Systematic Literature Review

Fraud Prevention with Cluster Analysis

Agung Budiwibowo(✉), Endang Siti Astuti, Muhammad Saifi, and Mohammad Iqbal

Faculty of Administrative Science, University of Brawijaya, Malang, Indonesia
agungbw.jp@gmail.com

Abstract. Fraud is a theme that continues to be the focus of researchers in the world because of changes in a person's character to commit fraud due to developments in technology and knowledge. The main focus is, among others, the existence of a corporate crisis due to the declining financial performance of the company due to fraud. In this study, the aim of this study is to classify studies that discuss fraud in companies and classify companies based on the company's fraud data. The literature used has the topic of fraud using cluster analysis. The results of the study show that the cluster clustering technique used by previous researchers to classify fraud tends to use clustering based on mature methods, where 41% of the clustering techniques encountered are partitional clusters. It is still difficult to find relatively new groupings such as ensemble grouping, large-scale grouping, multi-way grouping. Among the partition clustering techniques, k-means, c-means and their variants with Euclidian distance as the dissimilarity metric is the most common and popular to use. The hierarchical clustering technique was in second place used in a quarter of the papers surveyed. Interactive, clustering visualization techniques are also used but only in a small number of cases. Furthermore, the results of grouping companies using the Ward Linkage method with Manhattan Distance show that two clusters are formed, each with 8 and 24 members. Cluster 1 is the highest because most cluster 1 has the largest average value of the indicator compared to the others.

Keywords: Cluster Analysis · Euclidean · Fraud · Literature Review

1 Introduction

The existence of fraud can compromise a country's economic viability. Based on the 2018 Association of Certified Fraud Examiners (ACFE) report, this indicates that the organization's losses due to fraud are approximately 5% of the organization's total revenue. Most fraud is hidden and beyond being detected, most cases of fraud are never measured or reported, because well-known companies are afraid of damaging the company's reputation. On the other hand, most frauds carry considerable costs, which is why it is important to stay educated about the warning signs of job fraud or employee fraud, as well as the importance of having good controls. Every entity is exposed to fraud risk, but the impact of this threat can be reduced with good anti-fraud strategies and controls [1].

© The Author(s) 2023
A. F. Wijaya et al. (Eds.): AICoBPA 2022, ASSEHR 764, pp. 29–41, 2023.
https://doi.org/10.2991/978-2-38476-090-9_4

This issue has piqued the interest of shareholders and other stakeholders because of the negative consequences such as lowering the company's performance, tarnishing the image or reputation of the company, and most importantly causing monetary or asset loss. Low productivity, low job satisfaction, and high employee turnover all contribute to poor company performance, which results in unnecessary costs and lower long-term profits [2]. Company performance is one of the most important elements in the business world, because good or bad in company performance is an important indicator of the company's success. Company performance is indeed a complex thing, but if managed properly it will be very valuable. Companies with strong positive performance will attract consumers to use their products more than companies that have poor performance. Customers are more loyal and buy more products and services because the market believes that such companies will provide sustainable revenue and growth in the future, companies are believed to have higher price-earnings and market value multiples and lower costs of capital [3].

The Oxford English Dictionary defines fraud as a criminal act of deception by presenting untrue financial statements or other information with the aim of gaining an unfair advantage or imposing on the rights and interests of others. Meanwhile, Sorunke et al. [4] view fraud as a deliberate act of deception or concealing omissions or deviations from the truth such as breaking the law and acting unfairly.

There's many descriptions and interpretations of fraud. The Institute of Internal Auditors (IIA) defines fraud as "an array of irregularities and illegal acts characterized by intentional deception" which can be interpreted as a collection of illegal and illegal actions characterized by the presence of an instrument of intentional fraud. The expanded definition of fraud according to The Association of Certified Fraud Examiners (ACFE), fraud is any attempt to discover or manipulate other stakeholders in order to gain benefits.

Fraud or fraud can be said to be very detrimental to the victim when the value is material (large) and the value affects will cause significant differences in the information presented in the financial statements. The high level of materiality will be a gap for companies to commit fraud by changing or manipulating information in the form of posts or items of financial and non-financial information content.

