sciendo

Innovative approaches in forensic accounting: The role of data analytics

Amra Kapo
University Sarajevo, School of Economics and Business
amra.kapo@efsa.unsa.ba

Lejla Turulja
University Sarajevo, School of Economics and Business
lejla.turulja@efsa.unsa.ba

Zorana Vidačak
PhD candidate in Management, University of Sarajevo, School of Economics and Business
zorana.mo@hotmail.com

Abstract

The integration of data analytics into forensic accounting has revolutionized the detection and prevention of financial fraud. This paper conducts a comprehensive analysis of recent advancements in this field, highlighting the application of machine learning, data mining, and big data techniques in identifying fraudulent activities. By reviewing the latest research and examining case studies, we demonstrate the enhanced accuracy and efficiency these technologies offer over traditional methods. The findings suggest that financial institutions should adopt these advanced tools to mitigate fraud risks and improve overall financial security.

Paper type: Research article

Received: June 16, 2024
Accepted: June 30, 2024

Citation: Kapo, A., Turulja, L., & Vidačak, Z. (2024). Innovative approaches in forensic accounting: The role of data analytics. *Journal of Forensic Accounting Profession, 4*(1), pp. 1- 14

DOI: https://doi.org/10.2478/jfap-2024-0001

The paper also explores future research directions, emphasizing the need for developing hybrid models and real-time detection systems to further enhance fraud detection capabilities.

Keywords: *data analytics, forensic accounting, fraud detection, big data, audit analytics*

1. Introduction

The rapid advancement of information technology has significantly impacted various industries, including finance. Among the critical challenges faced by the financial sector is the increasing sophistication of financial fraud, which demands more advanced detection methods. Traditional approaches to fraud detection, while useful, are often inadequate in handling the complexity and scale of modern financial transactions. Consequently, the adoption of data analytics in forensic accounting has emerged as an important solution to enhance fraud detection and prevention. Data analytics encompasses a range of techniques, including data mining, machine learning, and big data analytics, which collectively offer robust tools for detecting and analyzing fraudulent activities. These methods enable the processing of large volumes of data, uncovering hidden patterns and anomalies that may indicate fraud. Forensic accounting, when integrated with these advanced techniques, significantly improves the accuracy and efficiency of fraud detection, providing financial institutions with powerful means to safeguard their operations.

Recent studies have demonstrated the effectiveness of various data analytics techniques in forensic accounting. For instance, Zhou et al. (2021) discuss various data mining techniques such as regression, decision trees, neural networks, and Bayesian networks, emphasizing the necessity for adaptive fraud detection mechanisms in the face of evolving fraudulent schemes (Zhou et al., 2021). Similarly, Ravisankar et al. (2011) investigate the application of multiple data mining techniques, including neural networks, support vector machines, and genetic programming, to identify financial statement fraud. They highlight the effectiveness of feature selection in improving detection accuracy, particularly in distinguishing fraudulent companies from non-fraudulent ones using financial ratios (Ravisankar et al., 2011). In another study, Glancy & Yadav (2011) introduce a computational fraud detection model (CFDM) that leverages text mining techniques to analyze textual data from corporate financial reports. Their model demonstrates

significant accuracy in identifying fraudulent filings by detecting deception cues in the Management's Discussion and Analysis (MDA) sections of 10-K filings (Glancy & Yadav, 2011). Bolton & Hand (2002) review statistical fraud detection methods, focusing on their application in various domains, including financial fraud. They discuss the limitations and challenges of existing techniques and emphasize the need for continuous improvement to adapt to new fraudulent strategies (Bolton & Hand, 2002). Dechow et al. (2011) explore the prediction of material accounting misstatements using financial statement analysis. Their study employs various financial ratios to predict the likelihood of misstatements, contributing to the broader understanding of indicators that signal potential fraud (Dechow et al., 2011).

