Artificial Intelligence Techniques to Detect and Prevent Corruption in Procurement: A Systematic Literature Review

Yeferson Torres Berru[1,2](✉), Vivian Félix López Batista[1](✉), Pablo Torres-Carrión[3](✉), and Maria Gabriela Jimenez[2]

[1] University of Salamanca, Plaza de la Merced, s/n, 37008 Salamanca, Spain
{ymtorresb,vivian}@usal.es
[2] Instituto Superior Tecnológico Loja, Av. Granada y Turunuma, 1101608 Loja, Ecuador
[3] Universidad Técnica Particular de Loja, San Cayetano Alto S/N, 1101608 Loja, Ecuador
pvtorres@utpl.edu.ec

Abstract. Transparency International estimates that the costs of corruption in public procurement reach between 20 and 25% of the contract value, sometimes reaching 40–50%. In this study, we analyzed differentness kinds of corruption like (bribery, collusion embezzlement, misappropriation, fraud, abuse of discretion, favoritism, nepotism), and six types of Artificial Intelligence techniques (classification, regression, clustering, prediction, outlier detection, and visualization). The methodology proposed by Torres-Carrion was used, and four research questions were raised, which allow knowing the types of research carried out, the characteristics of the organizations in which the investigations are carried out, the technological tools, and data mining methodologies and techniques. The search was done in the Scopus and Web of Science databases, getting 102 articles published between 2015 and 2019. The primary data mining techniques used are logistic models, neural networks, Bayesian networks, supported vector machines, and decision trees.

Keywords: Artificial Intelligence · Corruption · Data mining · Procurement · Systematic literature review

1 Introduction

Corruption expenditures are equivalent to 5% of global GDP, according to the G-20 [1], being the third most lucrative "industry" of all those in the world. Transparency International estimates that the costs of corruption in public contracts average 20–25% of the contract value, and can reach 40–50% in some cases [2]; public procurement accounted for 32.5% of government expenditure. The highest risk of corruption in hiring occurs during the planning stage, potential frauds in the procurement system take very diverse forms starting from bribery, collusion embezzlement, misappropriation, fraud, abuse of discretion, favoritism, nepotism [3].

M. Botto-Tobar et al. (Eds.): ICAT 2019, CCIS 1194, pp. 254–268, 2020.
https://doi.org/10.1007/978-3-030-42520-3_21

Data mining combines Artificial Intelligence techniques (classification, regression, clustering, prediction, outlier detection, and visualization) with statistical analysis techniques (clustering, dimensional analysis, etc.), which allow analyzing information such as data, text, images, audio. In automatic learning, algorithms can be classified as supervised, unsupervised, and reinforcement [4]. These fields of knowledge, discussed for this systematic review of the literature (SLR), are detailed in Fig. 1, which allows establishing the theoretical knowledge base on which the review has its basis to contribute to the knowledge about data mining research and artificial intelligence for the detection and prevention of anomalies in contracts.

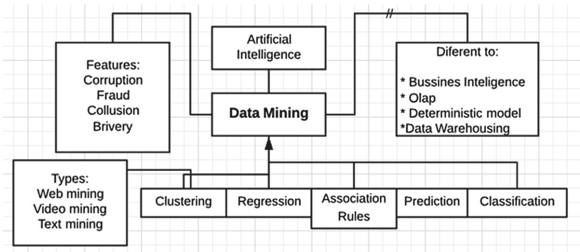

Fig. 1. Theoretical representation of data mining in contracts

1.1 Related Systematic Review of Literature

After the SLR, we proceed to look for previous literature review works related to the theoretical constructs: AI, data mining, and corruption. In the results obtained (see Table 1), the studies on the different types of fraud, at the private and governmental levels, are discussed, the different techniques of automatic learning and data mining are also considered in the process of fraud detection and prevention. All research is more than five years old and does not answer the research questions proposed in this study.

Table 1. Related reviews

Article	Analysis	#Papers
Fraud detection system: a survey	Survey about fraud prevention systems five areas of electronic frauds (e-fraud): credit card, telecommunication, health care insurance, automobile insurance and online auction in papers between 1994 and 2014	70
A survey of machine-learning and nature-inspired based credit card fraud detection techniques [5]	The survey revealed that various machine learning-based and nature-inspired algorithms had been used to handle credit card fraud detection between 2010 and 2014	47
The application of data mining techniques in financial fraud detection: a classification framework and an academic review of the literature [6]	Review data mining techniques in financial fraud, the articles were published between 1997 and 2008 analyzing fraud in bank, insurance and securities categories, the six data mining techniques (classification, regression, clustering, prediction, outlier detection, and visualization) and the main artificial intelligence methods logistic models, neural networks, the Bayesian belief network, and decision trees	49

Section two details the applied methodology, research questions, inclusion criteria, and semantic structure of the search.; in section three, research questions are answered based on the 102 articles found; finally, in section four, conclusions and future work are presented.

2 Method

2.1 Method for SLR

We used the method for a systematic review of the literature by Torres-Carrión [7] adapted from Kitchenham [8], which divides the process into three phases: planning, conducting the review, and reporting the review whit the PRISMA method [9]. As part of planning the search process, several general and specific inclusion and exclusion criteria were defined.

