo et al., 2021, p. 3; Druică et al., 2018, p. 78, Shi et al., 2018, pp. 879–880), being an important tool

for detecting financial fraud when it comes to observing the compliance of numbers that occur naturally in company systems (Bhattacharya, Xu, & Kumar, 2011, pp. 576–577). An argument about the unreliability of fraud detection by Benford's Law is based on the fact that potential criminals may be aware of the need for data to be distributed accordingly and, because of this, it is important that authors when developing their detection models consider alternative methods that can complement the distribution analysis (Shi et al., 2018, p. 879). Benford's Law, therefore, when used correctly, can be considered a powerful tool to identify anomalies and suspicious activities (Azevedo et al., 2021, p. 11).

However, this technique is not only infrequent among recent articles involving approaches to detecting anomalies in financial systems, but its use has also been poorly combined with machine learning, the most frequent technique, widely used in conjunction with other techniques under analysis. The clustered encoding “8A/8D” was used in only one article, in which four machine learning classification models were associated with Benford's Law number distribution conformity to identify patterns of deviation that could signal criminal money laundering activities (Badal-Valero et al., 2018, p. 24). The results obtained indicate that the hybrid model can identify anomalies and also assess compliance with the distribution of significant numbers of Benford's Law, considering the most appropriate machine learning algorithms for each situation (Badal-Valero et al., 2018, pp. 31–33). In this way, it is possible to identify the twelfth gap, namely:

G₁₂: What are the advantages of combining different techniques with machine learning in the anomaly detection process?

4.9. Number of strategies

The number of strategies refers to the number of techniques used in the same article to promote the development of a model for detecting anomalies in financial systems that is effective and relevant for studies on this topic. This classification reinforces the arguments already presented in the subsection “Technique used”, and clarifies the number of articles that perform hybrid models, combining detection techniques. The statistical results obtained from this classification were presented in Fig. 10.

Models that integrate two or more techniques have the potential to increase the quality of prediction performance (Lee, Han, & Kwon, 1996, p. 63; Liang, Chandler, & Han, 1990, p. 392). This is because hybrid approaches that combine different techniques are able to overcome

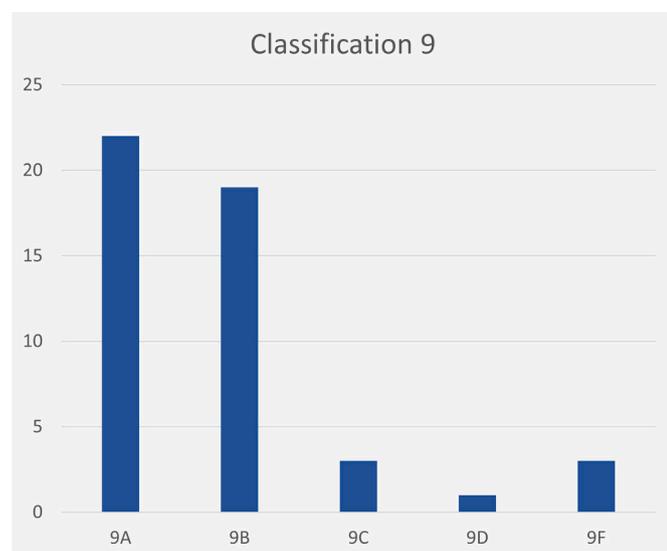


Fig. 10. 9 - Number of Strategies.

Notes: 10A - Health sciences; 10B - Life sciences; 10C - Physical sciences and engineering; and 10D - Social sciences and humanities.

limitations that would occur when using an isolated technique, and also manage to unite the advantages of each in order to improve the developed model (Kundu, Pal, & Mandal, 2020, p. 1; Ince & Trafalis, 2006, pp. 1060–1061). Considering the issue of anomaly detection, as the large amount of data from financial systems makes it impossible to manually detect abnormalities (Gianini et al., 2020, p. 549), the combination of techniques can also act strategically by promoting human-computer articulation and collaboration, in a hybrid model with potential to further enhance the capability of human decision-makers (Bayrak, McComb, Cagan, & Kotovsky, 2021, p. 1). Thus, the number of strategies used in a single model can increase the chances of success in detecting financial anomalies, both in computational terms and in terms of human management of the system.

Considering this, most articles were classified with the encoding “9A - One”, contemplating 45.83%, which means that it is very frequent that authors use only one strategy to build their study models, not considering the use of combined techniques. However, the encoding “9B - Two” also considers a significant percentage of selected articles, 39.58%, and the authors of such articles reveal in their studies the benefits arising from the use of hybrid models. To Baesens et al. (2021, p. 1), the investigation of suspicious activities, when considering, for example, the perspective of machine learning, can be seen as a problem of binary classification and, because of that, many techniques and models can be applied. Carneiro et al. (2017, p. 93), along the same line of reasoning, state that, for the model to be more effective, it is possible to combine solutions with supervised and unsupervised techniques. According to Jurgovsky et al. (2018, p. 244), this is because combinations can improve the accuracy of the detection model.

Furthermore, a hybrid model is capable of covering a fuller spectrum of disruptive activities (Zhai, Cao, Yao, Ding, & Li, 2017, p. 27) in order to improve the pattern and anomaly prediction process (Rtayli & Enneya, 2020, p. 2). With the combination of techniques, the accuracy is increased, because it is more of an abnormality recognition model in a single detection process (Chen et al., 2020, p. 3). In this way, models and techniques can strategically complement each other (Baesens et al., 2021, p. 2) to combat fraud, deviations, adulteration and inconsistency of data, by achieving more stable performances suitable for investigating anomalies (Zhai et al., 2017, p. 38). This characterizes the thirteenth research gap, namely:

G₁₃: The lack of more hybrid financial anomaly detection models that explore diverse combinations of techniques in order to identify which models are more efficient in detecting anomalies in financial systems.

Thus, future research may further examine the conduct of studies involving the mixing of techniques and models, aiming to increase detection accuracy levels strategically, and further clarify the benefits arising from the use of hybrid models in the financial sector.

4.10. Journal's subject category

The study of Finance is part of the field of knowledge of applied social sciences. Still, this classification aims to understand the categories in which the articles referring to the application of anomaly detection models in financial systems are inserted and if, really, because they are articles which encompass financial matters, these include the category of Social Sciences. This is because the citation analysis reveals that it is more likely when conducting scientific research in a specific base field that the citation of articles and journals is contained, mainly, within that base (Holsapple, Johnson, Manakyan, & Tanner, 1995, p. 360). Statistical data from the rankings can be seen in Fig. 11.

