International Journal of
Technology and Systems
(IJTS)

IMPLICATIONS FOR BIG DATA ANALYTICS ON CLAIMS FRAUD MANAGEMENT IN INSURANCE SECTOR

Dr. Ben Kajwang PhD

International Journal of Technology and Systems

ISSN 2518-881X (Online)

Vol.7, Issue 1, No.4, pp 60-71, 2022

www.iprjb.org

IMPLICATIONS FOR BIG DATA ANALYTICS ON CLAIMS FRAUD MANAGEMENT IN INSURANCE SECTOR

Dr. Ben Kajwang PhD, ACII (UK), AIRM (UK), FIIU, FIIK, CPT, Chartered Insurer
Chief Executive Officer, College of Insurance, Nairobi, Kenya
Corresponding Author Email: bkajwang@coi.ac.ke

ABSTRACT

Purpose: Because of the enormous financial burden that insurance fraud places on businesses, executives are moving quickly to implement big data analytics and other forms of cutting-edge technology in order to combat the issue. The purpose of the study is to assess the implications for Big data analytics on claims fraud management in insurance sector.

Methodology: This was accomplished through the use of a desktop literature review. The use of Google Scholar was utilized in order to locate seminal references and journal articles that were pertinent to the study. In order to meet the inclusion criteria, the papers had to be no more than ten years old.

Findings: The study concludes that Big Data Analytics in the insurance industry is becoming a promising field for gaining insight from very large data sets, enhancing outcomes, and lowering costs. It has tremendous potential, but there are still obstacles to overcome. The findings demonstrated that digital fraud detection had a positive and significant impact on insurers' underwriting procedures.

Unique contribution to theory, practice and policy: The research suggests that insurers should always strive to automate their claim processes. In addition, the study suggests that insurers implement elements of constructing digital insurance control mechanisms. Before incorporating new technologies and analytical tools, they recommend organizations to conduct a thorough cost-benefit analysis and scenario planning to address unintended outcomes.

Keywords: *Big Data, Analytics, Claims Fraud Management, Insurance Sector*

International Journal of Technology and Systems

ISSN 2518-881X (Online)

Vol.7, Issue 1, No.4, pp 60-71, 2022

www.iprjb.org

INTRODUCTION

Fraud is defined as "wrongful or unlawful deception intended to result in financial or personal advantage" according to the Oxford dictionary. Fraud can have a negative impact on a company in a number of different areas, including financial, operational, and psychological. Whilst the financial loss that can be attributed to fraud is substantial, the complete impact that fraud can have on a company can be astonishing. The damage to one's reputation, goodwill, and relationships with one's customers can be catastrophic. Since fraud can be committed by any worker within a company or by individuals from the outside, it is essential to establish a fraud management program that is both successful and well-implemented in order to protect the assets and reputation of the organization. Due to the ease with which people may access modern technology and communication, the rate of fraud is rising at an exponential rate, resulting in the loss of trillions of dollars around the world each year (Hargreaves and Singhania, 2015).

Automobile insurance fraud is a global concern. Dealing with fraudulent claims manually has always been an expensive endeavor for insurance firms. The use of data analytics is an efficient and effective technique to be more proactive in the battle against fraud and to identify transactions that suggest fraudulent activity or the increased risk of fraud. Data analytics has the potential to play a significant part in the process of fraud detection and can assist insurance firms in recognizing fraudulent activity. According to Bolton and Hand, the utilization of a graded system of investigation is the most appropriate overall technique for the detection of fraud. Accounts that have substantial but lower suspicion ratings should be observed more closely (without additional costs), whereas those with very high suspicion scores should be investigated immediately and thoroughly (at a high cost) (Bolton and Hand, 2002).

The term "big data" refers to a collection of datasets that are typically too large to be effectively managed by utilizing traditional data analysis techniques. The process of collecting, organizing, and analyzing a huge, diversified dataset that involves numerous categories such as structured and unstructured data, as well as streaming and batch, with sizes ranging from terabytes to zettabytes, in order to identify patterns and other useful information is referred to as "big data analytics." In the field of information security, often known as e-security or cyber security, big data analytics can be used to gather enormous volumes of digital information for the purpose of analyzing, visualizing, and deducing insights that can make it feasible to detect criminal activity. This has the potential to revolutionize security analytics by enhancing the management, storage, and analysis of various security-related data. Big data analytics create a unified view by correlating data taken from numerous sources, such as network traffic, log files, financial transactions, healthcare claims, and so on. This allows for the identification of anomalies and suspicious behaviors conducted by criminals. Big data is perfect for researching concerns related to information security, and finding data patterns that are not typical in log files is a significant part of the process of identifying criminal activity. The application of big data tools will make this type of research simpler, making it easier to find abnormalities that indicate a data breach has occurred. With such a potent

advantage, big data analytics might lead to the uncovering of major crimes, which would inevitably result in "major arrests" (Corlosquet-Habart and Janssen, 2018).

