Advances in Informatics, Management and Technology in Healthcare
J. Mantas et al. (Eds.)

doi:10.3233/SHTI220649

Artificial Intelligence Solutions to Detect Fraud in Healthcare Settings: A Scoping Review

Mohammad Sharique IQBAL[a], Alaa ABD-ALRAZAQ[a,b] and Mowafa HOUSEH[a,1]
[a] *Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar*
[b] *AI Center for Precision Health, Weill Cornell Medicine-Qatar, Doha, Qatar*

Abstract. Over the past decade, Artificial Intelligence (AI) technologies have quickly become implemented in protecting data, including detecting fraud in healthcare organizations. This scoping review aims to explore AI solutions utilized in fraud detection occurring in treatment settings. To find relevant literature, PubMed and Google Scholar were searched. Out of 183 retrieved studies, 31 met all inclusion criteria. This review found that AI has been used to detect different types of fraud such as identify theft and kickbacks in healthcare. Additionally, this review discusses how AI techniques used in network mapping fraud can detect and visualize the hacker's network. A proper system must be implemented in healthcare settings for successful fraud detection, which may overall improve the healthcare system.

Keywords. Fraud, artificial intelligence, deep learning, machine learning, healthcare settings

1. Introduction

Over the past twenty years, digital crime has rapidly increased. It is challenging and costly to eradicate these issues; it is estimated that more than $86 million is spent by the FBI to combat crime and fraud alike in the United States [1]. This affects large and small businesses alike, as the Association of Certified Fraud Examiners found that fraud costs businesses 5% of their annual turnover [2]. As healthcare organizations more frequently utilize electronic healthcare records and online payment systems, an efficient detection system or model may better assist in detecting and classifying any instances of fraud. As such, to mitigate issues of fraud, several organizations are implementing increasingly sophisticated resources to protect their networks and data. This includes adopting new technologies that utilize Artificial Intelligence (AI). To achieve implementing successful fraud detection, a better understanding of fraud and digital crime prevention strategies is essential to establish a more effective and successful learning strategy.

As big data evolves, detecting fraudulent activities within networks has become increasingly complex. However, AI technology such as deep learning and machine learning approaches can expedite awareness [3]. These approaches draw on data

[1] Corresponding Author, Dr. Mowafa Househ, Hamad Bin Khalifa University, College of Science and Engineering, Doha, Qatar; E-mail: MHouseh@hbku.edu.qa.

techniques to provide a holistic view of interdependencies within a network. Recently, deep learning approaches have made significant contributions to detecting fraudulent activities within healthcare networks. As such, fraud detection experts have recognized them as a solid, reliable, and promising anomaly detection technique [4]. While several studies on AI and fraud detection have been conducted, little research has summarized how novel AI approaches are utilized to mitigate fraud. Studies conducted before the year 2015 do not provide in-depth explorations into AI as the technology was not as developed at this time. To bridge this gap, this review aims to provide an overview of AI solutions used by previous studies to detect fraud in healthcare settings.

2. Methods

This scoping review was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Review and Meta-Analyses-extension for scoping reviews). Two databases were utilized to retrieve relevant studies: Google Scholar and PubMed. We used a combination of 3 groups of search terms related to fraud (e.g., fraud, crime, and forensic), AI (e.g., artificial intelligence, deep learning, and machine learning), and healthcare (e.g., health, medical). We included studies that used AI solutions for detecting fraud in healthcare settings whereas we excluded those that used non-AI solutions and not in healthcare settings. Any studies that were written in a language other than English or published before the year 2015 were not included in the review. Rayyan software was utilized to aid the study selection process. Study selection was conducted in three phases: removing duplicates, reviewing the titles and abstracts of articles, and then reviewing the full articles. The extracted data was then narratively synthesized using an Excel spreadsheet [2]. The study selection, data extraction, and data analysis were carried out by the first author only.