From the results, it is possible to observe that, although the number of articles with the coding “10D - Social science and humanities” is expressive, comprising 40% of the selected articles, this category is not the most frequent. The most common encoding was “10C - Physical sciences and engineering”, covering 56% of the articles. Therefore, studies about financial anomalies are being developed and published,

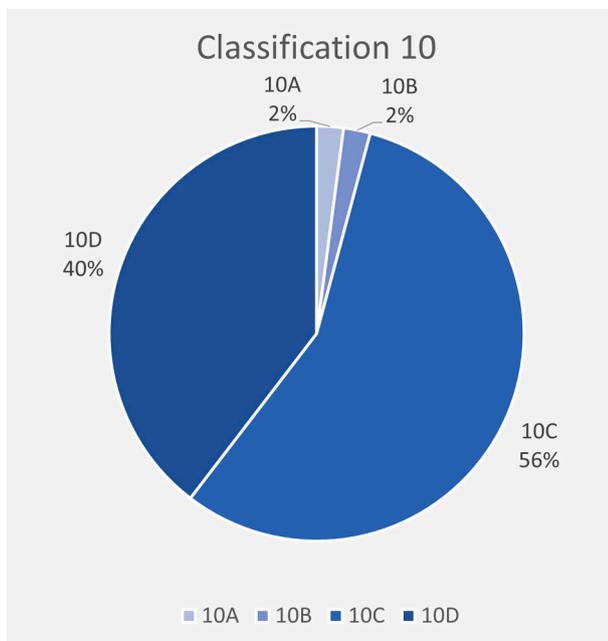


Fig. 11. 10 - Journal's Subject Category.

mainly, in journals related to physical sciences. This may be happening, since the articles selected for this literature review promote the construction of models, often quantitative, that detect data that diverge from a normal pattern, using computer simulations (Singh & Best, 2019, p. 4; Patil et al., 2018, p. 387), machine learning (Severino & Peng, 2021, p. 3; Climent et al., 2019, p. 888), Benford's Law (Arboleda et al., 2018, pp. 13–14; Druică et al., 2018, p. 77), statistics (Ardizzi et al., 2018, p. 108; Mamede & Malaquias, 2017, pp. 49–50), among other techniques that are normally related to computing and engineering. However, despite the fact that this issue of detecting anomalies has such characteristics, the studies still focus on the financial area, which characterizes the fourteenth research gap, namely:

G₁₄: Deepen an analysis of why a more humanistic view is not used to deal with abnormalities in finance, in addition to identifying publication patterns within the categories of Social Sciences and Physical Sciences, in order to investigate correlations between areas of knowledge.

Furthermore, despite the low percentages, the encodings “10A - Health sciences” and “10B - Life sciences”, presented classified articles (Asha and Suresh Kumar, 2021, p. 35; Yan et al., 2020, p. 115), even in the case of them being journal's unlikely categories considering the topic under analysis. Thus, the fifteenth research gap can be identified here, namely:

G₁₅: Develop researches to understand whether there is a need to further study the detection of financial anomalies in the areas of Health and Life sciences.

Future research may contemplate the development of articles within these sectors of studies, considering that anomalies can manifest themselves in different ways in different financial systems, including those related to the areas of health and the field of biological sciences.

5. Conclusions

Disruptive and anomalous behaviours in digital business financial systems arising from diversion, adulteration and data fraud seriously harm the capital market and the economy in general due to the spread of misleading information that generates negative consequences to different stakeholders, and thus, makes unfeasible the efficient allocation of economic resources (Hou, Wang, & Ma, 2021, p. 97), causing difficulties to the enhanced decision-making process. Anomalies, inconsistencies in the regular execution of data (Conti & Naldi, 2008, p.

420), can cause inefficiency in performance and maintenance, allowing potential errors in the systems to occur (Mukherjee, Gamble, & Parkinson, 1997, p. 232).

The current pandemic has rapidly changed technological scenarios, affecting business practices and data behaviour (Shaikh et al., 2020, p. 2). These changes have increased the digital threat, which may involve a range of actors, such as cybercriminals and victims, whether these are people or organizational systems (Wessels et al., 2021, p. 1). Therefore, the business model needs to be continuously maintained (Karki & Porras, 2021, p. 2) to create competitive advantage and to enhance the company's value to stakeholders (Marikyan et al., 2022, p. 1). Therefore, it is critically important to study tools that can identify the abnormalities, and thus, develop and improve the functional aspect of making decisions about them.

Based on this, the main objective of this study was to verify, analyse and systematize articles from the last five years about methods and techniques for detecting anomalies used in financial systems, aiming to understand and indicate the existing scientific knowledge base as well as gaps for future research to create a research agenda. This information is summarized in Fig. 12. For this, we used the five-step research method Jabbour (2013, p. 145) and Lage Junior and Godinho Filho (2010, p. 14) proposed: a systematic research of articles, develop a classification structure, classify pre-selected articles, design a results framework and discuss results to propose suggestions for future research.

Thus, using 10 categories, we identified a classification framework of 48 papers as relevant in research in the ScienceDirect® database, which was defined through encodings to systematize in a standardized way the existing knowledge about the subject under study. By adopting this method of research, we overcame the limitations found in earlier literature and uncovered combinations of features present in the papers that allow future researchers to improve their studies in a more accurate and informed way (Mallinguh & Zéman, 2020, p. 163), enabling the evolution of the knowledge base about anomaly detection methods and approaches, and thus offering benefits through a finance/economic lens.

Therefore, this paper contributes to developing the field of studies on detecting anomalies in financial systems. From the classification framework created and the statistical results obtained, it was possible to identify 15 main research gaps distributed throughout the text. Each of these gaps is important for scientific advancement in future research. However, the five-step method used, by defining research trends, made it possible to verify three gaps that stand out the most, as they are not present or not taken into account in most papers, but have characteristics relevant to the evolution of anomaly detection models. They are the lack of a common database, in which it is possible to use standardized data to promote effective comparison and measurement of detection models; the need to promote studies that test their models in databases of different sizes, seeking to mitigate the problems arising from the dimensionality of data; and the lack of indicators for measuring the effectiveness of the chosen technique for detecting abnormality patterns to further improve the comparative process between the models.

These three gaps are more relevant for the present research since, with the growth of digital business and the subsequent increase in financial transactions, bank transactions, and the flow of financial statements, among others, the anomaly detection approaches and the studies developed to test, adapt, and improve the techniques must be able to keep up with the digital transformation (Konopik et al., 2022, p. 1). This is because the rapid growth of these technologies in the financial sector contributes to the large accumulation of real-time data flowing in a network (Habeeb et al., 2019, p. 289), making efficient and scalable financial anomaly detection methods necessary for work with large data sets, considering the data's volume, velocity, variety and veracity (Fertier et al., 2020, p. 1).

