COIMBRA BUSINESS SCHOOL

iscac Politécnico de Coimbra

100 ANOS

COIMBRA BUSINESS SCHOOL 100 | since 1923

iscac Politécnico de Coimbra

Advanced analytical methods for fraud detection: a systematic literature review

Marta Sequeira Rego Fernandes

ISCAC | 2023

COIMBRA BUSINESS SCHOOL

iscac Politécnico de Coimbra 100 ANOS

Marta Sequeira Rego Fernandes

Advanced analytical methods for fraud detection: a systematic literature review

Coimbra, agosto de 2023

Marta Sequeira Rego Fernandes

Advanced analytical methods for fraud detection: a systematic literature review

Dissertação submetida ao Instituto Superior de Contabilidade e Administração de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de **Mestre em Auditoria Empresarial e Pública,** realizada sob a orientação da Professora Doutora Isabel Maria Mendes Pedrosa e Coorientação do Professor Doutor Raul Laureano.

Coimbra, agosto de 2023

TERMO DE RESPONSABILIDADE

Declaro ser a autora desta dissertação, que constitui um trabalho original e inédito, que nunca foi submetido a outra Instituição de ensino superior para obtenção de um grau académico ou outra habilitação. Atesto ainda que todas as citações estão devidamente identificadas e que tenho consciência de que o plágio constitui uma grave falta de ética, que poderá resultar na anulação da(o) presente dissertação.

PENSAMENTO

"Discipline is choosing between what you want now,

and what you want most."

Abraham Lincoln

DEDICATÓRIA

Ao Daniel,

Meu companheiro neste e em tantos outros êxitos.

AGRADECIMENTOS

Esta dissertação é o culminar de um glorioso percurso na Coimbra Business School – ISCAC. Por isso, agradeço a todos com quem tive a sorte de me cruzar. Aos professores, por todas as partilhas enriquecedoras; aos meus amigos e colegas, pelos momentos de descontração e companheirismo; e a todos os funcionários, que tanto me acarinharam durante estes anos.

Um agradecimento especial à minha orientadora, Professora Isabel Pedrosa, pela inspiração e motivação e por toda a paciência neste moroso processo.

Aos meus pais que me inspiram diariamente a ser a minha melhor versão e, claro, por todo o apoio e carinho ao longo destes anos de estudo e ausência.

Ao meu irmão pela amizade incondicional.

Ao Daniel por ter sempre acreditado em mim, dando-me força para continuar.

RESUMO

Os desenvolvimentos da era digital exigem novas formas de produzir bens e prestar serviços. Esta evolução acelerada nas empresas obriga a uma nova postura por parte dos auditores, que devem acompanhar a constante transformação. Com as características dinâmicas dos dados, é importante aproveitar a oportunidade para criar valor às empresas.

A necessidade de aplicar métodos mais robustos para deteção de fraudes é evidente.

Nesta tese será investigada a utilização de métodos analíticos avançados para deteção de fraude, através da análise da literatura existente sobre o tema.

Será feita uma revisão sistemática da literatura e serão aplicados indicadores bibliométricos numa pesquisa à base de dados mais adequada para medir a produção científica e as tendências atuais.

Este estudo pretende contribuir para as pesquisas académicas que têm vindo a ser realizadas, de forma a centralizar a informação existente sobre esta temática.

Palavras-chave: *Big Data*, Revisão Sistemática da Literatura, Indicadores Bibliométricos, Deteção de Fraude, Métodos Analíticos Avançados.

ABSTRACT

The developments of the digital era demand new ways of producing goods and rendering services. This fast-paced evolution in the companies implies a new approach from the auditors, who must keep up with the constant transformation. With the dynamic dimensions of data, it is important to seize the opportunity to add value to the companies.

The need to apply more robust methods to detect fraud is evident.

In this thesis the use of advanced analytical methods for fraud detection will be investigated, through the analysis of the existent literature on this topic.

Both a systematic review of the literature and a bibliometric approach will be applied to the most appropriate database to measure the scientific production and current trends.

This study intends to contribute to the academic research that have been conducted, in order to centralize the existing information on this topic.

Keywords: Big Data, Systematic Literature Review, Bibliometric Indicators, Fraud Detection, Advanced Analytical Methods.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES

List of abbreviations and acronyms

BA	Bibliometric Approach
BI	Bibliometric Indicator
CatBoost	Categorical Boosting.
COBIT	Control Objectives for Information and Related Technologies
CNN	Convolutional Neural Networks
CRISP-DM	Cross-Industry Standard Process for Data Mining
DOI	Digital Object Identifier
GBT	Gradient Boosted Trees
IEEE	Institute of Electrical and Electronics Engineers Inc.
ISACA	Information Systems Audit and Control Association
IT	Information Technology
MTM accounting	Mark-to-market accounting
PRISMA-P	Preferred Reporting Items for Systematic review and Meta-Analysis Protocols
PROSPERO	International Prospective Register of Ongoing Systematic Reviews
RNN	Recurrent Neutral Networks
SIQ	Specific Investigation Questions
SJR	SCImago Journal Rank
SLR	Systematic Literature Review
SOX Act	Sarbanes-Oxley Act
XGBoost	Extreme Gradient Boosting

INTRODUCTION

The data mining process arises from the irrefutable need to extract information from the growing databases of companies, with the aim of creating knowledge that increases their profits. Transforming data into information can mitigate risks and create opportunities. However, to do so, it is crucial to use the most appropriate tools, considering the speed, volume and variety of data.

The auditor's role is linked to the need of understanding the risks associated with predominantly technological environments and how it is possible to reduce the impact of these risks on financial information. To prevent the catastrophic consequences - already proven in the past - that can result from financial fraud, it is important to develop and implement mechanisms that detect signs and prevent their occurrence.

Short description of the methodology

The topic to be studied in this dissertation is the use of advanced analytical methods for detecting fraud. As this is a subject in vogue, there is a need to summarize and outline the studies that have been carried out on the subject. Thus, the proposed study methodology is a systematic review of the literature.

Additionally, a bibliometric approach will be adopted to measure the scientific production related to advanced analytical methods for fraud detection., through the analysis of bibliometric indicators.

This type of investigations on computer science topics is critical to understand the trends: not only for academic purposes; but also to lead future investigations on such a trendy subject as the use of technology for fraud detection.

Objectives

The main objective of this thesis is to identify the most relevant studies that propose the application of advanced analytical methods for the fraud detection.

Moreover, this study aims to analyse the categorization made for the different analytical methods advanced and for types of fraud, as well as analyse the relationships established between the various categories of analytical methods and types of fraud.

1 Context

1.1 Evolution of the Auditing Profession

Arens et al. (2014) define auditing as "the accumulation and evaluation of evidence about information to determine and report on the degree of correspondence between the information and established criteria", adding it "should be done by a competent, independent person".

The role of auditing has developed in the last twenty years due to the changes in the *modus operandi* of companies and, consequently, the regulatory requirements established to homogenize them.

Depending on the approaches to this theme, a variable branch of factors can be pointed out as the drivers for the evolutions of the role of the auditor. The specialists generally choose to rely on the Sarbanes-Oxley (SOX) Act of 2002 and its' requirements and on the use of technology in the auditing profession. Indeed, both factors are irrefutable and are linked to one another.

The publication of the SOX Act in July of 2002 came in response to the financial scandals that occurred in the previous years – such as:

1. Enron Corp., whose financial losses' were hidden using the MTM accounting technique - where the value of a security it's measured using its current market value, instead of it's book value (Segal, 2021). Enron's share prices came down to $0.26 and the firm declared bankruptcy in December 2nd of 2001. The scandal led to the dissolution of Enron's accounting firm, Arthur Andersen, who provided other services besides audit services, which raised questions about independence;
2. Tyco International, Ltd., involved in a long trial due to theft, tax evasion and ethical conflicts accusations of the company's CEO and Chairman, Dennis Kozlowski, and former corporate Chief Financial Officer, Mark Swartz;
3. WorldCom, who admitted to inflating its earnings by fraudulently capitalizing expenditures – rather than booking those entries as expenses – and booking

accounting entries related to fake revenues in "Corporate Unallocated" revenue accounts.

The motivations behind these scandals, as well as the cyclical recessions capable of shaking investor's confidence in their investments, led to the establishment of procedures to assure the transparency and reliability of the financial information.

1.2 The Sarbanes-Oxley Act

For a better understanding of the financial scandals mentioned above, it's important to insert these events in a timeline. The economy is cyclical, combining periods of expansion and growth with periods of stagnation or contraction. These fluctuations impact sharply companies' financial health, mainly during speculative bubbles. Below, the outlines of a specific bubble will be addressed.

The dot.com bubble, or the internet bubble, was a period between the late 1990s and early 2000s during which the value of internet-based companies increased. These companies' stock prices value was highly impacted by the potential and future growth expectations related to the beginning of the commercialization of the internet. Even though many of these companies were only start-ups with not so solid financial results, the investors chose to take a risk, leading to the companies' investment in their marketing and advertising, resulting in a speculative bubble.

Audit companies took the opportunity to start providing a wider range of services, making what was their core business a secondary slice of their honoraires charged.

Eventually, the money started to dry up, leading to the crash of the market. The burst of this bubble, in 2000, was driven by availability of venture capital, the mainstreaming of the internet and the hype regarding this type of companies.

Although the dot.com bubble was not the first asset bubble to occur, this crisis overlapped the beginning of the digital age.

1.3 The Big Data concept

ISACAs framework for information technology (IT) management and IT governance, COBIT 5 (ISACA, 2012), defines data "as something that is, or represents, a fact" and information as "data in context", meaning information is the main enabler for decision making at operation, management and governance levels.

For the same Association, big data "is a common term for a set of problems and techniques concerning the management and exploitation of very large sets of data". The term "big data" must be understood taking into consideration the enterprise's reality but always implies that the use of traditional techniques or tools is not efficient or useful to manage the amount of data available. This data is usually in a unstructured form but, when efficiently managed, can provide valuable findings and predictions to a wide range of stakeholders (e.g. consumers' behaviour; markets' latest trends; or competitors' strategy based on their public information). Data's value depends on the interested part on the data itself because it depends on the treatment and application of that information.

Initially, big data was categorised into three dimensions:

(1) Variety of information;
(2) Velocity of information creation and;
(3) Volume of information (ISACA, 2012).

Later, other two V's were added to the characteristics of big data:

(4) Value and;
(5) Veracity.

Ishwarappa and Anuradha J (2015) explain the link between (4) Veracity and (1) Variety, (2) Velocity and (3) Volume of data through the necessity of assuring the quality of data, when there is a high volume, variety and velocity of data to analyse. The accuracy of the analysed data can only be guaranteed through the veracity of the data source.

The same authors indicate (5) Value as "the most important aspect in the big data". The relevance of big data implies the creation of value to its stakeholders. There must exist a turnover in order to justify the investment in this type of IT infrastructures.

With the recognition and acceptance of the importance of data for enterprises, more categorizations with V words have emerged. Some authors add Variability - taking into consideration the inconsistency of the data flows - and Visibility or Visualization - the data comprehension is intrinsically linked to the way is displayed, meaning the way its presented is fundamental for the decision-making (Sami Owais & Sael Hussein, 2016; Zafar et al., 2021).

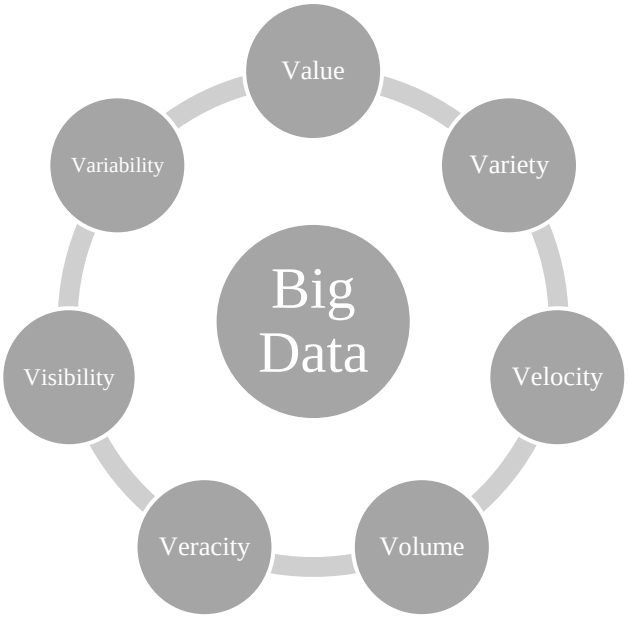


Figure 1 - The 7 Vs of Big Data

Besides the 7 Vs categorization explained above, other categorizations have been explored. Arockia Panimalar et al (2017) developed a categorization of 17 Vs, including some trendy classifications, such as Virality – defining the spreading speed of data – and Venue – data can be obtained through various platforms and sources, some of them being private and others being public (e.g. internet).

1.4 Data Mining and CRISP-DM

Data Mining can be defined as the process where statistical, mathematical and machine-learning techniques are used to extract and examine useful information from a database, creating knowledge to its stakeholders (Ko et al., 2011 *apud* Ngai et al., 2011).

IBM defines the Cross-Industry Standard Process for Data Mining (CRISP-DM) (2011) as process model that can be used as a guide for the data mining lifecycle. This process is represented in Figure 2, where the arrows indicate the most important and frequent dependencies between phases.

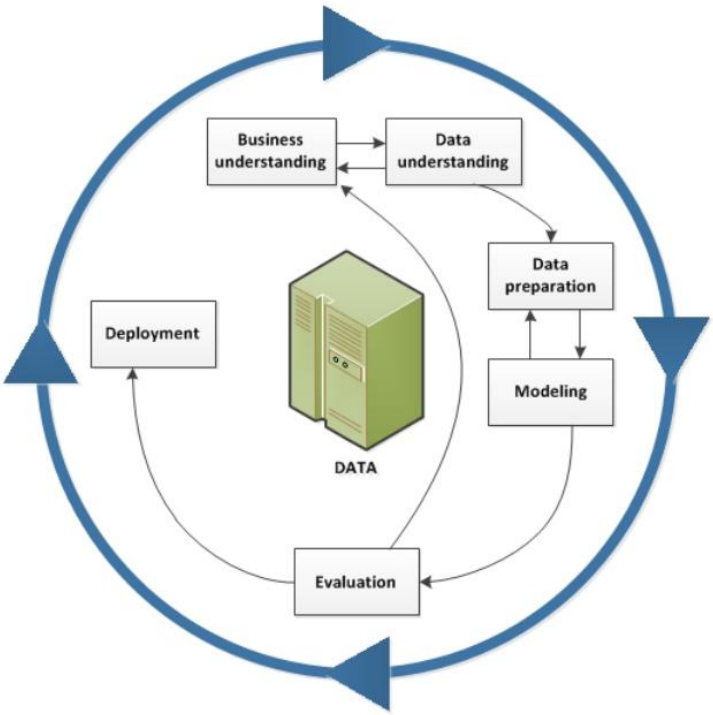


Figure 2 - The data mining life cycle *Source: IBM SPSS Modeler CRISP-DM Guide*, n.d.

The first phase of the process is "Business Understanding", which consists of the determination of the business objectives, through the clarification of the business problems, goals and available resources. It is also important to set the data mining goals by questioning the data mining problem:

1. Clustering - unsupervised learning approach, where the machine applies grouping rules based on similarities to existing data;

2. Classification - supervised learning approach, where a specific label/rule is setup in the machine to classify new observations;
3. Or Prediction – Based on the system's setup, the algorithm generates a model or a predictor. Then, when new data is provided to the system, a numerical output is generated. In most cases, regression analysis is the statistical methodology used. The model or predictor will predict a continuous-valued function or an ordered value.

Furthermore, it's important to go through the "Data Understanding" phase, which requires analysing the variety of data sources and describing the data available, though the following characteristics:

- amount of data – although bigger loads of data usually allow more reliability on the results, it is not always possible to invest a lot of time in processing these kinds of datasets;
- value types – variety of formats like numerical, categorical (string) or Boolean (true/false);
- and coding schemes – used to categorize data in a certain dataset (e.g. "F" and "M" generally stand for "Female" and "Male" respectively.

It might also be relevant to elaborate a data report, listing the data features abovementioned and testing data quality (e.g. missing data, data errors, coding inconsistencies).

After evaluating and understanding the dataset, researchers may proceed to the most time-consuming phase of this process: "Data Preparation", estimated to take up to 70% of the project's time. IBM (2011) notes the following tasks as the components of this phase:

• Merging data sets and/or records – taking into consideration their format;

• Selecting a sample subset of data – through inclusion or exclusion criteria;

• Aggregating records;

• Deriving new attributes ;

• Sorting the data for modelling;

• Removing or replacing blank or missing values;

• And splitting into training and test data sets.

Following the preparation of the data, data miners are able to prepare the Modeling phase. The most appropriate Modeling techniques are applied and several models with default parameters are tested. Then a fine-tune of those parameters is applied in order to achieve the data mining goals. After the refinement of the model, a comprehensive model assessment can be made to formalise its accuracy and results.

The "Evaluation" phases aims to measure the success of the data mining application in the business goals. Using the criteria established in the beginning of the process, findings can be discussed and presented. After that, data miners can determine the next steps: continue to the deployment phase or go back and define or replace the models (IBM, 2011).

"Deployment" is where the new insights are applied to the business. The results are monitored, and a project review can be produced.

2 Methodology

2.1 Types of Literature Review

As abovementioned, one of the main goals of this investigation is to define which advanced analytical methods are most used for fraud detection. To do so, the most common types of literature reviews were studied; and what was considered the most appropriate approach, was the selected one.

When it comes to literature reviews, one can categorize into three types:

- State-of-the-art review: it aims to define a broader concept in a certain period of time. It intends to summarize the current work of a specific field, by offering "interpretations of the historical progression of knowledge relating to a phenomenon (…)" (Barry et al., 2022);
- Scoping review: intends to present an overview of the investigations developed concerning a certain topic (Amendoeira et al., 2022). It consists in building maps of literature in order to do a preliminary assessment, and may be followed by systematic reviews;
- And systematic reviews: unlike state-of-the-art reviews, this type of reviews tend to concentrate on what was studied in the past and identify potentials for future research. Although the goal is not to focus on outdated literature, systematic reviews aim to reflect on the development of the studied field over time.

The selected methodology was Systematic Literature Review (SLR). Hereafter, this methodology and the required protocol will be explained. Even though SLR is not a new approach, there are not plenty of tutorials on how to implement such methodology. The perspective of Chitu Okoli and it's "Guide to Conducting a Standalone Systematic Literature Review" will be adapted ahead.

2.2 Systematic Literature Review

Okoli (2015) lists three characteristics through which is possible to define a Systematic Literature Review (SLR):

(1) systematic – in the methodology adopted.
(2) comprehensive and explicit – by including all the relevant contributes and transparent regarding the procedures used, respectively.
(3) and reproducible – in a way that the same methodology can be applied by other researchers.

This methodology aims to scrutinize a certain topic with such rigor and step-by-step technique that is acknowledged as a proper standalone investigation work. According to the same author, the systematic reviews of literature stand out from the ordinary reviews of literature due to being an objective exercise, rather than a subjective one.

Describing the available knowledge and identifying the "experts within a given field" are a few of the motivations sustained by Fink (2005 apud Okoli, 2015). On the other hand, this methodology doesn't add much value when it concerns a field in a early stage of investigation; when a similar and up-to-date summarization of literature can be found; or when the question behind the investigation is to vague, rather than concise and objective.

2.2.1 A guide to a Systematic Literature Review

It's possible to distinguish a systematic literature review from a conventional literature review by defining its scope and rigour. A SLR is most likely to be cited by other researchers because it's very clear and objective when it comes to its purpose. The first step to execute this kind of investigation is therefore to (1.) identify the purpose. It's important to determine the answer to the question "why do a literature review?" and, consequently, define its main goals. To do that, it's important to establish the necessity and the purpose of the review.

Still according to the same author, one should (2.) draft the protocol to be followed during the research, including the search's parameters and steps. This phase should start with the formulations of the investigation, which should define the

audience, the purpose and the use of the review. This document doesn't need to be finished before the research starts. It should be used as a guide to the researchers during the investigation and as a detailed explanation of the procedure adopted by the stakeholders of the SLR. The protocol should be the formalization of the planning stage.

Afterwards, (3.) a practical screen should be applied through the selection of a feasible number of studies to be analysed by the researchers. The intention shouldn't be to categorize the studies taking into account it's quality; instead, the researchers should decide if the studies that resulted from the initial search are worth of further analysis or if they don't fall into the scope. Multiple criteria can be used, such as the filters available in the databases: year of publication, journals, authors, keywords, between others. When in doubt in this phase, the study should be included.

The following step consists in the (4.) searching of the literature itself. For that, researchers must define the library in which the search will be applied. All sources should be considered, and its' inclusion or exclusion should be justifiable.

Regarding the (5.) data extraction, Okoli (2015) suggests the use of a form where the extraction and treatment of the data is fully explained.

After a practical screening is executed, it's important to assure that only relevant results are considered in the review. The author indicates (6.) the quality appraisal as the last step in the extraction phase. The purpose is to both categorize studies accordingly to their quality - regarding the search's scope - and exclude studies that don't follow the standards for the review. This scoring can be qualitative and/or quantitative.

