Volume04, Issue 07, 2024,
Publish Date: 29-07-2024
Doi https://doi.org/10.55640/ijbms-04-07-03

INTERNATIONAL JOURNAL OF BUSINESS AND MANAGEMENT SCIENCES (Open access)

A QUANTITATIVE ANALYSIS OF HEALTHCARE FRAUD AND UTILIZATION OF AI FOR MITIGATION

Md Abu Sayem
University of North Alabama, USA

Nazifa Taslima
University of North Alabama, USA

Gursahildeep Singh Sidhu
University of North Alabama, USA

Dr. Jerry W. Ferry
Professor Emeritus of Accounting, University of North Alabama, USA

ABSTRACT

Healthcare fraud is an emerging and prevalent problem that threatens the reputability of the healthcare system, leading to significant financial charges and disrupting the patient's care. Conventional fraud prevention techniques include manual audits and rule-based systems, which are no longer adequate in the contempt of sophisticated fraud schemes. The advent of advanced technologies like Artificial Intelligence contributes to new opportunities to confront healthcare fraud more effectively. AI-powered solutions include voice biometrics and scrutinizing distinctive identifiers like patterning voice to detect fraudulent activities with greater efficiency and accuracy in contrast to conventional methods. By leveraging Machine Learning algorithms, these systems could incessantly detect fraud patterns and curtail the risk of false positives, improving the overall effectiveness of fraud detection. The research has attempted to exemplify AI implementation in providing accessibility and availability for reliance aid in the healthcare system by gauging its effectiveness in fraud detection. The research presents a comprehensive quantitative scrutiny of AI facilitation over the healthcare system's security threats for mitigation. Since building large-scale labelled Medicare datasets, a data-centric approach empowers healthcare providers to reduce paperwork and time-consuming settlements for policyholders. The present research outcome has proven that applying AI-based fraud mitigation strategies could significantly influence the healthcare industry through quantitative analysis. Hence, by enhancing and automating fraud detection capabilities, healthcare organizations could maintain their capital resources, protect patient data, and retain public reliance. Moreover, the proposed results highlight AI's potential to transmit the prospect of healthcare fraud prevention, facilitating a more efficient and secure healthcare system.

KEYWORDS: Healthcare fraud prevention, Artificial Intelligence, Patient care, Machine Learning, Medicare.

INTRODUCTION

Healthcare fraud detection algorithms are typically based on machine learning techniques to disburse claims information about the transaction to detect fraudulent activities (Kumaraswamy et al., 2024). The prevalent issues of healthcare fraud adversely influence the patient's care and economic resources. The transformational impending of advanced technology like Artificial Intelligence (AI) approaches employed for detecting healthcare fraud. Conventional healthcare fraud detection techniques such as Manual audits, statistical analysis, and rule-based systems are labour-intensive, time-consuming, and futile in identifying fraud schemes. These technologies could not detect fraudulent activities accurately, leading the fraudsters to gain benefits for sustained periods.

Consequently, healthcare fraud deterrence has transformed into AI utilization, especially machine learning (ML). Complex datasets could be processed rapidly utilizing AI algorithms that could spot abnormalities and patterns in concurrent periods (Singh, 2024). The healthcare sector focused on regulating the growth rate by reducing and detecting extra expenses, especially fraudulent billing. Healthcare fraud exposed that the primary loss to medical infrastructure occurs because of the service providers. Service providers can perpetrate fraud in all ways, like billing for goods or services that are delivered, performing redundant operations, or prescribing redundant pharmaceuticals. These deceptions are arduous to examine; various healthcare sectors do not have standardized practices to examine fraudulent activities (Settipalli & Gangadharan, 2021). Figure 1 illustrates the healthcare fraud detection market size in the United States.

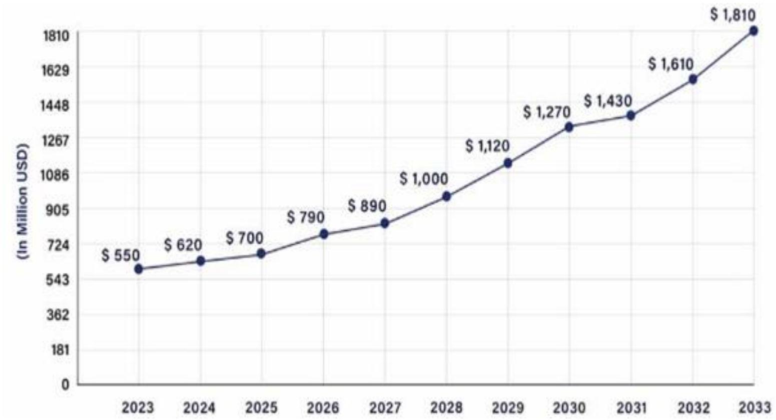

Figure.1 Healthcare fraud detection market size in the U.S. (Research, 2024)

Figure 1 denotes the Healthcare fraud detection market size in the United States, assessed at 550 million dollars in 2023. It is projected to attain 1,210 million dollars in the year 2033. The market is aimed to grow at a 12.70% CAGR from 2024 to 2033.