2.2 Research Questions

RQ1: What methods are being applied to investigate corruption in public procurement contracts?
RQ2: What are the characteristics of the organizations in which the research has been carried out?
RQ3: What technological tools are being used to investigate the detection and prevention of corruption?
RQ4: What algorithms, methodologies, and data analysis tools are used to detect corruption?

2.3 Quality Inclusion and Exclusion Protocols

As inclusion criteria, scientific articles published in the Web of Science (WOS) and Scopus databases during the years 2015–2019 are considered to be the type of document (article, review, editorial or conference proceedings), research area (Computer Science, Social Sciences and Decision Sciences); it excludes repeated documents, short papers, posters, and book chapters. As a quality criterion, a detailed review of the articles is carried out, filtering the studies that do not analyze corruption and its forms, or that do not apply data mining and artificial intelligence.

2.4 Semantic Search Structure

From the theoretical constructs (see Fig. 1), synonyms are sought in the scientific thesaurus; operators are applied (OR, AND, W/) to optimize the search, although Table 2 only shows the search query for Scopus, WOS search replaces the operators with the correspondents to its database. The procedure is detailed in levels with the resulting number of articles, as the inclusion, exclusion, and quality criteria are applied. One hundred two articles are obtained, identified as valid and explicitly related to the problem raised, and with which this SLR is working.

Table 2. Semantic search structure

Level	Thesaurus	SCOPUS script	Scopus	WOS
L1	Data mining			
	Mining W/4 (data or video or text or web)) Classificat* Cluster* Regression Association rules Detection Prediction Sequential Patterns	(Mining W/4 (data OR video OR text OR web)) OR classificat* OR cluster* OR regression OR (association W/2 rules) OR detection OR prediction OR (sequential W/2 patterns) OR (learning W/4 (machine OR deep OR reinforced))	6,962,346	8,072,286
L2	Corruption	Corruption OR	9,653	35,781
	Bribery, collusion, embezzlement, fraud, abuse of discretion, favoritism, nepotism	Bribery OR collusion OR (embezzlement OR misappropriation) OR fraud OR (abuse W/0 of W/0 discretion) OR favoritism OR nepotism)		
L3	Contracts	(Contract OR	691	1.811
	Contract, purchase, investment, procurement, acquisition, acquirement, tendering	Purchase OR investment OR procurement OR acquisition OR acquirement OR tendering)		
L4	Review protocol			
	(Last 5 years) from 2015		351	796
	(Research Areas) Computer Science, Social Sciences, Decision Sciences		222	275
	Article or review or editorial or conference proceedings		210	255
	Quality criteria		58	54
L5 Combination of results in Scopus and WOS (repeated = 5)			**102**	

3 Results

The results are presented in conformity with the research questions established in the methodology, and their corresponding variables and indicators.

RQ1: What methods are being applied to investigate corruption in public procurement contracts?

Knowing the methodology used by other researchers gives light to the planning of new studies. 87% of investigations work with a previously structured database; it is also observed that techniques such as web Scraping (3%) are rarely used for data collection (see Table 3). Quantitative research is the most used, as well as the statistical method of correlation. As for the time of validity of the data for the study, most of them are 1–3 years, and most of the research is an experimental type, with contributions to computer science. In this sense, the research highlights [10], which performs research with multivariate analysis and correlation, using web scraping and 4-year data.

Table 3. Research question 1.

Data collection instrument [11, 12]		*f*
Survey	[13–18]	6
Web Scraping	[13, 19]	2
Database	[10, 20–76]	57
Type of research [12]		
Qualitative	[77–79]	3
Quantitative	[10, 11, 13–102]	90
Mixt	[5, 103–110]	9
Statistics to evaluate results [12]		
Univariate	[14, 24, 26, 34, 64, 66, 80–90]	17
Multi-varied	[70, 91, 93]	6
Correlation	[10, 13, 15–23, 25, 27–33, 35–39, 41–48, 50–61, 65, 67–69, 71–77, 81, 89, 92, 95, 97–100, 111]	79
Data period		
<1 year	[26, 28, 29, 32, 40, 81]	6
1> & $\leq$ 3	[41, 43, 46, 47, 52, 53, 59, 61, 64–67, 71, 72, 75, 76, 95, 101, 102]	23
>3& $\leq$ 5	[10, 11, 13, 21, 27, 39, 49, 86, 97]	9
>5& $\leq$ 10	[92, 99]	2
>10	[20, 46, 52, 60, 82]	5
Research design [12]		
Experimental	[10, 11, 13, 16, 17, 19, 20, 22–26, 29–35, 37–76, 80, 86, 88, 91, 92, 95–97, 99, 100, 111]	70
No experimental	[5, 14, 23, 24, 36, 77–79, 82–85, 87–89, 94, 95, 98, 101, 104, 110]	23
Quasi experimental	[15, 18, 21, 81]	4
Field of science		
Computer Science	[11, 13, 17–26, 29–34, 36–40, 45–59, 62–76, 80, 81, 86–88, 90, 92, 94, 95, 99, 103, 104, 110, 111]	70
Economy	[27, 28, 35, 77–79, 82, 84, 85, 100]	10
Mathematic	[14, 34, 83, 89]	4
Statistics	[10, 16, 39, 61, 80, 91, 96]	6

RQ2: What are the characteristics of the organizations in which the research has been carried out?

50% of the works focus on the private sector, with a significant presence of experimental type works that relate the public and private sectors (17%). The main commercial activity of the organizations is the provision of services, having the banking sector as the most studied (64%) (see Table 4). In the government sector, tax fraud is prominent (33%), and public purchases account for only 6% of total investigations.