A financial report is a document released by a company that describes its profits, expenses, income, and loans [5]. Financial statement fraud, a type of corporate fraud, includes falsifying reports in order to make the company occur more advantageous. Fraudulent financial statements are committed for a variety of reasons, including increasing asset performance, reduced tax obligations, or an aim to overestimate performance as a outcome of managerial force [5]. Fraud against financial statements is difficult to detect due to the general lack of comprehension in the field, its rarity, and the fact that it is typically perpetrated by industry experts who are able to cover up their fraud [6].

One way to detect financial fraud mentioned above can be done with clustering techniques, both independent clustering techniques and combined clustering techniques, with classification techniques, especially neural networks and decision trees. Self-clustering methods are unsupervised data mining techniques, whereas combinatorial techniques are semi-supervised data mining techniques. The scientific literature classification systems and provides various classifications for classifying fraud. This study focuses on conducting an in-depth review of previous research that examines fraud; particularly fraud detection through the clustering method; and grouping companies based on fraud data in each company.

2 Literature Review

2.1 Fraud Theory

The Oxford English Dictionary defines fraud as a criminal act of deception by using falsification of financial statements and other information to gain unfair advantage or take by force the rights or interests of others. Meanwhile, Sorunke et al. [4] view fraud as an act or process of deception and deliberate concealment of omissions and deviations from the truth such as breaking the law and acting unfairly. There are lot of definitions and interpretations of fraud. Fraud is defined by The Institute of Internal Auditors (IIA) as "a variety of irregularities and illegal acts characterized by deliberate intent deceptionwhich can be defined as a collection of criminal and illegal actions characterized by the presence of an element of intentional fraud. The Association of Certified Fraud Examiners (ACFE) provides a more specific definition of fraud as "any attempt to discover or deceive other parties in order to obtain benefits."

Corporate fraud is a perplexing and repetitive problem. This problem continues to surround every company, industry and country of all sizes [7]. Fraud problems can include financial and non-financial which have been confirmed by various studies such as Goldman et al. [8] and Kuvvet [9] stated that are 6 well-known organizations that have lost 5% of their revenue each year due to fraud and have an estimated global loss of USD 3.7 trillion in Gross World Product (GWP). Similarly, in 2014 one in three such organizations reported fraud amounting to 22% of PWC's enterprise value. The social and economic consequences make stakeholders likely to emphasize examining factors that contribute to financial misreporting [10]. Equity-based compensation is one of these factors [11, 12], manager's reporting objectives are uncertain. According to Fischer and Verrecchia [13] and activity monitoring according to Povel et al. [14]. In addition, other empirical literatures also identify various internal and external factors that exacerbate corporate fraud. Intrinsic factors involve studies that show manager salaries and compensation are essential factors that motivate managers to manipulate financial reporting [15].

These findings support the current theoretical argument that executive compensation influences fraud tendencies [11, 12]. In addition, there are also studies that find corporate governance system characteristics such as board composition and expertise according to Beasley [16] and Klein [17], executive roles and social ties with board members and other executives according to Chidambaran et al. [18] and Khanna et al. [19] as a significant contributor to fraud. According to Wang et al. [20], according to Agrawal and Cooper [21], the research identifies external factors based on the external systems and channels, also including business conditions, the characteristics of the industry, as well as the role of regulators Although the existing theoretical and empirical literature tends to explain the factors that provide a significant understanding of the mechanisms of fraud, these factors are considered to have been neglected in comprehensive models. Previous studies have mostly concentrated on a number of factors that have the potential to increase fraudulent behavior.

Fraud or fraud can be said to be very detrimental to the victim when the value is material (large) and the value affects will cause significant differences in the information presented in the financial statements. The high level of materiality will be a gap for companies to commit fraud by changing or manipulating information in the form of posts or items of financial and non-financial information content.

A financial report is a document issued by a company that explains details such as their profits, income, expenses and loans [5]. Financial statement fraud, which is also part of corporate fraud, involves falsifying reports to make the organization appear more valuable. The reasons for committing fraudulent financial statements include increasing stock performance, reducing tax obligations, or as an effort to overestimate performance due to managerial pressure [5]. Financial statements fraud could be difficult to recognize due to the general lack of expertise in the field, its rarity, and the fact that it is typically perpetrated by knowledgeable industry professionals who are able to conceal their deception [6].