Healthcare fraud is a ubiquitous and inflated issue in the United States, with predicated annual impairments ranging from three hundred to thirty billion dollars for total healthcare expenses. The fraud scheme has outstripped conventional fraud prevention techniques, leading to the highest possibility of false favourable rates and inflexibility. Therefore, technologies like machine learning and artificial intelligence provide feasible solutions to resist healthcare fraud. Research emphasizes AI utilization for healthcare fraud detection utilizing Medicare claims data. The influence of data contextualization and different machine learning algorithms on fraud classification performance and inspection of AI-powered solutions comprised of voice biometrics in augmenting fraud detection competencies. Henceforth, the competence of AI technologies in the healthcare sector reinforces the detection of healthcare fraud. It is an agile and comprehensive technique for fraud prevention in the healthcare sector, as it can scrutinize unstructured data like textual material and medical notes that would offer informative data for detecting fraudulent activities. Therefore, Healthcare fraud prevention significantly impacts economic crime, leading to fraudulent activities. ML and AI play a vital role in mitigating healthcare fraud by examining copious amounts of healthcare data to classify patterns of fraudulent activities. These patterns comprised unintentional and intentional fraud, like fraudulent billings for services not provided (Filippello, 2022). ML and AI models could predict which claims are frequently considered fraudulent, allowing healthcare sectors to take preventive control measures. In addition, association among providers, payers, drug manufacturers, government agencies, and payers is an influential factor in fraud risk management. The study outcomes will render insights on AI and ML potential in improvising healthcare fraud mitigation and detection and notify the advancement of effective fraud prevention strategies to provide effective and proactive solutions for patient care, payers, and reliability of the healthcare sector.

1.1 Background of the study

Healthcare fraud is a ubiquitous, challenging factor that results in substantive capital forfeit in the healthcare sector. The Centres for Medicare & Medicaid Services (CMS) estimates that healthcare fraud in the United States outlays around three hundred billion dollars annually, considered the country's most substantial form of deception (Stowell et al., 2020). These multifactorial problems comprise healthcare insurance companies, patients, and providers. Healthcare deception includes various forms such as pharmaceutical fraud, fraudulent billing, Medicaid and Medicare fraud, and falsifying medical identity. Pharmaceutical fraud embraces the illegal sale and distribution of controlled substances as well as recommending false claims for prescription medications. Falsification of medical identity reveals the patient's details to an anonymous person. Fraudulent billing comprises false or inflated health care claims like unnecessary treatments and procedures. The significance of healthcare deception is devastating and wide-ranging. Financial expenses are the primary concern for healthcare fraud, influencing healthcare sectors, taxpayers, and companies. Another challenging factor for healthcare fraud is undermining the patient's care, which leads to the denial of necessary medical treatment. In addition, the healthcare deception results in less patient satisfaction and mistreatment procedures.

The conventional techniques for detecting healthcare fraud encompass manual audits, machine learning, data mining, and rule-based systems. The manual audit involves manually reviewing claims by healthcare providers. This time-consuming technique provides erroneous solutions (Beyer et al., 2024). The data mining method is employed to classify the anomalies and patterns in claims data, but these techniques are restricted to healthcare data complexity. Machine learning models become the most promising technique for detecting healthcare deception. Hence, these models need a large amount of high-quality data to enhance fraud detection accuracy, maximize efficiency, and improve patient care by preventing and detecting fraudulent activities. AI tools aid the providers in detecting fraudulent claims more effectively.

Through the automation process, the providers could focus on the patient's care rather than manual audits. AI models reduce false positives and ensure authorized claims are not unnecessarily identified as fraudulent (Ezeji, 2024). This rationalizes the administrative processes for healthcare providers. The mitigation and regulations of economic risks related to fraudulent activities are adhered to through AI assistance. It assists in maintaining an ethical billing practice, and collaborations among payers, law enforcement, and payers of the healthcare sector led to the sharing of data and insights that enhance comprehensive fraud prevention. Effective fraud prevention minimizes unnecessary costs that benefit patient's financial services. Allocating proper resources to patient care optimizes fraudulent activities and ensures the funds gained on medical needs. Therefore, AI utilization in healthcare to prevent fraudulent activities could benefit from early fraud detection and aid in rapidly detecting abnormal transactions or patterns to minimize the probability of financial expenses. Also, AI modernizes the manual data monitoring process and accurately examines many datasets in healthcare rather than conventional methods (Najjar, 2024). Using technologies like AI and ML identifies and averts fraudulent activities in the healthcare sector's reputation and finances.

1.2 Problem statement

Healthcare deception impacts the economy and patient care significantly. Conventional techniques cannot sustain advanced fraudulent schemes, demanding the advancement of practical AI-based approaches. Scholars focus on data-centric techniques, exploiting the inclusive Medicare claims data for supervised learning. Advanced AI algorithms such as Decision Trees, Random Forests, and Logistic Regression play a significant role in detecting fraudulent activities. Conversely, a holistic methodology that syndicates robust system design and cost-effective, secure, and augmented datasets is essential to alleviate healthcare care effectively.

1.3 Objective of the study

The present study's objectives are mentioned below.

- To overview the fraudulent cases reported in the health care system of the USA.
- To demonstrate the technologies employed in the mitigation of fraudulent billing.
- To analyse the beneficiaries and challenges for the implementation of AI in the health care system

- To evaluate the impact of machine learning technologies on the identification of healthcare frauds and elevation of security.
- To recommend the framework for the effective implementation of AI to eradicate fraud in the health care system and prevent financial losses to the nation.

1.4 Research Hypothesis

The research hypothesis of the present study has been emphasized below.

Hypothesis 1

$H_1$1: There has been a significant increase in fraudulent cases reported in the healthcare system of the USA over the past eight years.

$H_0$1: There is no significant increase in fraudulent cases reported in the healthcare system of the USA over the past eight years.

Hypothesis 2

$H_1$2: The implementation of advanced technologies such as artificial intelligence and machine learning can significantly reduce fraudulent billing in the USA healthcare system.

$H_0$2: The implementation of advanced technologies such as artificial intelligence and machine learning will not reduce fraudulent billing in the USA healthcare system.