Table 4. Research question 2.

Activity sector		*f*
Public	[10, 16, 18, 24, 34, 35, 39, 42, 45, 47, 48, 50, 54, 57, 65, 66, 68, 71, 76, 77, 81, 84–87, 89, 91, 92, 97, 100]	31
Private	[11, 13, 15, 16, 18, 19, 25–29, 31–33, 36–38, 43, 46, 51–53, 58–64, 70, 72, 75, 88, 90, 91, 101, 102, 104–107, 111]	49
Mixt	[10, 11, 17, 30, 44, 49, 78–80, 83, 92, 93]	17
Commercial activity		
Services	[5, 13, 18, 21–23, 25–29, 33, 36–38, 47, 50–53, 55–57, 59, 62, 63, 70–72, 76, 81, 88, 90, 92, 104, 111]	36
Commercial	[19, 31, 32, 49, 58, 64, 65, 94]	8
Government	[10, 11, 16, 24, 35, 39, 41–43, 45, 46, 48, 54, 61, 67, 68, 86, 87, 89, 91, 100]	22
Organization		
Bank	[5, 21, 22, 25, 26, 29, 33, 36, 37, 47, 51, 55, 59, 62, 63, 70–72, 76, 79, 88, 90, 96, 104, 111]	25
Buys online	[32, 65]	2
Public buys	[10, 16, 35, 42, 43, 54, 69]	7
Taxes	[41, 67, 87, 91, 100]	5

In the mixed activity sector, the research conducted by Dhurandhar [30] proposes a bigdata-based solution for risk analysis in public and private sector procurement; focused on the public sector, in [46] a framework for the detection of crimes in contracts is presented, supported by information provided by the World Bank; in the private sector in [90] an experimental analysis of data mining tools in the banking sector is conducted.

RQ3: What technological tools are being used to investigate the detection and prevention of corruption?

Twenty articles (20/112) are selected that in their methodology propose the use or development of a technological tool to evaluate corruption (see Table 5); 89% are desktop tools; in terms of web platforms, these analyze corruption in contracts in public procedures [30, 69] and in the process of buying medicines [13].

The degree of reliability in the detection or prevention of corruption, as the case may be, has as its greatest interval between the 80%-90%; Baader et al. [17] obtain the lowest detection rate with 48.6%, applying the red flag approach with mining process to reduce the number of false positives in fraud analysis, whereas Darwish [23] obtains the best detection rate with a 98,5% in the analysis of credit card transactions in the banking environment; in the government environment, [43] work detects fake suppliers from the analysis of satellite images of the locations of companies, obtaining the best index of detection (97%). The pattern of software engineering and the computer security standard used in the tools were also analyzed, but no coincidences were found in the analysis of the work, excepting Carminati [71], which, in addition to the analysis tool, presents a mechanism for preventing the Mimicry Attack.

Table 5. Research question 3.

Platform		*f*
Desktop	[19, 23, 24, 31, 38–40, 46, 48, 52, 56, 58, 61, 64, 86, 92, 95]	17
Web	[13, 30, 69]	3
Percent of detection/prevention		
<70	[17, 24]	2
70–80	[46, 56]	2
80–90	[10, 48, 52–54, 60, 63, 64, 68]	9
91–95	[13, 32, 38, 57]	4
96–99	[23, 25, 29, 43, 65, 67]	6

RQ4: What algorithms, methodologies, and data analysis tools are used to detect corruption?

In this question (see Table 6), we evaluated factors like the data source, type, kind of mining, preprocessing methods, outliers values processing, evaluation metrics, artificial intelligence techniques, types and learning techniques, and technological tools.

Table 6. Research question 4

Data source		*f*
Re. Public	[18–24, 26, 30, 41, 43, 45, 46, 54, 61, 80, 86]	17
Re. Private	[11, 13, 25, 26, 38, 46–49, 55–57, 59, 63, 65, 67, 72, 92, 95, 111]	30
Type of data		
Data	[10, 13, 16, 17, 20–23, 25–27, 30–33, 38, 40–46, 48, 50, 53, 55, 57–61, 65, 66, 68, 70, 72, 74, 75, 78–82, 87, 90–92, 95–100, 102, 103, 105, 106, 109]	59
Text	[31, 49, 62, 64]	4
Audio	[54]	1

(*continued*)

Table 6. (*continued*)