The 2016 National Report on Occupation Fraud and Abuse issued by the Association of Certified Fraud Examiners (ACFE) revealed that there were 78.9% of fraud cases that included corruption, misappropriation of assets, and falsification of financial statements, both in the public and private sectors. In addition, according to data from the Association of Certified Fraud Examiners (ACFE), fraud caused by fraudulent financial statements in 2016 increased by 9.6% from 9.0% in 2014. Previously, the Enron case that occurred in 2002 in the United States, which committed fraud in its financial statements, shocked the financial world. The company commits fraud by increasing the company's profits when in fact the company suffers a loss. As a result of the fraudulent financial statements at the Enron company, the economy and investor confidence declined and even affected other companies on the stock exchange.

3 Methodology

From a research definition perspective, a literature review on financial fraud identification using data mining clustering methods is represented. This timeline includes relevant recent research momentum where data mining techniques were developed as a fraud detection and prevention tool. Several criteria for searching and selecting articles have been established with an article classification framework as part of the research method. The stages in this literature review are as follows:

- In the early stages, Thomson Reuters Web of Science, IEEE Transactions, ScienceDirect Freedom Collection and Springer-Link Contemporary were searched for the keywords "fraud" and "clustering" in the article theme area.
- In the second stage, this search was carried out on September 2, 2022 utilizing keywords inside the form of search strategy relating to spiritual leadership, where keywords are investigated based on the article's title, keywords, and abstract. Based on predetermined keywords, the technique of searching for papers in online databases using predefined keywords is utilized as a sources of information. Scopus was used as the online databases in this study because it is the biggest credible scientific database currently available and contains a variety of peer-reviewed journal articles [22]. Thus, the quality obtained can be guaranteed.

- Furthermore, groupings linked to each major form of financial fraud were searched on Google Scholar. Based on the search, the first 100 entries were taken to be considered for inclusion, along with relevant bibliographic entries. The search expression contains the keywords "classification" and "fraud" combined with one of the following "credit cards", "money laundering", "insurance", and "corporate". No direct searches for "bank fraud" were made as this area of fraud is well covered by the keywords "credit card" and "money laundering".

At the cluster analysis stage, the following steps are carried out:

- Performing Manhattan distance calculations to determine the distance between research objects.
- Performing cluster analysis using Ward Linkage method which aims to minimize variance in the cluster.
- Formation of dendogram results of cluster analysis.

4 Results and Discussion

4.1 Study Literature Review

Based on the research conducted, some information was obtained, the results are presented in the form of figures and tables. In Fig. 1 below, we show information related to the journals that discuss a lot about clusters in the field of fraud.

The Fig. 1 above illustrates the number of research documents adjusted for importance to the themes of "Fraud" and "Cluster" published by each journal. The data lists the names of the most published journals and the interval of the number of documents published in a blue bar chart. The blue color indicates the quantity and relevance of the research theme, with the number of documents published by all journals ranging from 0 to 10.

Ceur Workshop Proceeding, Cluster Computing, IEEE Access, Lecture Noets in Networks and Systems, Neural Network World, Proceedings of The International Joint

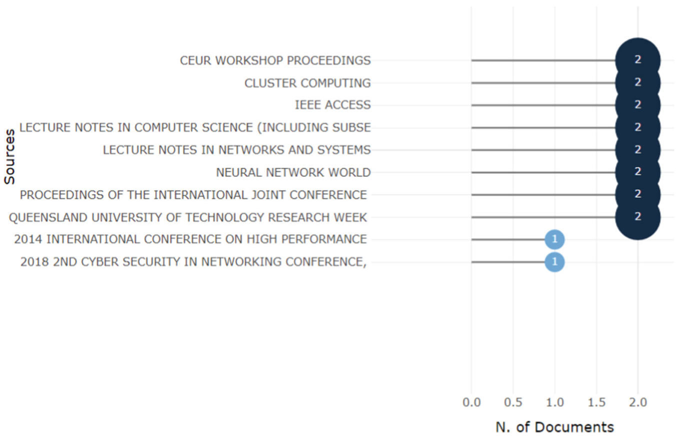

Fig. 1. Cluster in the field of fraud.