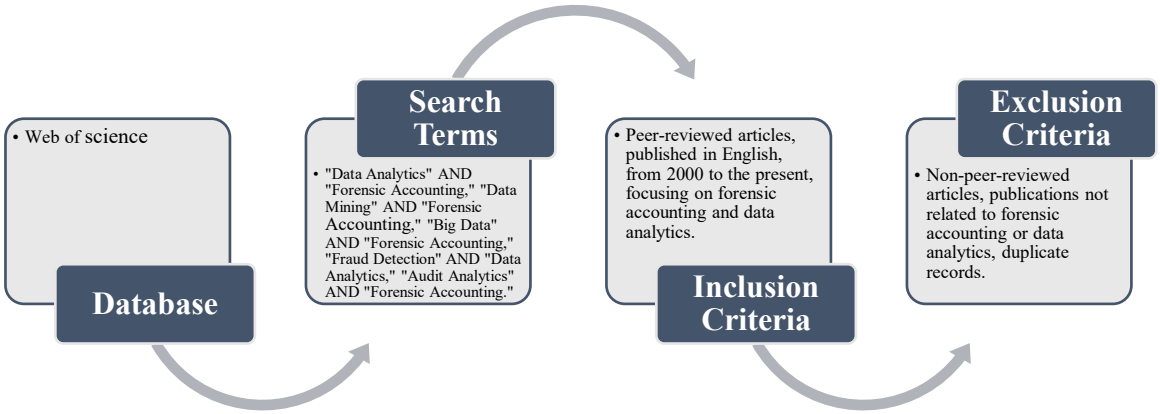
Furthermore, recent findings from Kroll's 2023 Fraud and Financial Crime Report highlight the increasing reliance on advanced technologies, such as artificial intelligence and machine learning, to combat the rising tide of financial crime. The report indicates that financial institutions invest significantly in these technologies to enhance their fraud detection capabilities and improve operational efficiency (Kroll, 2023).

This paper aims to provide a comprehensive overview of the current state of data analytics in forensic accounting, examining the latest research and practical applications in the field. By analyzing recent trends and case studies, we highlight the transformative impact of these technologies on fraud detection. Furthermore, we discuss the future directions for research, emphasizing the need for continuous improvement and innovation in developing more sophisticated detection models.

2. Methodology

This study employs bibliometric analysis to analyze the research narrative in data analytics in forensic accounting. Donthu et al. (2021) state that bibliometric analysis is interesting for a variety of reasons, such as to uncover emerging trends in article and journal performance, collaboration patterns, and research constituents, and to explore the intellectual structure of a specific domain in the extant literature. Figure 1 presents a methodological framework for bibliometric analysis with a focus on forensic accounting and data analytics. The process is presented through four key stages: database selection, definition of search terms, inclusion criteria, and exclusion criteria.

Figure 1: Methodological approach to data search



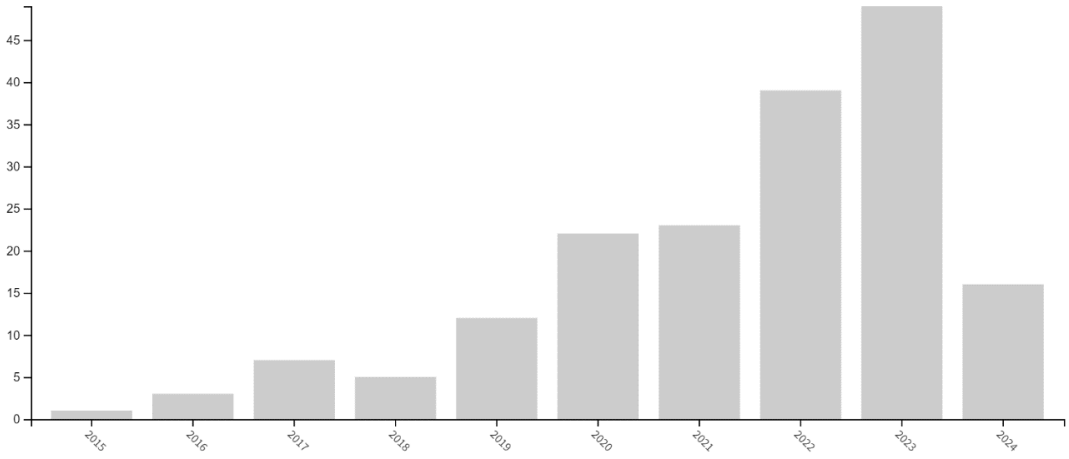
Source: Author's research

In the first phase, the database used for research is "Web of Science". This database was selected for its comprehensiveness and relevance to academic studies. The second phase includes defining the search terms that will be used to search the relevant literature. Terms used include: "Data Analytics" AND "Forensic Accounting", "Data Mining" AND "Forensic Accounting", "Big Data" AND "Forensic Accounting", "Fraud Detection" AND "Data Analytics", "Audit Analytics" AND " Forensic Accounting". The third stage concerns the inclusion criteria. To be included in the analysis, papers must be peer-reviewed, published in English, in the period from 2000 to the present, and must focus on forensic accounting and data analytics. The final stage includes the exclusion criteria. Non-peer-reviewed papers, publications unrelated to forensic accounting or data analytics, and duplicate records will be excluded from the analysis. A total of 197 papers are included in further analysis.