Thus, the critical importance of developing research on the three gaps mentioned can be inferred from this study. This is because when the dimensions of the actual data are much larger than the sample chosen for a detection model, the efficiency becomes limited (Zhang, Zhou,

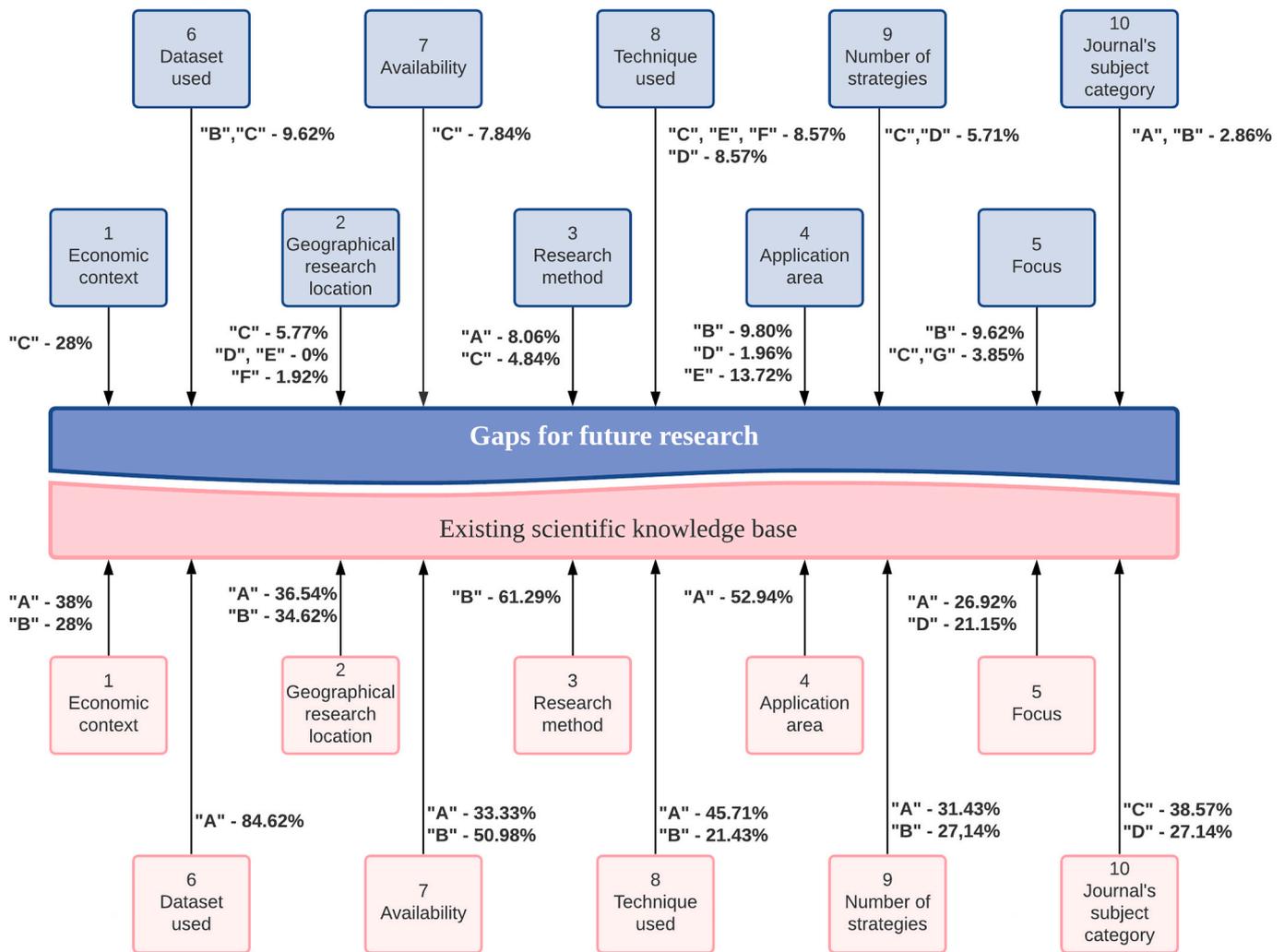


Fig. 12. Suggestions for research agenda.

Guo, & Zhu, 2021, p. 144), and the inability to compare models makes it difficult to interpret which ones are really more efficient and which are more restricted (Liu et al., 2010, p. 270) such as the lack of indicators that allow this measurement. Therefore, future studies can use the benefits of remedying such gaps to further enhance the way in which anomalies are dealt with in the contemporary world, consolidating the best techniques to mitigate illegal acts and transactional or registration errors and enabling the healthy evolution of financial systems in the digital world.

The scientific development and evolution of articles considering such gaps is critically important to improve decision-making systems, especially in financial terms. Thus, it would be interesting for editors to make a call for papers to promote constructing a special issue considering such proposals for the research agenda, thus emphasizing the research urgency these gaps cause that could be remedied. Therefore, our proposed framework and the gaps arising from the statistical analysis of the results obtained indicate that this area of study can still be much explored to improve the anomaly detection models, increasing the precision of the results and providing the construction of structures that create the necessary support for the decision-making process.

However, this literature review has some limitations, among which we can highlight the fact that only the ScienceDirect® database was used to survey the articles. Furthermore, it is possible that some relevant studies have been omitted, considering the keywords and search filters used. In this way, future research can explore the detection of anomalies in other financial digital businesses that did not appear among the

studies analysed in this paper, such as blockchain, which can create resources for organisations by changing the way involved parties connect in digital exchange. (Marikyan et al., 2022, p. 1) or consider the internet of things and the interconnection between physical and digital systems (Wessels et al., 2021, p. 1).

In addition, as it seeks to identify common traits from the classification, reading, understanding, and perception of research, the five-step technique becomes too complex to be completely automated or quantified, except when considering the results provided by the method. In this way, future research can complement analysis carried out based on other perspectives, especially in the case of quantitative research, such as the study developed by Alamsyah et al. (2013, p. 3), which discusses the existing connections within his subject of study, creating a taxonomy to recognize the limits of its theme and reinforcing the mapping of research efforts and interests through the development of a literature review. Still, with the number of articles selected and the results obtained, future researchers can rely on this literature review to identify the most relevant research gaps to be developed and analysed.

Declaration of Competing 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.