One of the advantages of big data analytics is the ability to avoid fraud. At least 10 percent of all insurance company payouts are false claims, and the global aggregate of these fraudulent payments might be billions or even trillions of dollars. There has been an increase in the severity of insurance fraud, and criminals have gotten more sophisticated, throughout the years (Srivastava and Gopalkrishnan, 2015).

When it comes to detecting insurance fraud, what role does big data analytics play? Insurance firms have a vested interest in detecting fraudulent activity as early as possible. In order to prevent fraudulent claims, firms can use both historical and real-time data to construct prediction models based on wage data, medical claim data, legal expenses, demographics, weather data, contact center notes and voice recordings. Fake medical bills or a faked accident could be part of a personal injury claim, for example. Fraudulent auto insurance and medical insurance scams have increased in sophistication in recent years, according to companies. Similar operations may be carried out in different locations of the country or by using different aliases for the claimants, depending on the type of ring (Srivastava and Gopalkrishnan, 2015).

Before a new claim gets too far into the process, big data analysis may immediately look for trends in previous claims and uncover parallels or raise issues about those claims. There is a consensus that big data analytics has the potential to be a game changer in the field of risk and fraud management at insurance companies, as well as actuary and underwriting executives and insurance business managers. The objective is to detect fraudulent claims as soon as a loss is reported, which is when an underwriter or actuary is most needed (Raghupathi and Raghupathi, 2014).

In order to combat the problem of insurance fraud, executives are fast implementing big data analytics and other innovative technology. Insurance companies aren't the only ones affected by these high expenses; customers who are forced to pay higher rates to cover the losses also bear the brunt of this burden. Large amounts of unstructured and structured data connected to insurance claims can be analyzed using big data analytics to spot patterns of fraudulent behavior. These businesses stand to gain greatly from their investments. With the implementation of a big data platform, they can now examine complicated information and accident scenarios in minutes rather than the days or months it took before (Gill, 2020).

To better forecast and even prevent insurance fraud, organizations can now utilize a variety of new "Big Data" possibilities.

1. A simple value or ranking of what to watch out for can be generated by predictive models, providing managers a clearer sense of what to expect, and therefore boosting possible prevention methods.

2. A neural network is a hardware and software system that is based on the structure of the human brain. These systems are capable of adapting to change and novelty. Neural network technology is capable of detecting new and systematic insurance scams when managed, customized, and implemented properly. If you use this technology, you can prevent catastrophic losses from newer fraud tactics before they take hold. Semantic processing and artificial intelligence are two examples of this type of technology in action.
3. Data mining: mountains of information are hidden in various sources, including documents, recordings, case notes, surveys, and other online and offline sources. The process of manually accessing specific information that can point to fraudulent activity is difficult, and it can also be inefficient and time-consuming. Applications and automated processes that can extract and synthesize relevant text, voice, or financial information can make this task much easier and, in less time, putting less burden on managers and analysts so that they can focus on pinpointing possible insurance fraud incidents and on the preparatory and preventative measures against them.

Detecting insurance fraud in the insurance industry is possible, and here are three data-driven methods to assist in the process: Industry-wide databases: . Any claim's evaluation will be enhanced if it can be linked to complete and reliable historical data. If you see a pattern of claims for the same item or ailment, you might be on the lookout for fraudulent claims and scammers. Data visualization: Making sense of the information you have at your disposal will aid you in making better judgments. In many cases, images and infographics are the most effective means of conveying enormous amounts of complex data. Those that are closest to the action will be able to use this information most effectively. Apps and software for investigation: Streamline the claims process by creating checklists and providing one-touch reporting. As a result, human error is less likely, and manual reports aren't as inefficient (Subramaniam and Mahmoud, 2021).

In contrast to a slew of disparate healthcare fraud detection tools that don't work together or worse, don't even "know" to check other Internet data sources, the beauty of big data is that it can all be mined and evaluated by a single tool or model. Another typical sort of fraud is claiming benefits on behalf of a deceased individual. This information is not available in an outdated system, but it is available in a system that is "connected into" big data (George et al., 2016).