3. Results

The analysis of the number of published papers by year clearly shows the growing trend in research and application of data analytics in forensic accounting. In particular, the last three years have been marked by a significant increase in interest and publications, indicating this field's growing importance in contemporary research and practice.

Figure 2: Number of published papers through time

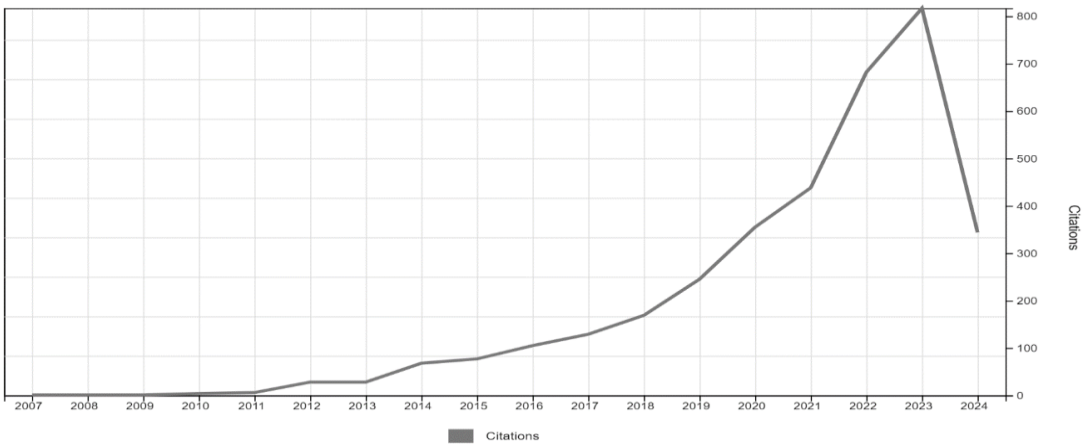


Source: Web of Science search, author's research

Based on the displayed search results, we can observe a significant growth trend in the number of published works in the field of Data Analytics in forensic accounting during the last years. In the period from 2015 to 2016, the number of published works was relatively low. In 2015, a very small number of works was recorded, while in 2016, the number of works increased slightly. This period can be considered the beginning of interest in applying data analytics in forensic accounting. In 2017, we noticed a significant increase in the number of published works, which indicates an increase in interest and recognition of the importance of this field. However, in 2018 there was a slight decrease in the number of papers, which may be the result of various factors, including potential changes in research focus or data availability. The period from 2019 to 2020 shows a steady growth in the number of published works. This growth suggests that researchers have increasingly recognized the importance of data analytics tools in financial fraud detection and forensic investigations. During these years, it becomes obvious that this field is gaining momentum and becoming more and more relevant. The period from 2021 to 2023 records the greatest growth in the number of published works. The jump in 2023 (49 published works) is particularly significant. Year, when the number of works reached its peak. This increase can be attributed to the increasing recognition of the role of big data and advanced analytical techniques in forensic accounting. Also, the increasing number of available tools and technologies probably contributed to this trend. The data recorded for the year 2024 show a slight decrease in the number of published

works compared to the previous year, but they are not particularly relevant considering that the publication of works in both fields is still expected by the end of the year.

Figure 3: Number of citations through time



Source: Web of Science search, author's research

The graph shows the total number of citations for works in the field of Data Analytics in forensic accounting by year, from 2007 to 2024. The total number of citations is 3,498, while the number of citations without self-citations is 3,222, with an average of 17.76 citations per paper. After a slow start until 2010, the number of citations gradually increased, and a significant jump was observed from 2014. The period from 2020 to 2022 records a peak with almost 800 citations in 2022, which may be the result of the accelerated development of technologies and increased interest in the application of advanced analytical tools in the detection of financial fraud.