Hypothesis 3

$H_1$3: Implementing AI will improvise the beneficiaries and challenges in healthcare system.

$H_0$3: Implementing AI will not improvise the beneficiaries and challenges in healthcare system.

Hypothesis 4

$H_1$4: Machine learning technologies will impact the identification of healthcare frauds and the elevation of security.

$H_0$4: Machine learning technologies will not impact the identification of healthcare frauds and the elevation of security.

1.5 Significance of the study

Deploying Artificial Intelligence (AI) in healthcare fraud prevention is advantageous. As instances of continuously evolving fraudulent activities, AI renders powerful tools for mitigation and detection. By scrutinizing copious amounts of healthcare data, refined AI algorithms could classify the indicative patterns of fraudulent activities like billing for not-provided services or duplicate claims. Predictive models allow hands-on measures, enabling the healthcare sector to safeguard patient's well-being and capital resources. AI's capacity and accuracy for predictive analytics significantly impact mitigating fraud risks and financial integrity.

1.6 Paper Organization

The paper is systematized in the subsequent sequence; section 1 provides a detailed introduction to healthcare fraud and the utilization of AI for mitigation. Besides, the introduction section exemplifies the significance of the research. In section 2, prevailing research works related to the current study will be reviewed. The present study's methodology will be illustrated in section 3. In section 4, the outcome of the analysis will be discussed. In section 5, the outcome of the analysis will be discussed with prevailing studies. Finally, in section 6, the brief conclusion regarding the current study will be discussed along with its limitations and future recommendations.

LITERATURE REVIEW

Healthcare deception is a significant threat in the healthcare sector. Machine learning models have influenced the identification and prevent healthcare fraud. The literature review emphasizes the significance of Artificial Intelligence and Machine Learning models in healthcare fraud and their potential to modernize fraud prevention in the healthcare sector. The existing research (Mohammed & Rahman, 2024) analyzed the utilization of Artificial intelligence in detecting deception in private organizations in

Saudi Arabia and understanding the possibilities and obstacles organizations encounter when employing AI technologies for fraud detection. It performed a mixed approach. Integrating AI, data analytics, and ML algorithms effectively mitigates and identifies fraudulent activities.

Similarly, the prevailing research (Zanke, 2023) described the AI-driven fraud detection mechanism as the emerging tool in preventing the healthcare sector from fraudulent activities. The study's outcomes highlight the development of machine learning algorithms, data analytics, and anomaly detection approaches directing the progression of fraud-detecting mechanisms. Also, it explores the implications of AI adoption on preventing deception strategies, patient trust, and organizational risk management. The healthcare sector is considered the influential sector where health and finance data can be assembled. The primary deception occurs in the healthcare sector because of enhanced electronic payment and credit card utilization. Monitoring credit card fraud has been considered a challenging factor in capital conditions for various service providers. The classical study (Mehbodniya et al., 2021) consists of different deep learning and machine learning techniques employed for detecting fraudulent in credit cards, and various algorithms like Logistic Regression, Sequential Convolutional Neural Networks, Random Forest, and Naive Bayes are employed for training the abnormal features and other standard of transaction for detecting the fraudulent in the credit card.

Healthcare deception is an international problem that impacts both developing and developed nations; the prevailing study (Amponsah et al., 2022) employed machine learning and blockchain technology to prevent and detect fraudulent activities in the healthcare sector. A decision tree classification algorithm is utilized to categorize the original claims dataset. The extracted data is programmed in the Ethereum blockchain to prevent and detect deception activities in the healthcare sector. Transforming a centralized mechanism to a decentralized blockchain-based mechanism could provide efficiency, high data integrity, and security in claims processing in the healthcare sector. Correspondingly, the existing research (Matloob, Khan, & Rahman, 2020) illustrated fraud deception in the healthcare sector as a significant risk factor attributable to the heterogeneous landscape of healthcare records. The scammers act like regular patients, and in the time interval, they change their planting fraudulent activities. This study adopted the framework for detecting fraudulent activities through prefix span sequence mining techniques and Bayes rule for inhabiting rare sequences and frequency in the sequence rule engine predicated on the patient time series trace of sequence database and specific patient traces for speciality (Matloob et al., 2020). The classical study (Mackey, Miyachi, Fung, Qian, & Short, 2020) developed a blockchain utilised to examine claims transactions and data in an unchangeable format and ensure the patients perform as a validating node for preventing and detecting healthcare fraud.

According to (Li, Lan, Xu, and Zhu, 2022), healthcare deception impacts the entire healthcare sector. Classifying the behavioural traits of fraudulent claims could assist the advancement of machine learning and artificial intelligence in detecting fraud in the healthcare sector. This study comprised a theoretical framework for identifying fraud in the medical sector from the following three aspects: expenses, time, and quantity. The structured equation modelling technique is implemented to authenticate the framework with large-scale healthcare records. The technique employed in this study is considered a powerful tool for detecting medical fraud in both private and public indemnity. The existing study (Ismail & Zeadally, 2021) considered medical health insurance fraud as a primary concern in the healthcare sector in the United States; health insurance resulted in ten billion losses annually because of healthcare deception. Certain fraud types of influence patient's health. This study implemented blockchain technology to examine healthcare claims in an immutable, transparent, and secure manner. According to (Zhang, Xiao, and Wu, 2020) research, fraudulent activities have predicted that 10% of healthcare mechanism expenses are exploited due to deception. It employed a neural network with totally connected sparse conversion and layers to calculate the disease drug association score on the anomaly detection that depends on the association score. The conventional study (Kapadiya et al., 2022) illustrated that healthcare fraud in health insurance aids individuals in covering the expenses of healthcare services in times of emergency period. It offers finance backup in contrast to the liability risk factor.