Statement of the Problem

Big data analytics is predicted to improve claims fraud management in insurance sector. However, in practice, claims fraud management has been poor in the insurance sector. This is driven by poor adoption of big data analytics tactics by the health insurance claims. This poses a huge challenge to the stakeholders and consequently a solution to the problem needs to be addressed. Previous research has been undertaken in try to identify the linkage between digitalization and performance of claims fraud management. However, they did not adequately illustrate the linkage between big data and analytics and performance of claims fraud management. With the existing scarcity in

literature, it is important that study is undertaken in try to determine the implications for Big data and analytics on claims fraud management in insurance sector.

LITERATURE REVIEW

Rational Choice Theory

Clarke (1997) argues that according to Rational Choice theory in criminology law breakers are rational people, who evaluate the outcomes of associating with in crime and the subsequent repercussions. The notion behind a criminal committing a crime is hinged to the fact that there are perceived benefits in committing a crime and less costly as opposed to noncriminal behaviour. The theory of Rational Choice theory in criminology evaluates the benefits of committing a crime and underlying risks. Criminals engage in fraudulent behaviours basically for economic gains. In this regard Rational Choice Theory explains causes and the source drive that result in criminal activities. Rational Choice theory in criminology enables analysis and in-depth understanding of crimes committed (Brown, Ebsen & Geis, 2010). Making of rational choices is widely used as an assumption of the behaviour of an individual which can be aggregated to a group.

With reference to insurance fraud the Rational Choice Theory the drive for fraudulent actions are triggered by the tradeoff between monetary benefits and the magnitude of punishment. In this regard the consumer may decide to commit or not to commit and this rational choice is related to one's demographic characteristics. Consumer insurance frauds are associated with high benefits and relatively low risks as the penalty for fraud is also relatively lenient compared to physical crimes like robberies.

Differential Association Theory

Edwin (1949) developed the theory of Differential Association that buttresses the explanation of white-collar crime as well as common street crime. Like robbery and burglary (Henry & Einstadter, 1998) Edwin (1949) crime occurs when a criminal is exposed to an enabling favourable environment. This enabling environment is reinforced by the criminal capacity to commit crime. This theory has been used to explain white collar crime, professional crime and street crime. The value proposition to commit crime by the offender is bolstered by the rewards attached to the success attached to committing crime (Ackers, 1985).The offender therefore ponders on the possible demystified risks that guarantees the success of committing a crime, that is normally motivated by substantial rewards(Arkers, 1998). The critics of this theory argue that theory of Differential Association is limited to consumer class of goods as well as its demerit of explaining defiance of criminals that is basically not learnt. Critics of Sutherland differential association theory emanate from the apparent inability of the theory to explain acts of deviance which are not necessarily learnt, as frauds are high reward and have relatively low penalty. This theory remains relevant to this study since it posits that crime is learnt and reinforced.

Empirical Review

Rizvi (2021) conducted a study with the intention of analyzing the function that big data plays in financial institutions for financial fraud. For this goal, secondary data were acquired through

literature review. According to the findings, unstructured big data have a beneficial function to play in the financial institutions for the prevention of financial fraud. Identification of criminal activities would be much improved by having the capability to manage higher volumes of data as well as to operate with new and unstructured kinds of data. Automated fraud detection in financial institutions has been implemented with the goal of accumulating useful information in order to reduce instances of financial fraud. Big data analytics and data mining techniques are the foremost consistent, best-turned approach for recognizing rules and designs of action that can help detect frauds or a financial crime. This is because it is the leading strategy to recognize similitude between individual or group behavior recognized from numerous sources and cross-checking value-based information behavior. The research suggests that researchers investigate further prospects, difficulties, tools, and approaches related to big data analytics. They can also investigate the number of firms that have incorporated big data and how big data contributes to the generation of sales.

Ngiri (2021) conducted a research study to investigate the impact that the incorporation of technology has had on the underwriting procedures of the top five insurance companies in Nairobi County. The methodology for the study was a survey descriptive approach. The population of interest consisted of 1,174 staff members working for the top five insurance businesses in Nairobi County. In the research, a method called stratified sampling was utilized to select 298 workers from the various insurance companies to take part in the survey. The study relies on first-hand accounts and observations. The usage of structured questionnaires was the method that was utilized for the collecting of the data. The findings indicated that the majority of insurers agreed that they face the risk of being unable to meet claims because of fraudulent claims from policyholders, which was discussed in relation to the effect that digital fraud detection has on the underwriting process among the top five insurers in Nairobi County. It was shown that there was a favorable but modest association between digital fraud detection and the underwriting process. The investigation also found that the detection of digital fraud had a beneficial and discernible impact on the underwriting process utilized by insurers in Nairobi County. According to the findings of the survey, insurance companies should make it a priority to find ways to automate their claim procedures. In addition, the research suggests that insurance companies should initiate certain initiatives toward constructing digital insurance control mechanisms.