Type of minning		
Predictive	[13, 21, 25, 46, 55–57, 59–61, 80, 88, 97]	13
Descriptive	[10, 11, 13, 16, 18–20, 22, 23, 25, 26, 29–32, 37–45, 47–49, 51–54, 63–65, 67–70, 72–75, 80, 86, 91, 92, 95, 96, 111]	51
Preprocessing of data		
Empirical assessment	[19, 25, 39, 41, 45–47, 55, 63, 75, 86]	13
Other	Cascade generalization [70] CKIP [48] Monroe and The [60]	3
Outlier values		
HBOS	[29, 45, 48, 57, 71, 80]	6
PSO	[48, 57, 59, 75]	4
Metric evaluation		
Confusion matrix	[13, 17, 18, 20, 23–25, 29, 32, 43, 49, 53, 60, 68, 80, 86, 111]	17
Curve ROC	[22, 26, 47, 54, 67, 71, 75]	7
Fraud Score	[53, 63]	3
K-fold	[57, 95]	2
Other	Matthews Correlation Coefficient [71]	
Artificial Intelligence technics		
Bayes based	[31, 38, 44, 45, 49, 52–54, 57, 60, 73, 99]	12
Neural network	[18, 19, 22, 37, 41, 47, 52, 68, 88]	9
SVM	[41, 44, 45, 48, 49, 52, 55, 87, 109]	9
Decision tree	[26, 29, 32, 51, 52, 55, 67]	7
Random Forest	[24, 25, 51, 55, 63, 67, 70]	7
Logistic- Linear regression	[14, 34, 41, 44, 45, 49, 51, 52, 67, 83, 87, 97]	12
Other	[11, 20, 36, 42, 44, 48–52, 56, 58, 59, 64, 70, 72, 86, 88]	17
Learning techniques		
Machine learning	[13, 18, 19, 23–26, 29–32, 37, 42, 46, 48–51, 53–57, 59, 64, 65, 67–70, 72–74, 76, 80, 81, 88, 91, 95, 111]	41
Deep learning	[38, 43, 47, 49]	4
Minning techniques		
Clustering	[18, 29, 30, 50, 59, 68, 70, 72, 73, 75, 76, 103]	41
Classification	[13, 18, 24–26, 29, 31, 32, 37, 38, 41, 43, 45–49, 52–54, 75, 76, 81, 95]	51
Regression	[41, 76]	2
Types of learning		
Supervised	[18–20, 25, 26, 31, 32, 37, 38, 42, 54, 55, 57, 68, 91]	15
No supervised	[18, 24, 30, 56, 59, 73]	6
Semi supervised	[29, 70, 71]	3

(*continued*)

Table 6. (*continued*)

Technological tools		
Java	[19, 64, 65, 95]	4
MatLab	[23, 41, 59, 64]	4
Python	[24, 31, 46, 65, 69, 80]	6
Weka	[44, 45, 52]	3
Others	[13, 38, 42, 50, 53, 57, 59, 61, 62, 70, 86, 92, 95]	13

Private datasets predominate in data collection sources 63% (see Table 6); this is because, as mentioned above, most work focuses on the private sector; the type of data used to generate and validate detection models in a large percentage is data (database or dataset). It should be noted that only 6% of jobs use text (documents), and only one job [54], the audio of telephone conversations are used to find fraud in public purchases.

The review focuses on the prevention and detection of corruption, with 79% of investigations of a descriptive type (detection) and 21% of a predictive type (prevention); in the techniques for the pre-processing of data, empirical assessment predominates with 81%, and amongst the methods are Cascade generalization [70] Chinese Knowledge Information Processing Group (CKIP) [48] Monroe and The [60]. In order to detect outliers in the data, the following method is used in similar percentages Particle Swarm Optimization (PSO) and the histogram-based outlier score (HBOS).

It is observed that the main techniques of artificial intelligence are those based on the theorem of Bayes, neural networks, Support Vector Machines (SVM), decision trees, Random Forest, and logistic and linear regression, reaching 76% among all of them; to a lesser extent technique such as: convolutional networks, tough set theory, graphs, natural language processing, kmeans, AdaBoost, genetic algorithms, bagging and logitBoos. The investigation with the highest number of applied techniques is [49], evaluating eleven artificial intelligence techniques to detect price manipulation in purchases.

The main evaluation metrics are: accuracy, efficiency, recalling using methods such as confusion matrix and ROC curves; in assessing of corruption, the significant contribution of the fraud score used as an evaluation indicator in some works, the confusion matrix is combined with the fraud score [53], and machine learning evaluation methods such as the Matthews Correlation Coefficient [71].

Concerning computer tools for programming, processing, and data storage Java, MatLab, Weka y Python have the highest percentage, with other tools such as R, RapidMiner, Hadoop, Spark, Neo4j, Casandra, Kafta, Visual Studio.

4 Conclusions and Future Work

Experimental quantitative research is the most widely used, as well as the statistical method of correlation. Most of the study is conducted with data over 1–3 years, with a significant contribution to computer science. The main commercial activity of the organizations is the provision of services, specifically the banking sector. In the government sphere, tax fraud is noteworthy, with a lesser presence of public procurement processes.

Web Scraping is a rarely used technique for obtaining data on corruption studies in contracts and can be used as a basis for future work. The few jobs related to contract analysis in public procurement use datasets and are not considered documents as the initial basis for data analysis. It is also evident that the computer tools created to carry out corruption analysis in contracts in both the public and private sectors are not considered computer security standards, and the percentage of tools in the web environment is very low.

The main artificial intelligence techniques found are logistic models, neural networks, bayesian networks, and supported vector machines. As future work, they would be enhanced with mixed learning methods. Fraud Score is proposed as a specific metric to assess the risk of corruption without leaving aside the metrics used to evaluate from the confusion matrix and ROC curves, with machine learning and supervised learning as the main types of technique.