References

- Abreu, E. S. D., Kimura, H., & Sobreiro, V. A. (2019). What is going on with studies on banking efficiency? *Research in International Business and Finance*, *47*, 195–219.
- Abu, F., Gholami, H., Saman, M. Z. M., Zakuan, N., Sharif, S., & Streimikiene, D. (2021). Pathways of lean manufacturing in wood and furniture industries: a bibliometric and systematic review. *European Journal of Wood and Wood Products*, *79*, 753–772.
- Ahmed, A., Deokar, A., & Lee, H. C. B. (2021). Vulnerability disclosure mechanisms: A synthesis and framework for market-based and non-market-based disclosures. *Decision Support Systems*, *148*, Article 113586 (Online first).
- Ain, N., Vaia, G., DeLone, W. H., & Waheed, M. (2019). Two decades of research on business intelligence system adoption, utilization and success - A systematic literature review. *Decision Support Systems*, *125*, Article 113113 (Online first).
- Alamsyah, A., Rahardjo, B., & Kuspriyanto. (2013). Social network analysis taxonomy based on graph representation. In R. R. M. A. Magetsari (Ed.), *The 5th Indonesia international conference on innovation, entrepreneurship, and small business (IICIES 2013)*, IICIES (pp. 1–11). Bandung, West Java, Indonesia: IICIES.
- Albayati, H., Kim, S. K., & Rho, J. J. (2020). Accepting financial transactions using blockchain technology and cryptocurrency: A customer perspective approach. *Technology in Society*, *62*, Article 101320 (Online first).
- Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying Data Mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, *40*, Article 100402 (Online first).
- Ali, W., Shamsuddin, S. M., & Ismail, A. S. (2012). Intelligent web proxy caching approaches based on machine learning techniques. *Decision Support Systems*, *53*, 565–579.
- Amani, F. A., & Fadlalla, A. M. (2017). Data Mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, *24*, 32–58.
- Amui, L. B. L., Jabbour, C. J. C., de Sousa Jabbour, A. B. L., & Kannan, D. (2017). Sustainability as a dynamic organizational capability: a systematic review and a future agenda toward a sustainable transition. *Journal of Cleaner Production*, *142*, 308–322.
- Andor, G., & Bohák, A. (2017). Identifying events in financial time series - A new approach with bipower variation. *Finance Research Letters*, *22*, 42–48.
- Arboleda, F. J. M., Guzman-Luna, J. A., & Torres, I. D. (2018). Fraud detection-oriented operators in a data warehouse based on forensic accounting techniques. *Computer Fraud & Security*, *2018*, 13–19.
- Ardizzi, G., Franceschis, P. D., & Giammatteo, M. (2018). Cash payment anomalies and money laundering: An econometric analysis of Italian municipalities. *International Review of Law and Economics*, *56*, 105–121.
- Arnott, D., Jirachiefpattana, W., & O'Donnell, P. (2007). Executive information systems development in an emerging economy. *Decision Support Systems*, *42*, 2078–2084.
- Asha, R. B., & Suresh Kumar, K. R. (2021). Credit card fraud detection using Artificial Neural Network. *Global Transitions Proceedings*, *2*, 35–41.
- Ausloos, M., Ficcacanti, V., Dhesi, G., & Shakeel, M. (2021). Benford's laws tests on S&P 500 daily closing values and the corresponding daily log-returns both point to huge non-conformity. *Physica A: Statistical Mechanics and its Applications*, *574*, Article 125699 (Online first).
- Azevedo, C. D. S., Gonçalves, R. F., Gava, V. L., & Spinola, M. D. M. (2021). A Benford's law based methodology for fraud detection in social welfare programs: Bolsa Família analysis. *Physica A: Statistical Mechanics and its Applications*, *567*, Article 125626 (Online first).
- Badal-Valero, E., Alvarez-Jareño, J. A., & Pavia, J. M. (2018). Combining Benford's law and machine learning to detect money laundering. An actual Spanish court case. *Forensic Science International*, *282*, 24–34.
- Baensens, B., Höppner, S., & Verdonck, T. (2021). Data Engineering for fraud detection. In *150. Decision Support Systems* (p. 113492) (Online first).
- Bajaj, Y., Kashiramka, S., & Singh, S. (2020). Application of capital structure theories: a systematic review. *Journal of Advances in Management Research*, *18*, 173–199.
- Baker, C. R., Cohanier, B., & Leo, N. J. (2017). Breakdowns in internal controls in bank trading information systems: The case of the fraud at Société Générale. *International Journal of Accounting Information Systems*, *26*, 20–31.
- Barboza, F., Kimura, H., Sobreiro, V. A., & Basso, L. F. C. (2016). Credit risk: from a systematic literature review to future directions. *Corporate Ownership and Control*, *13*, 326–346.
- Barroso, M., & Laborda, J. (2022). Digital transformation and the emergence of the Fintech sector: Systematic literature review. *Digital Business*, *2*, Article 100028 (Online first).
- Bayrak, A. E., McComb, C., Cagan, J., & Kotovsky, K. (2021). A strategic decision-making architecture toward hybrid teams for dynamic competitive problems. *Decision Support Systems*, *144*, Article 113490 (Online first).
- Bhattacharya, S., Xu, D., & Kumar, K. (2011). An ANN-based auditor decision support system using Benford's law. *Decision Support Systems*, *50*, 576–584.
- Bollaert, H., de Silanes, F. L., & Schwenbacher, A. (2021). Fintech and access to finance. *Journal of Corporate Finance*, *68*, Article 101941 (Online first).
- Borgersen, T. A., & King, R. M. (2014). Structural origins of debt-sustainability in mature and transition economies: Domar, Balassa-Samuels and Maastricht. *Structural Change and Economic Dynamics*, *30*, 101–119.
- Bose, I., Piramuthu, S., & Shaw, M. J. (2011). Quantitative methods for detection of financial fraud. *Decision Support Systems*, *50*, 557–558.
- Boubaker, S., Derouiche, I., & Lasfer, M. (2015). Geographic location, excess control rights, and cash holdings. *International Review of Financial Analysis*, *42*, 24–37.
- Boubakri, N., Guedhami, O., & Saffar, W. (2016). Geographic location, foreign ownership, and cost of equity capital: Evidence from privatization. *Journal of Corporate Finance*, *38*, 363–381.