Although there was a drop in the number of citations in 2023 and 2024, it is still significantly higher compared to earlier periods, which indicates the maturing of the research field. The overall citation statistics further confirm the importance of this field with a high average number of citations per paper and a significant number of citations without self-citations. This analysis shows the continuous growth of interest and impact of research in the field of Data Analytics in forensic accounting.

Table 1: List of most productive countries

Rank	Country	Record Count	% of 197
1	People's R China	59	29.95%
2	USA	37	18.78%
3	England	21	10.66%
4	India	19	9.65%
5	Saudi Arabia	14	7.11%
6	Jordan	7	3.55%
7	South Korea	6	3.05%
7	Malaysia	6	3.05%
9	South Africa	6	3.05%
10	Australia	5	2.54%

Source: Author's research

Based on the number of published works in the field of Data Analytics in forensic accounting, the following ten countries have the largest number of works. China stands out as the leading country with 59 works, which accounts for almost 30% of the total number of works. This high number shows the intensive research activity and importance that China attaches to the application of data analytics in forensic accounting.

The United States of America ranks second with 37 works, which represents 18,782% of the total number. This result indicates a significant research and practical application of data analytics in forensic accounting in the USA. England is in third place with 21 works, which makes 10,660% of the total number of works. England shows a high level of engagement in this area, which is a reflection of strong academic and practical interest. India is in fourth place with 19 works, which represents 9,645% of the total number. Indian researchers are actively involved in research in this field, which indicates a growing interest and development of research capacity.

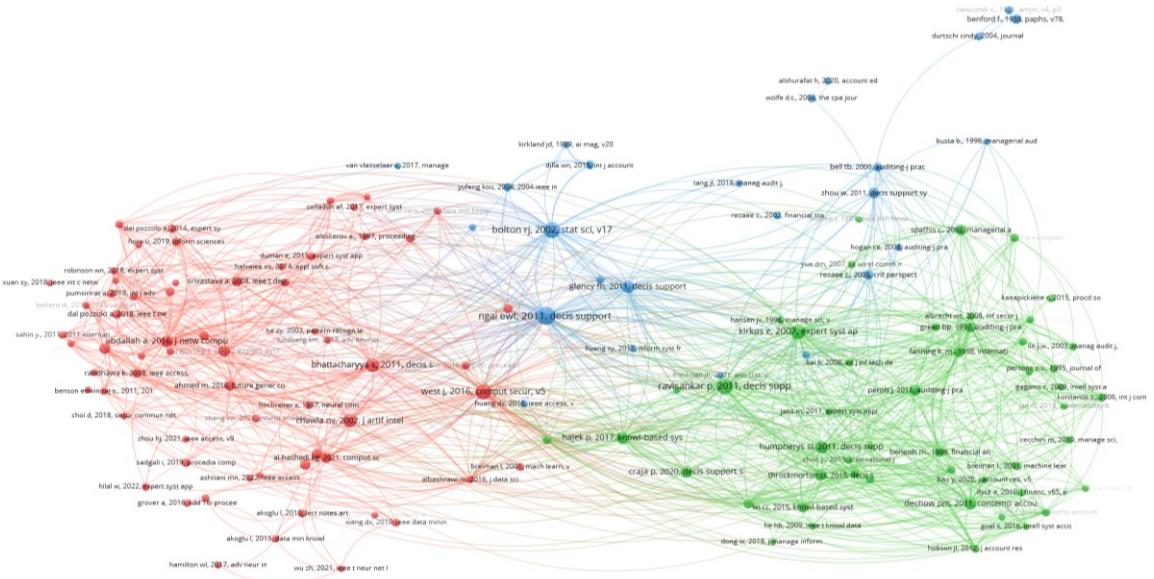
The distribution of the number of published works by country shows that China and the United States of America lead the research activity in the field of data analytics in forensic accounting, with almost 50% of the total number of works. England and India follow as significant contributors, while Saudi Arabia, Jordan, South Korea, Malaysia, South Africa, and Australia are also significantly represented. These results indicate global interest and research activity in this rapidly growing field.

The reviewed papers collectively highlight the growing importance of data analytics techniques in the field of forensic accounting. Techniques such as data mining, machine learning, and meta-learning are increasingly being applied to detect and prevent financial fraud. These advanced methods have been shown to improve accuracy and efficiency in identifying fraudulent activities compared to traditional methods. The application of these data analytics techniques is not limited to a specific context but spans various sectors and types of financial data.