References

1. Báez Gómez, J.E.: Relación entre el Índice de Control de la Corrupción y algunas variables sociales, económicas e institucionales. Nómadas. Rev. Crítica Ciencias Soc. y Jurídicas **38** (2013)
2. Volosin, N.A.: Datos abiertos, corrupción y compras públicas (2015)
3. Padhi, S.S., Mohapatra, P.K.J.: Detection of collusion in government procurement auctions. J. Purch. Supply Manag. **17**, 207–221 (2011)
4. Martin, R.: A review of the literature of the followership since 2008: the importance of relationships and emotional intelligence. SAGE Open **5**, 2158244015608421 (2015)
5. Adewumi, A.O., Akinyelu, A.A.: A survey of machine-learning and nature-inspired based credit card fraud detection techniques. Int. J. Syst. Assur. Eng. Manag. **8**, 937–953 (2017)
6. Ngai, E.W.T., Hu, Y., Wong, Y.H., Chen, Y., Sun, X.: The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature. Decis. Support Syst. **50**, 559–569 (2011)
7. Torres-Carrion, P.V., Gonzalez-Gonzalez, C.S., Aciar, S., Rodriguez-Morales, G.: Methodology for systematic literature review applied to engineering and education. IEEE Global Engineering Education Conference EDUCON, pp. 1364–1373, April 2018
8. Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering – a systematic literature review. Inf. Softw. Technol. **51**, 7–15 (2009)
9. Moher, D., et al.: PRISMA-P: preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P). Syst. Rev. 1–9 (2015)
10. Auriol, E., Straub, S., Flochel, T.: Public procurement and rent-seeking: the case of paraguay. World Dev. **77**, 395–407 (2016)
11. Lei, M., Yin, Z., Li, S., Li, H.: Detecting the collusive bidding behavior in below average bid auction. In: ICNC-FSKD 2017 - 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, pp. 1720–1727 (2018)
12. Hernández Sampieri, R., Fernández Collado, C., Baptista Lucio, M.: Metodología de la Investigación (2010)

13. Kose, I., Gokturk, M., Kilic, K.: An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance. Appl. Soft Comput. J. **36**, 283–299 (2015)
14. Charles Andoh, E., Ofosu-Hene, D.: Causes, effects and deterrence of insurance fraud: evidence from Ghana. J. Financ. Crime Iss. **5**, 39–44 (2016)
15. Huang, S.Y., Lin, C.C., Chiu, A.A., Yen, D.C.: Fraud detection using fraud triangle risk factors. Inf. Syst. Front. **19**, 1343–1356 (2017)
16. Seck, A.: Heterogeneous bribe payments and firms' performance in developing countries. J. African Bus. **21**, 1–20 (2019)
17. Baader, G., Krcmar, H.: Reducing false positives in fraud detection: combining the red flag approach with process mining. Int. J. Account. Inf. Syst. **31**, 1–16 (2018)
18. Choi, D., Lee, K.: An artificial intelligence approach to financial fraud detection under IoT environment: a survey and implementation. Secur. Commun. Netw. **2018** (2018)
19. Sadaoui, S., Wang, X.: A dynamic stage-based fraud monitoring framework of multiple live auctions. Appl. Intell. **46**, 197–213 (2017)
20. Yeh, C.C., Chi, D.J., Lin, T.Y., Chiu, S.H.: A hybrid detecting fraudulent financial statements model using rough set theory and support vector machines. Cybern. Syst. **47**, 261–276 (2016)
21. Ouenniche, J., Uvalle Perez, O.J., Ettouhami, A.: A new EDAS-based in-sample-out-of-sample classifier for risk-class prediction. Manag. Decis. **57**, 314–323 (2019)
22. Zakaryazad, A., Duman, E.: A profit-driven Artificial Neural Network (ANN) with applications to fraud detection and direct marketing. Neurocomputing **175**, 121–131 (2014)
23. Darwish, S.M.: An intelligent credit card fraud detection approach based on semantic fusion of two classifiers. Soft. Comput. **24**, 1243–1253 (2019)
24. Kehler, E., Paciello, J., Pane, J.: Anomaly detection in public procurements using the open contracting data standard (2019)
25. Van Vlasselaer, V., et al.: APATE: a novel approach for automated credit card transaction fraud detection using network-based extensions. Decis. Support Syst. **75**, 38–48 (2015)
26. Zareapoor, M., Shamsolmoali, P.: Application of credit card fraud detection: based on bagging ensemble classifier. Procedia Comput. Sci. **48**, 679–685 (2015)
27. Ngoc, B.H., Hai, D.B., Chinh, T.H.: Assessment of the should be effects of corruption perception index on foreign direct investment in ASEAN countries by spatial regression method. In: Anh, Ly H., Dong, L.S., Kreinovich, V., Thach, N.N. (eds.) ECONVN 2018. SCI, vol. 760, pp. 421–429. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-73150-6_33
28. Burböck, B., Macek, A., Podhovnik, E., Zirgoi, C.: Asymmetric influence of corruption distance on FDI. J. Financ. Crime (2018)
29. Carminati, M., Caron, R., Maggi, F., Epifani, I., Zanero, S.: BankSealer: a decision support system for online banking fraud analysis and investigation. Comput. Secur. **53**, 175–186 (2015)
30. Dhurandhar, A., Graves, B., Ravi, R., Maniachari, G., Ettl, M.: Big data system for analyzing risky procurement entities. In: Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1741–1750 August 2015
31. Hooi, B., et al.: BIRDNEST: Bayesian inference for ratings-fraud detection. In: 16th SIAM International Conference on Data Mining 2016, SDM 2016, pp. 495–503 (2016)
32. Snyder, P., Kanich, C.: Characterizing fraud and its ramifications in affiliate marketing networks. J. Cybersecur. **2**, 71–81 (2016)