3. Results

A total of 183 citations were retrieved from the two databases. Of these citations, 31 studies were found eligible for this scoping review [1-31]. A flow chart of the study selection process can be found in Appendix 1. Twenty-seven of the included studies were published journal articles, while the remaining 4 studies were papers presented at conferences. The included studies originated from 12 countries, with the largest number of studies published in the United States (n=13). All included articles were published between 2015 and 2020, with the largest number of studies published in 2020 (n=14). More details about the characteristics of the included studies are shown in Appendix 2.
The included studies utilized AI for fraud detection (n=17) [1, 2, 6, 8, 14, 17, 18, 19, 21, 22, 23, 24, 26, 27, 28, 30, 31], identifying and classifying detected fraud (n=8) [3, 4, 11, 12, 13, 15, 20, 25], and investigating and analyzing fraudulent data (n=6) [5, 7, 9, 10, 16, 29]. The most common algorithm used in the included studies was Convolutional Neural Network (CNN) (n=13), followed by Artificial Neural Network (ANN) (n=10). The most commonly used validation methods were 5-fold cross validation (n=9) and 10-fold cross validation (n=9). The size of dataset used in the studies ranged from 135 [5] to 4,310 [15]. Only three studies [16, 17, 18] utilized a dataset of 1,000 or less. The most common

[2] Appendices are available at GitHub: https://github.com/moiq33909/Research1.

metric used to assess performance of the model was accuracy (n=18), followed by sensitivity (n=15), specificity (n=15), and Area under ROC curve AUC (n=12).

4. Discussion

This review finds that the most utilized AI techniques to detect fraud include both deep learning and other AI detection systems such as Intrusion Detection systems, Neural Networks, and a Defendable Healthcare Networks Environment. Recent research has found that deep learning models can effectively identify patterns and distinguish features in various fraudulent activity more successfully than other techniques. Furthermore, deep learning has become a preferred technique because its algorithms can both more effectively protect medical data as well as prevent devices from being susceptible to malicious activity. Overall, many organizations rely on more traditional methods to protect themselves against cybercrime and fraud; however, these techniques are significantly less effective than those mentioned in this scoping review. This is because various advanced types of attacks and fraud can occur, such as advanced persistent threats (APT) carried out by highly skilled cyber fraud groups.

This review has some limitations. Most studies collected for this review were conducted in the United States, indicating that results were mostly limited to a specific population. This may have inadvertently led to missed information on fraud detection techniques utilized by different countries or cultures. In addition, only two databases, PubMed, and Google Scholar, were searched as other advanced databases (such as Web of Science, ProQuest, and others) were inaccessible. As a result of this limitation, relevant studies may have been missed. Moreover, this review restricted the article search to studies published in English; consequently, this review likely missed many relevant research studies written in other languages. Lastly, deep learning requires a large dataset that have been specifically designated for training and purchasing this data may not be feasible for smaller organizations.

5. Conclusions

This scoping review was performed to explore utilizing AI technology in detecting and identifying fraud and digital crimes occurring in healthcare settings. This review finds that safe, high quality and cost-effective systems must be developed to aid healthcare settings in effectively mitigating fraudulent activity. The applications and other AI techniques are beneficial to treatment settings, but can be challenging to implement and costly to maintain. It is recommended that hybrid technologies be developed to detect fraud alerts quickly and precisely, and that automatically provide alarms and support to designated staff and employees. Ultimately, providing a timely alert and relevant information on fraudulent activity, along with overall better quality of fraud detection, can greatly assist healthcare organizations wherein anonymous fraud may take place.

References

[1] Modugu KP, Anyaduba JO. Forensic accounting and financial fraud in Nigeria: An empirical approach. International Journal of Business and Social Science. 2013 Jul;4(7):281-9.