In their work, Kaur, Sharma, and Mittal (2018) provide a quick introduction to big data and its relevance in healthcare applications. It has been discovered that the utilization of big data architecture and methods are continuously assisting in managing the brisk growth of data that is occurring in the healthcare industry. To begin, an empirical study is carried out with the purpose of analyzing the role that big data plays in the healthcare industry. It has been noted that great progress has been made in the field of healthcare through the application of big data. The potential impact that machine learning and big data could have on the healthcare industry is difficult to conceptualize in the modern era. It has been seen that the majority of the authors who have adopted the use of machine learning and big data analytics in disease diagnosis have not given considerable

weightage to the privacy and security of the data. This is something that has been noted. For the purpose of managing the big data that the medical industry generates, the following article presents a novel architecture of a smart and safe healthcare information system that makes use of machine learning and advanced security mechanisms. The invention consists of the introduction of a data storage layer that is both optimum and secure, which is used to keep both security and privacy intact. A variety of security measures, including masking encryption, activity monitoring, granular access control, dynamic data encryption, and end point validation, have been integrated. The hybrid four-layer healthcare model that has been proposed appears to be more effective as a disease diagnostic big data system.

Mehta and Pandit (2018) conducted a study on the topic of the concurrent development of big data analytics and healthcare. The purpose of this comprehensive literature study is to assess the breadth of big data analytics in the healthcare industry, covering its applications as well as the problems associated with its adoption in the healthcare sector. In addition to that, it attempts to determine the methods that can be used to overcome the obstacles. The publications chosen for this study are descriptive in nature and focus on the usefulness of Big Data analytics in the medical and healthcare fields. Following the establishment of the inclusion criteria, a total of 58 articles were chosen for further investigation. The analyses of these articles came to the following conclusions: (1) researchers are unable to come to a consensus regarding the operational definition of Big Data in healthcare; (2) Big Data in healthcare comes from the internal sources within the hospitals or clinics as well as external sources including the government, laboratories, pharma companies, data aggregators, medical journals, and so on; (3) natural language processing (NLP) is the most widely used Big Data analytical technique for healthcare, and the majority of the processing is done through NLP; and (4) The lack of evidence of the practical benefits of using big data analytics in healthcare is the fifth and most significant barrier to its widespread adoption. This review study reveals that there is a dearth of information on proof of the application of big data analytics in real-world settings within the healthcare industry. This is due to the fact that the usability studies have only addressed the qualitative approach, which describes the possible benefits but does not take the quantitative study into consideration. In addition, the bulk of the studies originated in wealthy nations; this highlights the importance of fostering research on Healthcare Big Data analytics in countries that are still developing.

A research project on the application of big data analytics tools and techniques in the insurance industry was carried out by Banu (2022). This study aims to recognize what "big data" means in the insurance sector and how the application of "Big Data Analytics" has opened the door for new and innovative changes in the insurance industry. Specifically, the study will focus on how the application of "Big Data Analytics" has opened the door for new and innovative changes. This study provides an overview of the Big Data Analytics field as it relates to the insurance industry. It describes the field, discusses its benefits, outlines the tools, architectural framework, and method, describes applications generally and specifically, and provides a brief discussion of the opportunities and challenges. According to the findings of the study, Big Data Analytics in the

insurance industry is becoming a potential field for giving insight from extremely huge data sets, as well as improving results while simultaneously cutting costs. It has a lot of potential, but first there are several obstacles that need to be conquered.