The findings from these studies have significant practical implications for financial institutions. The enhanced accuracy and efficiency of data analytics tools can lead to better detection and prevention strategies, potentially saving organizations from substantial financial losses. Institutions are encouraged to adopt these advanced techniques to bolster their anti-fraud mechanisms and improve overall financial security.

The reviewed papers also suggest several recommendations for future research. There is a need for the development of more sophisticated hybrid models that combine various data analytics techniques to further improve detection accuracy. Additionally, real-time fraud detection systems that can process large volumes of data quickly and accurately are an important area for future exploration. The integration of new data sources and the continuous refinement of existing models will be crucial in advancing the field.

Figure 4: Bibliometric analysis - clusters



Source: Author's research

3.1. Red cluster: Fraud detection in financial transactions

The provided papers collectively explore advancements in the detection of financial fraud using various data-driven and algorithmic approaches. The rapid evolution of information technologies such as Big Data, the Internet of Things (IoT), Artificial Intelligence (AI), and Blockchain has significantly influenced the financial industry, introducing both opportunities and challenges. One critical challenge is the increasing sophistication of financial fraud, which necessitates advanced detection methods. Zhou et al. (2021) focus on Internet financial fraud detection by proposing a distributed Big Data approach utilizing the Node2Vec graph embedding algorithm. This method translates the topological features of financial networks into low-dimensional dense vectors, enabling efficient classification and prediction through deep neural networks. Their approach, implemented on Apache Spark and Hadoop, demonstrated superior performance in terms of precision, recall, F1-score, and F2-score, thus proving more effective than traditional rule-based systems and existing machine learning models in handling large-scale financial datasets. Carcillo et al. (2018) introduce SCARFF, a scalable credit card fraud detection framework designed to handle massive amounts of transactional data in real time. Their system leverages a combination of supervised and unsupervised learning techniques, including deep learning, to detect anomalous behavior indicative of fraud. The framework is notable for its adaptability and efficiency, which are critical for real-world applications where data volume and complexity are substantial. Bhattacharyya et al., 2011) focus on data mining techniques for fraud detection. Their research emphasizes the utility of machine learning algorithms, such as decision trees, neural networks, and Bayesian networks, in identifying fraudulent activities in financial transactions. They highlight the effectiveness of these methods in processing large volumes of data and detecting patterns that are indicative of fraud. Abdallah et al., (2016) explore the use of anomaly detection techniques in financial fraud detection. They propose a hybrid approach that combines statistical methods with machine learning to enhance detection accuracy. Their study demonstrates how integrating different analytical techniques can improve the robustness and reliability of fraud detection systems. Together, these papers highlight the critical role of advanced data analytics, machine learning, and distributed computing in enhancing the detection of financial fraud. The integration of these technologies offers promising solutions for addressing the

complexities and scale of modern financial transactions, ultimately contributing to a more secure and trustworthy financial ecosystem.

3.2. Green cluster: Advanced analytical techniques for financial fraud detection

The papers collectively underscore the significant role of data mining techniques in detecting financial statement fraud. The integration of various data mining methods, including Decision Trees, Neural Networks, Bayesian Belief Networks, Support Vector Machines (SVM), and Genetic Programming (GP), has shown promise in identifying fraudulent activities within financial statements. These techniques leverage financial ratios and other relevant indicators derived from financial statements to build predictive models.

Kirkos et al., (2007) demonstrated the effectiveness of Decision Trees, Neural Networks, and Bayesian Belief Networks in distinguishing between fraudulent and non-fraudulent financial statements. Their study found that Bayesian Belief Networks outperformed the other methods in terms of classification accuracy. Similarly, Ravisankar et al. (2011) employed several data mining techniques, including SVM and GP, and highlighted the superior performance of the Probabilistic Neural Network (PNN) in detecting financial statement fraud, particularly when feature selection was applied.

The studies collectively highlight that certain financial ratios, such as debt to equity, net profit to total assets, working capital to total assets, and the Z-score, are consistently significant indicators of fraud. These variables reflect different aspects of a firm's financial health, including leverage, profitability, solvency, and financial distress. The integration of these indicators into data mining models enhances the detection capabilities, providing auditors and stakeholders with robust tools to identify potential fraud.