- Bouzuenda, K. (2018). Emotional intelligence and financial decision making: Are we talking about a paradigmatic shift or a change in practices? *Research in International Business and Finance*, *44*, 273–284.
- Byers, J., Popova, I., & Simkins, B. (2021). Robust estimation of conditional risk measures using machine learning algorithm for commodity futures prices in the presence of outliers. *Journal of Commodity Markets*, *24*, Article 100174 (Online first).
- Cabral, C., & Dhar, R. L. (2019a). Ecotourism research in India: from an integrative literature review to a future research framework. *Journal of Ecotourism*, *19*, 23–49.
- Cabral, C., & Dhar, R. L. (2019b). Skill development research in India: a systematic literature review and future research agenda. *Bench-marking: An International Journal*, *26*, 2242–2266.
- Canhoto, A. I. (2021). Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective. *Journal of Business Research*, *131*, 441–452.
- Carmona, P., Climent, F., & Momparler, A. (2019). Predicting failure in the U.S. banking sector: An extreme gradient boosting approach. *International Review of Economics and Finance*, *61*, 304–323.
- Carneiro, N., Figueira, G., & Costa, M. (2017). A Data Mining based system for credit-card fraud detection in e-tail. *Decision Support Systems*, *95*, 91–101.
- Carta, S., Fenu, G., Recupero, D. R., & Saia, R. (2019). Fraud detection for e-commerce transactions by employing a prudential Multiple Consensus model. *Journal of Information Security and Applications*, *46*, 13–22.
- Chaturvedi, A. R., Hutchinson, G. K., & Nazareth, D. L. (1993). Supporting complex real-time decision making through machine learning. *Decision Support Systems*, *10*, 213–233.
- Chaudhry, S., Li, H., Xu, L., & Zhang, H. (2007). Decision support systems in emerging economies. *Decision Support Systems*, *42*, 1987–1988.
- Chen, K., Yadav, A., Khan, A., & Zhu, K. (2020). Credit fraud detection based on hybrid Credit Scoring Model. *Procedia Computer Science*, *167*, 2–8.
- Chen, M., Kanade, T., Pomerleau, D., & Rowley, H. A. (1999). Anomaly detection through registration. *Pattern Recognition*, *32*, 113–128.
- Chen, T. K. (2016). Does geography matter in a geographically small and culturally homogeneous country? Firm location and corporate credit risk. *International Review of Economics and Finance*, *44*, 323–348.
- Cho, T. Y., & Chen, Y. S. (2021). The impact of financial technology on China's banking industry: An application of the metafrontier cost Malmquist productivity index. *The North American Journal of Economics and Finance*, *57*, Article 101414 (Online first).
- Ciaccio, E. J., Wan, E. Y., Saluja, D. S., Acharya, U. R., Peters, N. S., & Garan, H. (2019). Addressing challenges of quantitative methodologies and event interpretation in the study of atrial fibrillation. *Computer Methods and Programs in Biomedicine*, *178*, 113–122.
- Climent, F., Momparler, A., & Carmona, P. (2019). Anticipating bank distress in the Eurozone: An extreme gradient boosting approach. *Journal of Business Research*, *101*, 885–896.
- Conti, P., & Naldi, M. (2008). Detection of anomalous bids in procurement auctions. *Decision Support Systems*, *46*, 420–428.
- Cruz, A. F. D., Kimura, H., & Sobreiro, V. A. (2019). What Do We Know About Corporate Cash Holdings? A Systematic Analysis. *The Journal of Corporate Accounting and Finance*, *30*, 77–143.
- Cumming, D., Groh, A. P., & Johan, S. (2018). Same rules, different enforcement: Market abuse in Europe. *Journal of International Financial Markets Institutions and Money*, *54*, 130–151.
- Daragmeh, A., Lentner, C., & Sági, J. (2021). FinTech payments in the era of COVID-19: Factors influencing behavioral intentions of "Generation X" in Hungary to use mobile payment. *Journal of Behavioral and Experimental Finance*, *32*, Article 100574 (Online first).
- Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems*, *49*, 498–506.
- Druica, E., Oancea, B., & Valsan, C. (2018). Benford's law and the limits of digit analysis. *International Journal of Accounting Information Systems*, *31*, 75–82.
- Duftscheid, G., & Miksch, S. (2001). Knowledge-based verification of clinical guidelines by detection of anomalies. *Artificial Intelligence in Medicine*, *22*, 23–41.
- Elsaify, M., & Hasan, S. (2021). Data exchanges among firms. *Digital Business*, *1*, 100010 (Online first).
- Ferraz, D., Falguera, F. P. S., Mariano, E. B., & Hartmann, D. (2021). Linking Economic Complexity, Diversification, and Industrial Policy with Sustainable Development: A Structured Literature Review. *Sustainability*, *13*, 1265.
- Ferreira, M. C. R. D. C., Sobreiro, V. A., Kimura, H., & Barboza, F. L. D. M. (2016). A systematic review of literature about finance and sustainability. *Journal of Sustainable Finance & Investment*, *6*, 112–147.
- Fertier, A., Barthe-Delanoë, A. M., Montarnal, A., Truptil, S., & Bénaben, F. (2020). A new emergency decision support system: The automatic interpretation and contextualisation of events to model a crisis situation in real-time. *Decision Support Systems*, *133*, Article 113260 (Online first).
- Florack, A., Keller, J., & Palcu, J. (2013). Regulatory focus in economic contexts. *Journal of Economic Psychology*, *38*, 127–137.
- Forough, J., & Momtazi, S. (2021). Ensemble of deep sequential models for credit card fraud detection. *Applied Soft Computing*, *99*, Article 106883 (Online first).
- Frömmel, M., Han, X., Li, Y., & Vigne, S. A. (2021). Low liquidity beta anomaly in China. *Emerging Markets Review*, *100832*. In press. (Online first).
- Gianini, G., Fossi, L. G., Mio, C., Caelen, O., Brunie, L., & Damiani, E. (2020). Managing a pool of rules for credit card fraud detection by a Game Theory based approach. *Future Generation Computer Systems*, *102*, 549–561.
- González-Sánchez, M. (2021). Is there a relationship between the time scaling property of asset returns and the outliers? Evidence from international financial markets. *Finance Research Letters*, *38*, Article 101510 (Online first).