A systematic literature analysis on the use of big data in healthcare was carried out by Senthilkumar et al. (2018). The review focused on publications published between 2010 and 2015. The study provided an overview of the definition, process, and applications of big data in medical administration. The growth of unstructured data is significantly outpacing that of both semi-structured and structured data. Data collecting, data storage, data administration, data analysis on data, and data visualization are the primary steps of big data management in the healthcare business. Ninety percent of big data is in the form of unstructured data. Recent research has focused on various tools for visualizing large amounts of data. The authors of this research conducted an analysis of the effective tools now being used for the visualization of large amounts of data and made suggestions for new visualization tools that may be utilized to better manage large amounts of data in the healthcare business. According to the findings of the study, the most difficult aspects of using big data in healthcare are understanding unstructured clinical notes in the appropriate context and gleaning potentially useful information from them. Other challenging aspects include data privacy, data leakage, data security, efficient handling of large volumes of medical imaging data, information confidentiality and security, incorrect use of health data or failure to safeguard the healthcare information, and understanding unstructured clinical notes. The authors recommend a few new data visualization tools for the healthcare analyst to use so that they can make more informed decisions. There is a significant possibility that big data may advance healthcare management and bring the healthcare business to the next level.

Yange (2019) developed a Fraud Detection System for the National Health Insurance Scheme (NHIS) in Nigeria as part of his research. This was done with the intention of addressing the fraudulent activities of some stakeholders in the NHIS; numerous studies have demonstrated that the absence of appropriate tools to address this problem has negatively impacted service providers as well as the beneficiaries of this Scheme. This was done in light of the fact that this Scheme lacked the appropriate tools to address the problem. In order to accomplish the purpose of this study, an examination of organizational documents, direct observation, and the collection of previously collected data from NHIS accredited health facilities and Health Maintenance Organizations in Nigeria were all carried out. According to the findings of the research, there are four different modules that make up the system: user management, enrollment processing, referral processing, and claims processing. As a result of this, various types of fraudulent activity within the NHIS were uncovered, including double billing, billing for services that were not actually rendered, ghost patients, identity theft, self-referral, collusion with providers, and kickback schemes. The purpose of this paper was to develop a method for the detection of fraudulent activities carried out by actors of the NHIS. This system made use of data from the Nigerian National Health Insurance Scheme (NHIS), which was arranged in the following categories:

enrollment, referral, and claim data, and stored in the following file formats: pdf, jpg, png, csv, and excel.

Zhang and Vishnevsky (2020) conducted a literature review on the most recent research pertaining to Big Data Analytics (BDA) and its applications in the medical field. They discussed the ways in which the advancement of technologies and analytical tools has significantly altered the landscape of the healthcare industry. When we look into the future, we anticipate that the Healthcare industry will continue to be affected by the innovative disruptions that are brought about by the developments in Big Data Analytics. Because the healthcare industry is most accurately portrayed as an ecosystem consisting of a great number of interconnected stakeholders, it is essential to maintain a comprehensive and fluid perspective when analyzing the rapidly advancing technologies and business applications associated with big data analytics. In addition, researchers and practitioners have a responsibility to be vigilant and cautious regarding new risks associated with Big Data Analytics. Some examples of these new risks include cyber risk and issues related to regulatory compliance. They strongly recommend that organizations conduct thorough cost-benefit analyses as well as scenario planning before implementing new analytical tools and new technological advancements. This will allow for the mitigation of unintended consequences.

A technology evaluation using Satellite Big Data Analytics was carried out by Nagendra et al. (2022) for the Agri-Insurance sector of India. They present a case study on the assessment of using satellite big data as a technology deployed in Northern India to solve the afore-mentioned problems between the stakeholders in the agri-insurance claim settlement process in this article. The technology was deployed in order to solve the problems between the stakeholders in the agri-insurance claim settlement process. Agri-insurers are able to conduct an assessment for the purpose of calculating the indemnity payments by using satellite big data-based analytics because it provides an independent data source and decision-making platform. The findings demonstrate how the increased transparency made possible by satellite big data analytics helps to prevent the plausible exploitation of the claim settlement process and ultimately results in increased efficiency and effectiveness in the process of settling farmer claims.

METHODOLOGY

The study employed a desktop methodology. Desk research refers to secondary data or that which can be acquired without fieldwork. Desk research is essentially concerned in gathering data from existing resources so it is frequently considered a low-cost strategy as opposed to field research, as the main expense is involved in executive's time, telephone costs and directories. Thus, the study relied on already published studies, reports and statistics. This secondary data was easily obtained through the online journals and library.

RESULTS

The results were grouped into various research gap categories namely as knowledge and methodological gaps.

Knowledge Gaps

Studies by Rizvi (2021), Ngiri (2021), Kaur, Sharma and Mittal (2018), Mehta and Pandit (2018), Banu (2022), Senthilkumar et al. (2018), Yange (2019), Nagendra et al. (2022) and Zhang and Vishnevsky (2020) had knowledge gap. In addition, all the mentioned studies did not establish the implications for big data analytics on claims fraud management in insurance sector. Therefore, the current study seeks to address these knowledge gaps.