Similarly, the prevailing research (Ahmed, Xi, Hou, Shah, & Hameed, 2023) discussed healthcare insurance and numerous benefits that could encounter different problems with privacy, fraudulent activities, and security. To evade the issues, the study implemented Big Data Analytics (BDA), which is effective in the healthcare sector, resulting in mounting growth of data and technology advancements. Data integration from various sources and employing modernized approaches significantly impact healthcare by enhancing diagnostic accuracy, aiding tailored prescriptions, and improving patient outcomes. Classical research (Johnson & Khoshgoftaar, 2023) utilized techniques for detecting deception in healthcare providers; they could safeguard healthcare costs by billions of dollars and enhance the quality of patient care comprehensively. The research implemented data-centric techniques to increase the reliability and performance of healthcare fraud classification through Medicare claims data. The study outcomes exhibited that the new augmented datasets reliably overtake the original Medicare dataset, which is presently utilized by utilizing the search. Also, the outcome emboldens the data-centric machine learning workflow and contributes a robust framework for data insights and training approaches for machine learning techniques in the healthcare sector to prevent fraudulent activities.

2.1 Research gaps

- The existing study (Mohammed & Rahman, 2024) needed further exploration to understand the specific AI competencies and skills in which internal auditors and fraud examiners should gain the knowledge to address data privacy and cyber security challenges.
- Classical research (Mackey et al., 2020) defined verified claims transaction logs as validated and efficient workflows. The mechanism of the patient's validation layer is not available in the legacy claims reimbursement system.
- According to (Zhang et al., 2020), the study failed to attain enough training data comprehensively due to security and privacy problems. The medical record history is not in evident, which would lead to misapprehended results.
- The prevailing study (Kapadiya et al., 2022) should focus more on the private and public healthcare sectors to examine real-world HIC data. The availability of real-world HIC datasets has been constricted because of competitive concerns and legal privacy.
- Classical research (Ahmed et al., 2023) failed to discuss pivotal components of healthcare, such as hospital manageability and patients' specific problems.

METHODOLOGY

3.1 Research design

The research framework provides a comprehensive design of the research work. The method for providing an outlined framework on which the study will be carried out will be defined. The approach is followed in present study to gauge research objectives by congregating and analysing the collected data. The present research will actuate with quantitative technique for collecting data in regard of research questions and study variables. As well, the predefined data of the quantitative approach will be employed to assemble data from the secondary sources. Quantitative data for the research are sourced from secondary sources, including healthcare sector reports, project databases, scholarly literature, and kaggle. The research strategy comprised keywords such as healthcare fraudulent cases, USA healthcare fraud prevention, and fraudulent billing, ensuring a complete retrieval of related research published between 2020 - 2024 (David & Bommu, 2024). Figure.2 illustrate the research framework employed in present study.

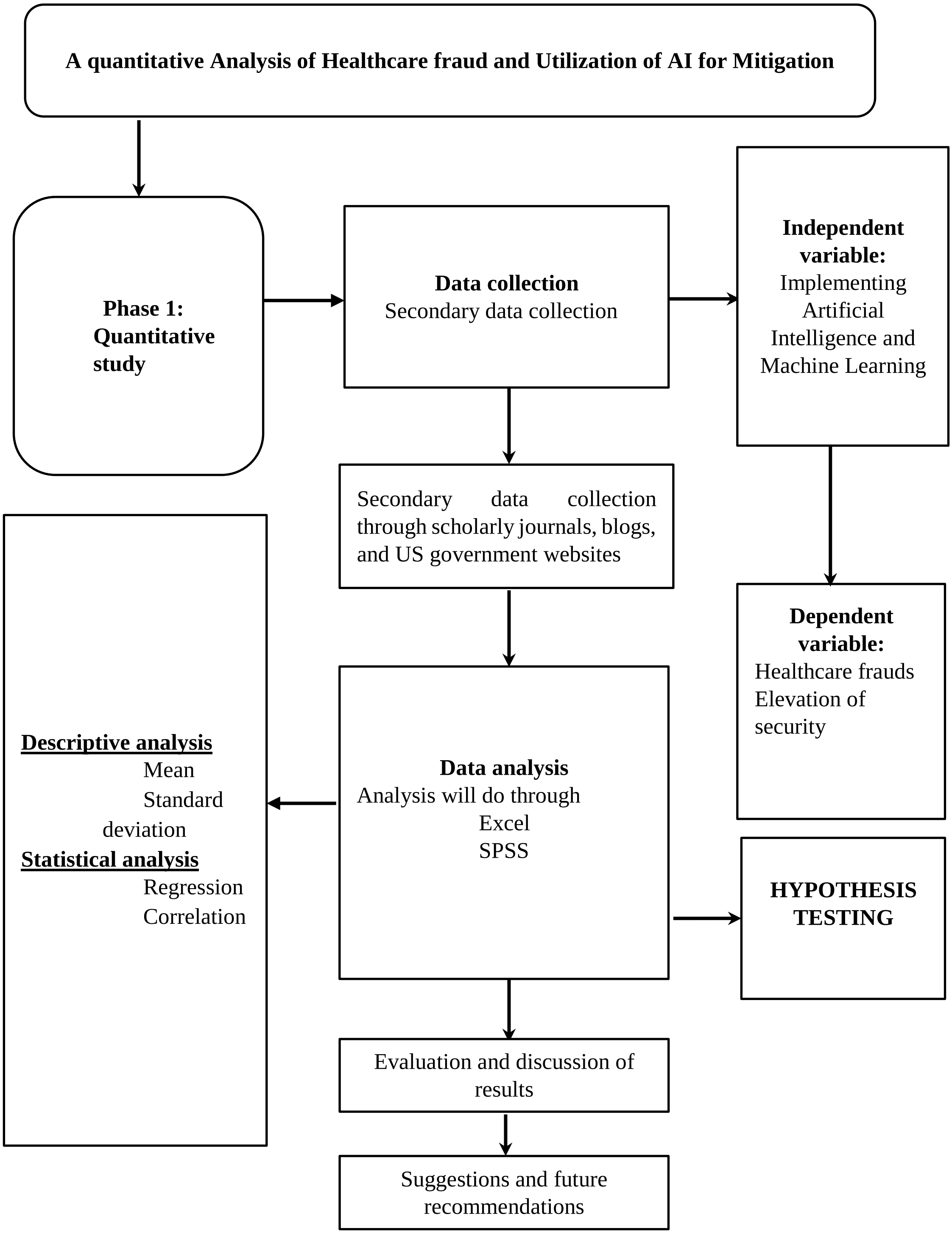
A quantitative Analysis of Healthcare fraud and Utilization of AI for Mitigation
Phase 1: Quantitative study
Data collection
Secondary data collection
Independent variable:
Implementing Artificial Intelligence and Machine Learning
Secondary data collection through scholarly journals, blogs, and US government websites
Dependent variable:
Healthcare frauds
Elevation of security
Descriptive analysis
Mean
Standard deviation
Statistical analysis
Regression
Correlation
Data analysis
Analysis will do through
Excel
SPSS
HYPOTHESIS TESTING
Evaluation and discussion of results
Suggestions and future recommendations