Conference, Queensland University of Technology Research Week, International Conference on High Performance 2014, 2nd Cyber Security in Networking Conference 2018 is a medium that is in first place, with two published documents shown in a dark blue bar graph in comparison to other journal entries. This is because the journal is relevant to the theme discussed.

This study also discusses the development of journals that are the source of research information on the theme of Cluster Fraud. The curve in Fig. 2 above shows the development of each journal from 2000 to 2022, to get an idea of whether the journal has increased or decreased with such a curve line during the study years, especially in publishing manuscripts on the topic of Cluster Fraud. The graph illustrates that research on the theme of Cluster Fraud tends to experience fluctuating growth in its publication.

The Fig. 3 below shows the number of research documents that published by each author based on the level of relevance to fraud cluster theme. A blue bar chart displays a list of the most commonly published authors' names and the number of documents issued in intervals. The darker the blue, the greater the amount and contribution of the theme research; the total amount of papers published by all journals ranged from 0 to 3.0. The 10 authors are listed on the most pertinent data sources. Author Benchaji I is the author who is in the top position with the number of published documents is 3.0 documents displayed in the graph bars are dark blue compared to other journal bars. This is significant because the journal is relevant to the topic under discussion.

The research methodology resulted in the selection of 16 articles for inclusion. They have been classified according to application domain, clustering technique, and case study dataset. Papers ordered by year of publication and grouping techniques are in Table 1.

The papers in the table use various methods of grouping (cluster). Namely hierarchical and partitional grouping. Several papers employ a stand-alone single clustering

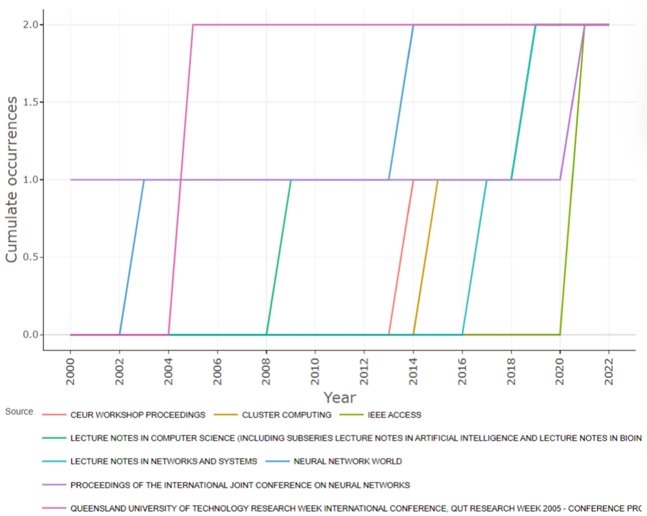

Fig. 2. List of the names of the top published journals.

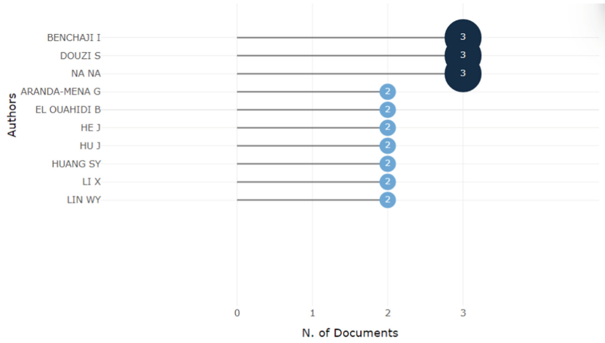

Fig. 3. Number of research documents that published by each author.

technique which is used as the only data mining method (1), (4), (7), (10), (11), (14), (16). On the other hand, we find data mining techniques implemented in complex data mining (9), (12). Also present is a comparison technique of 2 different clustering methods (6), (8), (13).