33. Moalosi, M., Hlomani, H., Phefo, O.S.D.: Combating credit card fraud with online behavioural targeting and device fingerprinting. Int. J. Electron. Secur. Digit, Forensics (2019)
34. Anh, N.N., Minh, N.N., Tran-Nam, B.: Corruption and economic growth, with a focus on Vietnam. Crime, Law Soc. Change **45**, 307–324 (2016)
35. Ferwerda, J., Deleanu, I., Unger, B.: Corruption in public procurement: finding the right indicators. Eur. J. Crim. Policy Res. **23**, 245–267 (2017)
36. Amanze, B.C., Onukwugha, C.G.: Credit card fraud detection system in nigeria banks using adaptive data mining and intelligent agents: A review. Int. J. Sci. Technol. Res. **7**, 175–184 (2018)
37. Zanin, M., Romance, M., Moral, S., Criado, R.: Credit card fraud detection through parenclitic network analysis. Complexity **2018** (2018)
38. Randhawa, K., Loo, C.K., Seera, M., Lim, C.P., Nandi, A.K.: Credit card fraud detection using AdaBoost and majority voting. IEEE Access **6**, 14277–14284 (2018)
39. Ausloos, M., Cerqueti, R., Mir, T.A.: Data science for assessing possible tax income manipulation: the case of Italy. Chaos, Solitons Fractals **104**, 238–256 (2017)
40. Helmy, T.H., Zaki, M., Salah, T., Badran, K.: Design of a monitor for detecting money laundering and terrorist financing. J. Theor. Appl. Inf. Technol. **85**, 425–436 (2016)
41. Rahimikia, E., Mohammadi, S., Rahmani, T., Ghazanfari, M.: Detecting corporate tax evasion using a hybrid intelligent system: a case study of Iran. Int. J. Account. Inf. Syst. **25**, 1–17 (2017)
42. Van Erven, G.C.G., Carvalho, R.N., De Holanda, M.T., Ralha, C.: Graph database: a case study for detecting fraud in acquisition of Brazilian Government (Banco de Dados em Grafo: Um Estudo de Caso em Detecção de Fraudes no Governo Brasileiro). In: Iberian Conference on Information Systems and Technologies CISTI, pp. 1–6 (2017)
43. Wacker, J., Ferreira, R.P., Ladeira, M.: Detecting fake suppliers using deep image features (2018)
44. Kim, Y.J., Baik, B., Cho, S.: Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. Expert Syst. Appl. **62**, 32–43 (2016)
45. Dutta, I., Dutta, S., Raahemi, B.: Detecting financial restatements using data mining techniques. Expert Syst. Appl. **90**, 374–393 (2017)
46. Grace, E., Rai, A., Redmiles, E., Ghani, R.: Detecting fraud, corruption, and collusion in international development contracts: the design of a proof-of-concept automated system (2016)
47. Gómez, J.A., Arévalo, J., Paredes, R., Nin, J.: End-to-end neural network architecture for fraud scoring in card payments. Pattern Recognit. Lett. **105**, 175–181 (2018)
48. Chen, Y.J., Wu, C.H., Chen, Y.M., Li, H.Y., Chen, H.K.: Enhancement of fraud detection for narratives in annual reports. Int. J. Account. Inf. Syst. **26**, 32–45 (2017)
49. Wang, Q., Xu, W., Huang, X., Yang, K.: Enhancing intraday stock price manipulation detection by leveraging recurrent neural networks with ensemble learning. Neurocomputing **347**, 46–58 (2019)
50. Tan, M., Lee, W.-L.: Evaluation and improvement of procurement process with data analytics. Int. J. Adv. Comput. Sci. Appl. **6**, 70–80 (2015)
51. Correa Bahnsen, A., Aouada, D., Stojanovic, A., Ottersten, B.: Feature engineering strategies for credit card fraud detection. Expert Syst. Appl. **51**, 134–142 (2016)
52. Li, H., Wong, M.-L.: Financial fraud detection by using grammar-based multi-objective genetic programming with ensemble learning (2015)
53. Throckmorton, C.S., Mayew, W.J., Venkatachalam, M., Collins, L.M.: Financial fraud detection using vocal, linguistic and financial cues. Decis. Support Syst. **74**, 78–87 (2015)