Finished the stages of planning, selection and extraction, the researchers are now able to execute. To (7.) synthetize the studies means to transition "from an author- to a concept- centric focus" (Webster & Watson, 2002). Like the quality appraisal, this could be done in a quantitative or qualitative way. After knowing and understanding the literature, in this phase the researchers must synthesize and evaluate the search.

At last, the process ends when the researchers (8.) write the review. Depending on the criteria and standards established in the previous stages, this phase could be more

or less time consuming. The goal is to communicate the results and make them available by publishing them.

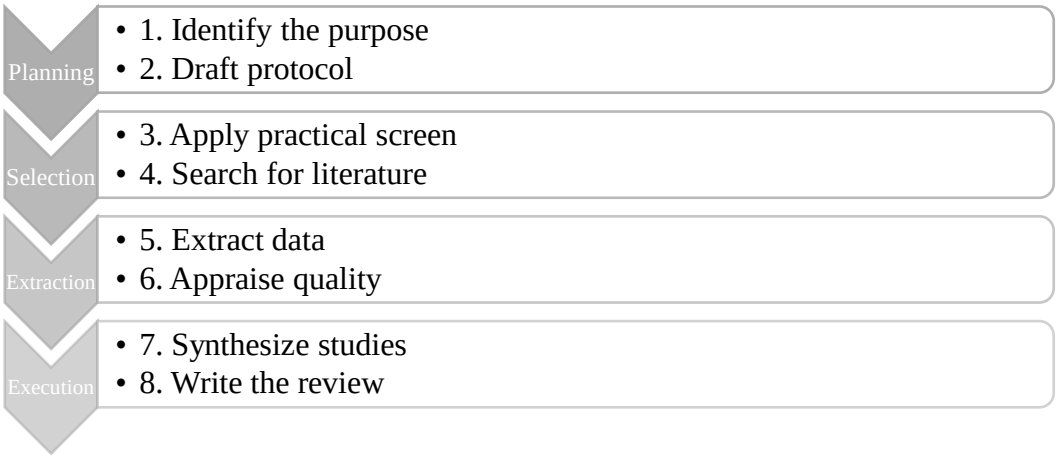


Figure 3 - The steps of a systematic review of literature Source: adapted from Okoli (2015)

2.2.2 PROSPERO and PRISMA-P

In an effort to provide guidance and increase the use of this methodology, the Centre for Reviews and Dissemination, of the University of York, created an international register system called PROSPERO (International Prospective Register of Ongoing Systematic Reviews). With the availability of a protocols' database, the duplication of effort in this phase of the investigation would be avoided, and the publication bias would be reduced (Moher et al., 2016).

In order to help researchers "to create a clear and complete document of their *a priori* methods", Moher et al. (2016) created a reporting guideline called PRISMA-P (Preferred Reporting Items for Systematic reviews and Meta-Analyses for Protocols).

Besides enabling the steps of a Systematic Literature Review, a protocol empowers transparency and trustworthiness on the investigation (Amendoeira et al., 2022).

The following protocol was adapted from the most recent PRISMA-P Checklist, from 2020 (in Annex 1 and 2).

2.3 Protocol for Systematic Literature Review

Below it is established the protocol used as a guideline during the research.

2.3.1 Review objective

Identify studies that propose the application of advanced analytical methods for detecting fraud. Analyse the categorization made for the different advanced analytical methods and for the types of fraud. Analyse the relationships established between the various categories of analytical methods and types of fraud.

2.3.2 Review question

How have advanced analytical methods been applied in fraud detection?

2.3.3 Specific investigation questions

i. What advanced analytical methods are there?
ii. What types of fraud are there? How are they categorized?
iii. How can analytical methods for fraud detection be categorized?
iv. How are advanced analytical methods selected to detect each type of fraud?
v. How to apply advanced analytical methods in fraud detection?
vi. What types of fraud signs can be detected using advanced analytical methods? And what types of fraud?
vii. What are the most used advanced analytical methods in fraud detection?
viii. How to measure the reliability of the use of analytical methods in the detection of fraud? What are the most reliable analytical methods for detecting fraud?

2.3.4 Inclusion criteria

- Free access publications, available through *Scopus* database;
- Recent publications, published between 2003 and 2022;

- Publications containing "Fraud Detection" as one of the keywords (Indexed Keyword or Author Keyword), and at least one advanced analytical method as another keyword.

2.3.5 Exclusion criteria

- Publications unavailable, through *Scopus*, other databases or Journal's website;
- Publications unrelated to the theme (i.e. combining advanced analytical methods with fraud detection);
- Publications with little scientific value for the study (i.e. studies that couldn't answer the review question).

2.3.6 Search strategy

Application of the query "KEY ("Fraud Detection")" in *Scopus* "Advanced document search".

Depending on the number of results and the characteristics of the results, proceed to the application of filters to obtain publications that:

- answer to the review question;
- and comply with the inclusion and exclusion criteria previously established.

2.4 Data collection

The database selected for this research was *Scopus*, after concluding that other databases (namely Web of Science, IEEE, Research Gate, Springer, Arxiv, dblp, EBSCO) would not add value to the search, since the documents indexed in those databases were also available in *Scopus*. Therefore, using additional databases would create duplicate results.

The query "KEY ("Fraud Detection")" was applied in *Scopus* "Advanced document search", resulting in 3073 documents.

Afterwards, the following filters were applied:

- "Document type" was limited to "Article", resulting in 1032 document results;
- "Subject Area" was limited to "Computer Science" and "Business, Management and Accounting", resulting in 815 document results;
- "Keywords" were limited to all advanced analytical methods used in more than 9 documents[1], resulting in 191 documents:
    - "Decision Trees" (75);
    - "Outlier Detection" (41);
    - "Clustering Algorithms" (29);
    - "Logistic Regression" (23);
    - "Clustering" (19);
    - "Random Forests" (17);
    - "Adaptive Boosting" (16);
    - "Signal Detection" (15);
    - "Bayesian Networks" (13);
    - "Random Forest" (12);
    - "Benford's Law" (9).

[1] In parenthesis on the list are the number of documents with the mentioned keywords.

- The results without open access were excluded, leaving 60 document results.

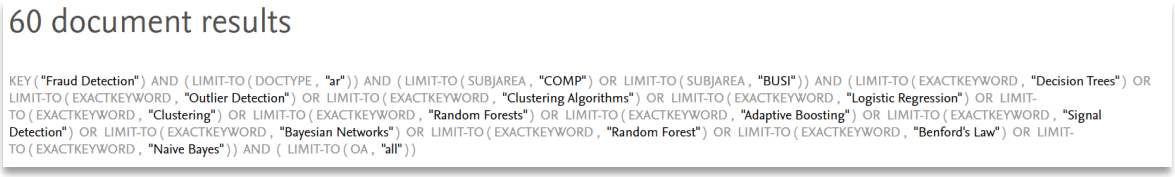


Figure 4 – SLR Search results (Extracted from Scopus)

Due to access restrictions, 55 out of the 60 documents were analysed.

3 Advanced analytical methods in fraud detection

The 55 documents published in *Scopus* between 2003 and 1 of February 2023 were analysed and categorized regarding the advanced analytical method(s) applied and the type of fraud(s) object of study. The most representative ones will be discussed ahead.

After classifying the 55 documents, 110 results were found, as each article studied one or more advanced analytical method in one or more types of fraud (Annex 3).

3.1 Advanced analytical methods

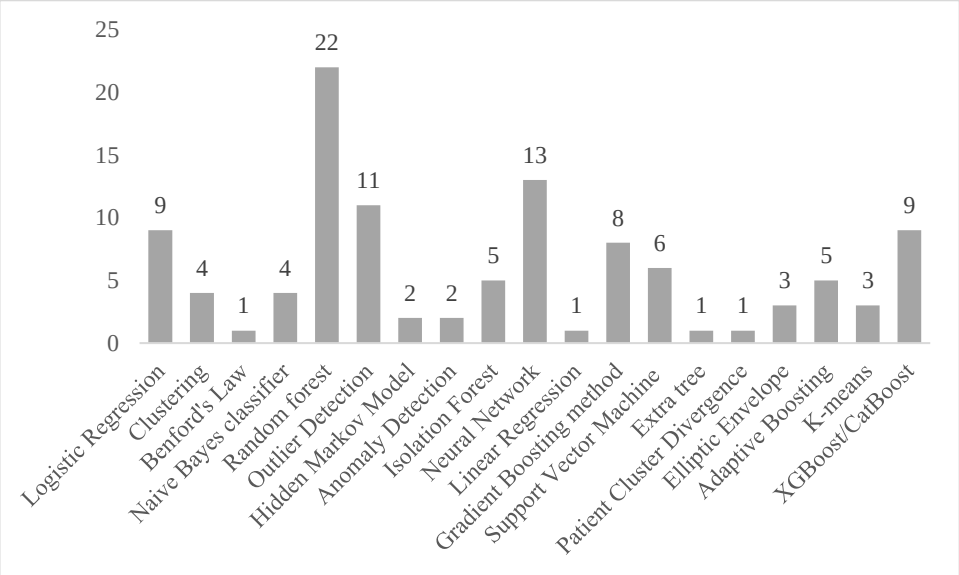
The advanced analytical methods with higher occurrence where split in the aforementioned data mining approaches - (1) Clustering, (2) Classification or (3) Prediction. Some methods were broken down into more specific methods (e.g. Isolation Forest and Random Forest are based on the Decision Tree Algorithm so they were classified as more specific methods).

These relations were schematized in the table below:

Table 1 - Relation between data mining modalities and advanced analytical methods

Data Mining Modalities	Advanced Analytic Methods	Specific Methods
Clustering	Clustering	
	Patient Clustering Divergence	
	Elliptic Envelope	
	K-means	
Classification	Decision Tree	Isolation Forest
		Random Forest
	Neural Networks	
	Bayesian Networks	
	Support Vector Machine	
	Gradient Boosting	Adaptive boosting
		XGBoost
		CatBoost
Prediction	Linear Regression	
	Logistic Regression	
	Benford's Law	
	Hidden Markov Model (HMM)	

In Graph 1 are represented the advanced analytical methods found of in the results.



Graph 1 - Advanced analytical methods in SLR results

The classification techniques with higher occurrence are: (i) Random Forest; (ii) Neural Network), (iii) Outlier Detection; (iv) XGBoost and CatBoost and (v) Logistic Regression.

(i) Random Forest is a machine learning technique that integrates several decision trees (Ashfaq et al., 2022), combining their output in an unique result, based on the majority vote. Decision tree, in its turn, is a technique that splits data into different categories and classifies it "from the root to the leaf node" while "highlights the structural information in the data" (Valavan & Rita, 2023). Easily understandable by different stakeholders, Decision Trees can manage missing information while handling high volumes of data with a lot of attributes. These techniques can " ensemble a learning model for classification, regression, and other tasks" (Lin & Jiang, 2021). Random forests are far used because they work with both categorical/qualitative and numerical/quantitative data (Sharma et al., 2021; Valavan & Rita, 2023). The bias of Random Forest is the same of any individual Decision Tree, making this a more robust model over Decision Trees;

(ii) Neural Network is a series of algorithms composed by neurons who "rely on training data to learn and improve their accuracy over time" (IBM, n.d.). This

subset of machine learning is inspired in human nervous system, as the signal obtained in the input layer is carried out to the hidden layers (Nguyen et al., 2022). These deep learning methods can be categorized in Convolutional Neural Networks (CNN) and Recurrent Neutral Networks (RNN) (Dong et al., 2020). They differentiate from one another through the ability to process temporal information, that can only be done using RNN. On the other hand, CNN can be useful to analyse spatial data (e.g. images);

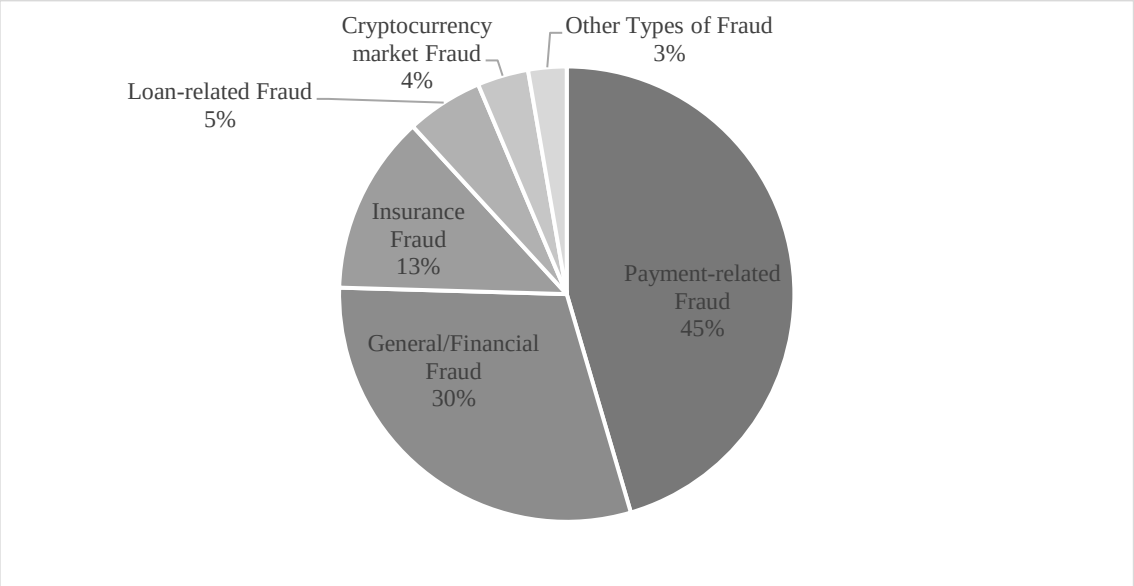
(iii) Outlier Detection is technique that aims to identify data points that differ significantly from the remaining dataset (Elmogy et al., 2021). As these algorithms detect variations and abnormal behaviours, they have been exhaustively developed and applied in different realities. Although the causes might be related to simple changes in the environment, instrumentation error or human error (Hassaan et al., 2021), they have been crucial to detect malicious activity. Detecting outliers is a quite useful task when preparing large datasets and is normally applied in preliminary stages of machine learning models, to highlight and help to understand information;

(iv) XGBoost and CatBoost are two methods based on Gradient Boosted Trees. Both GBT and Random Forests use Decision Trees; yet the two algorithms differ in the way individual trees are built and in the way the results are combined. Boosting collects weak learners (i.e. predictors with poor accuracy) and transforms them in a strong learner (i.e. predictor with high accuracy). In GBT. the trees are built consecutively so that new trees learn from the previous ones; in Random Forests, trees are built independently and combined in parallel. XGBoost, or Extreme Gradient Boosting, generates sequential trees "and each successive tree aims to reduce the error of the previous tree and update the residual error" (Ashfaq et al., 2022). Whereas this method creates asymmetric trees (i.e. splitting condition for each node across the same depth can differ), CatBoost, or Categorical Boosting, creates symmetric trees or balanced trees. This algorithm, developed after XGBoost, is seen as an improved version of the other boosted trees algorithms (Nguyen et al., 2022) since the splitting condition is consistent across all nodes at the same depth of the tree.

(v) Logistic Regression is a machine learning algorithm from the same group as Linear Regression. Logistic Regression provides a discreet output, predicting if the evaluated data is "True" or "False" (i.e. binary classification). On the other hand, Linear Regression provides a linear output, evaluating the correlation between data and determining if the value found is statistically significant. Although Logistic Regression also provides a statistic output, it is mainly used for classification (e.g. if the percentage of a value is high, then is classified as True). This algorithm is known for its "efficiency of detecting frauds based on its ability to isolate the data that belong to different binary classes" (Alenzi & Aljehane, 2020).

3.2 Types of Fraud

In Graph 2 are represented the classifications of the Types of Fraud mentioned in the articles studied.



Graph 2 - Types of fraud in SLR results

Payment-related fraud is the type of fraud with more occurrences: 45% of the results. This category aggregate types of fraud related with payments, namely credit card fraud (Alenzi & Aljehane, 2020; Alfaiz & Fati, 2022; Carneiro et al., 2022; Esenogho et al., 2022; Y. Fang et al., 2019; Ileberi et al., 2021; Khan et al., 2022; Lin &

Jiang, 2021; Lucas et al., 2020; Muaz et al., 2020; Randhawa et al., 2018; Santosh & Ramesh, 2020; Sasikala et al., 2022; Zhang et al., 2022), card payment fraud (Nguyen et al., 2022) and online payment fraud (Chang et al., 2022; Hajek et al., 2022; Nasr et al., 2022).

General/Financial Fraud category represents 30% of the results, yet it doesn't represent a specific type of fraud. The researchers which articles rely on this tier tested data mining approaches:

- In specific situations with low representation among the results (e.g. frauds and anomalies related to financial data, e-commerce, fintech applications) (Li et al., 2021; Liu et al., 2020; Stojanović & Božić, 2022);
- Or in nonspecific datasets (Rubaidi et al., 2022).

Insurance Fraud category comprises studies where the data mining model aimed to detect insurance related anomalies (Dhieb et al., 2020; Palacio, 2019). The insurance category targeted by the largest number of studies was health insurance (Kotekani & Ilango, 2022; Kotekani & Velchamy, 2020; Sun et al., 2019).

Researchers also identified loans as a sensitive accounting item in which fraud occurs and for which data mining techniques can be used to detect it (W. Fang et al., 2021).

In the last years, publications on cryptocurrency markets fraud were published. Although these still have low representation in this study results, it is worth noting the investment in the research of this trendy topic (Ashfaq et al., 2022; Mittal & Bhatia, 2021; Nerurkar et al., 2021)

Other types of fraud include non-financial types of fraud (e.g. medical prescriptions fraud) (Aral et al., 2012) and studies in which the models examined had as objective identify fraudsters (Bhargava et al., 2003).

3.3 Review conclusions

According to the results obtained with this systematic review of literature, it is concluded that the type of fraud for which there is more articles published is payment-related frauds, namely credit card fraud.

The advanced analytical methods most used to develop fraud and anomaly detection models are Random Forests, Neutral Networks, Outlier Detection, Logistic Regression and XGBoost/CatBoost.

It is also possible to conclude that the most recurrent advanced analytical methods are Decision Trees ensembles, in particular Random Forests, XGBoost and CatBoost. Also, Isolation Forests, Extra Trees and Adaptive Boosting were found among the results.

The majority of the methods – both in the top 5 and in the overall results – are Supervised Learning methods. In this approach, input and output data are provided to the model with the goal of training it so it can predict results when new data is given. On the other side, Unsupervised Learning uses machine learning methods to cluster unlabelled datasets and discover patterns and outliers without human intervention.

Within Supervised machine learning, the category of methods in which researchers tend to investigate is Classification, since the ultimate goal is to classify if a certain transaction is fraudulent or non-fraudulent.

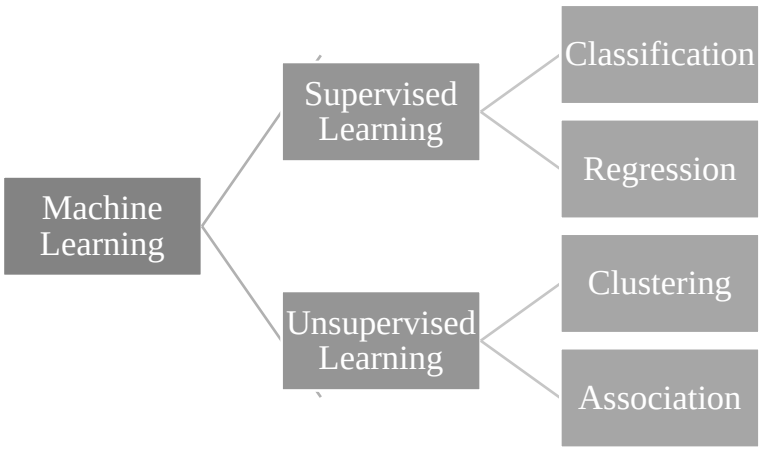


Figure 5 - Machine Learning Approaches

4 Bibliometric Approach

In Chapter 3 the results of a standard database search were explained, narrowing them down by using the filters available in the search engine *Scopus*.

In the present Chapter, the results will be selected through an advanced search in the same database using a query, in order to measure the scientific production related to advanced analytical methods for fraud detection.

The study on the scientific production of the topic advanced analytical methods for fraud detection is relevant for academic purposes of future thesis but also to acknowledge the authors and journals who have been studying this matter. Therefore, highlighting the trends on advanced analytical methods used to detect fraud will ease future scientific investigations.

The specific objectives of this research are:

i. Analyse bibliometric indicators of the articles' characteristics, namely: the year of publications, the number of citations, the journal in which it was published the articles and the topics investigated; and
ii. Analyse bibliometric indicators regarding the articles' authors, namely: their productivity, the authorship type of the article and the geographic affiliation.