Most of the cases where single and independent clustering methods are used utilize k-means, c-means and their variations for outlier detection. Many cases, Euclidean distance is employed as a dissimilarity metric. Research conducted by Issa et al. [26] implemented k-means with the intention of identifying counterfeit money returns in telecommunications companies with fraudulent transactions considered as outliers. Whereas a recent study conducted by Li [38] used c-means to automate fraud screening during audits. Claims with comparable characteristics were small population groups and grouped together were flagged for further investigation.

The hierarchical clustering technique forms another case group. Murcia et al. [24] used the hierarchical method. The auditors in the sample are members of the Instituto dos Auditores Independentes do Brasil (IBRACON). This study included the participation of 33 auditors. According to the findings, 95.56% of red flags presented during the detection of fraudulent financial reporting were ignored a “intermediate risk” or “high risk.” On a 1–5 scale, all six clusters had an average fraud risk of 3.35 or higher. Other researchers, for example Yazid and Fiananta [35] use the *support vector machine* (SVM) techniques to detect fraud based on *outliers*/anomalies in transaction data. The sample data is 100 data, using the account number attribute as a label, month, and transaction nominal as training data. Simulating transactions on the form with the account number, month, and transaction nominal is used for testing. The system will determine whether or not the transaction with the account number belongs to the class specified in the model. The transaction is paused if you leave the class.

Hierarchical agglomerative clustering was used with stable endpoints for all clustering tests resulting in two stable clusters. Anowar and Sadaoui [37] using hierarchical agglomerative clustering as the outlier ranking, they construct the actual training data set, but the data is unlabeled. First, they appropriately label a sample of shill offerings by combining a strong hierarchical clustering method and a semi-automated labeling method. Because the shill bidding data set is unstable, they evaluate advanced over-sampling, under-sampling, and hybrid-sampling methods and compare their performance to that of

Table 1. Previous papers.

No	Author	Year	Application	Cluster type	Dataset
1	Virdhagriswaran et al. [23]	2006	Accounting fraud	K means	Quarterly and annual financial Reports data
2	Murcia et al. [24]	2008	Financial reporting fraud	Hierarchical clustering	Annual financial reports data
3	Deng et al. [25]	2009	Financial statement fraud	Hybrid k means	Financial statements data
4	Issa et al. [26]	2011	Refund fraud/financial fraud	K means	Refund transaction data
5	Glancy et al. [27]	2011	Financial reporting fraud	Hierarchical clustering	Annual financial reports data
6	Li et al. [28]	2012	ATM phone scams	Bayesian, plot visualization	ATM phone scams data
7	Wakoli et al. [29]	2014	Medical claims fraud	K means	Insurance companies data
8	Carneiro et al. [30]	2015	Internet credit card	Naïve Bayes, plot visualization	Real internet credit card transactions
9	Li [31]	2015	Financial reporting fraud	Data mining	Annual financial reports data
10	Tangod and Kulkarni [32]	2015	Financial statement fraud	K means	Annual financial reports data
11	Meenatkshi and Sivaranjani [33]	2016	Financial statement fraud	K means	Financial statements data
12	Aluko [34]	2017	Electronic banking fraud	Data mining	Electronic banking data
13	Yazid and Fiananta [35]	2017	Credit card fraud	Hierarchical clustering, SVM	Credit card data
14	Min and Lin [36]	2018	Crime of telecom fraud	K means	Call records data
15	Anowar and Sadaoui [37]	2019	Auction fraud	Hierarchical clustering	Commercial sites
16	Li [38]	2021	Financial statement fraud	C Means	Financial statements data

several classification algorithms. The optimized shill offer classifier detects and classifies fraudulent activity with high accuracy and low misclassification rates.

Although k-means and hierarchical clustering are the most popular techniques in existing studies, other clustering techniques also exist. Yazid and Fiananta [35] uses another method, namely the *support vector machine* (SVM) to detect fraud based on *outliers*/anomalies in transaction data. Another researcher, namely Carneiro et al. [30] also used two kinds of methods in their research, namely Naïve Bayes and Plot visualization. The data of this study were normalized. The results obtained from the comparison of these two methods were the use of nave Bayes and the same visualization plot, indicating that *neuronal inputs can be reduced by clustering attributes.*

4.2 Clustering

Based on the results of cluster analysis, the results of grouping such as the dendogram are formed in the following Fig. 4.