- GSMA. (2017). *State of the Industry Report on Mobile Money*. Report published by GSMA Mobile Money.
- Guo, Y., & Liang, C. (2016). Blockchain application and outlook in the banking industry. *Financial Innovation*, 2, 1–12.
- Habeeb, R. A. A., Nasaruddin, F., Gani, A., Hashem, I. A. T., Ahmed, E., & Imran, M. (2019). Real-time big data processing for anomaly detection: A Survey. *International Journal of Information Management*, 45, 289–307.
- Hanafizadeh, P., & Kim, S. (2020). Digital Business: A new forum for discussion and debate on digital business model and digital transformation. *Digital Business*, 1, Article 100006.
- Hanafizadeh, P., & Marjaie, S. (2021). Exploring banking business model types: A cognitive view. *Digital Business*, 1, Article 100012.
- Handley, L. (2019). Nearly three quarters of the world will use just their smartphones to access the internet by 2025. In *Report published by World Advertising Research Center*.
- Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, 124, 226–251.
- Henriques, I. C., Sobreiro, V. A., & Kimura, H. (2018). Science and technology park: Future challenges. *Technology in Society*, 53, 144–160.
- Henriques, I. C., Sobreiro, V. A., Kimura, H., & Mariano, E. B. (2020). Two-stage DEA in banks: Terminological controversies and future directions. *Expert Systems with Applications*, 161, Article 113632 (Online first).
- Hess, R. L., Rubin, R. S., & West, L. A. (2004). Geographic information systems as a marketing information system technology. *Decision Support Systems*, 38, 197–212.
- Hohberger, J., & Wilden, R. (2022). Geographic diversity of knowledge inputs: The importance of aligning locations of knowledge inputs and inventors. *Journal of Business Research*, 145, 705–719.
- Holsapple, C. W., Johnson, L. E., Manakyan, H., & Tanner, J. (1995). An empirical assessment and categorization of journals relevant to DSS research. *Decision Support Systems*, 14, 359–367.
- Hooi, B., Shin, K., Song, H. A., Beutel, A., Shah, N., & Faloutsos, C. (2017). Graph-Based Fraud Detection in the Face of Camouflage. *ACM Transactions on Knowledge Discovery from Data*, 11, 1–26.
- Höppner, S., Baesens, B., Verbeke, W., & Verdonck, T. (2022). Instance-dependent cost-sensitive learning for detecting transfer fraud. *European Journal of Operational Research*, 297, 291–300.
- Hosseini, S., & Ivanov, D. (2020). Bayesian networks for supply chain risk, resilience and ripple effect analysis: A literature review. *Expert Systems with Applications*, 161, Article 113649 (Online first).
- Hou, X., Wang, T., & Ma, C. (2021). Economic policy uncertainty and corporate fraud. *Economic Analysis and Policy*, 71, 97–110.
- Houdek, P. (2020). Detecting fraud beyond cooked books. *Organizational Dynamics*, 49, 100734 (Online first).
- Ince, H., & Trafalis, T. B. (2006). A hybrid model for exchange rate prediction. *Decision Support Systems*, 42, 1054–1062.
- Jabbour, C. J. C. (2013). Environmental training in organisations: From a literature review to a framework for future research. *Resources, Conservation and Recycling*, 74, 144–155.
- Jabbour, C. J. C., Sobreiro, V. A., de Sousa Jabbour, A. B. L., de Souza Campos, L. M., Mariano, E. B., & Renwick, D. W. S. (2017). An analysis of the literature on humanitarian logistics and supply chain management: paving the way for future studies. *Annals of Operations Research*, 283, 289–307.
- Jones, B. H., & Chin, A. G. (2015). On the efficacy of smartphone security: A critical analysis of modifications in business students' practices over time. *International Journal of Information Management*, 35, 561–571.
- Jung, S. H., & Jeong, Y. J. (2020). Twitter data analytical methodology development for prediction of start-up firms' social media marketing level. *Technology in Society*, 63, Article 101409 (Online first).
- Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P. E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications*, 100, 234–245.
- Kankanhalli, A., Tan, B. C., Wei, K. K., & Holmes, M. C. (2004). Cross-cultural differences and information systems developer values. *Decision Support Systems*, 38, 183–195.
- Karki, B. R., & Porras, J. (2021). Digitalization for sustainable maintenance services: A systematic literature review. *Digital Business*, 1, Article 100011.
- Kaur, R., & Singh, S. (2016). A survey of Data Mining and social network analysis based anomaly detection techniques. *Egyptian Informatics Journal*, 17, 199–216.
- Khan, A. I., & Al-Badi, A. (2020). Open source machine learning frameworks for industrial Internet of Things. *Procedia Computer Science*, 170, 571–577.
- Khan, M. A., Gu, L., Khan, M. A., & Meyer, N. (2021). The effects of national culture on financial sector development: Evidence from emerging and developing economies. *Borsa Istanbul Review*, 1–10. in press. (Online first).
- Khudzari, J. M., Kurian, J., Tartakovsky, B., & Raghavan, G. (2018). Bibliometric analysis of global research trends on microbial fuel cells using Scopus database. *Biochemical Engineering Journal*, 136, 51–60.
- Konopik, J., Jahn, C., Schuster, T., Hoßbach, N., & Pfäum, A. (2022). Mastering the digital transformation through organizational capabilities: A conceptual framework. *Digital Business*, 2, Article 100019 (Online first).
- Koyuncugil, A. S., & Ozgulbas, N. (2012). Financial early warning system model and data mining application for risk detection. *Expert Systems with Applications*, 39, 6238–6253.
- Kundu, D., Pal, R. K., & Mandal, D. P. (2020). Preference enhanced hybrid expertise retrieval system in community question answering services. *Decision Support Systems*, 129, Article 113164 (Online first).
- Kuzey, C., Uyar, A., & Delen, D. (2014). The impact of multinationality on firm value: A comparative analysis of machine learning techniques. *Decision Support Systems*, 59, 127–142.