Moreover, the research emphasizes the need for continual improvement and validation of these models. While Neural Networks and Bayesian Belief Networks show high accuracy, the inclusion of qualitative data, such as management characteristics and auditor qualifications, could further enhance fraud detection accuracy. The use of advanced data mining techniques offers a promising avenue for reducing the incidence of financial statement fraud and ensuring greater transparency and reliability in financial reporting.

In conclusion, the integration of data mining techniques in auditing practices presents a significant advancement in the detection of financial statement fraud. By leveraging financial ratios and other relevant indicators, these methods provide a robust framework for identifying fraudulent activities, thereby aiding auditors and enhancing the overall integrity of financial reporting (Dechow et al., 2011; Hobson et al., 2012; Kirkos et al., 2007; Perols, 2011; Ravisankar et al., 2011).

3.3. Blue cluster: Comprehensive financial fraud detection techniques

In the realm of financial fraud detection, a diverse array of methodologies has been explored, focusing on both quantitative and qualitative approaches. Zhou & Kapoor (2011) discuss various data mining techniques such as regression, decision trees, neural networks, and Bayesian networks, emphasizing the necessity for adaptive fraud detection mechanisms in the face of evolving fraudulent schemes. They propose a self-adaptive framework utilizing domain knowledge to enhance the effectiveness of these techniques (Zhou & Kapoor, 2011). Similarly, Ravisankar et al. (2011) investigate the application of multiple data mining techniques, including neural networks, support vector machines, and genetic programming, to identify financial statement fraud. They highlight the effectiveness of feature selection in improving detection accuracy, particularly in distinguishing fraudulent companies from non-fraudulent ones using financial ratios (Ravisankar et al., 2011). In another study, Glancy and Yadav (2011) introduce a computational fraud detection model (CFDM) that leverages text mining techniques to analyze textual data from corporate financial reports. Their model demonstrates significant accuracy in identifying fraudulent filings by detecting deception cues in the Management's Discussion and Analysis (MDA) sections of 10-K filings. This approach underscores the potential of automated text analysis in early fraud detection (Glancy & Yadav, 2011). Bolton and Hand (2002) review statistical fraud detection methods, focusing on their application in various domains, including financial fraud. They discuss the limitations and challenges of existing techniques and emphasize the need for continuous improvement to adapt to new fraudulent strategies (Bolton & Hand, 2002). Dechow et al. (2011) explore the prediction of material accounting misstatements using financial statement analysis. Their study employs various financial ratios to predict the likelihood of misstatements, contributing to the broader understanding of indicators that signal potential fraud (Dechow et al.,

2011). Collectively, these studies illustrate the multi-faceted approach required to tackle financial fraud, integrating advanced data mining techniques, text analysis, and financial ratio analysis to enhance detection and prevention strategies in an increasingly complex financial landscape.

4. Conclusion

The integration of data analytics into forensic accounting has emerged as a transformative approach to the detection and prevention of financial fraud. This paper has comprehensively reviewed the current state of research and practical applications in this field, highlighting significant advancements brought by techniques such as data mining, machine learning, and big data analytics. These advanced analytical tools offer superior accuracy and efficiency compared to traditional fraud detection methods. For instance, machine learning algorithms can process large datasets and identify complex patterns indicative of fraudulent activities (Zhou et al., 2021; Kroll, 2023). Data mining techniques have proven particularly effective in analyzing financial statements and transactional data, enabling auditors and investigators to detect anomalies with greater precision (Ravisankar et al., 2011; Glancy & Yadav, 2011).

The bibliometric analysis and literature review conducted in this study demonstrate a growing interest and substantial progress in the application of data analytics within forensic accounting. The findings underscore the critical role of these technologies in enhancing the reliability and effectiveness of fraud detection systems (Bolton & Hand, 2002; Dechow et al., 2011). Furthermore, the study identifies several areas for future research, including the development of hybrid models that combine multiple analytical techniques and the implementation of real-time fraud detection systems. These innovations are essential for addressing the evolving nature of financial fraud and ensuring the continued advancement of forensic accounting practices.

In conclusion, the adoption of data analytics in forensic accounting represents a significant leap forward in combating financial fraud. By leveraging these advanced tools, financial institutions can improve their fraud detection capabilities, ultimately contributing to a more secure and trustworthy financial ecosystem. Continued research and innovation in this field will be crucial in maintaining the effectiveness of fraud prevention strategies and adapting to new challenges in the financial environment (ComplyAdvantage, 2023; Kroll, 2023).