To follow the protocol abovementioned in Chapter 2, the list of specific investigation questions was prepared (in Table 2):

Table 2 - Specific investigation questions used for BA

Specific objectives	Specific investigation questions	
	Reference	Description
i.	SIQ #1	Was there an increase in the number of scientific articles regarding advanced analytical methods for fraud detection?
	SIQ #2	Was there an increase in the number of citations of the articles about advanced analytical methods for fraud detection?
	SIQ #3	Is the research on advanced analytical methods for fraud detection more relevant on journals of which area?
	SIQ #4	Which are the most used keywords on articles about advanced analytical methods for fraud detection?
ii.	SIQ #5	Is there any investigator who prevails in the publication of scientific articles on advanced analytical methods for fraud detection?
	SIQ #6	Is there a prevalence of collective authorship over individual authorship in scientific articles on advanced analytical methods for

		fraud detection?
	SIQ #7	Is there any country and/or continent that prevails in the affiliation of authors in the investigation about advanced analytical methods for fraud detection?

4.1 Bibliometric Indicators – An Overview

Cole and Eales were the pioneers of the bibliometric analysis, when presented in 1917 "a quantitative picture of progress in a field of research", with their statistical analysis of the history of comparative anatomy (Okubo, 1997).

When executing a bibliometric research, it is important to define the criteria to rank the different journals and, consequently, the papers published by that journal, this is, to apply the most appropriate indicators in order to measure the scientific activity.

Sengupta (1986) discussed how the ranking of scientific periodicals should not be strictly dependent of the citations counting, but also rely on "(1) scientific interest of a journal in relation to total number of articles published; (2) compactness of information content in a scientific periodical; and (3) scientific value of the papers in relation to compactness of presentation".

The most widely used bibliometric indicators (hereafter, BIs) classification is qualitative and quantitative. Qualitative BIs focus on measuring the quality of the journal and/or the authors, ending up being judgemental indicators more than objective indicators. On the other hand, quantitative BIs are numerical and aim to measure co-relations between authors and journals' scientific activity (García-Villar & García-Santos, 2021).

4.2 Bibliometric Research - Methodology

In this analysis, the production of articles published in journals regarding advanced analytical methods used for fraud detection will be analysed. To do so, several bibliometric indicators that can be divided into indicators of (i) scientific quality, (ii) scientific activity, (iii) scientific impact and (iv) thematic associations will be applied.

The database selected for this research was, again, *Scopus*, since it ensures a broad scientific coverage.

The following query was designed:

(Audit* OR "Fraud Detection" OR "Fraud Prevention" OR "Risk Management") AND ("Data Forensic" OR "Data Analytics" OR "Predictive Analytics" OR "Data Mining" OR "Text Mining" OR "Machine Learning" OR "Deep Learning" OR Big Data OR Social Network* OR "Artificial Intelligence")

The strings that compose this query are related to (1) fraud/risk related terms, (2) data analytics related terms and (3) financial related terms. The Boolean operator "AND" was used to assure each one of the upper mentioned topics were included in the results.

Within each string, the Boolean operator "OR" was used to guarantee the results contained any of the terms. (Scopus Search Guide, 2019)

Although Scopus default search is in Keywords, Title and Abstract, the query was only applied to Keywords, since they are used as the "key" to the article. This item must respect the journal's scope in order to be accepted and published. Consequently, keywords are considering the marketing to appeal to the article's target audience. (The Importance of Using Strategic Keywords in Research Papers, n.d.)

The results were then limited to the subject areas "Business, Management and Accounting" and "Computer Science" and to the document type "Article".

The sample of 337 documents, published between 2000 and 11 of July 2023, were considered eligible for the study.

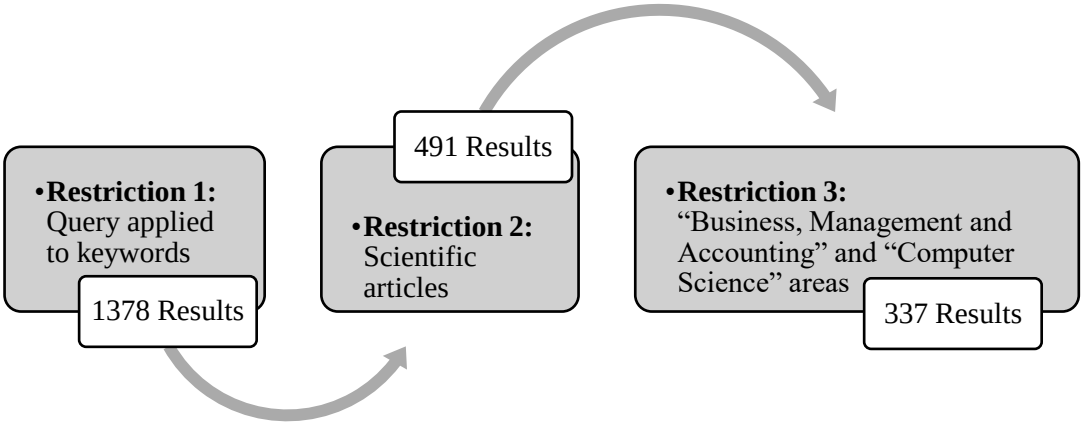


Figure 6 - Flowchart of the application of restrictions in BA

A *csv file with the abovementioned results was exported from *Scopus*. This file - containing the information on all results Author(s), Document Title, Year, Source Title, Citation Count, Digital Object Identifier (DOI), Affiliations e Author Keywords - was used to process and analyse the data in *Excel* and in *VOSviewer*, a software developed in 2010 by Nees Jan van Eck and Ludo Waltman in University of Leiden.

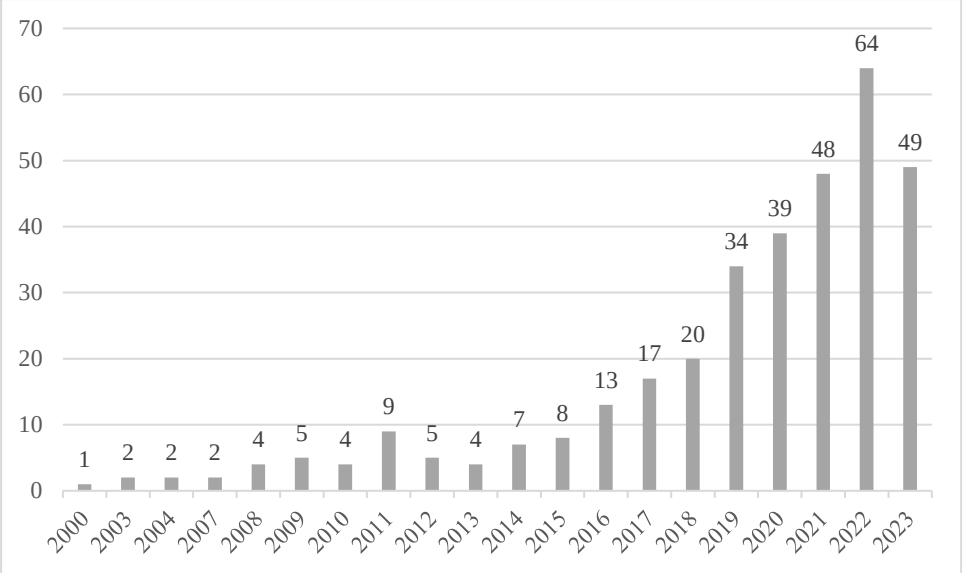
The metrics provided by *Scopus* (i.e. "Analyze results" feature) were also analysed, including h-index.

4.3 Indicators of scientific activity

The indicators of scientific activity that will be analysed in this study are: the evolution of the number of articles published by the researcher over the years, their productivity, collaboration in the authorship of the studies and geographic collaboration.

4.3.1 Number of articles

As can be seen from Graph 3, this search comprises a time horizon of 23 years, with the first article being published in 2003 and the most recent ones in the current year, 2023.



Graph 3 - Number of publications per year in BA results

The number of publications regarding advanced analytical methods for fraud detection has been increasing over the years, presenting consistent growth rates in the last 10 years.

Most of the articles has been published in the last 5 years, considering the number of articles from 2018 to 2022 (since 2023 data does not represent a full year). The year with the highest number of publications is last year, 2022.

In conclusion, the answer to SIQ #1 is positive, as there has been a clear increase in the number of scientific articles regarding advanced analytical methods for fraud detection.

4.3.2 Authors' productivity

The authors' productivity indicator attempts to clarify which researchers contribute the most for the developments on a topic.

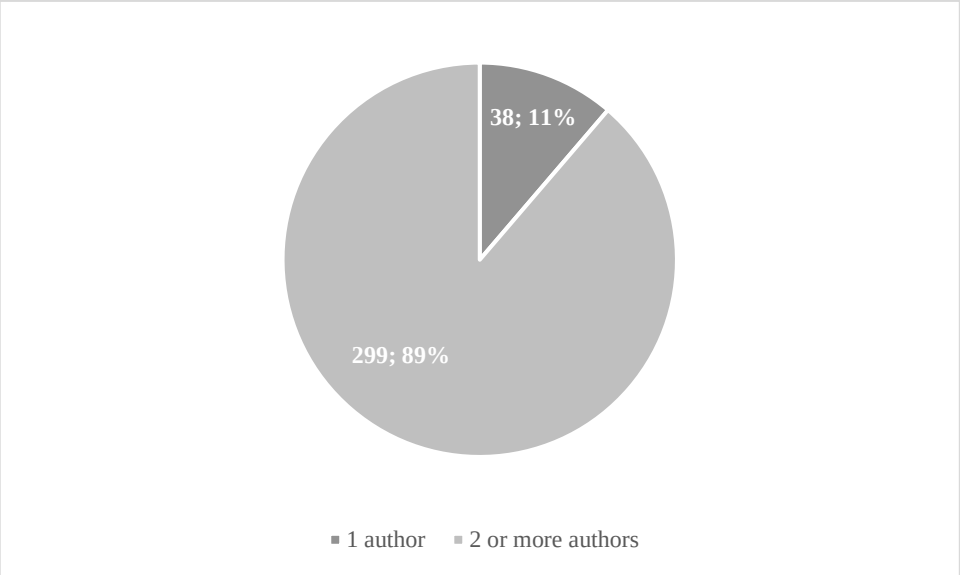
In this case, none of the authors has published more than 2 articles of the sample analysed.

The answer to the SIQ #5 question is no, as there were not an investigator who prevailed in the publication of scientific articles on advanced analytical methods for fraud detection.

4.3.3 Collaboration in the authorship of the studies

An article can be published by only one author (i.e. individual authorship) or by more than one author (i.e. collaboration).

In the Graph 4 is represented the distribution of type of authorship of the analysed sample: 38 articles (11%) were written by a single author and the remaining 299 (89%) were result of a authorship collaboration.



Graph 4 - Types of authorship in BA results

It is possible to conclude that yes, there is a prevalence of collective authorship over individual authorship in scientific articles on advanced analytical methods for fraud detection, as an answer to SIQ #6.

In order to understand the relationship between the authors, it would be interesting to create a network authorship map. However, since the maximum number of articles per author is very low (i.e. 2 articles), there is not enough data in this sample to analyse such relation.

4.3.4 Geographic affiliation

Authors' affiliation was identified through their university's/institution's country at the date of publication.

121 (36%) articles were published in China, followed by United Sates with 69 (20%) and United Kingdom with 27 (8%) articles.

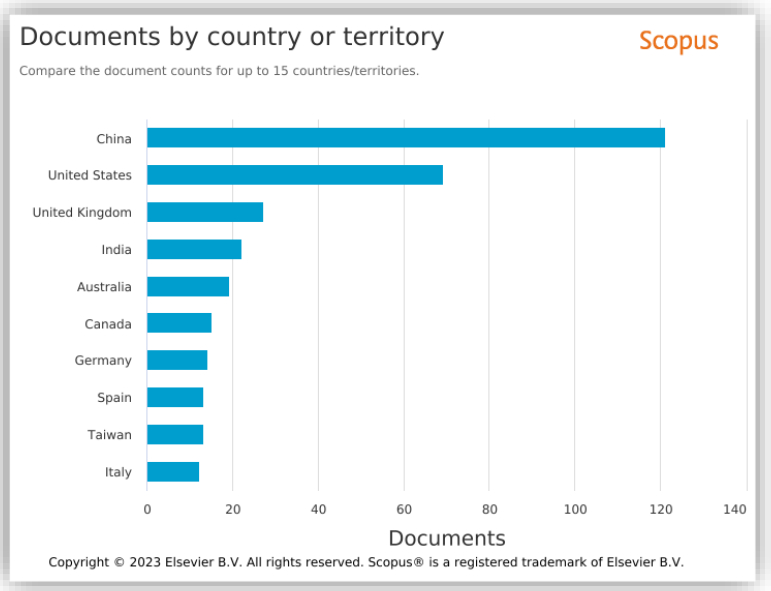


Figure 7 - Tope 10 countries with higher number of articles published in BA results

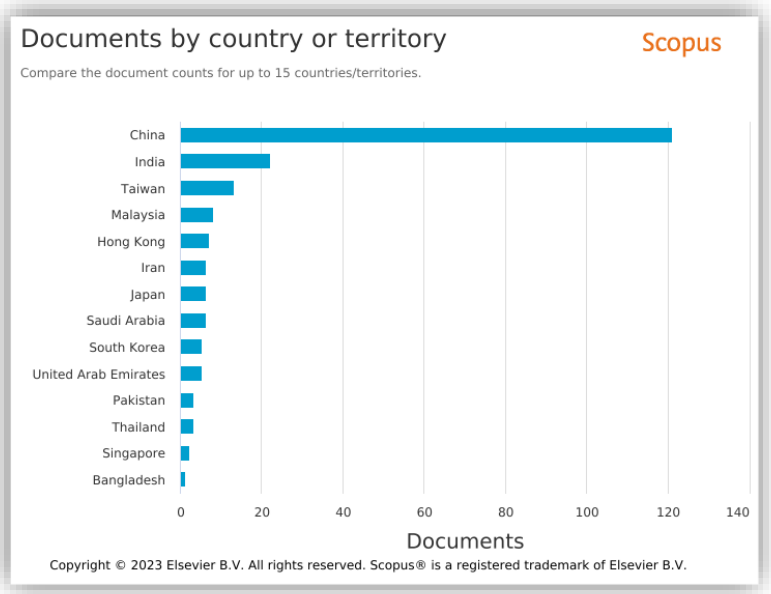


Figure 8 - Number of publications in Asia in BA results

The majority of the articles were published in Asia, with 208 (62%) out of the 337 publications. In response to the question SIQ #7, yes, there is a country that prevails in the affiliation of authors: China. Also, Asia is the continent with higher number of publications about advanced analytical methods for fraud detection.

4.4 Indicators of scientific impact

After analysing the activity production regarding the advanced analytical methods used for fraud detection, it is important to assess the impact of that production.

The indicators of scientific impact can refer to impact indicators on the scientific community or impact indicators on publication sources (Costa et al., 2012).

In this study, to analyse the scientific impact, the h-index indicator will be used. Besides this analysis on the number of citations of each article, the CiteScore will be analysed in order to measure the impact of each journal.

4.4.1 Number of citations

The more citations an article has, the higher is the influence it has in the scientific community, which can be used to understand the trends and define future investigations.

As of July 28 of 2023, there are 25 articles with more than 100 citations. The article with the higher number of citations is "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature" with 669 citations to date.

Table 3 - Articles with more than 100 citations in BA results

Title	Authors	Cited by
The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature	Ngai E.W.T.; Hu Y.; Wong Y.H.; Chen Y.; Sun X.	669
The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients	Yeh I.-C.; Lien C.-h.	432
Data Mining techniques for the detection of fraudulent financial statements	Kirkos E.; Spathis C.; Manolopoulos Y.	410
A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0	Ivanov D.; Dolgui A.	395
On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms	Yamanishi K.; Takeuchi J.-I.; Williams G.; Milne P.	299

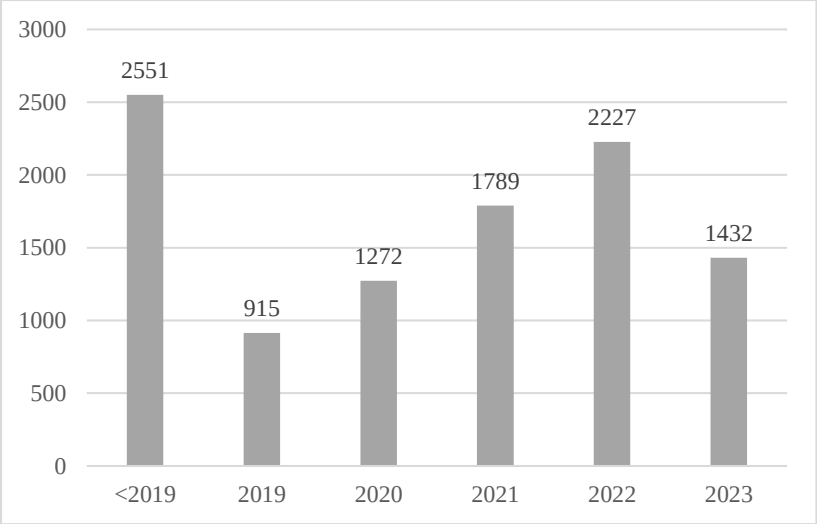
Detection of financial statement fraud and feature selection using data mining techniques	Ravisankar P.; Ravi V.; Raghava Rao G.; Bose I.	295
Using generative adversarial networks for improving classification effectiveness in credit card fraud detection	Fiore U.; De Santis A.; Perla F.; Zanetti P.; Palmieri F.	281
Feature engineering strategies for credit card fraud detection	Correa Bahnsen A.; Aouada D.; Stojanovic A.; Ottersten B.	210
Effective detection of sophisticated online banking fraud on extremely imbalanced data	Wei W.; Li J.; Cao L.; Ou Y.; Chen J.	203
A comparison of models for predicting early hospital readmissions	Futoma J.; Morris J.; Lucas J.	187
Survey of data management and analysis in disaster situations	Hristidis V.; Chen S.-C.; Li T.; Luis S.; Deng Y.	171
Recent Development in Big Data Analytics for Business Operations and Risk Management	Choi T.-M.; Chan H.K.; Yue X.	168
A data mining based system for credit-card fraud detection in e-tail	Carneiro N.; Figueira G.; Costa M.	161
Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events	Gaspar R.; Pedro C.; Panagiotopoulos P.; Seibt B.	139
Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods	Hajek P.; Henriques R.	135
Software project risk analysis using Bayesian networks with causality constraints	Hu Y.; Zhang X.; Ngai E.W.T.; Cai R.; Liu M.	130
SCARFF: A scalable framework for streaming credit card fraud detection with spark	Carcillo F.; Dal Pozzolo A.; Le Borgne Y.-A.; Caelen O.; Mazzer Y.; Bontempi G.	128
Fraud detection: A systematic literature review of graph-based anomaly detection approaches	Pourhabibi T.; Ong K.-L.; Kam B.H.; Boo Y.L.	124
An intraday market risk management approach based on textual analysis	Groth S.S.; Muntermann J.	116
Automatic identification of eyewitness messages on twitter during disasters	Zahra K.; Imran M.; Ostermann F.O.	110
Enabling Cloud Computing in Emergency Management Systems	Qiu M.; Ming Z.; Wang J.; Yang L.T.; Xiang Y.	110
Detecting evolutionary financial statement fraud	Zhou W.; Kapoor G.	110
An overview of social network analysis	Oliveira M.; Gama J.	106
Machine Learning Algorithms for Construction Projects Delay Risk Prediction	Gondia A.; Siam A.; El-Dakhakhni W.; Nassar A.H.	101

In Graph 5 it is represented the evolution of the number of citations over the year.

To note that the number of citations previous to 2019 are aggregated in a single sum. Also, to note that this analysis was performed during the year of 2023, therefore not being considered.

Nevertheless, the answer to SIQ #2 is positive, as there was an increase in the number of citations of the articles about advanced analytical methods for fraud detection over the years.

It is possible to conclude that the increase in the number of citations confirms that tis topic is becoming more relevant every year.



Graph 5 - Evolution of the number of citations in BA results

4.4.1.1 H-index

The h-index attempts to measure the productivity and the impact of the published work from a certain author. It can be applied to a single publication or to a set of publications, journals, countries, between other factors.

Through the "Analyse results" *Scopus* database's feature it is possible to apply the h-index in the analysed sample, resulting in 50. This means that, for the 337 analysed articles, 50 were cited at least 50 times.