- Lage Junior, M., & Godinho Filho, M. (2010). Variations of the Kanban System: Literature review and classification. *International Journal of Production Economics*, 125, 13–21.
- Lee, J. K. Y., Gholami, H., Saman, M. Z. M., Ngadiman, N. H. A. B., Zakuan, N., Mahmood, S., & Omain, S. Z. (2021). Sustainability-oriented application of value stream mapping: a review and classification. *IEEE Access*, 9, 68414–68434.
- Lee, K. C., Han, I., & Kwon, Y. (1996). Hybrid neural network models for bankruptcy predictions. *Decision Support Systems*, 18, 63–72.
- León, C. (2020). Detecting anomalous payments networks: A dimensionality-reduction approach. *Latin American journal of central Banking*, 1, 100001 (Online first).
- León, C., Barucca, P., Acero, O., Gage, G., & Ortega, F. (2020). Pattern recognition of financial institutions' payment behavior. *Latin American journal of central Banking*, 1, Article 100011 (Online first).
- Liang, T. P., Chandler, J. S., & Han, I. (1990). Integrating statistical and inductive learning methods for knowledge acquisition. *Expert Systems with Applications*, 1, 391–401.
- Lin, F., Lin, L. J., Yeh, C. C., & Wang, T. S. (2018). Does the board of directors as fat cats exert more earnings management? Evidence from Benford's law. *The Quarterly Review of Economics and Finance*, 68, 158–170.
- Liu, H., Lin, F., He, J., & Cai, Y. (2010). New approach for the sequential pattern mining of high-dimensional sequence databases. *Decision Support Systems*, 50, 270–280.
- Lo, F. Y., Rey-Martí, A., & Botella-Carrubi, D. (2020). Research methods in business: Quantitative and qualitative comparative analysis. *Journal of Business Research*, 115, 221–224.
- Lozano, M. G., Schreiber, J., & Brynielsson, J. (2017). Tracking geographical locations using a geo-aware topic model for analyzing social media data. *Decision Support Systems*, 99, 18–29.
- Lucas, Y., Portier, P. E., Laporte, L., He-Guelton, L., Caelen, O., Granitzer, M., & Calabretto, S. (2020). Towards automated feature engineering for credit card fraud detection using multi-perspective HMMs. *Future Generation Computer Systems*, 102, 393–402.
- Maass, W., & Storey, V. C. (2021). Pairing conceptual modeling with machine learning. *Data & Knowledge Engineering*, 134, 101909 (Online first).
- Mallinguh, E. B., & Zéman, Z. (2020). Financial Distress, Prediction, and Strategies by Firms: A Systematic Review of Literature. *Periodica Polytechnica, Social and Management Sciences*, 28, 162–176.
- Mamede, S. D. P. N., & Malaquias, R. F. (2017). Monday effect in Brazilian hedge funds with immediate redemption. *Research in International Business and Finance*, 39, 47–53.
- Mariano, E. B., Sobreiro, V. A., & Do Nascimento Rebelatto, D. A. (2015). Human development and data envelopment analysis: A structured literature review. *Omega*, 54, 33–49.
- Marikyan, D., Papagiannidis, S., Rana, O. F., & Ranjan, R. (2022). Blockchain: A business model innovation analysis. *Digital Business*, 2, Article 100033.
- Masudin, I., & Fernanda, F. W. (2019). A review of literature on types, stages of recovery and humanitarian logistics operations in the tsunami and earthquake disaster in Indonesia. *IOP Conference Series: Materials Science and Engineering*, 674, Article 012043 (Online first).
- Misra, S., Thakur, S., Ghosh, M., & Saha, S. K. (2020). An autoencoder based model for detecting fraudulent credit card transaction. *Procedia Computer Science*, 167, 254–262.
- Mukherjee, R., Gamble, R. F., & Parkinson, J. A. (1997). Classifying and detecting anomalies in hybrid knowledge-based systems. *Decision Support Systems*, 21, 231–251.
- Muniyandi, A. P., Rajeswari, R., & Rajaram, R. (2012). Network anomaly detection by cascading k-means clustering and C4.5 decision tree algorithm. *Procedia Engineering*, 30, 174–182.
- Nazário, R. T. F., Silva, J. L. E., Sobreiro, V. A., & Kimura, H. (2017). A literature review of technical analysis on stock markets. *The Quarterly Review of Economics and Finance*, 66, 115–126.
- Nikkel, B. (2020). Fintech forensics: Criminal investigation and digital evidence in financial technologies. *Forensic Science International Digital Investigation*, 33, Article 200908 (Online first).
- Ojha, P. (2021). The Upsurge of Graphical Abstracts: A Neuroscience Perspective. *Neuroscience*, 457, 218–219.
- Okoli, C. (2015). A Guide to Conducting a Standalone Systematic Literature Review. *Communications of the Association for Information Systems*, 37, 879–910.
- Parnes, D. (2020). Exploring economic anomalies in the S&P 500 index. *The Quarterly Review of Economics and Finance*, 76, 292–309.
- Patil, S., Nemade, V., & Soni, P. K. (2018). Predictive modelling for credit card fraud detection using data analytics. *Procedia Computer Science*, 132, 385–395.
- Perdisci, R., Ariu, D., Fogla, P., Giacinto, G., & Lee, W. (2009). McPAD: A multiple classifier system for accurate payload-based anomaly detection. *Computer Networks*, 53, 864–881.
- Pham, N. T., Hoang, H. T., & Phan, Q. P. T. (2019). Green human resource management: a comprehensive review and future research agenda. *International Journal of Manpower*, 41, 845–878.
- Pick, J. B., Turetken, O., Deokar, A. V., & Sarkar, A. (2017). Location analytics and decision support: Reflections on recent advancements, a research framework, and the path ahead. *Decision Support Systems*, 99, 1–8.
- Picoto, W. N., & Pinto, I. (2021). Cultural impact on mobile banking use – A multi-method approach. *Journal of Business Research*, 124, 620–628.
- Pinheiro, M. A. P., Seles, B. M. R. P., Fiorini, P. D. C., Jugend, D., de Sousa Jabbour, A. B. L., da Silva, H. M. R., & Latan, H. (2019). The role of new product