According to Fig. 4, It is clear that there is no cluster which have the same stem length, so that the results of clusters with significantly different groupings are obtained. Cluster cutting according to the length of the longest rod. So, many clusters optimum with ward linkage method and Manhattan distance is 2 clusters. Table of the number of members in each cluster presented in Table 2.

Based on the results listed in Table 2, it can be concluded that cluster 1 is a group that commits fraud and consists of 8 companies, while cluster 2 is a group that does not commit fraud and consists of 24 companies. Companies included in cluster 1 are Ace Hardware Indonesia Tbk (a company in the property sector), Alam Sutera Realty Tbk (a company in the construction and management of housing developments), Bank Capital Indonesia Tbk (banking), Bukit Darmo Property Tbk (a real estate development company). PT Eureka Prima Jakarta Tbk (a real estate company), Media Nusantara Citra Tbk (a media company), Perdana Karya Perkasa Tbk (a company in development,

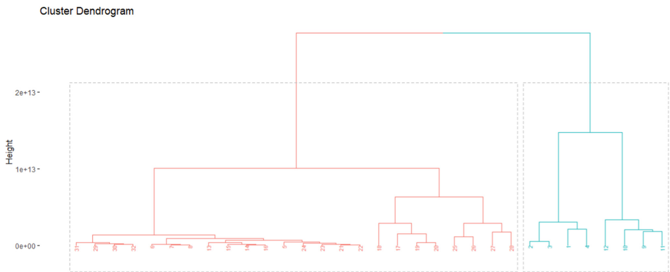

Fig. 4. Ward method Dendrogram *Linkage with Manhattan* distance.

Table 2. Many members of each cluster ward linkage method with Manhattan distance.

Cluster to-	Cluster members
1	8
2	24

industry, and mining), and J Resources Asia Pacific Tbk (a company in gold mining industry). Furthermore, the companies that are members of cluster 2 are well-known companies that are sampled for not committing fraud, consisting of CD Tbk (a company in the construction sector), NKE Tbk (a company in the construction sector), ITM Tbk (a company in the mining and sales sector). coal), JM Tbk (service provider company), SN Tbk (electronic manufacturing company), SA Tbk (oil palm and rubber plantation company), GEE Tbk (coal mining company), WK Tbk (state-owned company in the building construction sector), AE Tbk (a coal mining company), BAP Tbk (a real estate company), BSD Tbk (a real estate development company), BTPN Tbk (a banking company), BR Tbk (a in the coal producer sector), E Tbk (pipe threading service company), HMR Tbk (hospitality company), IE Tbk (investment company), IJP Tbk (coal company a loading and unloading), KIA Tbk (a trading company), SI Tbk (a coal mining company), TM Tbk (a general service provider company), TI Tbk (a general trading company), VMF Tbk (investment company), YH Tbk (manufacturing company), and KBRI Tbk (paper production company).

When the average indicators of every cluster are compared, it is discovered that cluster 1 has the highest average value of metrics when tried to compare to the others, indicating that it is a high cluster. Cluster 2 is a low cluster because it has the lowest average value of the indicators when compared to the others.

5 Implications

With this research, it is hoped that it can provide an overview related to clustering in classifying fraud and become a reference for further research.

6 Conclusion

The cluster grouping technique used by previous researchers to classify fraud tends to use clustering based on mature methods. It is still difficult to find relatively new groupings such as ensemble grouping, large-scale grouping, multi-way grouping.

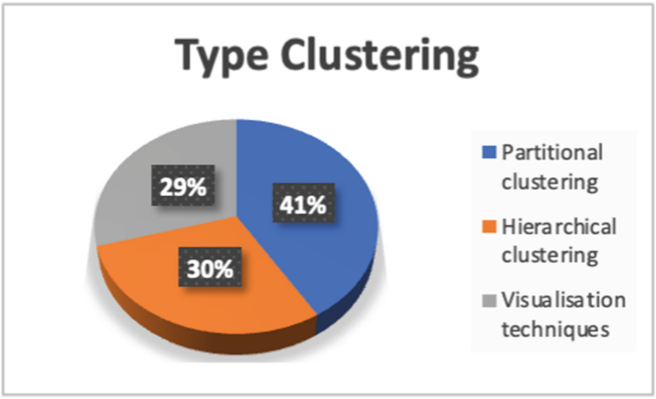

Fig. 5. Type clustering techniques.