References

Abbasi, A., Albrecht, C., Vance, A., & Hansen, J. (2012). Metafraud: A meta-learning framework for detecting financial fraud. *MIS Quarterly: Management Information Systems*, *36*(4). https://doi.org/10.2307/41703508

Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, *68*, pp. 90–113.

https://doi.org/10.1016/j.jnca.2016.04.007

Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, *50*(3), pp. 602–613. https://doi.org/10.1016/j.dss.2010.08.008

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, *17*(3), pp. 235–255. https://doi.org/10.1214/ss/1042727940

Carcillo, F., Dal Pozzolo, A., Le Borgne, Y. A., Caelen, O., Mazzer, Y., & Bontempi, G. (2018). SCARFF: A scalable framework for streaming credit card fraud detection with spark. *nformation Fusion*, *41*, pp. 182–194. https://doi.org/10.1016/j.inffus.2017.09.005

Davies, A. (2023). *Four Trends Shaping the State of Financial Crime in 2023.* Retrieved from: https://complyadvantage.com/insights/four-trends-shaping-the-state-of-financial-crime-in-2023/

Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting Material Accounting Misstatements. *Contemporary Accounting Research*, *28*(1), pp. 17–82. https://doi.org/10.1111/j.1911-3846.2010.01041.x

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, pp. 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

Glancy, F. H., & Yadav, S. B. (2011). A computational model for financial reporting fraud detection. *Decision Support Systems*, *50*(3), pp. 595–601. https://doi.org/10.1016/j.dss.2010.08.010

Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance*, *32*, 100577. https://doi.org/10.1016/j.jbef.2021.100577

Hobson, J. L., Mayew, W. J., & Venkatachalam, M. (2012). Analyzing speech to detect financial misreporting. *Journal of Accounting Research*, *50*(2), pp. 349–392. https://doi.org/10.1111/j.1475-679X.2011.00433.x

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data Mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*, *32*(4), pp. 995–1003.

https://doi.org/10.1016/j.eswa.2006.02.016

Kroll. (2023). *2023 Fraud and Financial Crime Report*. Retrieved from: https://www.kroll.com/en/insights/publications/fraud-and-financial-crime-report

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 the literature. *Decision Support Systems*, *50*(3), pp. 559–569. https://doi.org/10.1016/j.dss.2010.08.006

Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing*, *30*(2), pp. 19–50. https://doi.org/10.2308/ajpt-50009

Ravisankar, P., Ravi, V., Raghava Rao, G., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems*, *50*(2), pp. 491–500. https://doi.org/10.1016/j.dss.2010.11.006

Zhou, H., Sun, G., Fu, S., Wang, L., Hu, J., & Gao, Y. (2021). Internet financial fraud detection based on a distributed Big Data approach with Node2vec. *IEEE Access*, *9*, pp. 43378–43386. https://doi.org/10.1109/ACCESS.2021.3062467

Zhou, W., & Kapoor, G. (2011). Detecting evolutionary financial statement fraud. *Decision Support Systems*, *50*(3), pp. 570–575. https://doi.org/10.1016/j.dss.2010.08.007

Sažetak

Integracija analitike podataka u forenzičko računovodstvo revolucionirala je otkrivanje i prevenciju finansijskih prevara. Ovaj rad nudi sveobuhvatnu analizu naučne literature u ovoj oblasti, ističući primjenu mašinskog učenja, rudarenja podataka i velikih podataka u identifikaciji prijevara. Pregledom najnovijih istraživanja i studija slučaja, demonstrirana je unaprijeđena tačnost i efikasnost koje ove tehnologije nude u odnosu na tradicionalne metode. Nalazi sugeriraju da finansijske institucije trebaju usvojiti ove napredne alate kako bi smanjile rizik od prevara i poboljšale ukupnu finansijsku sigurnost. Rad također istražuje smjerove budućih istraživanja, naglašavajući potrebu za razvojem hibridnih modela i sistema za real-time otkrivanje kako bi se dodatno poboljšale mogućnosti otkrivanja prevara.

Ključne riječi*: analitika podataka, forenzično računovodstvo, otkrivanje prijevara, veliki podaci, revizijska analitika*