Figure.2 Research Framework

3.2 Secondary data

The secondary data collection indicates accumulating information that is available in open source. The data was earlier collected data, has endured with statistical analysis, and the data is not maintained by the scholars. These data are typically gathered from primary sources and later provided as open-sourced data. Specifically, the secondary data is information that is collected from third parties. The scholars might collect data from various sources. The prevailing data is usually organised and congregate the data for enhancing comprehensive research efficacies (Mazhar, Anjum, Anwar, & Khan, 2021).

3.3 Research instrument

Secondary data for research will be congregated from various sources such as healthcare sector publications, U.S. government reports, and scholarly articles. The data will provide in-depth insights into the present state of healthcare fraud, which influences the healthcare sector. It encompasses statistics on the frequency of healthcare fraud, economic losses, and types of fraudulent activities caused by healthcare providers. The data will include statistics on present approaches to prevent and detect fraud using artificial intelligence and machine learning. The secondary data also consists of effective fraud prevention and detection strategies and the limitations and challenges of these techniques.

QUANTITATIVE DATA ANALYSIS

4.1 Data collection

Data collection is collecting information from all the pertinent sources that provide solutions to the research problem (Ohme et al., 2024). The data is congregated from secondary data, which different scholars have already gathered. This data is available and can be used to support analysis or research without requiring new data collection. Secondary data could save resources and time rather than congregating primary data. Moreover, the relevance and accuracy of secondary data differ based on the purpose and source of the research.

4.2 Data analysis method

Data analysis inspects, interprets, cleans, and transforms raw data to retrieve significant insights. Quantitative analysis is the methodical phenomenon of collecting data and performing mathematical, statistical, and computational approaches (Pilcher & Cortazzi, 2024). The quantitative approach congregated data from scholarly articles, healthcare reports, and blogs. The outcome of the quantitative approach is determined numerically. The numerical values are interpreted, and the impending research is forecasted with suitable modifications. The quantitative data analysis techniques employed for examined data have been congregated from the secondary source—the software tool known as SPSS is applied to examine the study variable in an M.S. Excel sheet. The study's finding is assessed using four approaches: Correlation, regression, descriptive, and graph analysis.

RESULTS

5.1 Presentation of findings

The statistical outcomes utilizing quantitative research methodology have been illustrated in this section. The secondary data are analyzed using SPSS. These formulated outcomes are represented in graphs and charts in this section.

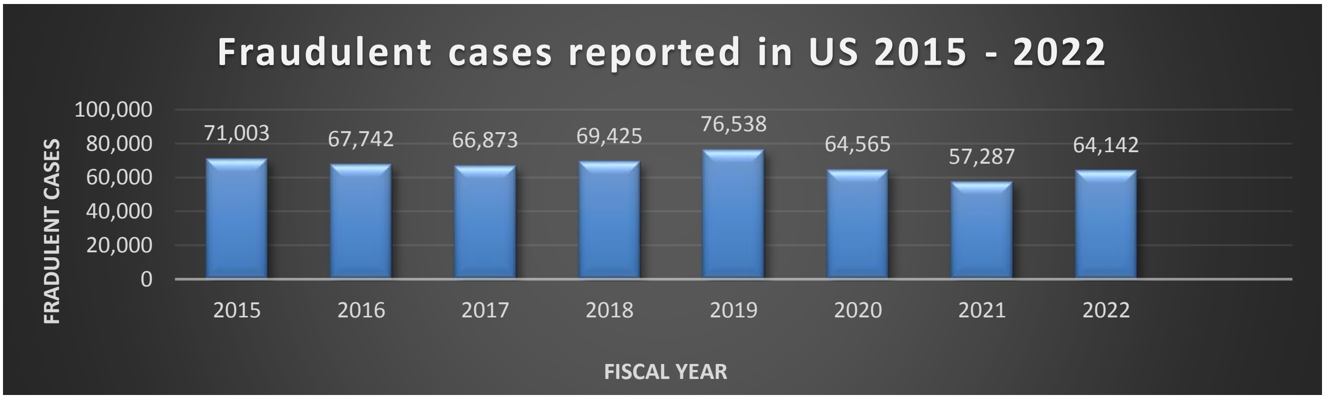

Figure.3 Fraudulent cases in healthcare sector United States 2015 – 2022

From the comprehensive observation of the graph demonstrated in Figure.3, fraudulent cases were reported in the United States. In 2015, 71,003 cases were reported in the US. However, 67,742 and 66,873 cases were reported in 2016 and 2017, respectively. In 2018, 69,425 cases were reported. Similarly, 64,565 cases were reported in 2020. From the above statistical data, fraudulent cases were more highly reported in 2019 than in other years. In 2021, fewer cases were reported in the United States.