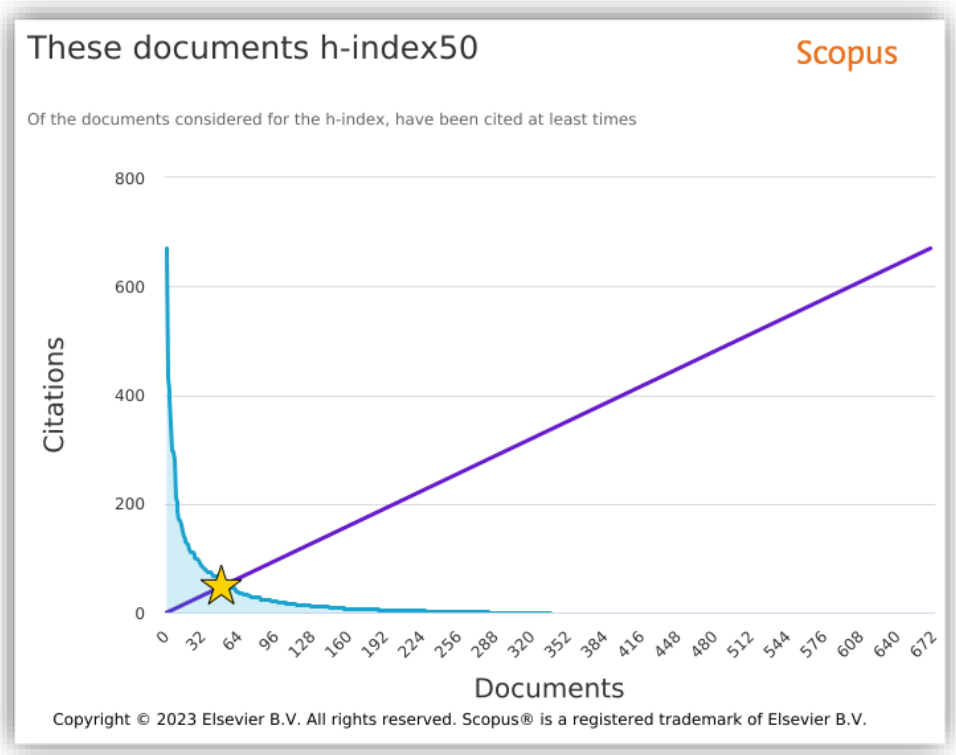


Figure 9 - H-index in BA results

4.4.2 Journals' influence and reputation

The impact factor of a journal is measured through the SCImago Journal Rank (hereafter, SJR). This indicator is a measure of prestige pf scholarly journals, impacted by both the number of citations received by a journal and the prestige of the journal where the citations came from. The rating goes from Q1, the most prestigious journals in the field with the higher number of citations, to Q4, that gathers journals with lower impact factors (*Scimago Journal & Country Rank*, n.d.).

The 337 articles sampled in this study were published by 192 different journals.

To better understand the journals that publish more articles about advanced analytical methods for fraud detection, were considered journals with more than 5 (1%) articles from the sample:

Table 4 - Journals with more than 5 articles published in BA results

Journal	**N.º Articles**	**%**
IEEE Access	25	7%
Expert Systems with Applications	17	5%
IEEE Journal of Biomedical and Health Informatics	11	3%
Decision Support Systems	10	3%
International Journal of Advanced Computer Science and Applications	7	2%

25 (7%) articles were published by IEEE Access, a professional association for electronics engineering, electrical engineering, and other related disciplines. 11 (3%) articles were published by a journal of the same association: IEEE Journal of Biomedical and Health Informatics. Both these journals present a Q1 as of 2022.

The other three journals with more than 5 articles published are: Expert Systems with Applications, published by Elsevier; Decision Support Systems, also published by Elsevier; and International Journal of Advanced Computer Science and Applications, published by the Science and Information Organization. Both Elsevier journals are Q1 journals in 2022 ranking, whereas the Science and Information Organization journal is a Q3.

In conclusion, the answer to the SIQ #3 is yes, the research on advanced analytical methods for fraud detection is more relevant on journals of Computer Science area.

Although the search was restricted to both "Business, Management and Accounting" and "Computer Science" areas, the second one presents more articles and in prestigious journals.

4.5 Indicators of thematic associations

Lastly, it is relevant to analyse thematic associations of the sampled articles. To do so, the most frequent keywords were analysed.

4.5.1 Keywords analysis

In order to understand which are the advanced analytical methods for fraud detection subject to more investigations, *VOSviewer* was used to create a network map. In this map, each circle represents a theme – which means, the bigger the circle, the higher is the number of occurrences in the sample. Therefore, more relevant it is -, and each colour represents a cluster. The lines between the circles represent links, and the a smaller distance between circles represents a higher level of co-relation between the items (van Eck & Waltman, 2018).

The threshold chosen to create this network map was the minimum of 5 occurrences of a keyword. From the 1113 keywords, 26 meet the criteria. These will be analysed in order to answer the SIQ #4:“Which are the most used keywords on articles about advanced analytical methods for fraud detection?”.

To avoid that similar keywords would appear in different circles, the input data of the network map was manipulated. This way, some keywords were merged and/or replaced by similar ones, for example “neural network” and neural networks”.

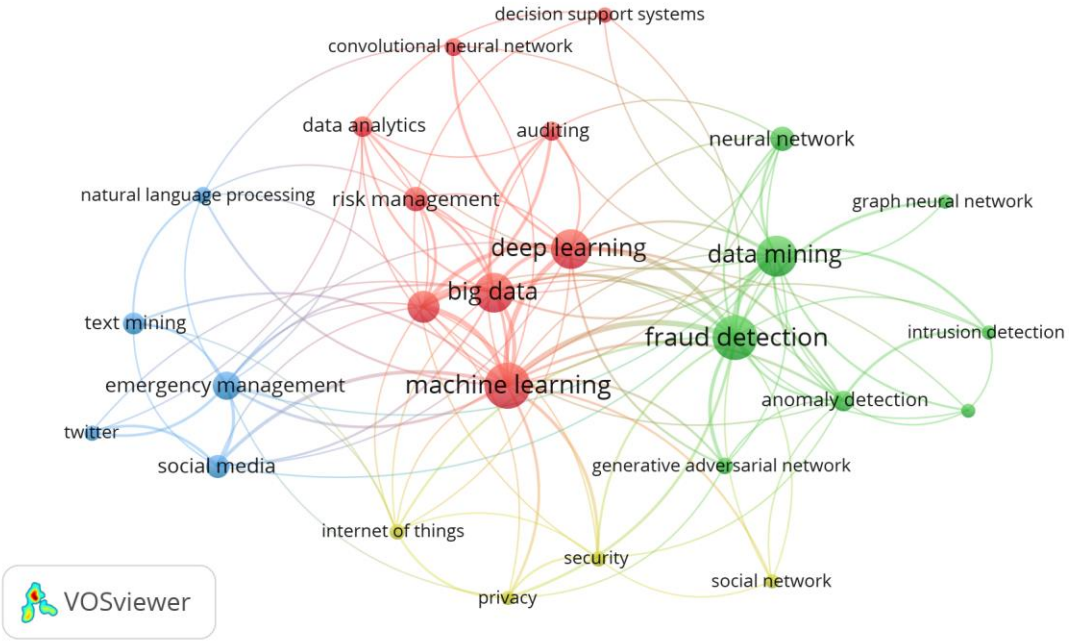


Figure 10 - VOSviewer authors' keywords network map in BA results

The keywords “deep learning”, “big data” and “machine learning” have the higher number of occurrences, with a total of 39, 40 and 53 occurrences, respectively, and connect all the clusters.

Regarding the advanced analytical method with the higher number of occurrences is “neural network”, with 15 occurrences. This keyword is associated with “farud detection”, “data mining” and “anomaly detection”. So, as a conclusion, taking into consideration the analysis on studies performed over the last 20 years, neural networks are the most appropriate advanced analytical method for fraud detection.

This method was already explored in Chapter 3, as it was one of the advanced analytical methods resulting from the systematic literature review.

The advanced analytical method "outlier detection" presents 5 occurrences as a authors' keyword and was also subject to study in Chapter 3.

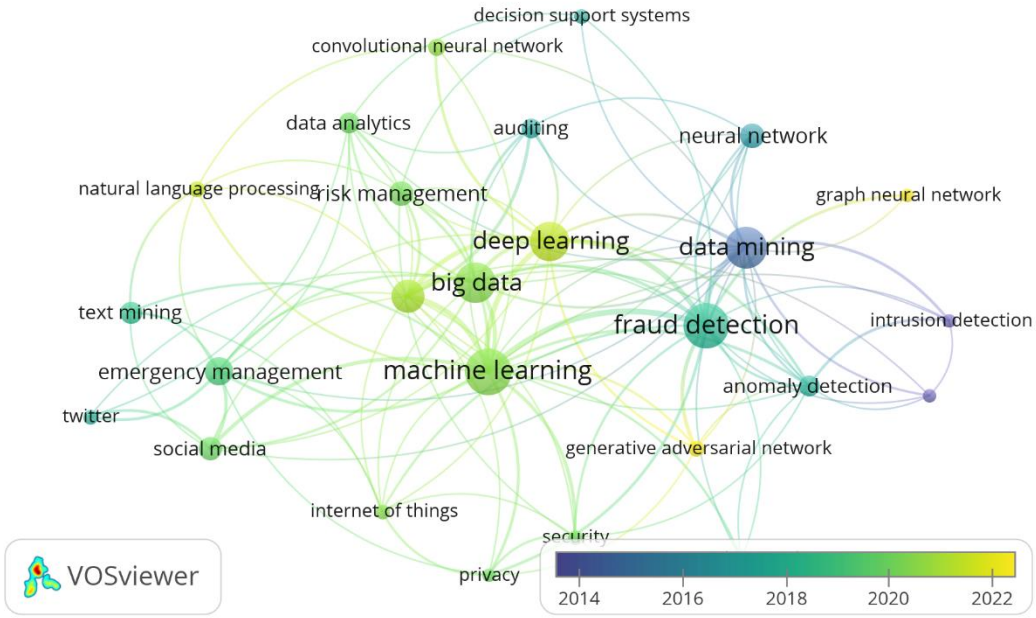


Figure 11 - VOSviewer Authors' keywords - Evolution per year in BA results

The keywords "convolutional neural network", "generative adversarial network" and "graph neural network" were also used in 5 or more sampled articles, namely used in 8, 7 and 5 results, respectively. These are variations and adaptations of neural networks and were not combined with the main keyword to highlight the diversity of neural network approaches being investigated over the last years, as visible in Figure 11.

CONCLUSION

This thesis had as main objective to answer the question "How have advanced analytical methods been applied in fraud detection?". By performing a Systematic Literature Review, reinforced with a Bibliometric Indicators approach, the present study contributes to a better understating of the literature published in the last 20 years regarding advanced analytical methods and its application in the detection of fraud.

In the beginning of the dissertation, a theoretical framework was defined, through the concept of big data, data mining and the role of the auditor in the detection of fraud.

Afterwards, it was given an explanation on the methodology, a protocol of Systematic Literature Review was defined, with resource to the PRISMA 2020 checklist. The applications of the necessary filters and restrictions was presented, culminating in the search results' analysis.

The advanced analytical methods and types of fraud mentioned in the 55 articles were summarised and the conclusions were explained.

In the last Chapter, a Bibliometric approach was taken. Firstly, the bibliometric indicators' framework was justified, and then the Specific Investigation Questions (hereafter, SIQs) were justified. Then indicators of scientific activity, indicators of scientific impact and indicators of thematic associations were applied to a larger branch of results than the SLR and each SIQ was answered.

Research Contribution

The contributions of the SLR are based on the answer to the research question, this is, the results show which advanced analytical methods have been applied in the detection of fraud.

In accordance with the results, the investigations performed in the last 20 years have been focusing on payment-related frauds, namely credit card frauds. In terms of advanced analytical methods, researchers have been using Random Forests, Neutral Networks, Outlier Detection, Logistic Regression and XGBoost/CatBoost to the develop mechanisms to detect fraud.

The bibliometric approach reinforced some of the conclusions already stated in the SLR, and enriched the characterization of the publications: (i) the number of scientific articles regarding advanced analytical methods for fraud detection has been increasing over the past years; (ii) the same is true for the number of citations; (iii) the most relevant journals with published articles on this topic are Computer Science related journals; (iv) considering the keywords used by the authors, the advanced analytical method that presents more occurrences is neural networks; (v) there is not a prevalent author publishing investigations' conclusions on this topic; (vi) there is a prevalence of co-authorship over individual authorship; and (vii) the majority of the articles were published in Asia, namely in China.

Research Limitations

During this study, the most relevant limitation found was related to the keywords used by the authors.

In the beginning it was made the decision to perform the SLR and BA searches solely on the keywords. Although the number of results would increase if the query was also applied in the articles' title and abstract, the authors' keywords should be selected in a way that other researchers should feel invited to read the article and get to know the investigation. However, the present investigation revealed that authors commonly choose broader keywords, making some investigations harder to perform.

In this thesis, finding investigations that handle the treatment and application of data in the detection of anomalies in such a comprehensive way made it difficult to fulfil the main objective.

Some of the bibliometric indicators could not be analysed due to lack of occurrences in the search results. For example, if there was one or more outstanding authors in terms of number of publications in the analysed sample, it would be interesting to assess the relationship between those authors and the articles with higher number of citations.

Future Work

In future work it would be interesting to perform this investigation in a more in-depth way, to better understand what type of studies have been published regarding the use of advanced analytical methods in the detection of fraud. To do so, a preliminary analysis of the databases could be useful, to learn how have these studies been published, and which characteristics of the articles call for further analysis.

It could also be valuable to perform similar studies in collaboration with specialists in data science and/or information systems managers. That way, a case study could be developed using a combination of methods or creating a specific model.

REFERENCES

Alenzi, H. Z., & Aljehane, N. O. (2020). Fraud Detection in Credit Cards using Logistic Regression. *International Journal of Advanced Computer Science and Applications*, *11*(12), 540–551. https://doi.org/10.14569/IJACSA.2020.0111265

Alfaiz, N. S., & Fati, S. M. (2022). Enhanced Credit Card Fraud Detection Model Using Machine Learning. *Electronics (Switzerland)*, *11*(4). https://doi.org/10.3390/electronics11040662

Amendoeira, J., Silva, M. R. da, Ferreira, M. R., & Dias, H. (2022). *Revisão sistemática de literatura: a scoping review*. Instituto Politécnico de Santarém. https://repositorio.ipsantarem.pt/handle/10400.15/3784

Aral, K. D., Güvenir, H. A., Sabuncuoğlu, I., & Akar, A. R. (2012). A prescription fraud detection model. *Computer Methods and Programs in Biomedicine*, *106*(1), 37–46. https://doi.org/10.1016/j.cmpb.2011.09.003

Arens, A. A., Elder, R. J., & Beasley, M. S. (2014). Auditing and Assurance Services - An Integrated Approach. In *Book1*. Prentice Hall.

Arockia Panimalar, S., Varnekha Shree, S., & Veneshia Kathrine, A. (2017). The 17 V's of Big Data. *International Research Journal of Engineering and Technology(IRJET)*, *4*(9), 3–6. https://irjet.net/archives/V4/i9/IRJET-V4I957.pdf

Ashfaq, T., Khalid, R., Yahaya, A. S., Aslam, S., Azar, A. T., Alsafari, S., & Hameed, I. A. (2022). A Machine Learning and Blockchain Based Efficient Fraud Detection Mechanism. *Sensors*, *22*(19). https://doi.org/10.3390/s22197162

Barry, E. S., Merkebu, J., & Varpio, L. (2022). State-of-the-art literature review methodology: A six-step approach for knowledge synthesis. *Perspectives on Medical Education*, *11*(5), 281–288. https://doi.org/10.1007/s40037-022-00725-9

Bhargava, B., Zhong, Y., & Lu, Y. (2003). Fraud formalization and detection. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 2737). https://doi.org/10.1007/978-3-540-45228-7_33

Carneiro, E. M., Forster, C. H. Q., Mialaret, L. F. S., Dias, L. A. V., & da Cunha, A. M. (2022). High-Cardinality Categorical Attributes and Credit Card Fraud Detection. *Mathematics*, *10*(20). https://doi.org/10.3390/math10203808

Chang, V., Doan, L. M. T., Di Stefano, A., Sun, Z., & Fortino, G. (2022). Digital payment fraud detection methods in digital ages and Industry 4.0. *Computers and Electrical Engineering*, *100*. https://doi.org/10.1016/j.compeleceng.2022.107734

Costa, T., Lopes, S., Fernández-Llimós, F., & Amante, M., & Lopes, P. (2012). A Bibliometria e a Avaliação da Produção Científica: indicadores e ferramentas. *ACTAS – Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas*, *11*, 1–7.

Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A Secure AI-Driven Architecture for Automated Insurance Systems: Fraud Detection and Risk Measurement. *IEEE Access*, *8*, 58546–58558. https://doi.org/10.1109/ACCESS.2020.2983300

Dong, M., Yao, L., Wang, X., Benatallah, B., Huang, C., & Ning, X. (2020). Opinion fraud detection via neural autoencoder decision forest. *Pattern Recognition Letters*, *132*, 21–29. https://doi.org/10.1016/j.patrec.2018.07.013

Elmogy, A., Rizk, H., & Sarhan, A. M. (2021). OFCOD: On the fly clustering based outlier detection framework. *Data*, *6*(1), 1–20. https://doi.org/10.3390/data6010001

Esenogho, E., Mienye, I. D., Swart, T. G., Aruleba, K., & Obaido, G. (2022). A Neural Network Ensemble with Feature Engineering for Improved Credit Card Fraud Detection. *IEEE Access*, *10*, 16400–16407. https://doi.org/10.1109/ACCESS.2022.3148298

Fang, W., Li, X., Zhou, P., Yan, J., Jiang, D., & Zhou, T. (2021). Deep Learning Anti-Fraud Model for Internet Loan: Where We Are Going. *IEEE Access*, *9*, 9777–9784. https://doi.org/10.1109/ACCESS.2021.3051079

Fang, Y., Zhang, Y., & Huang, C. (2019). Credit card fraud detection based on machine learning. *Computers, Materials and Continua*, *61*(1), 185–195.

https://doi.org/10.32604/cmc.2019.06144

Fink, A. (2005). *Conducting research literature reviews: From the Internet to paper* (2nd editio). SAGE Publications, Inc.

García-Villar, C., & García-Santos, J. M. (2021). Bibliometric indicators to evaluate scientific activity. In *Radiologia* (Vol. 63, Issue 3, pp. 228–235). https://doi.org/10.1016/j.rx.2021.01.002

Hajek, P., Abedin, M. Z., & Sivarajah, U. (2022). Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-022-10346-6

Hassaan, M., Maher, H., & Gouda, K. (2021). A Fast and Efficient Algorithm for Outlier Detection Over Data Streams. *International Journal of Advanced Computer Science and Applications*, *12*(11), 749–756. https://doi.org/10.14569/IJACSA.2021.0121185

IBM. (n.d.). *What are neural networks?* https://www.ibm.com/topics/neural-networks

IBM. (2011). IBM SPSS Modeler CRISP-DM Guide. *Career: Data and Analytics*, 53.

IBM SPSS Modeler CRISP-DM Guide. (n.d.).

Ileberi, E., Sun, Y., & Wang, Z. (2021). Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost. *IEEE Access*, *9*, 165286–165294. https://doi.org/10.1109/ACCESS.2021.3134330

ISACA. (2012). *Enabling Information Superiority*. http://www.globalservices.bt.com/static/assets/pdf/Insights and Ideas/DSiC_WhitePaper_190115.pdf

Ishwarappa, & Anuradha, J. (2015). A brief introduction on big data 5Vs characteristics and hadoop technology. *Procedia Computer Science*, *48*(C), 319–324. https://doi.org/10.1016/j.procs.2015.04.188

Khan, S., Alourani, A., Mishra, B., Ali, A., & Kamal, M. (2022). Developing a Credit Card Fraud Detection Model using Machine Learning Approaches. *International Journal of Advanced Computer Science and Applications*, *13*(3), 411–418.

https://doi.org/10.14569/IJACSA.2022.0130350

Ko, R. K. L., Jagadpramana, P., Mowbray, M., Pearson, S., & Kirchberg, M. (2011). TrustCloud : A Framework for Accountability and Trust in Cloud Computing TrustCloud : A Framework for Accountability and Trust in Cloud Computing. *Proceedings - 2011 IEEE World Congress on Services, SERVICES 2011*, *July*, 584–588. https://doi.org/10.1109/SERVICES.2011.91

Kotekani, S. S., & Ilango, V. (2022). HEMClust: An Improved Fraud Detection Model for Health Insurance using Heterogeneous Ensemble and K-prototype Clustering. *International Journal of Advanced Computer Science and Applications*, *13*(3), 127–139. https://doi.org/10.14569/IJACSA.2022.0130318

Kotekani, S. S., & Velchamy, I. (2020). An Effective Data Sampling Procedure for Imbalanced Data Learning on Health Insurance Fraud Detection. *Journal of Computing and Information Technology*, *28*(4), 269–285. https://doi.org/10.20532/cit.2020.1005216

Li, T., Kou, G., Peng, Y., & Yu, P. S. (2021). An Integrated Cluster Detection, Optimization, and Interpretation Approach for Financial Data. *IEEE Transactions on Cybernetics*, *52*(12), 13848–13861. https://doi.org/10.1109/TCYB.2021.3109066

Lin, T. H., & Jiang, J. R. (2021). Credit card fraud detection with autoencoder and probabilistic random forest. *Mathematics*, *9*(21). https://doi.org/10.3390/math9212683

Liu, J., Gu, X., & Shang, C. (2020). Quantitative Detection of Financial Fraud Based on Deep Learning with Combination of E-Commerce Big Data. *Complexity*, *2020*. https://doi.org/10.1155/2020/6685888

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. https://doi.org/10.1016/j.future.2019.08.029

Mittal, R., & Bhatia, M. P. S. (2021). Detection of Suspicious or UnTrusted Users in

Crypto-Currency Financial Trading Applications. *International Journal of Digital Crime and Forensics*, *13*(1), 79–93. https://doi.org/10.4018/IJDCF.2021010105

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., Estarli, M., Barrera, E. S. A., Martínez-Rodríguez, R., Baladia, E., Agüero, S. D., Camacho, S., Buhring, K., Herrero-López, A., Gil-González, D. M., Altman, D. G., Booth, A., … Whitlock, E. (2016). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Revista Espanola de Nutricion Humana y Dietetica*, *20*(2), 148–160. https://doi.org/10.1186/2046-4053-4-1

Muaz, A., Jayabalan, M., & Thiruchelvam, V. (2020). A comparison of data sampling techniques for credit card fraud detection. *International Journal of Advanced Computer Science and Applications*, *11*(6), 477–485. https://doi.org/10.14569/IJACSA.2020.0110660

Nasr, M. H., Farrag, M. H., & Nasr, M. M. (2022). A Proposed Fraud Detection Model based on e-Payments Attributes a Case Study in Egyptian e-Payment Gateway. *International Journal of Advanced Computer Science and Applications*, *13*(5), 179–186. https://doi.org/10.14569/IJACSA.2022.0130522

Nerurkar, P., Bhirud, S., Patel, D., Ludinard, R., Busnel, Y., & Kumari, S. (2021). Supervised learning model for identifying illegal activities in Bitcoin. *Applied Intelligence*, *51*(6), 3824–3843. https://doi.org/10.1007/s10489-020-02048-w

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, *50*(3), 559–569. https://doi.org/10.1016/j.dss.2010.08.006

Nguyen, N., Duong, T., Chau, T., Nguyen, V. H., Trinh, T., Tran, D., & Ho, T. (2022). A Proposed Model for Card Fraud Detection Based on CatBoost and Deep Neural Network. *IEEE Access*, *10*, 96852–96861. https://doi.org/10.1109/ACCESS.2022.3205416

Okoli, C. (2015). A guide to conducting a standalone systematic literature review.