Based on the papers obtained (Fig. 5), 41% of the clustering methods encountered are partitional clusters. Few papers are counted using partition and hierarchical grouping techniques because they do use both types of grouping. Among the partition clustering methods, grouping k-means, c-means and their variants with Euclidian distance as the dissimilarity metric is the most common and popular to use. The hierarchical clustering technique was in second place used in a quarter of the papers surveyed. Interactive, clustering visualization techniques are also used but only in a small number of cases.

In terms of how clustering techniques are mixed or utilized conjunction with other data mining methods, the researchers have reviewed were divided into three categories: standalone clustering techniques with only one clustering algorithm used, combined clustering techniques with two or more clustering algorithms used, and hybrid clustering techniques integrates clustering and other data mining algorithms, almost all classifiers based on decision trees, neural networks, and decision tree networks.

In this study, the Ward Linkage method combined with Manhattan distance produced two company clusters. Cluster 1 is a group of companies that commit fraud, while Cluster 2 is a group of companies that do not commit fraud.

Acknowledgment. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

References

1. Fagerberg, J.: Occupational Fraud: Auditors' Perceptions of Red Flags and Internal Control, (2008).
2. Oyewumi, O.R., Ogunmeru, O.A., Oboh, C.S.: Investment in corporate social responsibility, disclosure practices, and financial performance of banks in Nigeria. Future Business Journal, 4(2), 195–205 (2018).
3. Eccles, R.G., Newquist, S.C., Schatz, R.: Reputation and its risks. Harvard Business Review, 85(2), 104 (2007).
4. Sorunke, O.A.: Personal ethics and fraudster motivation: The missing link in fraud triangle and fraud diamond theories. International Journal of Academic Research in Business and Social Sciences, 6(2), 159–165 (2016).
5. Ravisankar, P., Ravi, V., Rao, G.R., Bose, I.: Detection of financial statement fraud and feature selection using data mining techniques. Decision support systems, 50(2), 491–500 (2011).
6. Maes, S., Tuyls, K., Vanschoenwinkel, B., Manderick, B.: Credit card fraud detection using Bayesian and neural networks. In: Proceedings of the 1st international naiso congress on neuro fuzzy technologies. p. 270 (2002).
7. Mohamed, N., Handley-Schachler, M.: Roots of responsibilities to financial statement fraud control. Procedia Economics and Finance, 28 46–52 (2015).
8. Goldman, H.H., Karakus, M., Frey, W., Beronio, K.: Economic grand rounds: financing first-episode psychosis services in the United States. Psychiatric Services, 64(6), 506–508 (2013).
9. Kuvvet, E.: Corporate fraud and liquidity. The Journal of Trading, 10(3), 65–71 (2015).
10. Erickson, M., Hanlon, M., Maydew, E.L.: Is there a link between executive equity incentives and accounting fraud? Journal of accounting research, 44(1), 113–143 (2006).