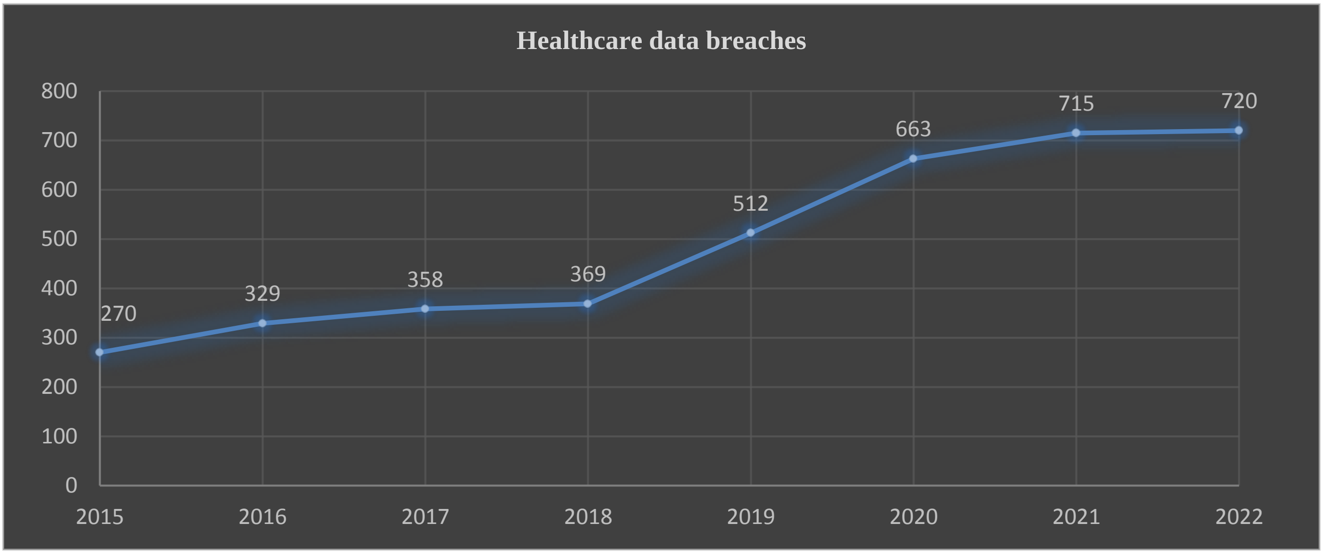

Figure.4 Healthcare data breaches

Figure.4 illustrated in Figure 4 represents the data breaches in the healthcare sector from 2015 to 2022. In 2015, 270 data breaches occurred, and in 2016, 329 occurred. Likewise, 358,369,512,663,715 data breaches occurred in 2017, 2018, 2019, 2020, 2021, and 2022. According to the statistical data provided, healthcare data breaches are gradually increasing in the United States.