Communications of the Association for Information Systems, *37*(1), 879–910. https://doi.org/10.17705/1cais.03743

Okubo, Y. (1997). *Bibliometric Indicators and Analysis of Research Systems METHODS AND EXAMPLES, OECD Science, Technology and Industry Working Papers, 1997/01, OECD Publishing.* 70.

Palacio, S. M. (2019). Abnormal pattern prediction: Detecting fraudulent insurance property claims with semi-supervised machine-learning. *Data Science Journal*, *18*(1). https://doi.org/10.5334/dsj-2019-035

Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. (2018). Credit Card Fraud Detection Using AdaBoost and Majority Voting. *IEEE Access*, *6*, 14277–14284. https://doi.org/10.1109/ACCESS.2018.2806420

Rubaidi, Z. S., Ammar, B. Ben, & Aouicha, M. Ben. (2022). Fraud Detection Using Large-scale Imbalance Dataset. *International Journal on Artificial Intelligence Tools*, *31*(8). https://doi.org/10.1142/s0218213022500373

Sami Owais, S., & Sael Hussein, N. (2016). Extract Five Categories CPIVW from the 9V's Characteristics of the Big Data. *International Journal of Advanced Computer Science and Applications*, *7*(3), 254–258. https://doi.org/10.14569/ijacsa.2016.070337

Santosh, T., & Ramesh, D. (2020). Machine learning approach on apache spark for credit card fraud detection. *Ingenierie Des Systemes d'Information*, *25*(1), 101–106. https://doi.org/10.18280/isi.250113

Sasikala, G., Laavanya, M., Sathyasri, B., Supraja, C., Mahalakshmi, V., Mole, S. S. S., Mulerikkal, J., Chidambaranathan, S., Arvind, C., Srihari, K., & Dejene, M. (2022). An Innovative Sensing Machine Learning Technique to Detect Credit Card Frauds in Wireless Communications. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/2439205

Scimago Journal & Country Rank. (n.d.). https://www.scimagojr.com/aboutus.php

Sengupta, I. N. (1986). Three new parameters in bibliometric research and their

application to rerank periodicals in the field of biochemistry. *Scientometrics*, *10*(5–6), 235–242. https://doi.org/10.1007/BF02016772

Sharma, P., Banerjee, S., Tiwari, D., & Patni, J. C. (2021). Machine learning model for credit card fraud detection-A comparative analysis. *International Arab Journal of Information Technology*, *18*(6), 789–796. https://doi.org/10.34028/iajit/18/6/6

Stojanović, B., & Božić, J. (2022). Robust Financial Fraud Alerting System Based in the Cloud Environment. *Sensors*, *22*(23). https://doi.org/10.3390/s22239461

Sun, C., Li, Q., Li, H., Shi, Y., Zhang, S., & Guo, W. (2019). Patient Cluster Divergence Based Healthcare Insurance Fraudster Detection. *IEEE Access*, *7*, 14162–14170. https://doi.org/10.1109/ACCESS.2018.2886680

Valavan, M., & Rita, S. (2023). Predictive-Analysis-based Machine Learning Model for Fraud Detection with Boosting Classifiers. *Computer Systems Science and Engineering*, *45*(1), 231–245. https://doi.org/10.32604/csse.2023.026508

van Eck, N. J., & Waltman, L. (2018). VOSviewer Manual - version 1.6.8. *Univeristeit Leiden*, *April*, 1–51. http://www.vosviewer.com/documentation/Manual_VOSviewer_1.5.4.pdf

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), xiii–xxiii.

Zafar, S., Kayani, H. ur R., Haq, H. B. ul, Khalid, I., & Nasir, A. (2021). Big Data: Challenges, Popular Tools Of Big Data -Benefits And Applications. *International Journal of Computer Science and Network Security*, *20*(May), 64–75. https://www.researchgate.net/publication/351820412_Big_Data_Challenges_Popular_Tools_Of_Big_Data_-Benefits_And_Applications

Zhang, Y. F., Lu, H. L., Lin, H. F., Qiao, X. C., & Zheng, H. (2022). The Optimized Anomaly Detection Models Based on an Approach of Dealing with Imbalanced Dataset for Credit Card Fraud Detection. *Mobile Information Systems*, *2022*. https://doi.org/10.1155/2022/8027903

ANNEXES

ANNEX 1 - PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	

Section and Topic	Item #	Checklist item	Location where item is reported
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	
Study characteristics	17	Cite each included study and present its characteristics.	
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	
Results of	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	

Section and Topic	Item #	Checklist item	Location where item is reported
syntheses	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	
	23b	Discuss any limitations of the evidence included in the review.	
	23c	Discuss any limitations of the review processes used.	
	23d	Discuss implications of the results for practice, policy, and future research.	
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71 http://www.prisma-statement.org/

ANNEX 2 – PRISMA 2020 for Abstracts Checklist

Section and Topic	Item #	Checklist item
TITLE		
Title	1	Identify the report as a systematic review.
BACKGROUND		
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.
METHODS		
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.
Synthesis of results	6	Specify the methods used to present and synthesise results.
RESULTS		
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).
DISCUSSION		
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).
Interpretation	10	Provide a general interpretation of the results and important implications.
OTHER		
Funding	11	Specify the primary source of funding for the review.
Registration	12	Provide the register name and registration number.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71 http://www.prisma-statement.org/

ANNEX 3 – List of analysed articles – Systematic Literature Review

N.º	Title	Authors	Year	DOI
1	Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework	Hajek P., Abedin M.Z., Sivarajah U.	2022	10.1007/s10796-022-10346-6
2	A Proposed Model for Card Fraud Detection Based on CatBoost and Deep Neural Network	Nguyen N., Duong T., Chau T., Nguyen V.-H., Trinh T., Tran D., Ho T.	2022	10.1109/ACCESS.2022.3205416
3	Feature Engineering and Resampling Strategies for Fund Transfer Fraud with Limited Transaction Data and a Time-Inhomogeneous Modi Operandi	Hsin Y.-Y., Dai T.-S., Ti Y.-W., Huang M.-C., Chiang T.-H., Liu L.-C.	2022	10.1109/ACCESS.2022.3199425
4	The Optimized Anomaly Detection Models Based on an Approach of Dealing with Imbalanced Dataset for Credit Card Fraud Detection	Zhang Y.-F., Lu H.-L., Lin H.-F., Qiao X.-C., Zheng H.	2022	10.1155/2022/8027903
5	Developing a Credit Card Fraud Detection Model using Machine Learning Approaches	Khan S., Alourani A., Mishra B., Ali A., Kamal M.	2022	10.14569/IJACSA.2022.0130350
6	A Neural Network Ensemble with Feature Engineering for Improved Credit Card Fraud Detection	Esenogho E., Mienye I.D., Swart T.G., Aruleba K., Obaido G.	2022	10.1109/ACCESS.2022.3148298
7	Deep Learning Anti-Fraud Model for Internet Loan: Where We Are Going	Fang W., Li X., Zhou P., Yan J., Jiang D., Zhou T.	2021	10.1109/ACCESS.2021.3051079
8	Opinion fraud detection via neural autoencoder decision forest	Dong M., Yao L., Wang X., Benatallah B., Huang C., Ning X.	2020	10.1016/j.patrec.2018.07.013
9	Credit card fraud detection based on machine learning	Fang Y., Zhang Y., Huang C.	2019	10.32604/cmc.2019.06144
10	Credit Card Fraud Detection Using AdaBoost and Majority Voting	Randhawa K., Loo C.K., Seera M., Lim C.P., Nandi A.K.	2018	10.1109/ACCESS.2018.2806420
11	A Credit Card Fraud Model Prediction Method Based on Penalty Factor Optimization AWTadaboost	Ning W., Chen S., Qiang F., Tang H., Jie S.	2023	10.32604/cmc.2023.035558
12	Predictive-Analysis-based Machine Learning Model for Fraud Detection with Boosting Classifiers	Valavan M., Rita S.	2023	10.32604/csse.2023.026508
13	Robust Financial Fraud Alerting System Based in the Cloud Environment	Stojanović B., Božić J.	2022	10.3390/s22239461
14	Fraud Detection Using Large-scale Imbalance Dataset	Rubaidi Z.S., Ammar B.B., Aouicha M.B.	2022	10.1142/S0218213022500373
15	An Integrated Cluster Detection, Optimization, and Interpretation Approach for Financial Data	Li T., Kou G., Peng Y., Yu P.S.	2022	10.1109/TCYB.2021.3109066
16	High-Cardinality Categorical Attributes and Credit Card Fraud Detection	Carneiro E.M., Forster C.H.Q., Mialaret L.F.S., Dias L.A.V., da Cunha A.M.	2022	10.3390/math10203808
17	A Machine Learning and Blockchain Based Efficient Fraud Detection Mechanism	Ashfaq T., Khalid R., Yahaya A.S., Aslam S., Azar A.T., Alsafari S., Hameed I.A.	2022	10.3390/s22197162

N.º	Title	Authors	Year	DOI
18	Digital payment fraud detection methods in digital ages and Industry 4.0	Chang V., Doan L.M.T., Di Stefano A., Sun Z., Fortino G.	2022	10.1016/j.compeleceng.2022.107734
19	Tax evasion risk management using a Hybrid Unsupervised Outlier Detection method	Savić M., Atanasijević J., Jakovetić D., Krejić N.	2022	10.1016/j.eswa.2021.116409
20	Enhanced Credit Card Fraud Detection Model Using Machine Learning	Alfaiz N.S., Fati S.M.	2022	10.3390/electronics11040662
21	An Effective Ensemble-based Framework for Outlier Detection in Evolving Data Streams	Hassan A.F., Barakat S., Rezk A.	2022	10.14569/IJACSA.2022.0131135
22	An Innovative Sensing Machine Learning Technique to Detect Credit Card Frauds in Wireless Communications	Sasikala G., Laavanya M., Sathyasri B., Supraja C., Mahalakshmi V., Mole S.S.S., Mulerikkal J., Chidambaranathan S., Arvind C., Srihari K., Dejene M.	2022	10.1155/2022/2439205
23	Mixed Quantum-Classical Method for Fraud Detection With Quantum Feature Selection	Grossi M., Ibrahim N., Radescu V., Loredo R., Voigt K., Von Altrock C., Rudnik A.	2022	10.1109/TQE.2022.3213474
24	A Proposed Fraud Detection Model based on e-Payments Attributes a Case Study in Egyptian e-Payment Gateway	Nasr M.H., Farrag M.H., Nasr M.M.	2022	10.14569/IJACSA.2022.0130522
25	E-Commerce Fraud Detection Model by Computer Artificial Intelligence Data Mining	Li J.	2022	10.1155/2022/8783783
26	HEMClust: An Improved Fraud Detection Model for Health Insurance using Heterogeneous Ensemble and K-prototype Clustering	Kotekani S.S., Ilango V.	2022	10.14569/IJACSA.2022.0130318
27	Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms	Alarfaj F.K., Malik I., Khan H.U., Almusallam N., Ramzan M., Ahmed M.	2022	10.1109/ACCESS.2022.3166891
28	An In-Depth Study and Improvement of Isolation Forest	Chabchoub Y., Togbe M.U., Boly A., Chiky R.	2022	10.1109/ACCESS.2022.3144425
29	Improving Tax Audit Efficiency Using Machine Learning: The Role of Taxpayer's Network Data in Fraud Detection	Baghdasaryan V., Davtyan H., Sarikyan A., Navasardyan Z.	2022	10.1080/08839514.2021.2012002
30	Machine learning model for credit card fraud detection-A comparative analysis	Sharma P., Banerjee S., Tiwari D., Patni J.C.	2021	10.34028/iajit/18/6/6
31	Credit card fraud detection with autoencoder and probabilistic random forest	Lin T.-H., Jiang J.-R.	2021	10.3390/math9212683
32	Supervised learning model for identifying illegal activities in Bitcoin	Nerurkar P., Bhirud S., Patel D., Ludinard R., Busnel Y., Kumari S.	2021	10.1007/s10489-020-02048-w
33	Follow the trail: Machine learning for fraud detection in fintech applications	Stojanović B., Božić J., Hofer-Schmitz K., Nahrgang K., Weber A., Badii A., Sundaram M., Jordan E., Runevic J.	2021	10.3390/s21051594

N.º	Title	Authors	Year	DOI
34	Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost	Ileberi E., Sun Y., Wang Z.	2021	10.1109/ACCESS.2021.3134330
35	A Fast and Efficient Algorithm for Outlier Detection Over Data Streams	Hassaan M., Maher H., Gouda K.	2021	10.14569/IJACSA.2021.0121185
36	Financial Fraud Detection in Healthcare Using Machine Learning and Deep Learning Techniques	Mehbodniya A., Alam I., Pande S., Neware R., Rane K.P., Shabaz M., Madhavan M.V.	2021	10.1155/2021/9293877
37	OFCOD: On the fly clustering based outlier detection framework	Elmogy A., Rizk H., Sarhan A.M.	2021	10.3390/data6010001
38	Detection of Suspicious or UnTrusted Users in Crypto-Currency Financial Trading Applications	Mittal R., Bhatia M.P.S.	2021	10.4018/IJDCF.2021010105
39	Statistical hierarchical clustering algorithm for outlier detection in evolving data streams	Krleža D., Vrdoljak B., Brčić M.	2021	10.1007/s10994-020-05905-4
40	Deshelling the Shell Companies Using Benford's Law: An Emerging Market Study	Aggarwal V., Dharni K.	2020	10.1177/0256090920979695
41	A comparative evaluation of novelty detection algorithms for discrete sequences	Domingues R., Michiardi P., Barlet J., Filippone M.	2020	10.1007/s10462-019-09779-4
42	Machine learning approach on apache spark for credit card fraud detection	Santosh T., Ramesh D.	2020	10.18280/isi.250113
43	An Effective Data Sampling Procedure for Imbalanced Data Learning on Health Insurance Fraud Detection	Kotekani S.S., Velchamy I.	2020	10.20532/cit.2020.1005216
44	Fraud Detection in Credit Cards using Logistic Regression	Alenzi H.Z., Aljehane N.O.	2020	10.14569/IJACSA.2020.0111265
45	Quantitative Detection of Financial Fraud Based on Deep Learning with Combination of E-Commerce Big Data	Liu J., Gu X., Shang C.	2020	10.1155/2020/6685888
46	A comparison of data sampling techniques for credit card fraud detection	Muaz A., Jayabalan M., Thiruchelvam V.	2020	10.14569/IJACSA.2020.0110660
47	A Secure AI-Driven Architecture for Automated Insurance Systems: Fraud Detection and Risk Measurement	Dhieb N., Ghazzai H., Besbes H., Massoud Y.	2020	10.1109/ACCESS.2020.2983300
48	Towards automated feature engineering for credit card fraud detection using multi-perspective HMMs	Lucas Y., Portier P.-E., Laporte L., He-Guelton L., Caelen O., Granitzer M., Calabretto S.	2020	10.1016/j.future.2019.08.029
49	Evaluating the benefits of using proactive transformed-domain-based techniques in fraud detection tasks	Saia R., Carta S.	2019	10.1016/j.future.2018.10.016
50	Fraud detection using machine learning in e-commerce	Saputra A., Suharjito	2019	10.14569/ijacsa.2019.0100943
51	Abnormal pattern prediction: Detecting fraudulent insurance property claims with semi-supervised machine-learning	Palacio S.M.	2019	10.5334/dsj-2019-035

N.º	Title	Authors	Year	DOI
52	Patient Cluster Divergence Based Healthcare Insurance Fraudster Detection	Sun C., Li Q., Li H., Shi Y., Zhang S., Guo W.	2019	10.1109/ACCESS.2018.2886680
53	Abnormal Group-Based Joint Medical Fraud Detection	Sun C., Yan Z., Li Q., Zheng Y., Lu X., Cui L.	2019	10.1109/ACCESS.2018.2887119
54	Universal outlier hypothesis testing	Li Y., Nitinawarat S., Veeravalli V.V.	2014	10.1109/TIT.2014.2317691
55	Fraud formalization and detection	Bhargava B., Zhong Y., Lu Y.	2003	10.1007/978-3-540-45228-7_33

ANNEX 4 – Results categorization: Advanced Analytical Methods/Types of Fraud

Advanced Analytical Methods/Types of Fraud	Payments	"General" fraud	Cryptocurrency market	Loan	Insurance	Non-financial	Fraudsters	TOTAL
Logistic Regression	5	2			2			9
Clustering		3			1			4
Benford's law		1						1
Naive Bayes classifier	1	2			1			4
Decision trees - Random forest	10	4	3	2	2	1		22
Outlier Detection	2	6			2	1		11
Hidden Markov Model (HMM)	1	1						2
Anomaly Detection	1						1	2
Isolation Forest	3	2						5
Neural Network	7	3		1	2			13
Linear Regression				1				1
Gradient Boosting method	5	2		1				8
Support Vector Machine	4	2						6
Extra tree	1							1
Patient Cluster Divergence					1			1
Elliptic Envelope	2	1						3
Adaptive Boosting	4	1						5
K-means	1				2			3
XGBoost/CatBoost	3	3	1	1	1			9
Pattern Mining								0
TOTAL	50	33	4	6	14	2	1	110

ANNEX 5 – List of analysed articles – Bibliometric Approach

N.º	Title	Authors	Year	DOI
1	The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature	Ngai E.W.T.; Hu Y.; Wong Y.H.; Chen Y.; Sun X.	2011	10.1016/j.dss.2010.08.006
2	The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients	Yeh I.-C.; Lien C.-h.	2009	10.1016/j.eswa.2007.12.020
3	Data Mining techniques for the detection of fraudulent financial statements	Kirkos E.; Spathis C.; Manolopoulos Y.	2007	10.1016/j.eswa.2006.02.016
4	A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0	Ivanov D.; Dolgui A.	2021	10.1080/09537287.2020.1768450
5	On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms	Yamanishi K.; Takeuchi J.-I.; Williams G.; Milne P.	2004	10.1023/B:DAMI.0000023676.72185.7c
6	Detection of financial statement fraud and feature selection using data mining techniques	Ravisankar P.; Ravi V.; Raghava Rao G.; Bose I.	2011	10.1016/j.dss.2010.11.006
7	Using generative adversarial networks for improving classification effectiveness in credit card fraud detection	Fiore U.; De Santis A.; Perla F.; Zanetti P.; Palmieri F.	2019	10.1016/j.ins.2017.12.030
8	Feature engineering strategies for credit card fraud detection	Correa Bahnsen A.; Aouada D.; Stojanovic A.; Ottersten B.	2016	10.1016/j.eswa.2015.12.030
9	Effective detection of sophisticated online banking fraud on extremely imbalanced data	Wei W.; Li J.; Cao L.; Ou Y.; Chen J.	2013	10.1007/s11280-012-0178-0
10	A comparison of models for predicting early hospital readmissions	Futoma J.; Morris J.; Lucas J.	2015	10.1016/j.jbi.2015.05.016
11	Survey of data management and analysis in disaster situations	Hristidis V.; Chen S.-C.; Li T.; Luis S.; Deng Y.	2010	10.1016/j.jss.2010.04.065
12	Recent Development in Big Data Analytics for Business Operations and Risk Management	Choi T.-M.; Chan H.K.; Yue X.	2017	10.1109/TCYB.2015.2507599
13	A data mining based system for credit-card fraud detection in e-tail	Carneiro N.; Figueira G.; Costa M.	2017	10.1016/j.dss.2017.01.002
14	Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events	Gaspar R.; Pedro C.; Panagiotopoulos P.; Seibt B.	2016	10.1016/j.chb.2015.11.040
15	Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods	Hajek P.; Henriques R.	2017	10.1016/j.knosys.2017.05.001
16	Software project risk analysis using Bayesian networks with causality constraints	Hu Y.; Zhang X.; Ngai E.W.T.; Cai R.; Liu M.	2013	10.1016/j.dss.2012.11.001
17	SCARFF: A scalable framework for streaming credit card fraud detection with spark	Carcillo F.; Dal Pozzolo A.; Le Borgne Y.-A.; Caelen O.; Mazzer Y.; Bontempi G.	2018	10.1016/j.inffus.2017.09.005
18	Fraud detection: A systematic literature review of graph-based anomaly detection approaches	Pourhabibi T.; Ong K.-L.; Kam B.H.; Boo Y.L.	2020	10.1016/j.dss.2020.113303