Methodology Gaps

Studies by Rizvi (2021), Ngiri (2021), Kaur, Sharma and Mittal (2018), Mehta and Pandit (2018), Banu (2022), Senthilkumar et al. (2018), Yange (2019), Nagendra et al. (2022) and Zhang and Vishnevsky (2020) had methodological gap. In addition, all the mentioned studies did not employ desktop review methodology. Therefore, the current study seeks to address these methodology gaps.

CONCLUSION AND RECOMMENDATIONS

Conclusions

Experts in risk and fraud at insurance companies, as well as executives in actuarial and underwriting, and insurance business managers, are unanimous in their belief that big data analytics has the potential to deliver a significant benefit by assisting in the prevention of attempted fraud and reducing the number of fraudulent claims. According to the findings of the study, Big Data Analytics in the insurance industry is becoming a promising field for providing insight from extremely large data sets, as well as improving outcomes while simultaneously reducing costs. It has a lot of potential, but first there are some obstacles that need to be conquered. The analysis of large amounts of data allows for a rapid search for patterns in previous claims and the identification of similarities or the raising of questions in a new claim before the process has progressed too far. According to the findings, big data analytics had a beneficial and noticeable impact on the process of claims fraud management in the insurance sector.

Recommendations

According to the findings of the study, insurance companies should make it a priority to find ways to automate their claim procedures. In addition, the research suggests that insurance companies should initiate certain steps toward constructing digital insurance control mechanisms. They strongly recommend that organizations, before adopting new technologies and new analytic tools, conduct thorough cost and benefit analyses as well as scenario planning in order to address any unintended consequences that may arise.

REFERENCES

Banu, A. (2022). Big Data Analytics–Tools and Techniques–Application in the Insurance Sector. In *Big Data: A Game Changer for Insurance Industry* (pp. 191-212). Emerald Publishing Limited.

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical science*, *17*(3), 235-255.

Corlosquet-Habart, M., & Janssen, J. (Eds.). (2018). *Big data for insurance companies*. John Wiley & Sons.

George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big data and data science methods for management research. *Academy of Management Journal*, *59*(5), 1493-1507.

Gill, J. K. (2020). Health insurance fraud detection.

Hargreaves, C. A., & Singhania, V. (2015). Analytics for Insurance Fraud Detection: An Empericial Study. *American Journal of Mobile Systems, Applications and Services*, *1*(3), 223-232.

Kaur, P., Sharma, M., & Mittal, M. (2018). Big data and machine learning based secure healthcare framework. *Procedia computer science*, *132*, 1049-1059.

Mehta, N., & Pandit, A. (2018). Concurrence of big data analytics and healthcare: A systematic review. *International journal of medical informatics*, *114*, 57-65.

Nagendra, N. P., Narayanamurthy, G., Moser, R., Hartmann, E., & Sengupta, T. (2022). Technology Assessment Using Satellite Big Data Analytics for India's Agri-Insurance Sector. *IEEE Transactions on Engineering Management*.

Ngiri, S. M. (2021). *Effect of Technology Adoption on Underwriting Processes among Top Five Insurance Companies in Nairobi County* (Doctoral dissertation, United States International University-Africa).

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems*, *2*(1), 1-10.

Rizvi, S. R. Z. (2021). Role of Big Data in Financial Institutions for Financial Fraud. *Available at SSRN 3800777*.

Senthilkumar, S. A., Rai, B. K., Meshram, A. A., Gunasekaran, A., & Chandrakumarmangalam, S. (2018). Big data in healthcare management: a review of literature. *American Journal of Theoretical and Applied Business*, *4*(2), 57-69.

Srivastava, U., & Gopalkrishnan, S. (2015). Impact of big data analytics on banking sector: Learning for Indian banks. *Procedia Computer Science*, *50*, 643-652.

Subramaniam, G., & Mahmoud, M. A. (2021). Fraud Detection in Shipping Industry using K-NN Algorithm. *International Journal of Advanced Computer Science and Applications*, *12*(4).

Yange, T. S. (2019). A Fraud Detection System for Health Insurance in Nigeria. *Journal of Health Informatics in Africa*, *6*(2), 64-73.

Zhang, X., & Vishnevsky, B. (2020). Big Data Analytics Applied to Healthcare.