- development in underpinning the circular economy. *Management Decision*, 57, 840–862.
- Plc, R. (2021). *Annual report and financial statements 2020*. Report published by Elsevier®.
- Pourhabibi, T., Ong, K. L., Kam, B. H., & Boo, Y. L. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems*, 133, Article 113303 (Online first).
- Raj, A., Agrahari, A., & Srivastava, S. K. (2020). Do pressures foster sustainable public procurement? An empirical investigation comparing developed and developing economies. *Journal of Cleaner Production*, 266, Article 122055 (Online first).
- Rajaean, M. M., Cater-Steel, A., & Lane, M. (2017). A systematic literature review and critical assessment of model-driven decision support for IT outsourcing. *Decision Support Systems*, 102, 42–56.
- Rocha-Salazar, J. D. J., Segovia-Vargas, M. J., & del Mar Camacho-Miñano, M. (2021). Money laundering and terrorism financing detection using neural networks and an abnormality indicator. *Expert Systems with Applications*, 169, Article 114470 (Online first).
- Rtayli, N., & Enneya, N. (2020). Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization. *Journal of Information Security and Applications*, 55, Article 102596 (Online first).
- Rubio, J., Barucca, P., Gage, G., Arroyo, J., & Morales-Resendiz, R. (2020). Classifying payment patterns with Artificial Neural Networks: An autoencoder approach. *Latin American journal of central Banking*, 1, Article 100013 (Online first).
- Sá, A. G. D., Pereira, A. C., & Pappa, G. L. (2018). A customized classification algorithm for credit card fraud detection. *Engineering Applications of Artificial Intelligence*, 72, 21–29.
- Sabetti, L., & Heijmans, R. (2021). Shallow or deep? Training an autoencoder to detect anomalous flows in a retail payment system. *Latin American journal of central Banking*, 2, Article 100031 (Online first).
- Saleh, K., Boujarwah, A. A., & Al-Dallal, J. (2001). Anomaly detection in concurrent Java programs using dynamic data flow analysis. *Information and Software Technology*, 43, 973–981.
- Salim, N., Rahman, M. N. A., & Wahab, D. A. (2019). A systematic literature review of internal capabilities for enhancing eco-innovation performance of manufacturing firms. *Journal of Cleaner Production*, 209, 1445–1460.
- Sankarrao, L., Ghose, D. K., & Rathinsamy, M. (2021). Predicting land-use change: Intercomparison of different hybrid machine learning models. *Environmental Modelling & Software*, 145, Article 105207 (Online first).
- Savage, D., Zhang, X., Yu, X., Chou, P., & Wang, Q. (2014). Anomaly detection in online social networks. *Social Networks*, 39, 62–70.
- Serrano-Silva, Y. O., Villuendas-Rey, Y., & Yáñez-Márquez, C. (2018). Automatic feature weighting for improving financial Decision Support Systems. *Decision Support Systems*, 107, 78–87.
- Seuring, S. (2013). A review of modeling approaches for sustainable supply chain management. *Decision Support Systems*, 54, 1513–1520.
- Severino, M. K., & Peng, Y. (2021). Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata. *Machine Learning with Applications*, 5, Article 100074 (Online first).
- Shaikh, A. A., Sharma, R., & Karjaluo, H. (2020). Digital innovation & enterprise in the sharing economy: An action research agenda. *Digital Business*, 1, Article 100002 (Online first).
- Shi, J., Ausloos, M., & Zhu, T. (2018). Benford's law first significant digit and distribution distances for testing the reliability of financial reports in developing countries. *Physica A: Statistical Mechanics and its Applications*, 492, 878–888.
- Shtanko, A., & Kulik, S. (2021). Increasing the effectiveness of intelligent module by enlarging training dataset from real data. *Procedia Computer Science*, 190, 712–716.
- Silva, A. D. B., Gomes, M. M., Costa, C. A., Righi, R. D. R., Barbosa, J. L. V., Pessin, G., ... Federizzi, G. (2020). Intelligent personal assistants: A systematic literature review. *Expert Systems with Applications*, 147, Article 113193 (Online first).
- Silva, W., Kimura, H., & Sobreiro, V. A. (2017). An analysis of the literature on systemic financial risk: A survey. *Journal of Financial Stability*, 28, 91–114.
- Singh, K., & Best, P. (2019). Anti-money laundering: Using data visualization to identify suspicious activity. *International Journal of Accounting Information Systems*, 34, Article 100418 (Online first).
- Singh, S., Bhardwaj, N., Sharma, G. D., Kaya, T., Mahendru, M., & Erkut, B. (2019). Research in market-calibrated option pricing analysis. *Qualitative Research in Financial Markets*, 12, 159–176.
- Smith, S. D., Dixon, M. J., Tays, W. J., & Bulman-Fleming, M. (2004). Anomaly detection in the right hemisphere: The influence of visuospatial factors. *Brain and Cognition*, 55, 458–462.
- Sokolova, K., Perez, C., & Lemerrier, M. (2017). Android application classification and anomaly detection with graph-based permission patterns. *Decision Support Systems*, 93, 62–76.
- Soulé, E., Michonneau, P., Michel, N., & Bockstaller, C. (2021). Environmental sustainability assessment in agricultural systems: A conceptual and methodological review. *Journal of Cleaner Production*, 325, Article 129291 (Online first).
- Stefanelli, N. O., Teixeira, A. A., Oliveira, J. H. C. D., Ferreira, M. A., & Sehnem, S. (2019). Environmental training: a systematic review of the state of the art of the theme. *Benchmarking: An International Journal*, 27, 2048–2076.
- Stripling, E., Baesens, B., Chizi, B., & van den Broucke, S. (2018). Isolation-based conditional anomaly detection on mixed-attribute data to uncover workers' compensation fraud. *Decision Support Systems*, 111, 13–26.
- Subudhi, S., & Panigrahi, S. (2020). Use of optimized Fuzzy C-Means clustering and supervised classifiers for automobile insurance fraud detection. *Journal of King Saud University - Computer and Information Sciences*, 32, 568–575.
- Talan, G., & Sharma, G. (2019). Doing Well by Doing Good: A Systematic Review and Research Agenda for Sustainable Investment. *Sustainability*, 11, 353.
- Tardelli, S., Avvenuti, M., Tesconi, M., & Cresci, S. (2021). Detecting inorganic financial campaigns on Twitter. *Information Systems*, 103, Article 101769 (Online first).
- Thompson, B. S. (2017). Can Financial Technology Innovate Benefit Distribution in Payments for Ecosystem Services and REDD? *Ecological Economics*, 139, 150–157.
- Timmons, S., McGowan, F. P., & Lunn, P. D. (2019). Setting defaults for online banking transactions: Experimental evidence from personal loan repayment terms. *Journal of Behavioral and Experimental Finance*, 23, 161–165.
- Trujillo, J., Davis, K. C., Du, X., Damiani, E., & Storey, V. C. (2021). Conceptual modeling in the era of Big Data and Artificial Intelligence: Research topics and introduction to the special issue. *Data & Knowledge Engineering*, 135, Article 101911 (Online first).
- Tsai, C. F., & Chen, Z. Y. (2014). Towards high dimensional instance selection: An evolutionary approach. *Decision Support Systems*, 61, 79–92.
- Tung, L. L., & Quaddus, M. (2002). Cultural differences explaining the differences in results in GSS: Implications for the next decade. *Decision Support Systems*, 33, 177–199.
- Turner, A. (2021). *How many smartphones are in the world?*. Report published by BankMyCell.
- Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European Journal of Operational Research*, 218, 211–229.
- Waite, N. (1999). A real-time system-adapted anomaly detector. *Information Sciences*, 115, 221–259.
- Wang, H., & Weigend, A. S. (2004). Data mining for financial decision making. *Decision Support Systems*, 37, 457–460.
- Wang, Y., & Xu, W. (2018). Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decision Support Systems*, 105, 87–95.
- Wessels, M., van den Brink, P., Verburch, T., Cadet, B., & van Ruijven, T. (2021). Understanding incentives for cybersecurity investments: Development and application of a typology. *Digital Business*, 1, Article 100014.
- Wimmer, H., Yoon, V. Y., & Sugumaran, V. (2016). A multi-agent system to support evidence based medicine and clinical decision making via data sharing and data privacy. *Decision Support Systems*, 88, 51–66.
- Wu, N., & Zhang, J. (2006). Factor-analysis based anomaly detection and clustering. *Decision Support Systems*, 42, 375–389.
- Wu, Y., Ngai, E. W., Wu, P., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. *Decision Support Systems*, 132, Article 113280 (Online first).
- Yan, C., Li, Y., Liu, W., Li, M., Chen, J., & Wang, L. (2020). An artificial bee colony-based kernel ridge regression for automobile insurance fraud identification. *Neurocomputing*, 393, 115–125.
- Yasami, Y., & Safaei, F. (2017). A statistical infinite feature cascade-based approach to anomaly detection for dynamic social networks. *Computer Communications*, 100, 52–64.
- Zaffar, M. A., Kumar, R. L., & Zhao, K. (2019). Using agent-based modelling to investigate diffusion of mobile-based branchless banking services in a developing country. *Decision Support Systems*, 117, 62–74.
- Zhai, J., Cao, Y., Yao, Y., Ding, X., & Li, Y. (2017). Computational intelligent hybrid model for detecting disruptive trading activity. *Decision Support Systems*, 93, 26–41.
- Zhang, J. T., Zhou, B., Guo, J., & Zhu, T. (2021). Two-sample Behrens-Fisher problems for high-dimensional data: A normal reference approach. *Journal of Statistical Planning and Inference*, 213, 142–161.
- Zhang, K. Z., & Benyoucef, M. (2016). Consumer behavior in social commerce: A literature review. *Decision Support Systems*, 86, 95–108.
- Zoričák, M., Gnip, P., Drotár, P., & Gazda, V. (2020). Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets. *Economic Modelling*, 84, 165–176.