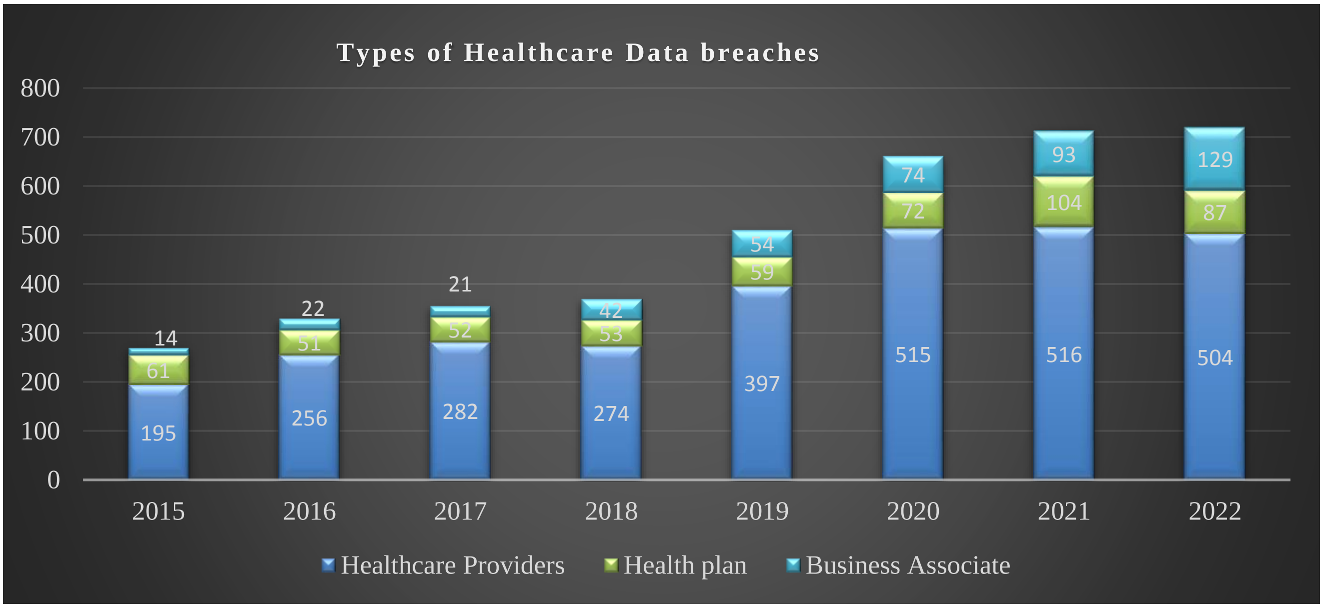

Figure.5 Types of Healthcare data breaches

Figure.5 exemplified the data breach types in healthcare sector from the year 2015 -2022. Some of the data breaches are categorised as healthcare providers, health plan and business associates. In 2022, the highest number of fraudulent activities were done by healthcare providers. Similarly, 104 health plans impact the healthcare sector, and in 2022, 129 business associates influenced the healthcare sector with fraudulent activities in the US.

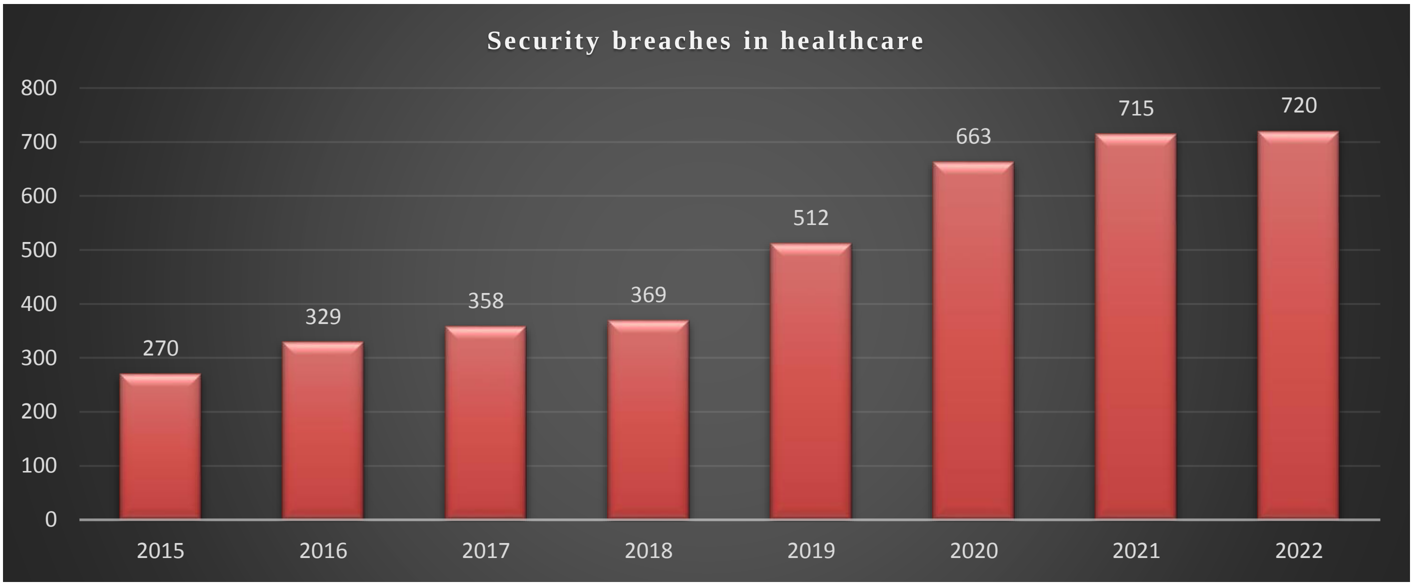

Figure. 6 Security data breaches in the healthcare sector

Figure.6 represented the security data breaches in the healthcare sector from the year 2015 -2022. In 2015, 270 security threats occurred in the healthcare sector. The highest security threat happened in the year 2022, and the lowest security threat occurred in the year 2015 in the US healthcare sector.