[2] Schreyer M, Sattarov T, Schulze C, Reimer B, Borth D. Detection of accounting anomalies in the latent space using adversarial autoencoder neural networks. arXiv preprint arXiv:1908.00734. 2019 Aug 2.
[3] Arora S, Bhatia MP. Biometrics for forensic identification in web applications and social platforms using deep learning. InForensic Investigations and Risk Management in Mobile and Wireless Communications 2020 (pp. 80-113). IGI Global.
[4] Ghasemi M, et al. The Application of Machine Learning to a General Risk–Need Assessment Instrument in the Prediction of Criminal Recidivism. Criminal Justice and Behavior. 2021 Apr;48(4):518-38.
[5] Pourhabibi T, Ong KL, Kam BH, Boo YL. Fraud detection: A systematic literature review of graph-based anomaly detection approaches. Decision Support Systems. 2020 Jun 1;133:113303.
[6] Iorliam A. Cybersecurity in Nigeria: A Case Study of Surveillance and Prevention of Digital Crime. Springer; 2019 Mar 15.
[7] Chandrakala T, Rajini SN, Dharmarajan K, Selvam K. Development of crime and fraud prediction using data mining approaches. Technology. 2020;11(12):1450-70.
[8] Hassija V, Chamola V, Gupta V, Jain S, Guizani N. A survey on supply chain security: Application areas, security threats, and solution architectures. IEEE Internet of Things Journal. 2020 Sep 22;8(8):6222-46.
[9] Legotlo TG, Mutezo A. Understanding the types of fraud in claims to South African medical schemes. South African Medical Journal. 2018 May 8;108(4):299-303.
[10] Villegas-Ortega, J., Bellido-Boza, L., & Mauricio, D. (2021). Fourteen years of manifestations and factors of health insurance fraud, 2006–2020: a scoping review. Health & Justice, 9(1), 1-23.
[11] Li J, Huang KY, Jin J, Shi J. A survey on statistical methods for health care fraud detection. Health care management science. 2008 Sep;11(3):275-87.
[12] Edwards JR, et al. National Healthcare Safety Network (NHSN) report, data summary for 2006, issued June 2007. American journal of infection control. 2007 Jun 1;35(5):290-301.
[13] Stowell NF, Schmidt M, Wadlinger N. Healthcare fraud under the microscope: improving its prevention. Journal of Financial Crime. 2018 Oct 1.
[14] Mishevski M. (2019). Utilization of Artificial Intelligence for Network Security (Doctoral dissertation, Utica College).
[15] Abeshu A, Chilamkurti N. Deep learning: the frontier for distributed attack detection in fog-to-things computing. IEEE Communications Magazine. 2018 Feb 13;56(2):169-75.
[16] Parra GD, Rad P, Choo KK, Beebe N. Detecting Internet of Things attacks using distributed deep learning. Journal of Network and Computer Applications. 2020 Aug 1;163:102662.
[17] Gonçalves CP. Cyberspace and Artificial Intelligence: The New Face of Cyber-Enhanced Hybrid Threats. InCyberspace 2019 Aug 9. IntechOpen.
[18] Ye Y, Li T, Adjeroh D, Iyengar SS. A survey on malware detection using data mining techniques. ACM Computing Surveys (CSUR). 2017 Jun 29;50(3):1-40.
[19] Ayala L. Active medical device cyber-attacks. InCybersecurity for hospitals and healthcare facilities 2016 (pp. 19-37). Apress, Berkeley, CA.
[20] Zhang Z, Zhou X, Zhang X, Wang L, Wang P. A model based on convolutional neural network for online transaction fraud detection. Security and Communication Networks. 2018 Aug 6;2018.
[21] Naagas MA, et al. DEH-DoSv6: A defendable security model against IPv6 extension headers denial of service attack. Bulletin of Electrical Engineering and Informatics. 2021 Feb 1;10(1):274-82.
[22] Bollé T, Casey E, Jacquet M. The role of evaluations in reaching decisions using automated systems supporting forensic analysis. Forensic Science International: Digital Investigation. 2020;34:301016.
[23] Mena J. Machine learning forensics for law enforcement, security, and intelligence. CRC Press; 2016 Apr 19.
[24] Mena J. Machine learning forensics for law enforcement, security, and intelligence. CRC Press; 2016 Apr 19.
[25] Abakarim Y. An efficient real time model for credit card fraud detection based on deep learning. InProceedings of the 12th international conference on intelligent systems: theories and applications 2018.
[26] Sun Y, Lo FP, Lo B. Security and privacy for the internet of medical things enabled healthcare systems: A survey. IEEE Access. 2019 Dec 18;7:183339-55.
[27] Rudd EM, et al. A survey of stealth malware attacks, mitigation measures, and steps toward autonomous open world solutions. IEEE Communications Surveys & Tutorials. 2016 Dec 8;19(2):1145-72.
[28] Larson BJ. False Positive Reduction in Credit Card Fraud Prediction: an Evaluation of Machine Learning Methodology on Imbalanced Data (Doctoral dissertation, Capitol Technology University).
[29] Cuzzocrea A, et al. Improving Machine Learning Tools with Embeddings: Applications to Big Data Security. In2018 IEEE International Conference on Big Data (Big Data) 2018 Dec 10 (pp. 5086-5092).
[30] Ch R, Gadekallu TR, Abidi MH, Al-Ahmari A. Computational system to classify cyber crime offenses using machine learning. Sustainability. 2020 Jan;12(10):4087.
[31] Abuhamad M, et al. Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey. IEEE Internet of Things Journal. 2020 Aug 28;8(1):65-84.