54. Arief, H.A.A., Saptawati, G.A.P., Asnar, Y.D.W.: Fraud detection based-on data mining on Indonesian E-Procurement System (SPSE). In: Proceedings of 2016 International Conference on Data and Software Engineering, ICoDSE 2016 (2017)
55. Vimala Devi, J., Kavitha, K.S.: Fraud detection in credit card transactions by using classification algorithms. In: International Conference on Current Trends in Computer, Electrical, Electronics and Communication, CTCEEC 2017, pp. 125–131 (2018)
56. Zhou, H., Chai, H.F., Qiu, M.L.: Fraud detection within bankcard enrollment on mobile device based payment using machine learning. Front. Inf. Technol. Electron. Eng. **19**, 1537–1545 (2018)
57. Hooda, N., Bawa, S., Rana, P.S.: Fraudulent firm classification: a case study of an external audit. Appl. Artif. Intell. **32**, 48–64 (2018)
58. Fu, Y., Liu, G., Papadimitriou, S., Xiong, H., Li, X., Chen, G.: Fused latent models for assessing product return propensity in online commerce. Decis. Support Syst. **91**, 77–88 (2016)
59. Demiriz, A., Ekizoğlu, B.: Fuzzy rule-based analysis of spatio-temporal ATM usage data for fraud detection and prevention1. J. Intell. Fuzzy Syst. **31**, 805–813 (2016)
60. Chimonaki, C., Papadakis, S., Vergos, K., Shahgholian, A.: Identification of financial statement fraud in Greece by using computational intelligence techniques. In: Mehandjiev, N., Saadouni, B. (eds.) FinanceCom 2018. LNBIP, vol. 345, pp. 39–51. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-19037-8_3
61. Correa, M.A.O.S., Galindo Leal, A.: Identification of overpricing in the purchase of medication by the Federal Government of Brazil, using text mining and clustering based on ontoiogy (2018)
62. Alzaidi, A.A.: Impact of use of big data in decision making in banking sector of Saudi Arabia. Int. J. Comput. Sci. Netw. Secur. **18**, 72–80 (2018)
63. Kasa, N., Dahbura, A., Ravoori, C., Adams, S.: Improving credit card fraud detection by profiling and clustering accounts (2019)
64. Chen, Y.-J., Wu, C.-H.: On big data-based fraud detection method for financial statements of business groups (2017)
65. Weng, H., et al.: Online e-commerce fraud: a large-scale detection and analysis (2018)
66. Torres, C.F., Schütte, J., State, R.: Osiris: hunting for integer bugs in ethereum smart contracts (2018)
67. Lismont, J.: Predicting tax avoidance by means of social network analytics. Decis. Support Syst. **108**, 13–24 (2018)
68. Zhang, H., Wang, L.: Prescription fraud detection through statistic modeling (2018)
69. Martínez-Plumed, F., Casamayor, J.C., Ferri, C., Gómez, J.A., Vendrell Vidal, E.: SALER: a data science solution to detect and prevent corruption in public administration. In: Alzate, C., et al. (eds.) ECML PKDD 2018. LNCS (LNAI), vol. 11329, pp. 103–117. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-13453-2_9
70. Carcillo, F., Dal Pozzolo, A., Le Borgne, Y.A., Caelen, O., Mazzer, Y., Bontempi, G.: SCARFF: a scalable framework for streaming credit card fraud detection with spark. Inf. Fusion. **41**, 182–194 (2018)
71. Carminati, M., Polino, M., Continella, A., Lanzi, A., Maggi, F., Zanero, S.: Security evaluation of a banking fraud analysis system. ACM Trans. Priv. Secur. **21**, 1–31 (2018)
72. Robinson, W.N., Aria, A.: Sequential fraud detection for prepaid cards using hidden Markov model divergence. Expert Syst. Appl. **91**, 235–251 (2018)
73. Ekin, T., Ieva, F., Ruggeri, F., Soyer, R.: Statistical medical fraud assessment: exposition to an emerging field. Int. Stat. Rev. **86**, 379–402 (2018)
74. Fauzan, A.C., Sarno, R., Ariyani, N.F.: Structure-based ontology matching of business process model for fraud detection. In: ICTS 2017, pp. 221–225 (2018)

75. Saghehei, E., Memariani, A.: Suspicious behavior detection in debit card transactions using data mining: a comparative study using hybrid models. Inf. Resour. Manag. J. **28**, 1–14 (2015)
76. El-kaime, H., Hanoune, M., Eddaoui, A.: The data mining: a solution for credit card fraud detection in banking. In: Mizera-Pietraszko, J., Pichappan, P., Mohamed, L. (eds.) RTIS 2017. AISC, vol. 756, pp. 332–341. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-91337-7_31
77. Schlenther, B.O.: Addressing illicit financial flows in Africa: how broad is the whole of government approach supposed to be? J. Financ. Crime (2016)
78. Sadaf, R., Oláh, J., Popp, J., Máté, D.: An investigation of the influence of theworldwide governance and competitiveness on accounting fraud cases: a cross-country perspective. Sustain **10**, 1–11 (2018)
79. Wang, H., Chen, H.M.: Deterring bidder collusion: auction design complements antitrust policy. J. Compet. Law Econ. **12**, 31–68 (2016)
80. Wahid, A., Rao, A.C.S.: A distance-based outlier detection using particle swarm optimization technique. In: Fong, S., Akashe, S., Mahalle, Parikshit N. (eds.) Information and Communication Technology for Competitive Strategies. LNNS, vol. 40, pp. 633–643. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-0586-3_62
81. Coma-Puig, B., Carmona, J.: A quality control method for fraud detection on utility customers without an active contract. In: Proceedings of the 33rd Annual ACM Symposium on Applied Computing, pp. 495–498 (2018)
82. Fazekas, M., Cingolani, L.: Breaking the cycle? How (not) to use political finance regulations to counter public procurement corruption. Slav. East Eur. Rev. **95**, 76–116 (2017)
83. Lehne, J., Shapiro, J.N., Vanden Eynde, O.: Building connections: political corruption and road construction in India. J. Dev. Econ. **131**, 62–78 (2018)
84. Cieślik, A., Goczek, Ł.: Control of corruption, international investment, and economic growth – Evidence from panel data. World Dev. **103**, 323–335 (2018)
85. Lourenço, I.C., Rathke, A., Santana, V., Branco, M.C.: Corruption and earnings management in developed and emerging countries. Corp. Gov. **18**, 35–51 (2018)
86. van Erven, G.C.G., Holanda, M., Carvalho, Rommel N.: Detecting evidence of fraud in the brazilian government using graph databases. In: Rocha, Á., Correia, A.M., Adeli, H., Reis, L.P., Costanzo, S. (eds.) WorldCIST 2017. AISC, vol. 570, pp. 464–473. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56538-5_47
87. Rad, M.S., Shahbahrami, A.: Detecting high risk taxpayers using data mining techniques. In: 2016 2nd International Conference of Signal Processing and Intelligent Systems ICSPIS 2016, pp. 14–15 (2017)
88. Monirzadeh, Z., Habibzadeh, M., Farajian, N.: Detection of violations in Credit Cards of Banks and financial institutions based on artificial neural network and Metaheuristic optimization algorithm. Int. J. Adv. Comput. Sci. Appl. **9**, 176–182 (2018)
89. Bramoullé, Y., Goyal, S.: Favoritism. J. Dev. Econ. **122**, 16–27 (2016)
90. Saxena, A., Sharma, N., Saxena, K., Parikh, Satyen M.: Financial data mining: appropriate selection of tools, techniques and algorithms. In: Deshpande, A.V., Unal, A., Passi, K., Singh, D., Nayak, M., Patel, B., Pathan, S. (eds.) SmartCom 2017. CCIS, vol. 876, pp. 244–251. Springer, Singapore (2018). https://doi.org/10.1007/978-981-13-1423-0_27
91. Bogdanov, D., Jõemets, M., Siim, S., Vaht, M.: How the Estonian tax and customs board evaluated a tax fraud detection system based on secure multi-party computation. In: Böhme, R., Okamoto, T. (eds.) FC 2015. LNCS, vol. 8975, pp. 227–234. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-47854-7_14