N.º	Title	Authors	Year	DOI
19	An intraday market risk management approach based on textual analysis	Groth S.S.; Muntermann J.	2011	10.1016/j.dss.2010.08.019
20	Automatic identification of eyewitness messages on twitter during disasters	Zahra K.; Imran M.; Ostermann F.O.	2020	10.1016/j.ipm.2019.102107
21	Enabling Cloud Computing in Emergency Management Systems	Qiu M.; Ming Z.; Wang J.; Yang L.T.; Xiang Y.	2014	10.1109/MCC.2014.71
22	Micro Data analytics: a test for analytical procedures	Santosuosso P.	2022	10.1108/MEDAR-02-2020-0767
23	Detecting evolutionary financial statement fraud	Zhou W.; Kapoor G.	2011	10.1016/j.dss.2010.08.007
24	An overview of social network analysis	Oliveira M.; Gama J.	2012	10.1002/widm.1048
25	Machine Learning Algorithms for Construction Projects Delay Risk Prediction	Gondia A.; Siam A.; El-Dakhakhni W.; Nassar A.H.	2020	10.1061/(ASCE)CO.1943-7862.0001736
26	Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings	Zhong H.; Miao C.; Shen Z.; Feng Y.	2014	10.1016/j.neucom.2013.02.054
27	Relational Deep Learning Detection with Multi-Sequence Representation for Insider Threats	Alshehri A.	2022	10.14569/IJACSA.2022.0130587
28	Extreme learning machines for credit scoring: An empirical evaluation	Bequé A.; Lessmann S.	2017	10.1016/j.eswa.2017.05.050
29	Incorporating domain knowledge into data mining classifiers: An application in indirect lending	Sinha A.P.; Zhao H.	2008	10.1016/j.dss.2008.06.013
30	DGHNL: A new deep genetic hierarchical network of learners for prediction of credit scoring	Pławiak P.; Abdar M.; Pławiak J.; Makarenkov V.; Acharya U.R.	2020	10.1016/j.ins.2019.12.045
31	Remote Music Teaching Classroom Based on Machine Learning and 5G Network Station	Seng W.	2022	10.1155/2022/7569763
32	The digital transformation of external audit and its impact on corporate governance	Manita R.; Elommal N.; Baudier P.; Hikkerova L.	2020	10.1016/j.techfore.2019.119751
33	Participatory sensing-based semantic and spatial analysis of urban emergency events using mobile social media	Xu Z.; Zhang H.; Sugumaran V.; Choo K.-K.R.; Mei L.; Zhu Y.	2016	10.1186/s13638-016-0553-0
34	A data mining-based framework for supply chain risk management	Er Kara M.; Oktay Fırat S.Ü.; Ghadge A.	2020	10.1016/j.cie.2018.12.017
35	Decentralized Big Data Auditing for Smart City Environments Leveraging Blockchain Technology	Yu H.; Yang Z.; Sinnott R.O.	2019	10.1109/ACCESS.2018.2888940
36	A survey of intrusion detection and prevention systems	Patel A.; Qassim Q.; Wills C.	2010	10.1108/09685221011079199
37	Comprehensive review of deep reinforcement learning methods and applications in economics	Mosavi A.; Faghan Y.; Ghamisi P.; Duan P.; Ardabili S.F.; Salwana E.; Band S.S.	2020	10.3390/MATH8101640
38	An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance	Kose I.; Gokturk M.; Kilic K.	2015	10.1016/j.asoc.2015.07.018
39	An information granulation based data mining approach for classifying	Chen M.-C.; Chen L.-S.; Hsu C.-C.; Zeng W.-R.	2008	10.1016/j.ins.2008.03.018

N.º	Title	Authors	Year	DOI
	imbalanced data			
40	SMARTDIAB: A communication and information technology approach for the intelligent monitoring, management and follow-up of type 1 diabetes patients	Mougiakakou S.G.; Bartsocas C.S.; Bozas E.; Chaniotakis N.; Iliopoulou D.; Kouris I.; Pavlopoulos S.; Prountzou A.; Skevofilakas M.; Tsoukalis A.; Varotsis K.; Vazeou A.; Zarkogianni K.; Nikita K.S.	2010	10.1109/TITB.2009.2039711
41	Application of Security Algorithm in Audit Data Asset Valuation Based on Distributed Machine Learning	Liu C.	2023	10.1007/s11277-023-10497-y
42	Characterization and detection of taxpayers with false invoices using data mining techniques	Castellón González P.; Velásquez J.D.	2013	10.1016/j.eswa.2012.08.051
43	Dynamic Audit of Internet Finance Based on Machine Learning Algorithm	Zhang J.	2022	10.1155/2022/7072955
44	Enhanced cyber-physical security in internet of things through energy auditing	Li F.; Shi Y.; Shinde A.; Ye J.; Song W.	2019	10.1109/JIOT.2019.2899492
45	DLSeF: A dynamic key-length-based efficient real-time security verification model for big data stream	Puthal D.; Nepal S.; Ranjan R.; Chen J.	2016	10.1145/2937755
46	An efficient algorithm for distributed density-based outlier detection on big data	Bai M.; Wang X.; Xin J.; Wang G.	2016	10.1016/j.neucom.2015.05.135
47	Social and geographical disparities in Twitter use during Hurricane Harvey	Zou L.; Lam N.S.N.; Shams S.; Cai H.; Meyer M.A.; Yang S.; Lee K.; Park S.-J.; Reams M.A.	2019	10.1080/17538947.2018.1545878
48	Generative adversarial network based telecom fraud detection at the receiving bank	Zheng Y.-J.; Zhou X.-H.; Sheng W.-G.; Xue Y.; Chen S.-Y.	2018	10.1016/j.neunet.2018.02.015
49	Cybersecurity in industrial control systems: Issues, technologies, and challenges	Asghar M.R.; Hu Q.; Zeadally S.	2019	10.1016/j.comnet.2019.106946
50	50 years of data mining and OR: Upcoming trends and challenges	Baesens B.; Mues C.; Martens D.; Vanthienen J.	2009	10.1057/jors.2008.171
51	Homesound: Real-time audio event detection based on high performance computing for behaviour and surveillance remote monitoring	Alsina-Pagès R.M.; Navarro J.; Alías F.; Hervás M.	2017	10.3390/s17040854
52	Impacts of digitization on auditing: A Delphi study for Germany	Tiberius V.; Hirth S.	2019	10.1016/j.intaccaudtax.2019.100288
53	The Intertwine of Brain and Body: A Quantitative Analysis on How Big Data Influences the System of Sports	Patel D.; Shah D.; Shah M.	2020	10.1007/s40745-019-00239-y
54	Ensemble of deep sequential models for credit card fraud detection	Forough J.; Momtazi S.	2021	10.1016/j.asoc.2020.106883
55	End-to-end neural network architecture for fraud scoring in card payments	Gómez J.A.; Arévalo J.; Paredes R.; Nin J.	2018	10.1016/j.patrec.2017.08.024

N.º	Title	Authors	Year	DOI
56	Towards secure FinTech: A survey, taxonomy, and open research challenges	Mehrban S.; Khan M.A.; Nadeem M.W.; Hussain M.; Ahmed M.M.; Hakeem O.; Saqib S.; Kiah M.L.M.; Abbas F.; Hassan M.	2020	10.1109/ACCESS.2020.2970430
57	VFChain: Enabling Verifiable and Auditable Federated Learning via Blockchain Systems	Peng Z.; Xu J.; Chu X.; Gao S.; Yao Y.; Gu R.; Tang Y.	2022	10.1109/TNSE.2021.3050781
58	A big data mining approach of PSO-Based BP neural network for financial risk management with IoT	Zhou H.; Sun G.; Fu S.; Liu J.; Zhou X.; Zhou J.	2019	10.1109/ACCESS.2019.2948949
59	Biometric-Based Key Generation and User Authentication Using Acoustic Characteristics of the Outer Ear and a Network of Correlation Neurons	Sulavko A.	2022	10.3390/s22239551
60	E2mC: Improving emergency management service practice through social media and crowdsourcing analysis in near real time	Havas C.; Resch B.; Francalanci C.; Pernici B.; Scalia G.; Fernandez-Marquez J.L.; Van Achte T.; Zeug G.; Mondardini M.R.R.; Grandoni D.; Kirsch B.; Kalas M.; Lorini V.; Rüping S.	2017	10.3390/s17122766
61	An Arabic social media based framework for incidents and events monitoring in smart cities	Alkhatib M.; El Barachi M.; Shaalan K.	2019	10.1016/j.jclepro.2019.02.063
62	A Secure AI-Driven Architecture for Automated Insurance Systems: Fraud Detection and Risk Measurement	Dhieb N.; Ghazzai H.; Besbes H.; Massoud Y.	2020	10.1109/ACCESS.2020.2983300
63	A data mining framework for detecting subscription fraud in telecommunication	Farvaresh H.; Sepehri M.M.	2011	10.1016/j.engappai.2010.05.009
64	RETRACTED ARTICLE: Intelligent city emergency intelligence perception model based on social media big data	Xiong G.	2022	10.1007/s12652-021-03065-4
65	A prescription fraud detection model	Aral K.D.; Güvenir H.A.; Sabuncuoğlu T.; Akar A.R.	2012	10.1016/j.cmpb.2011.09.003
66	Data mining for financial prediction and trading: Application to single and multiple markets	Chun S.-H.; Kim S.H.	2004	10.1016/S0957-4174(03)00113-1
67	Identifying convergence fields and technologies for industrial safety: LDA-based network analysis	Song B.; Suh Y.	2019	10.1016/j.techfore.2018.08.013
68	Stochastic approximation vis-à-vis online learning for big data analytics	Slavakis K.; Kim S.-J.; Mateos G.; Giannakis G.B.	2014	10.1109/MSP.2014.2345536
69	Anomaly detection in wide area network meshes using two machine learning algorithms	Zhang J.; Gardner R.; Vukotic I.	2019	10.1016/j.future.2018.07.023
70	How Do the Global Stock Markets Influence One Another? Evidence from Finance Big Data and Granger Causality Directed Network	Tang Y.; Xiong J.J.; Luo Y.; Zhang Y.-C.	2019	10.1080/10864415.2018.1512283
71	Application of computational intelligence technologies in emergency management:	Chen N.; Liu W.; Bai R.; Chen A.	2019	10.1007/s10462-017-9589-8

N.º	Title	Authors	Year	DOI
	a literature review			
72	Boosting a weather monitoring system in low income economies using open and non-conventional systems: Data quality analysis	Strigaro D.; Cannata M.; Antonovic M.	2019	10.3390/s19051185
73	Learning relational policies from electronic health record access logs	Malin B.; Nyemba S.; Paulett J.	2011	10.1016/j.jbi.2011.01.007
74	High temporal resolution rainfall–runoff modeling using long-short-term-memory (LSTM) networks	Li W.; Kiaghadi A.; Dawson C.	2021	10.1007/s00521-020-05010-6
75	Risk prediction using natural language processing of electronic mental health records in an inpatient forensic psychiatry setting	Le D.V.; Montgomery J.; Kirkby K.C.; Scanlan J.	2018	10.1016/j.jbi.2018.08.007
76	A Neural Network Ensemble with Feature Engineering for Improved Credit Card Fraud Detection	Esenogho E.; Mienye I.D.; Swart T.G.; Aruleba K.; Obaido G.	2022	10.1109/ACCESS.2022.3148298
77	VOST: A case study in voluntary digital participation for collaborative emergency management	Fathi R.; Thom D.; Koch S.; Ertl T.; Fiedrich F.	2020	10.1016/j.ipm.2019.102174
78	An automatic system to identify heart disease risk factors in clinical texts over time	Chen Q.; Li H.; Tang B.; Wang X.; Liu X.; Liu Z.; Liu S.; Wang W.; Deng Q.; Zhu S.; Chen Y.; Wang J.	2015	10.1016/j.jbi.2015.09.002
79	Embracing textual data analytics in auditing with deep learning	Sun T.; Vasarhelyi M.A.	2018	10.4192/1577-8517-v18_3
80	Social media data analytics for business decision making system to competitive analysis	Yang J.; Xiu P.; Sun L.; Ying L.; Muthu B.	2022	10.1016/j.ipm.2021.102751
81	Spatio-temporal prediction of crop disease severity for agricultural emergency management based on recurrent neural networks	Xu W.; Wang Q.; Chen R.	2018	10.1007/s10707-017-0314-1
82	A cyberGIS-enabled multi-criteria spatial decision support system: A case study on flood emergency management	Zhang Z.; Hu H.; Yin D.; Kashem S.; Li R.; Cai H.; Perkins D.; Wang S.	2019	10.1080/17538947.2018.1543363
83	Towards domain invariant heart sound abnormality detection using learnable filterbanks	Humayun A.I.; Ghaffarzadegan S.; Ansari M.I.; Feng Z.; Hasan T.	2020	10.1109/JBHI.2020.2970252
84	Sound Event Recognition Using Auditory-Receptive-Field Binary Pattern and Hierarchical-Diving Deep Belief Network	Wang C.-Y.; Wang J.-C.; Santoso A.; Chiang C.-C.; Wu C.-H.	2018	10.1109/TASLP.2017.2738443
85	Enterprise Intelligent Audit Model by Using Deep Learning Approach	Ding R.	2022	10.1007/s10614-021-10192-9
86	Credit Card Fraud Detection through Parenclitic Network Analysis	Zanin M.; Romance M.; Moral S.; Criado R.	2018	10.1155/2018/5764370
87	Predicting corporate acquisitions: An application of uncertain reasoning using rule induction	Ragothaman S.; Naik B.; Ramakrishnan K.	2003	10.1023/B:ISFI.0000005653.53641.b3

N.º	Title	Authors	Year	DOI
88	Information fusion for future COVID-19 prevention: continuous mechanism of big data intelligent innovation for the emergency management of a public epidemic outbreak	Yin S.; Zhang N.; Xu J.	2021	10.1080/23270012.2021.1945499
89	Objective Detection of Eloquent Axonal Pathways to Minimize Postoperative Deficits in Pediatric Epilepsy Surgery Using Diffusion Tractography and Convolutional Neural Networks	Xu H.; Dong M.; Lee M.-H.; O'Hara N.; Asano E.; Jeong J.-W.	2019	10.1109/TMI.2019.2902073
90	A tamper-proof audit and control system for the doctor in the loop	Kieseberg P.; Malle B.; Frühwirt P.; Weippl E.; Holzinger A.	2016	10.1007/s40708-016-0046-2
91	Deep feature learning for disease risk assessment based on convolutional neural network with intra-layer recurrent connection by using hospital big data	Usama M.; Ahmad B.; Wan J.; Hossain M.S.; Alhamid M.F.; Hossain M.A.	2018	10.1109/ACCESS.2018.2879158
92	Don't mention it? Analyzing user-generated content signals for early adverse event warnings	Abbasi A.; Li J.; Adjeroh D.; Abate M.; Zheng W.	2019	10.1287/isre.2019.0847
93	Prediction and Analysis of Financial Default Loan Behavior Based on Machine Learning Model	Chen H.	2022	10.1155/2022/7907210
94	Scientometric review of articles published in ASCE's journal of construction engineering and management from 2000 to 2018	Jin R.; Zuo J.; Hong J.	2019	10.1061/(ASCE)CO.1943-7862.0001682
95	On oversampling imbalanced data with deep conditional generative models	Fajardo V.A.; Findlay D.; Jaiswal C.; Yin X.; Houmanfar R.; Xie H.; Liang J.; She X.; Emerson D.B.	2021	10.1016/j.eswa.2020.114463
96	Data-intensive analytics for predictive modeling	Apte C.V.; Hong S.J.; Natarajan R.; Pednault E.P.D.; Tipu F.A.; Weiss S.M.	2003	10.1147/rd.471.0017
97	Backup Battery Analysis and Allocation against Power Outage for Cellular Base Stations	Wang F.; Fan X.; Wang F.; Liu J.	2019	10.1109/TMC.2018.2842733
98	Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms	Alarfaj F.K.; Malik I.; Khan H.U.; Almusallam N.; Ramzan M.; Ahmed M.	2022	10.1109/ACCESS.2022.3166891
99	Corruption risk in contracting markets: a network science perspective	Wachs J.; Fazekas M.; Kertész J.	2021	10.1007/s41060-019-00204-1
100	Fraud detection within bankcard enrollment on mobile device based payment using machine learning	Zhou H.; Chai H.-F.; Qiu M.-L.	2018	10.1631/FITEE.1800580
101	Complexities in financial network topological dynamics: Modeling of emerging and developed stock markets	Tang Y.; Xiong J.J.; Jia Z.-Y.; Zhang Y.-C.	2018	10.1155/2018/4680140
102	Intrusion detection system: A review	Sharma S.; Gupta R.K.	2015	10.14257/ijsia.2015.9.5.07
103	Predictive Risk Management for Dynamic Tree Trimming Scheduling for Distribution Networks	Dokic T.; Kezunovic M.	2018	10.1109/TSG.2018.2868457

N.º	Title	Authors	Year	DOI
104	Data-driven financial and operational risk management: Empirical evidence from the global tramp shipping industry	Bai X.; Cheng L.; Iris Ç.	2022	10.1016/j.tre.2022.102617
105	Exploring Trends and Patterns of Popularity Stage Evolution in Social Media	Kong Q.; Mao W.; Chen G.; Zeng D.	2020	10.1109/TSMC.2018.2855806
106	Future of artificial intelligence and its influence on supply chain risk management – A systematic review	Deiva Ganesh A.; Kalpana P.	2022	10.1016/j.cie.2022.108206
107	Cryptocurrencies and artificial intelligence: Challenges and opportunities	Sabry F.; Labda W.; Erbad A.; Malluhi Q.	2020	10.1109/ACCESS.2020.3025211
108	A decision support system for fraud detection in public procurement	Velasco R.B.; Carpanese I.; Interian R.; Paulo Neto O.C.G.; Ribeiro C.C.	2021	10.1111/itor.12811
109	An assertive reasoning method for emergency response management based on knowledge elements C4.5 decision tree	Han L.; Li W.; Su Z.	2019	10.1016/j.eswa.2018.12.042
110	Edge computing enabled non-technical loss fraud detection for big data security analytic in Smart Grid	Han W.; Xiao Y.	2020	10.1007/s12652-019-01381-4
111	From associations to sarcasm: Mining the shift of opinions regarding the Supreme Court on twitter	Sandhu M.; Vinson C.D.; Mago V.K.; Giabbanelli P.J.	2019	10.1016/j.osnem.2019.100054
112	Automated Data Slicing for Model Validation: A Big Data - AI Integration Approach	Chung Y.; Kraska T.; Polyzotis N.; Tae K.H.; Whang S.E.	2020	10.1109/TKDE.2019.2916074
113	Assessing relevance of tweets for risk communication	Liu X.; Kar B.; Zhang C.; Cochran D.M.	2019	10.1080/17538947.2018.1480670
114	Graph Neural Network for Fraud Detection via Spatial-Temporal Attention	Cheng D.; Wang X.; Zhang Y.; Zhang L.	2022	10.1109/TKDE.2020.3025588
115	A Public Auditing Protocol for Cloud Storage System with Intrusion-Resilience	Ding R.; Xu Y.; Cui J.; Zhong H.	2020	10.1109/JSYST.2019.2923238
116	Enterprise financial risk management platform based on 5 G mobile communication and embedded system	Qiu W.	2021	10.1016/j.micpro.2020.103594
117	Automatic Detection of Target Engagement in Transcutaneous Cervical Vagal Nerve Stimulation for Traumatic Stress Triggers	Gurel N.Z.; Wittbrodt M.T.; Jung H.; Ladd S.L.; Shah A.J.; Vaccarino V.; Bremner J.D.; Inan O.T.	2020	10.1109/JBHI.2020.2981116
118	Efficient and secure auditing scheme for outsourced big data with dynamicity in cloud	Gan Q.; Wang X.; Fang X.	2018	10.1007/s11432-017-9410-9
119	E-Commerce Enterprise Supply Chain Financing Risk Assessment Based on Linked Data Mining and Edge Computing	Qu Q.; Liu C.; Bao X.	2021	10.1155/2021/9938325
120	Behavioral data-driven analysis with Bayesian method for risk management of financial services	Lin E.M.H.; Sun E.W.; Yu M.-T.	2020	10.1016/j.ijpe.2020.107737