11. Goldman, E., Slezak, S.L.: An equilibrium model of incentive contracts in the presence of information manipulation. Journal of Financial Economics, 80(3), 603–626 (2006).
12. Peng, L., Röell, A.: Manipulation and equity-based compensation. American Economic Review, 98(2), 285–290 (2008).
13. Fischer, P.E., Verrecchia, R.E.: Reporting bias. The Accounting Review, 75(2), 229–245 (2000).
14. Povel, P., Singh, R., Winton, A.: Booms, busts, and fraud. The Review of Financial Studies, 20(4), 1219–1254 (2007).
15. Bergstresser, D., Philippon, T.: CEO incentives and earnings management. Journal of financial economics, 80(3), 511–529 (2006).
16. Beasley, M.S., Carcello, J. V, Hermanson, D.R., Lapides, P.D.: Fraudulent financial reporting: Consideration of industry traits and corporate governance mechanisms. Accounting horizons, 14(4), 441–454 (2000).
17. Klein, A.:Audit committee, board of director characteristics, and earnings management. Journal of accounting and economics, 33(3), 375–400 (2002).
18. Chidambaran, N.K., Kedia, S., Prabhala, N.: CEO director connections and corporate fraud. Fordham University Schools of Business Research Paper, (1787500), (2011).
19. Khanna, V., Kim, E.H., Lu, Y.: CEO connectedness and corporate fraud. The Journal of Finance, 70(3), 1203–1252 (2015).
20. Wang, Y.-H., Chuang, C.-C., Lee, S.-Y.: Impact of compositions and characteristics of board of directors and earnings management on fraud. African Journal of Business Management, 4(4), 496 (2010).
21. Agrawal, A., Cooper, T.: Accounting scandals in IPO firms: do underwriters and VCs help? Journal of Economics & Management Strategy, 19(4), 1117–1181 (2010).
22. Chhatoi, B.P., Sahoo, S.P., Nayak, D.P.: Assessing the Academic Journey of 'Financial Inclusion'from 2000 to 2020 through Bibliometric Analysis. Journal of Scientometric Research, 10(2), 148–159 (2021).
23. Virdhagriswaran, S., Dakin, G.:Camouflaged fraud detection in domains with complex relationships. In: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 941–947 (2006).
24. Murcia, F.D.-R., Borba, J.A., Schiehll, E.: Relevance Of Red Flags In The Evaluatinh The Risk Of Financial Statement Fraud: Perceptions Of Brazilian Independent Auditors. Revista Universo Contábil, 4(1), 25 (2008).
25. Deng, Q., Mei, G.: Combining self-organizing map and K-means clustering for detecting fraudulent financial statements. In: 2009 IEEE international conference on granular computing. pp. 126–131. IEEE (2009).
26. Issa, H., Vasarhelyi, M.A.: Application of anomaly detection techniques to identify fraudulent refunds. Available at SSRN 1910468, (2011).
27. Glancy, F.H., Yadav, S.B.: A computational model for financial reporting fraud detection. Decision Support Systems, 50(3), 595–601 (2011).
28. Li, S.-H., Yen, D.C., Lu, W.-H., Wang, C.: Identifying the signs of fraudulent accounts using data mining techniques. Computers in Human Behavior, 28(3), 1002–1013 (2012).
29. Wakoli, L.W., Orto, A., Mageto, S.: Application of The K-Means Clustering Algorithm In Medic al Claims Fraud/Abuse Detection. (2014).
30. Carneiro, E.M., Dias, L.A.V., Da Cunha, A.M., Mialaret, L.F.S.: Cluster analysis and artificial neural networks: A case study in credit card fraud detection. In: 2015 12th International Conference on Information Technology-New Generations. pp. 122–126. IEEE (2015).
31. Li, R.: Detection of financial reporting fraud based on clustering algorithm of automatic gained parameter K value. International Journal of Database Theory and Application, 8(1), 157–168 (2015).

32. Tangod, K., Kulkarni, G.: Detection of financial statement fraud using data mining technique and performance analysis. International Journal of Advanced Research in Computer and Communication Engineering, 4(7), 312–320 (2015).
33. Meenatkshi, R., Sivaranjani, K.: Fraud detection in financial statement using data mining technique and performance analysis. JCTA, 9 407–413 (2016).
34. Aluko, S.E.: Electronic Banking Fraud Detection using Data Mining Techniques and R for implementing Machine Learning Algorithms in prevention of Fraud, (2017).
35. Yazid, Y., Fiananta, A.: Mendeteksi Kecurangan Pada Transaksi Kartu Kredit Untuk Verifikasi Transaksi Menggunakan Metode SVM. Indonesian Journal of Applied Informatics, 1(2), 61–66 (2017).
36. Min, X., Lin, R.: K-means algorithm: fraud detection based on signaling data. In: 2018 IEEE World congress on services (SERVICES). pp. 21–22. IEEE (2018).
37. Anowar, F., Sadaoui, S.: Detection of auction fraud in commercial sites. Journal of theoretical and applied electronic commerce research, 15(1), 81–98 (2020).
38. Li, J.: False financial statement identification based on fuzzy c-means algorithm. Complexity, 2021 1–11 (2021).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.