92. Kültür, Y., Çağlayan, M.U.: Hybrid approaches for detecting credit card fraud. Expert Syst. **34**, 1–13 (2017)
93. Indrajani, Prabowo, H., Meyliana: Learning fraud detection from big data in online banking transactions: a systematic literature review. J. Telecommun. Electron. Comput. Eng. **8**, 127–131 (2016)
94. Hutchings, A.: Leaving on a jet plane: the trade in fraudulently obtained airline tickets. Crime Law Soc. Change **70**, 461–487 (2018)
95. Saia, R., Boratto, L., Carta, S.: Multiple behavioral models: a divide and conquer strategy to fraud detection in financial data streams (2015)
96. Lee, P.S., Owda, M., Crockett, K.: Novel methods for resolving false positives during the detection of fraudulent activities on stock market financial discussion boards. Int. J. Adv. Comput. Sci. Appl. **9**, 1–10 (2018)
97. Fazekas, M.: Red tape, bribery and government favouritism: evidence from Europe. Crime Law Soc. Change **68**, 403–429 (2017)
98. Yaseen, M., et al.: Secure sensors data acquisition and communication protection in eHealthcare: review on the state of the art. Telemat. Inform. **35**, 702–726 (2018)
99. Jetter, M., Parmeter, C.F.: Sorting through global corruption determinants: institutions and education matter – Not culture. World Dev. **109**, 279–294 (2018)
100. Jagger, P., Shively, G.: Taxes and Bribes in Uganda. J. Dev. Stud. **51**, 66–79 (2015)
101. Williams, M.J.: The political economy of unfinished development projects: corruption, clientelism, or collective choice? Am. Polit. Sci. Rev. **114**, 705–723 (2017)
102. Kussainov, D.S.: The problems of qualification of illegal alienation of ownership of residential premises. Asian Soc. Sci. **11**, 188 (2015)
103. Ahmed, M., Mahmood, A.N., Islam, M.R.: A survey of anomaly detection techniques in financial domain. Future Gener. Comput. Syst. **55**, 278–288 (2016)
104. Moro, S., Cortez, P., Rita, P.: Business intelligence in banking: a literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. Expert Syst. Appl. **42**, 1314–1324 (2015)
105. Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., Bontempi, G.: Credit card fraud detection: a realistic modeling and a novel learning strategy. IEEE Trans. Neural Netw. Learn. Syst. **29**, 3784–3797 (2018)
106. Kumar, P., Iqbal, F.: Credit card fraud identification using machine learning approaches (2019)
107. Mahmoudi, N., Duman, E.: Detecting credit card fraud by modified fisher discriminant analysis. Expert Syst. Appl. **42**, 2510–2516 (2015)
108. Abdallah, A., Maarof, M.A., Zainal, A.: Fraud detection system: a survey. J. Netw. Comput. Appl. **68**, 90–113 (2016)
109. Rajak, I., Mathai, K.J.: Intelligent fraudulent detection system based SVM and optimized by danger theory. In: IEEE International Conference on Computer, Communication and Control IC4 2015, pp. 2–5 (2016)
110. Xu, J.J., Lu, Y., Chau, M.: P2P lending fraud detection: a big data approach. In: Chau, M., Wang, G.A., Chen, H. (eds.) PAISI 2015. LNCS, vol. 9074, pp. 71–81. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18455-5_5
111. Hajek, P., Henriques, R.: Mining corporate annual reports for intelligent detection of financial statement fraud – a comparative study of machine learning methods. Knowl.-Based Syst. **128**, 139–152 (2017)