N.º	Title	Authors	Year	DOI
121	Utilizing a multilayer perceptron artificial neural network to assess a virtual reality surgical procedure	Alkadri S.; Ledwos N.; Mirchi N.; Reich A.; Yilmaz R.; Driscoll M.; Del Maestro R.F.	2021	10.1016/j.compbiomed.2021.104770
122	Corporate network centrality score: Methodologies and informativeness	Debreceny R.S.; Rahman A.; Wang T.	2017	10.2308/isys-51797
123	Technical framework design of safety production information management platform for chemical industrial parks based on cloud computing and the internet of things	Lele Q.; Lihua K.	2016	10.14257/ijgdc.2016.9.6.28
124	Distributed morality, privacy, and social media in natural disaster response	Hayes P.; Kelly S.	2018	10.1016/j.techsoc.2018.05.003
125	Building a normative decision support system for clinical and operational risk management in hemodialysis	Cornalba C.; Bellazzi R.G.; Bellazzi R.	2008	10.1109/TITB.2008.920781
126	Explainability in supply chain operational risk management: A systematic literature review	Nimmy S.F.; Hussain O.K.; Chakrabortty R.K.; Hussain F.K.; Saberi M.	2022	10.1016/j.knosys.2021.107587
127	A new emergency management dynamic value assessment model based on social media data: a multiphase decision-making perspective	Shan S.; Liu X.; Wei Y.; Xu L.; Zhang B.; Yu L.	2020	10.1080/17517575.2020.1722251
128	Sensorineural hearing loss identification via nine-layer convolutional neural network with batch normalization and dropout	Wang S.-H.; Hong J.; Yang M.	2020	10.1007/s11042-018-6798-3
129	Risk Assessment of Private Information Inference for Motion Sensor Embedded IoT Devices	Huang Y.; Guan X.; Chen H.; Liang Y.; Yuan S.; Ohtsuki T.	2020	10.1109/TETCI.2019.2902866
130	Deep Learning Anti-Fraud Model for Internet Loan: Where We Are Going	Fang W.; Li X.; Zhou P.; Yan J.; Jiang D.; Zhou T.	2021	10.1109/ACCESS.2021.3051079
131	A distributed approach of big data mining for financial fraud detection in a supply chain	Zhou H.; Sun G.; Fu S.; Fan X.; Jiang W.; Hu S.; Li L.	2020	10.32604/CMC.2020.09834
132	Introduction to graph databases	Larriba-pey J.L.; Martínez-bazán N.; Domínguez-sal D.	2014	10.1007/978-3-319-10587-1_4
133	Rough sets in economy and finance	Podsiadło M.; Rybiński H.	2014	10.1007/978-3-642-54756-0_6
134	"The innovation governance dilemma: Alternatives to the precautionary principle"	Hemphill T.A.	2020	10.1016/j.techsoc.2020.101381
135	Regional response to large-scale emergency events: Building on historical data	Romanowski C.; Raj R.; Schneider J.; Mishra S.; Shivshankar V.; Ayengar S.; Cueva F.	2015	10.1016/j.ijcip.2015.07.003
136	A for effort? Using the crowd to identify moral hazard in New York City restaurant hygiene inspections	Mejia J.; Mankad S.; Gopal A.	2019	10.1287/isre.2019.0866
137	Leveraging cross-media analytics to detect events and mine opinions for emergency management	Xu W.; Liu L.; Shang W.	2017	10.1108/OIR-08-2015-0286
138	Research on big data audit based on financial sharing service model using	Jiang S.	2021	10.3233/JIFS-189646

N.º	Title	Authors	Year	DOI
	fuzzy AHP			
139	Classification of auditory brainstem responses through symbolic pattern discovery	Molina M.E.; Perez A.; Valente J.P.	2016	10.1016/j.artmed.2016.05.001
140	CATCHM: A novel network-based credit card fraud detection method using node representation learning	Van Belle R.; Baesens B.; De Weerdt J.	2023	10.1016/j.dss.2022.113866
141	Discovering rare categories from graph streams	Zhou D.; Karthikeyan A.; Wang K.; Cao N.; He J.	2017	10.1007/s10618-016-0478-6
142	Survey of data-mining techniques used in fraud detection and prevention	Thiruvadi S.; Patel S.C.	2011	10.3923/itj.2011.710.716
143	Analysis of financial statement and prewarning of audit risks based on artificial neural network	Zhang Y.	2021	10.46300/9106.2021.15.109
144	An Application of Data Envelopment Analysis and Machine Learning Approach to Risk Management	Jomthanachai S.; Wong W.-P.; Lim C.-P.	2021	10.1109/ACCESS.2021.3087623
145	Aberrant Brain Connectivity in Schizophrenia Detected via a Fast Gaussian Graphical Model	Zhang A.; Fang J.; Liang F.; Calhoun V.D.; Wang Y.-P.	2019	10.1109/JBHI.2018.2854659
146	Systematic identification and analysis of different fraud detection approaches based on the strategy ahead	Zandian Z.K.; Keyvanpour M.	2017	10.3233/KES-170357
147	An information system for assessing the likelihood of child labor in supplier locations leveraging Bayesian networks and text mining	Thöni A.; Taudes A.; Tjoa A.M.	2018	10.1007/s10257-018-0368-0
148	BigPromises: using organisational mindfulness to integrate big data in emergency management decision making	Amaye A.; Neville K.; Pope A.	2016	10.1080/12460125.2016.1187419
149	A multi-faceted and automatic knowledge elicitation system (MAKES) for managing unstructured information	Cheung C.F.; Lee W.B.; Wang W.M.; Wang Y.; Yeung W.M.	2011	10.1016/j.eswa.2010.10.033
150	BTG: A Bridge to Graph machine learning in telecommunications fraud detection	Hu X.; Chen H.; Liu S.; Jiang H.; Chu G.; Li R.	2022	10.1016/j.future.2022.07.020
151	A comparative study of time frequency representation techniques for freeze of gait detection and prediction	Mostafa T.A.; Soltaninejad S.; McIsaac T.L.; Cheng I.	2021	10.3390/s21196446
152	A deep convolutional neural network model for hand gesture recognition in 2D near-infrared images	Can C.; Kaya Y.; Kilic F.	2021	10.1088/2057-1976/ac0d91
153	BOD: An efficient algorithm for distributed outlier detection	Wang X.-T.; Shen D.-R.; Bai M.; Nie T.-Z.; Kou Y.; Yu G.	2016	10.11897/SP.J.1016.2016.00036
154	Early warning system for financially distressed hospitals via data mining application	Koyuncugil A.S.; Ozgulbas N.	2012	10.1007/s10916-011-9694-1
155	Artificial Intelligence and Statistics: Just the Old Wine in New Wineskins?	Faes L.; Sim D.A.; van Smeden M.; Held U.; Bossuyt P.M.; Bachmann L.M.	2022	10.3389/fdgth.2022.833912

N.º	Title	Authors	Year	DOI
156	Improving Tax Audit Efficiency Using Machine Learning: The Role of Taxpayer's Network Data in Fraud Detection	Baghdasaryan V.; Davtyan H.; Sarikyan A.; Navasardyan Z.	2022	10.1080/08839514.2021.2012002
157	From human resources to human rights: Impact assessments for hiring algorithms	Yam J.; Skorburg J.A.	2021	10.1007/s10676-021-09599-7
158	Industry 4.0 contribution to asset management in the electrical industry	Biard G.; Nour G.A.	2021	10.3390/su131810369
159	A deep structured model for video captioning	Vinodhini V.; Sathiyabhama B.; Sankar S.; Somula R.	2020	10.4018/IJGCMS.2020040103
160	Telecom traffic pumping analytics via explainable data science	Irarrázaval M.E.; Maldonado S.; Pérez J.; Vairetti C.	2021	10.1016/j.dss.2021.113559
161	Prospects of Artificial Intelligence and Machine Learning Application in Banking Risk Management	Milojević N.; Redzepagic S.	2021	10.2478/jcbtp-2021-0023
162	Agent-based modeling from the perspectives of FinTech	Cui Y.; Xiong X.; Wei L.; He S.	2020	10.12011/1000-6788-2018-1798-09
163	Piecewise evolutionary segmentation for feature extraction in time series models	Glezakos T.J.; Tsiligiridis T.A.; Yialouris C.P.	2014	10.1007/s00521-012-1212-y
164	MAFI: GNN-Based Multiple Aggregators and Feature Interactions Network for Fraud Detection over Heterogeneous Graph	Jiang N.; Duan F.; Chen H.; Huang W.; Liu X.	2022	10.1109/TBDATA.2021.3132672
165	Design of a Network Security Audit System Based on Log Data Mining	Xing Y.	2022	10.1155/2022/6737194
166	Audit and tax in the context of emerging technologies: A retrospective analysis, current trends, and future opportunities	Atayah O.F.; Alshater M.M.	2021	10.4192/1577-8517-v21_4
167	Fraud Detection in Online Product Review Systems via Heterogeneous Graph Transformer	Tang S.; Jin L.; Cheng F.	2021	10.1109/ACCESS.2021.3084924
168	An emergency management system for government data security based on artificial intelligence	Luo H.	2020	10.18280/isi.250208
169	Identifying High-Risk Intersections for Walking and Bicycling Using Multiple Data Sources in the City of San Diego	Hasani M.; Jahangiri A.; Sener I.N.; Munira S.; Owens J.M.; Appleyard B.; Ryan S.; Turner S.M.; Ghanipoor MacHiani S.	2019	10.1155/2019/9072358
170	A topic evolution mining algorithm of news text based on feature evolving	Zhao X.-J.; Yang C.-M.; Li B.; Zhang H.; Jin P.-Q.; Yue L.-H.; Dai W.-K.	2014	10.3724/SP.J.1016.2014.00819
171	Detecting network intrusions by data mining and variable-length sequence pattern matching	Xinguang T.; Miyi D.; Chunlai S.; Xin L.	2009	
172	Network intrusion detection based on system calls and data mining	Tian X.; Cheng X.; Duan M.; Liao R.; Chen H.; Chen X.	2010	10.1007/s11704-010-0570-9
173	Adding Redundancy to LoRaWAN for Emergency Communications at the Factory Floor	Sisinni E.; Carvalho D.F.; Ferrari P.; Flammini A.; Gidlund M.	2022	10.1109/TII.2021.3124054

N.º	Title	Authors	Year	DOI
174	Uncertainty of key performance indicators for Industry 4.0: A methodology based on the theory of belief functions	Souifi A.; Boulanger Z.C.; Zolghadri M.; Barkallah M.; Haddar M.	2022	10.1016/j.compind.2022.103666
175	Reconfiguring a hierarchical supply chain model under pandemic using text mining and social media analysis	Wu K.J.; Bin Y.; Ren M.; Tseng M.-L.; Wang Q.; Chiu A.S.F.	2022	10.1108/IMDS-06-2021-0358
176	Research on Supply Chain Financial Risk Assessment Based on Blockchain and Fuzzy Neural Networks	Wang Y.	2021	10.1155/2021/5565980
177	Emotion recognition with residual network driven by spatial-frequency characteristics of EEG recorded from hearing-impaired adults in response to video clips	Bai Z.; Liu J.; Hou F.; Chen Y.; Cheng M.; Mao Z.; Song Y.; Gao Q.	2023	10.1016/j.compbiomed.2022.106344
178	An Economic Decision-Making Model for Drugs Using Big Data and Convolution Neural Network in Healthcare	Yuan J.	2022	10.1155/2022/2034685
179	Big data and artificial intelligence in the fields of accounting and auditing: a bibliometric analysis	Agustí M.A.; Orta-Pérez M.	2022	10.1080/02102412.2022.2099675
180	Construction and Simulation of Financial Audit Model Based on Convolutional Neural Network	Zhang X.	2021	10.1155/2021/1182557
181	Conceptualizing the use of the term financial risk by non-academics and academics using twitter messages and ScienceDirect paper abstracts	Kwak E.J.; Grable J.E.	2021	10.1007/s13278-020-00709-9
182	Person-identification using familiar-name auditory evoked potentials from frontal EEG electrodes	Jijomon C.M.; Vinod A.P.	2021	10.1016/j.bspc.2021.102739
183	Big data analysis of e-commerce loan risk of college students in the context of network finance	Haitao S.	2020	10.1007/s10257-019-00424-9
184	Scientific and Technological Strategies Proposal for the Construction of Digital Public Health Emergency Management System in China; [中国数字化公共卫生应急管理体系建设的科技策略建议]	Zhang X.; Lin H.; Wang J.; Xu C.; Hu M.; Meng B.; Liu D.; Xu M.; Zhu C.; Wang G.; Cao C.; Luo J.; Xiao G.; Lu Y.; Yang Y.; Zhi G.	2020	10.13203/j.whugis20200151
185	A novel hierarchical machine learning model for hospital-acquired venous thromboembolism risk assessment among multiple-departments	Ma H.; Sheng W.; Li J.; Hou L.; Yang J.; Cai J.; Xu W.; Zhang S.	2021	10.1016/j.jbi.2021.103892
186	STARS: Defending against Sockpuppet-Based Targeted Attacks on Reviewing Systems	Liu R.; Liu R.; Pugliese A.; Subrahmanian V.S.	2020	10.1145/3397463
187	Nonlaboratory-Based Risk Assessment Model for Type 2 Diabetes Mellitus Screening in Chinese Rural Population: A Joint Bagging-Boosting Model	Zhang L.; Wang Y.; Niu M.; Wang C.; Wang Z.	2021	10.1109/JBHI.2021.3077114

N.º	Title	Authors	Year	DOI
188	Power Grid Enterprise Intelligent Risk Identification Model Considering Multi-Attribute and Low Correlation Data	Zhou L.; Cai L.; Jiang L.; Chen L.	2019	10.1109/ACCESS.2019.2933754
189	Supply chain risk identification: a real-time data-mining approach	Deiva Ganesh A.; Kalpana P.	2022	10.1108/IMDS-11-2021-0719
190	Analyzing CSP Trustworthiness and Predicting Cloud Service Performance	Maeser R.	2020	10.1109/OJCS.2020.2994095
191	The role of data analytics within operational risk management: A systematic review from the financial services and energy sectors	Cornwell N.; Bilson C.; Gepp A.; Stern S.; Vanstone B.J.	2023	10.1080/01605682.2022.2041373
192	A proof-of-concept and feasibility analysis of using social sensors in the context of causal machine learning-based emergency management	Sahoh B.; Choksuriwong A.	2022	10.1007/s12652-021-03317-3
193	Time-Frequency Analysis of Scalp EEG With Hilbert-Huang Transform and Deep Learning	Zheng J.; Liang M.; Sinha S.; Ge L.; Yu W.; Ekstrom A.; Hsieh F.	2022	10.1109/JBHI.2021.3110267
194	Fraud Detection Using Neural Networks: A Case Study of Income Tax	Murorunkwere B.F.; Tuyishimire O.; Haughton D.; Nzabanita J.	2022	10.3390/fi14060168
195	Estimating Rainfall Intensity Using an Image-Based Deep Learning Model	Yin H.; Zheng F.; Duan H.-F.; Savic D.; Kapelan Z.	2023	10.1016/j.eng.2021.11.021
196	How emergency managers engage Twitter users during disasters	Xu Z.	2020	10.1108/OIR-08-2019-0275
197	E-Commerce Fraud Detection Model by Computer Artificial Intelligence Data Mining	Li J.	2022	10.1155/2022/8783783
198	Automated Risk Management based Software Security Vulnerabilities Management	Althar R.R.; Samanta D.; Kaur M.; Singh D.; Lee H.	2022	10.1109/ACCESS.2022.3185069
199	An Internet of Things based scalable framework for disaster data management	Ding Z.; Jiang S.; Xu X.; Han Y.	2022	10.1016/j.jnlssr.2021.10.005
200	Network distribution and sentiment interaction: Information diffusion mechanisms between social bots and human users on social media	Cai M.; Luo H.; Meng X.; Cui Y.; Wang W.	2023	10.1016/j.ipm.2022.103197
201	Leveraging the synergies between design science and behavioral science research methods	Sutton S.G.; Arnold V.; Collier P.; Leech S.A.	2021	10.1016/j.accinf.2021.100536
202	Generative adversarial networks for data augmentation and transfer in credit card fraud detection	Langevin A.; Cody T.; Adams S.; Beling P.	2022	10.1080/01605682.2021.1880296
203	Multimodal Data Fusion of Electromyography and Acoustic Signals for Thai Syllable Recognition	Jong N.S.; De Herrera A.G.S.; Phukpattaranont P.	2021	10.1109/JBHI.2020.3034158
204	A generic paradigm for mining human mobility patterns based on the GPS trajectory data using complex network analysis	Wang S.; Mei G.; Cuomo S.	2021	10.1002/cpe.5335

N.º	Title	Authors	Year	DOI
205	Detection of violations in Credit Cards of Banks and financial institutions based on Artificial neural network and Metaheuristic optimization algorithm	Monirzadeh Z.; Habibzadeh M.; Farajian N.	2018	10.14569/IJACSA.2018.090124
206	Risk words suggestion for information security audit by Bayesian inference	Satoh N.; Samejima M.	2019	10.1002/ecj.12133
207	Detection of fraud risks in retailing sector using MLP and SVM techniques	Pehlivanli D.; Eken S.; Ayan E.	2019	10.3906/elk-1902-18
208	Spatio-temporal mining of keywords for social media cross-social crawling of emergency events	Autelitano A.; Pernici B.; Scalia G.	2019	10.1007/s10707-019-00354-1
209	Tie me to the mast: artificial intelligence & reputation risk management	Hirsch P.B.	2018	10.1108/JBS-11-2017-0160
210	Big Data-Artificial Intelligence Fusion Technology in Education in the Context of the New Crown Epidemic	Zhao J.; Li Q.	2022	10.1089/big.2021.0245
211	Time-Aware Attention-Based Gated Network for Credit Card Fraud Detection by Extracting Transactional Behaviors	Xie Y.; Liu G.; Yan C.; Jiang C.; Zhou M.	2023	10.1109/TCSS.2022.3158318
212	Detection of Suspicious or UnTrusted Users in Crypto-Currency Financial Trading Applications	Mittal R.; Bhatia M.P.S.	2021	10.4018/IJDCF.2021010105
213	An Efficient Domain-Adaptation Method using GAN for Fraud Detection	Hwang J.; Kim K.	2020	10.14569/IJACSA.2020.0111113
214	Turkish sign language recognition based on multistream data fusion	Gündüz C.; Polat H.	2021	10.3906/ELK-2005-156
215	When Two are Better Than One: Synthesizing Heavily Unbalanced Data	Ferreira F.; Lourenco N.; Cabral B.; Fernandes J.P.	2021	10.1109/ACCESS.2021.3126656
216	The role of emerging banking technologies for risk management and mitigation to reduce non-performing assets and bank Frauds in the Indian Banking System	Bhasin N.K.; Rajesh A.	2022	10.4018/IJeC.290293
217	Applying deep learning to audit procedures: An illustrative framework	Sun T.S.	2019	10.2308/acch-52455
218	Quantitative Detection of Financial Fraud Based on Deep Learning with Combination of E-Commerce Big Data	Liu J.; Gu X.; Shang C.	2020	10.1155/2020/6685888
219	Walk2Map: Extracting Floor Plans from Indoor Walk Trajectories	Mura C.; Pajarola R.; Schindler K.; Mitra N.	2021	10.1111/cgf.142640
220	Design of Internet of Things and big data analytics-based disaster risk management	Zhou L.; Huang H.; Muthu B.A.; Sivaparthipan C.B.	2021	10.1007/s00500-021-05953-5
221	Big data quality prediction informed by banking regulation	Wong K.Y.; Wong R.K.	2021	10.1007/s41060-021-00257-1
222	Detecting problematic transactions in a consumer-to-consumer e-commerce network	Kodate S.; Chiba R.; Kimura S.; Masuda N.	2020	10.1007/s41109-020-00330-x
223	Co-Check: Collaborative Outsourced Data Auditing in Multicloud Environment	Mao J.; Tian W.; Zhang Y.; Cui J.; Ma H.; Bian J.; Liu J.; Zhang J.	2017	10.1155/2017/2948025

N.º	Title	Authors	Year	DOI
224	Density-Based Local Outlier Detection on Uncertain Data	Cao K.-Y.; Luan F.-J.; Sun H.-L.; Ding G.-H.	2017	10.11897/SP.J.1016.2017.02231
225	Machine learning- A nd evidence theory-based fraud risk assessment of China's box office	Qiu S.; He H.-Q.	2018	10.1109/ACCESS.2018.2883487
226	Adverse drug reaction early warning using user search data	Shang W.; Chen H.; Livoti C.	2017	10.1108/OIR-10-2015-0341
227	Risk based government audit planning using naïve bayes classifiers	Balaniuk R.; Bessiere P.; Mazer E.; Cobbe P.	2012	10.3233/978-1-61499-105-2-1313
228	Simulated neural dynamics of decision-making in an auditory delayed match-to-sample task	Wen S.; Ulloa A.; Husain F.; Horwitz B.; Contreras-Vidal J.L.	2008	10.1007/s00422-008-0234-0
229	"Stock-touting" through spam e-mails: A data mining case study	Zaki M.; Theodoulidis B.; Sols D.D.	2011	10.1108/17410381111149639
230	Advancing Stuttering Detection via Data Augmentation, Class-Balanced Loss and Multi-Contextual Deep Learning	Sheikh S.A.; Sahidullah M.; Hirsch F.; Ouni S.	2023	10.1109/JBHI.2023.3248281
231	Forecasting corporate credit ratings using big data from social media	Chen Y.-J.; Chen Y.-M.	2022	10.1016/j.eswa.2022.118042
232	Intelligent Risk Management in Construction Projects: Systematic Literature Review	Chenya L.; Aminudin E.; Mohd S.; Yap L.S.	2022	10.1109/ACCESS.2022.3189157
233	A Novel Domain Adversarial Networks Based on 3D-LSTM and Local Domain Discriminator for Hearing-Impaired Emotion Recognition	Tian Z.; Li D.; Yang Y.; Hou F.; Yang Z.; Song Y.; Gao Q.	2023	10.1109/JBHI.2022.3212475
234	Deep learning in the stock market—a systematic survey of practice, backtesting, and applications	Olorunnimbe K.; Viktor H.	2023	10.1007/s10462-022-10226-0
235	An evaluation of deep learning models for chargeback Fraud detection in online games	Wei Y.-C.; Lai Y.-X.; Wu M.-E.	2023	10.1007/s10586-022-03674-4
236	A real-time approach to recognition of Turkish sign language by using convolutional neural networks	Güney S.; Erkuş M.	2022	10.1007/s00521-021-06664-6
237	BotSpot++: A Hierarchical Deep Ensemble Model for Bots Install Fraud Detection in Mobile Advertising	Zhu Y.; Wang X.; Li Q.; Yao T.; Liang S.	2022	10.1145/3476107
238	A Survey on GAN Techniques for Data Augmentation to Address the Imbalanced Data Issues in Credit Card Fraud Detection	Strelcenia E.; Prakoonwit S.	2023	10.3390/make5010019
239	Deep Learning Models for Single-Channel Speech Enhancement on Drones	Mukhutdinov D.; Alex A.; Cavallaro A.; Wang L.	2023	10.1109/ACCESS.2023.3253719
240	Design of financial big data audit model based on artificial neural network	Zhang Z.; Wang Z.	2021	10.1007/s13198-021-01258-w
241	An empirical study of methods, metrics and evaluation of data mining techniques in credit card fraudulence detection	Karthika J.; Senthilselvi A.	2020	10.5373/JARDCS/V12I7/20202016
242	Research on the information construction of accounting audit based on the big data	Lin D.	2017	10.4018/IJITWE.2017070107

N.º	Title	Authors	Year	DOI
	of computer			
243	Innovative risk early warning model based on internet of things under big data technology	Wang C.; Liu S.	2021	10.1109/ACCESS.2021.3095503
244	Exploring Social Relationships in Text Streams	Wang Y.	2016	10.4108/eai.9-8-2016.151631
245	Financial volatility forecasting: A sparse multi-head attention neural network	Lin H.; Sun Q.	2021	10.3390/info12100419
246	Active Learning for Human-in-the-Loop Customs Inspection	Kim S.; Mai T.; Han S.; Park S.; Nguyen T.; So J.; Singh K.; Cha M.	2022	10.1109/TKDE.2022.3144299
247	Arabic Sign Language Recognition using Lightweight CNN-based Architecture	Al-Khuraym B.Y.; Ismail M.M.B.	2022	10.14569/IJACSA.2022.0130438
248	The BP Neural Network with Adam Optimizer for Predicting Audit Opinions of Listed Companies	Wu H.-P., Member IAENG; Li L.	2021	
249	Fraud prediction for credit card using classification method	Monika E.; Kaur E.A.	2018	
250	Theory and Method of Time-varying Computational Experiments for the Fully Mechanized Mining Process in an Artificial System Environment	Feng Z.; Zhu S.; Wu J.; Guo H.	2019	10.1109/ACCESS.2019.2954591
251	Digital technologies and artificial intelligence's present and foreseeable impact on lawyering, judging, policing and law enforcement	Nissan E.	2017	10.1007/s00146-015-0596-5
252	Accurately detecting source code of attacks that increase privilege	Cunningham R.K.; Stevenson C.S.	2015	
253	Fraud detection in financial statement using data mining technique and performance analysis	Meenatkshi R.; Sivaranjani	2016	
254	Artificial intelligence to counteract "KPI overload" in business process monitoring: the case of anti-corruption in public organizations	Caruso S.; Bruccoleri M.; Pietrosi A.; Scaccianoce A.	2023	10.1108/BPMJ-11-2022-0578
255	Research on road extraction of remote sensing image based on convolutional neural network	Jiang Y.	2019	10.1186/s13640-019-0426-7
256	Towards Improving Causality Mining using BERT with Multi-level Feature Networks	Ali W.; Zuo W.; Ali R.; Rahman G.; Zuo X.; Ullah I.	2022	10.3837/tiis.2022.10.002
257	An Intelligent Mechanism to Automatically Discover Emerging Technology Trends: Exploring Regulatory Technology	Huang S.M.; Yen D.C.; Yan T.J.; Yang Y.T.	2022	10.1145/3485187
258	Adaptation of a robotic dialog system for medication reminder in elderly care	Su Z.; Sheng W.; Yang G.; Bishop A.; Carlson B.	2022	10.1016/j.smhl.2022.100346
259	Social IoT Approach to Cyber Defense of a Deep-Learning-Based Recognition System in front of Media Clones Generated by Model Inversion Attack	Khosravy M.; Nakamura K.; Nitta N.; Dey N.; Gonzalez Crespo R.; Herrera-Viedma E.; Babaguchi N.	2023	10.1109/TSMC.2022.3220080

N.º	Title	Authors	Year	DOI
260	On generating high InfoQ with Bayesian networks	Kenett R.S.	2016	10.1080/16843703.2016.1189182
261	Modeling the dynamic brain network representation for autism spectrum disorder diagnosis	Cao P.; Wen G.; Liu X.; Yang J.; Zaiane O.R.	2022	10.1007/s11517-022-02558-4
262	Probabilistic time series forecasting with deep non-linear state space models	Du H.; Du S.; Li W.	2023	10.1049/cit2.12085
263	A social media event detection framework based on transformers and swarm optimization for public notification of crises and emergency management	Dahou A.; Mabrouk A.; Ewees A.A.; Gaheen M.A.; Abd Elaziz M.	2023	10.1016/j.techfore.2023.122546
264	Unsupervised learning for financial statement fraud detection using manta ray foraging based convolutional neural network	Singh Yadav A.K.; Sora M.	2022	10.1002/cpe.7340
265	A Proposed Fraud Detection Model based on e-Payments Attributes a Case Study in Egyptian e-Payment Gateway	Nasr M.H.; Farrag M.H.; Nasr M.M.	2022	10.14569/IJACSA.2022.0130522
266	Fixing shelf out-of-stock with signals in point-of-sale data	Chuang H.H.-C.	2018	10.1016/j.ejor.2017.10.059
267	On the dynamics of credit history and social interaction features, and their impact on creditworthiness assessment performance	Muñoz-Cancino R.; Bravo C.; Ríos S.A.; Graña M.	2023	10.1016/j.eswa.2023.119599
268	Quantile correlative deep feedforward multilayer perceptron for crop yield prediction	Sivanantham V.; Sangeetha V.; Alnuaim A.A.; Hatamleh W.A.; Anilkumar C.; Hatamleh A.A.; Sweidan D.	2022	10.1016/j.compeleceng.2022.107696
269	Disentangled and Side-Aware Unsupervised Domain Adaptation for Cross-Dataset Subjective Tinnitus Diagnosis	Li Y.; Liu Z.; Yao L.; Monaghan J.J.M.; McAlpine D.	2023	10.1109/JBHI.2022.3225089
270	Strategic Supply Chain Risk Management Artificial Intelligence and Big Data to Support Strategic Supply Chain Risk Management; [Strategisches Supply-Chain-Risikomanagement Einsatz von Künstlicher Intelligenz und Big Data zur Unterstützung des strategischen Supply-Chain-Risikomanagements]	Kramer K.J.; Mousavi D.; Schmidt M.	2022	10.1515/zwf-2022-1055
271	Typhoon Disaster Network Emotion Analysis Method based on Semantic Rules and Word Vector; [基于语义规则和词向量的台风灾害网络情感分析方法]	Lin X.; Wu S.	2022	10.12082/dqxxkx.2022.210575
272	Using community information for natural disaster alerts	Chen C.C.; Wang H.-C.	2022	10.1177/0165551520979870
273	Complex nonlinear neural network prediction with IOWA layer	Hussain W.; Merigó J.M.; Gil-Lafuente J.; Gao H.	2023	10.1007/s00500-023-07899-2
274	A Fast and Efficient Algorithm for Outlier Detection Over Data Streams	Hassaan M.; Maher H.; Gouda K.	2021	10.14569/IJACSA.2021.0121185

N.º	Title	Authors	Year	DOI
275	NFAD: fixing anomaly detection using normalizing flows	Ryzhikov A.; Borisyak M.; Ustyuzhanin A.; Derkach D.	2021	10.7717/PEERJ-CS.757
276	Research challenges and future directions towards medical data processing	Ampavathi A.; Vijaya Saradhi T.	2022	10.1080/21681163.2021.2018665
277	Research on financial network big data processing technology based on fireworks algorithm	Luo T.	2019	10.1186/s13638-019-1443-z
278	Cloud Network and Mathematical Model Calculation Scheme for Dynamic Big Data	Chen Y.; Qiu Z.	2020	10.1109/ACCESS.2020.3009675
279	A social media-based over layer on the edge for handling emergency-related events	Tundis A.; Melnik M.; Naveed H.; Mühlhäuser M.	2021	10.1016/j.compeleceng.2021.107570
280	Data distribution-based cost-sensitive broad learning system; [基于数据分布特性的代价敏感宽度学习系统]	Xu P.-F.; Wang M.; Liu J.-P.; Tang Z.-H.; Ma T.-Y.	2021	10.13195/j.kzyjc.2019.1484
281	Towards provenance cloud security auditing based on association rule mining	Tu S.; Huang X.	2019	10.31534/engmod.2019.2-4.ri.01d
282	Security challenges and cybercrime	Cook E.; Kearney P.	2015	
283	High Risk Tree Mining Method for Analysis of Power System Risk Device Set	Wu R.; Chen W.; Tang L.; Fan J.	2017	10.7500/AEPS20161228011
284	A novel density-based outlier detection approach for low density datasets	Guan D.; Chen K.; Yuan W.; Han G.	2017	10.6138/JIT.2017.18.7.20170804
285	Cross-Subject Tinnitus Diagnosis Based on Multi-Band EEG Contrastive Representation Learning	Wang C.-D.; Zhu X.-R.; Zhou X.; Li J.; Lan L.; Huang D.; Zheng Y.; Cai Y.	2023	10.1109/JBHI.2023.3264521
286	Financial Default Risk Prediction Algorithm Based on Neural Network under the Background of Big Data	Xie T.; Zhang J.	2022	10.1155/2022/8743778
287	Tracking down financial statement fraud by analyzing the supplier-customer relationship network	Li J.; Chang Y.; Wang Y.; Zhu X.	2023	10.1016/j.cie.2023.109118
288	Insights from hashtag #supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research	Chae B.	2015	10.1016/j.ijpe.2014.12.037
289	Empirical Analysis of Enterprise Financial Management Risk Prediction in View of Associative Memory Neural Network	Cheng H.; Zhang X.	2022	10.1155/2022/7825000
290	A Big Data-Driven Financial Auditing Method Using Convolution Neural Network	Zhao H.; Wang Y.	2023	10.1109/ACCESS.2023.3269438
291	Medical Big Data Risk Management: A Systematic Management Approach Based on Bayesian Belief Networks	Zhang X.; Liu X.; Zhou S.; Ma N.	2023	10.1155/2023/9507349
292	Edge Technologies for Disaster Management: A Survey of Social Media and Artificial Intelligence Integration	Aboualola M.; Abualsaud K.; Khattab T.; Zorba N.; Hassanein H.S.	2023	10.1109/ACCESS.2023.3293035

N.º	Title	Authors	Year	DOI
293	Film and television art innovation in network environment by using collaborative filtering recommendation algorithm	Lai X.; Chen J.	2023	10.1007/s00500-023-08134-8
294	Coupled Attention Networks for Multivariate Time Series Anomaly Detection	Xia F.; Chen X.; Yu S.; Hou M.; Liu M.; You L.	2023	10.1109/TETC.2023.3280577
295	The AILA Methodology for Automated and Intelligent Likelihood Assignment in Risk Assessment	Bella G.; Daniele C.; Raciti M.	2023	10.1109/ACCESS.2023.3245333
296	Artificial Intelligence Enterprise Management Using Deep Learning	Liu X.; Han L.	2022	10.1155/2022/2422434
297	Artificial Intelligence Model for Risk Management in Healthcare Institutions: Towards Sustainable Development	Darwiesh A.; El-Baz A.H.; Abualkishik A.Z.; Elhoseny M.	2023	10.3390/su15010420
298	Characterization of Synthetic Health Data Using Rule-Based Artificial Intelligence Models	Lenatti M.; Paglialonga A.; Orani V.; Ferretti M.; Mongelli M.	2023	10.1109/JBHI.2023.3236722
299	A novel framework for online transaction fraud detection system based on deep neural network	Kanika; Singla J.	2022	10.3233/JIFS-212616
300	MRFS: Mining Rating Fraud Subgraph in Bipartite Graph for Users and Products	Yu W.; Wang W.; Xu G.; Wu H.; Li H.; Wang J.; Li X.; Liu J.	2023	10.1109/TCSS.2022.3233821
301	Forecasting the volatility of stock price index	Hyup Roh T.	2007	10.1016/j.eswa.2006.08.001
302	User profiling and classification for fraud detection in mobile communications networks	Hollmén J.	2000	
303	Fraud detection from paper texture using Siamese networks	Emiroğlu E.E.; Şahin E.; Vural F.T.Y.	2023	10.1007/s11760-023-02558-3
304	Measuring Ethical Values with AI for Better Teamwork	Altuntas E.; Gloor P.A.; Budner P.	2022	10.3390/fi14050133
305	A Novel Model for Ship Trajectory Anomaly Detection Based on Gaussian Mixture Variational Autoencoder	Xie L.; Guo T.; Chang J.; Wan C.; Hu X.; Yang Y.; Ou C.	2023	10.1109/TVT.2023.3284908
306	Time Series Impact Through Topic Modeling	Cendrero J.; Gonzalo J.; Galletero M.; Zapata I.	2022	10.1109/ACCESS.2022.3202960
307	Graph-based ship traffic partitioning for intelligent maritime surveillance in complex port waters	Xin X.; Liu K.; Loughney S.; Wang J.; Li H.; Yang Z.	2023	10.1016/j.eswa.2023.120825
308	Big Data and precision agriculture: a novel spatio-temporal semantic IoT data management framework for improved interoperability	San Emeterio de la Parte M.; Martínez-Ortega J.-F.; Hernández Díaz V.; Martínez N.L.	2023	10.1186/s40537-023-00729-0
309	A contingency approach for time-cost trade-off in construction projects based on machine learning techniques	Wang P.; Wang K.; Huang Y.; Fenn P.	2023	10.1108/ECAM-11-2022-1104
310	E-ware: a big data system for the incremental discovery of spatio-temporal events from microblogs	Afyouni I.; Khan A.; Alghbari Z.	2022	10.1007/s12652-022-04104-4

N.º	Title	Authors	Year	DOI
311	One-Class Adversarial Fraud Detection Nets with Class Specific Representations	Peng H.; Zhao J.; Li L.; Ren Y.; Zhao S.	2023	10.1109/TNSE.2023.3273543
312	Ethical scaling for content moderation: Extreme speech and the (in)significance of artificial intelligence	Udupa S.; Maronikolakis A.; Wisiorek A.	2023	10.1177/20539517231172424
313	A forensics and compliance auditing framework for critical infrastructure protection	Henriques J.; Caldeira F.; Cruz T.; Simões P.	2023	10.1016/j.ijcip.2023.100613
314	Venice Was Flooding ⋯ One Tweet at a Time	Lorini V.; Rufolo P.; Castillo C.	2022	10.1145/3555107
315	Social Media Multimodal Information Analysis based on the BiLSTM-Attention-CNN-XGBoost Ensemble Neural Network	Jixian L.; Gang A.; Zhihao S.; Xiaoqiang S.	2022	10.14569/IJACSA.2022.0131215
316	A unique color-coded visualization system with multimodal information fusion and deep learning in a longitudinal study of Alzheimer's disease	Eslami M.; Tabarestani S.; Adjouadi M.	2023	10.1016/j.artmed.2023.102543
317	Prediction of sign language recognition based on multi layered CNN	Arun Prasath G.; Annapurani K.	2023	10.1007/s11042-023-14548-1
318	Mining spatio-temporal information on microblogging streams using a density-based online clustering method	Lee C.-H.	2012	10.1016/j.eswa.2012.02.136
319	Moving Emergency Response Forward: Leveraging Machine-Learning Classification of Disaster-Related Images Posted on Social Media	Johnson M.; Murthy D.; Robertson B.W.; Smith W.R.; Stephens K.K.	2023	10.1080/07421222.2023.2172778
320	Fraud detection in financial statements using data mining and GAN models	Aftabi S.Z.; Ahmadi A.; Farzi S.	2023	10.1016/j.eswa.2023.120144
321	Anomaly Identification Model for Telecom Users Based on Machine Learning Model Fusion	Lin J.; Wang P.; Wu C.	2022	10.20532/cit.2022.1005459
322	Unsupervised machine learning for managing safety accidents in railway stations	Alawad H.; Kaewunruen S.	2023	10.1109/ACCESS.2023.3264763
323	Finding the needle: A risk-based ranking of product listings at online auction sites for non-delivery fraud prediction	Almendra V.	2013	10.1016/j.eswa.2013.02.027
324	Detecting Credit Card Fraud by Generative Adversarial Networks and Multi-head Attention Neural Networks	Meng Z.; Xie Y.; Sun J.	2023	
325	Machine Learning Techniques for Detecting Phishing URL Attacks	Mosa D.T.; Shams M.Y.; Abohany A.A.; El-Kenawy E.-S.M.; Thabet M.	2023	10.32604/cmc.2023.036422
326	An Audit Risk Model Based on Improved BP Neural Network Data Mining Algorithm	Niu W.; Zhao L.; Jia P.; Chu J.	2022	10.1155/2022/9977292
327	Management of gross negligence manslaughter liability construction for professionals and lessons learned	Liao M.-C.; Hsieh T.-Y.; Wang W.-H.	2023	10.1108/ECAM-11-2022-1059

N.º	Title	Authors	Year	DOI
328	Flower pollination optimization algorithm with stacked temporal convolution network-based classification for financial anomaly fraud detection	Krishnavardhan N.; Govindarajan M.; Achutha Rao S.V.	2023	10.1007/s00500-023-08732-6
329	Structural Analysis of the Evolution Mechanism of Online Public Opinion and its Development Stages Based on Machine Learning and Social Network Analysis	Liu Z.; Wu X.	2023	10.1007/s44196-023-00277-8
330	SPQER: Speech quality evaluation using word recognition for VoIP communication in lossy and mobile networks	Schuetz B., schuetz@uos.de; Aschenbruck N.	2020	10.1109/OJCS.2020.3011392
331	The revolution in spectrum allocation	Greenstein S.	2009	10.1109/MM.2009.49
332	Trust-based federated learning for network anomaly detection	Chen N.; Jin Y.; Li Y.; Cai L.	2021	10.3233/WEB-210475
333	Applying datamining techniques to predict hearing aid type for audiology patients	Aljabery M.A.; Kurnaz S.	2020	10.6688/JISE.202003_36(2).0002
334	The use of Artificial Intelligence and its relation with Auditing: the case of companies quoted on the PSI-20; [O uso da Inteligência Artificial e a sua relação com a Auditoria: o caso das empresas cotadas no PSI-20]	Matias R.; Bonsón E.; Pedrosa I.	2021	
335	Comparative Analysis of a Deep Learning Approach with Various Classification Techniques for Credit Score Computation	Pandey A.; Shukla S.; Mohbey K.K.	2021	10.2174/2666255813999200721004720
336	Designing an expert system for fraud detection in private telecommunications networks	Hilas C.S.	2009	10.1016/j.eswa.2009.03.031
337	Cost-effective provable secure cloud storage self-auditing scheme for big data in WMSNS	Zhang X.; Zhao J.; Mu L.; Zhang X.	2019	10.1504/IJESDF.2019.102566