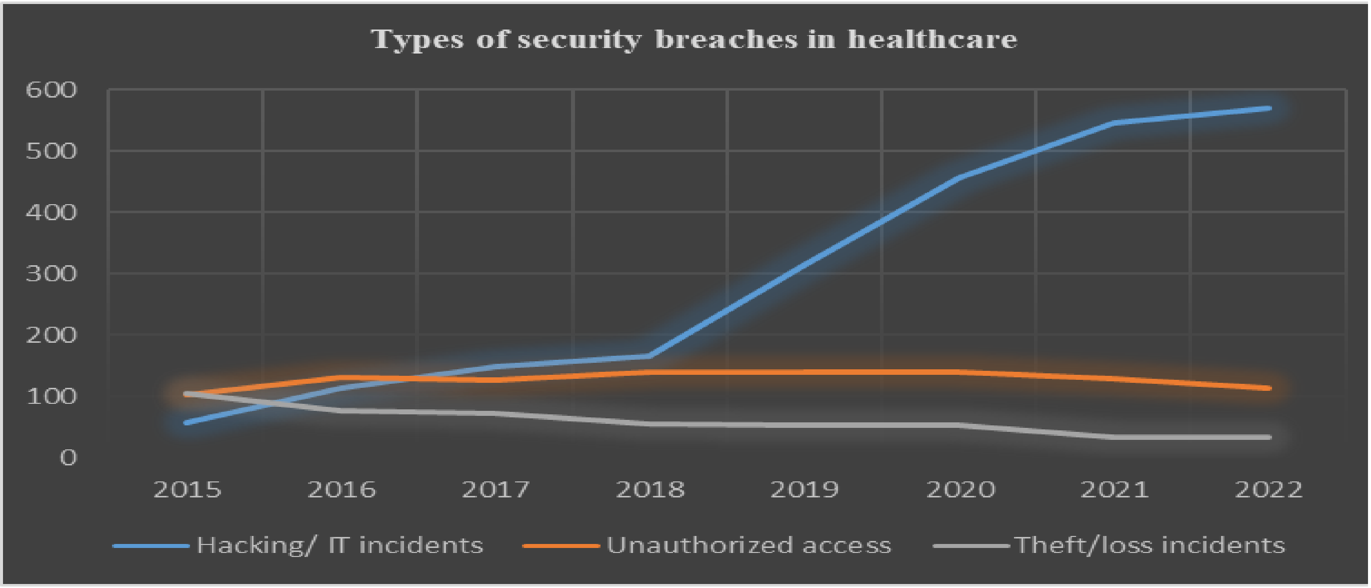

Figure.7 Types of security data breaches in the healthcare sector

Figure.7 illustrated the security data breaches types in the healthcare sector from 2015 to 2022. Some of the data breaches are classified as hacking or IT incidents, theft or loss incidents, and unauthorized access. In 2022, 571 of the highest threats occurred through hacking/ IT incidents. Correspondingly, 140 and 105 threats occurred through unauthorized access and theft/ loss incidents, which impact the healthcare sector in the US.

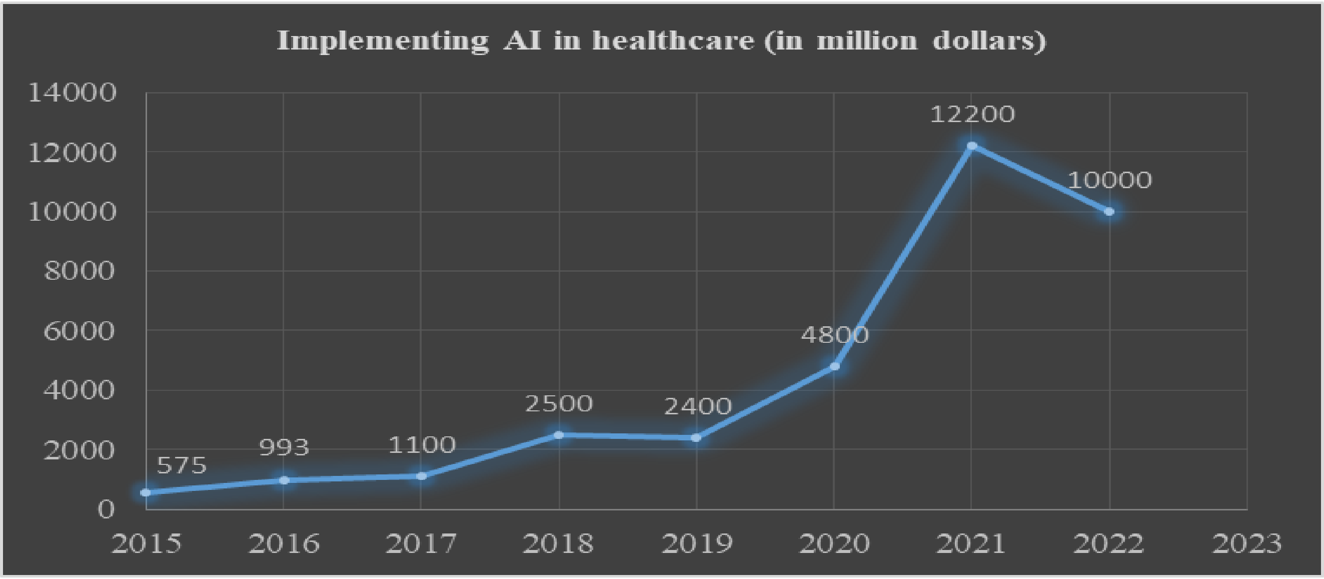

Figure.8 Implementing AI in the healthcare sector (in a million dollars)

Figure.8 exhibited the statistical report on implementing AI technology in the healthcare sector from the year 2015-2022. The highest cost of implementing AI technology in the healthcare sector comprised 12200 million dollars in the year 2021.

5.2 Statistical significance

Data were scrutinized utilizing SPSS (Statistical Package for Social Science). Descriptive statistics were performed for all variables. T- Test were utilized to compare means between the groups. Pearson's correlation coeff were used to examine the association between variables. The statistical significance was estimated as $p < 0.05$. Likewise, the regression analysis was performed to find the relationship between independent variables (Hacking/IT incidents and implementing AI technologies in healthcare) and dependent variables (security data breaches in the healthcare sector)

DISCUSSIONS

6.1 Interpretation of Results

Results extracted and formulated from SPSS tool has been listed in this section. With the objective of examining the prevailing records of the healthcare sector in the United States. Specific statistical analyses such as frequency distribution, correlation, regression, and ANOVA analysis have been implemented.

Hypothesis 1

$H_1$1: There has been a significant increase in fraudulent cases reported in the healthcare system of USA over the past eight years.

$H_0$1: There is no significant increase in fraudulent cases reported in the healthcare system of the USA over the past eight years.

Descriptive Statistics

Descriptive Statistics shows the statistical characterisation on collected datasets in which data are quantified with a central tendency value from the congregated data.

Table.1 Descriptive Statistics

	N	Min	Max	M	Std.dev
Fiscal year	8	2015	2022	2018.50	2.449
Fraudulent cases	8	57287	76538	67196.88	5626.144
Healthcare data breaches	8	270	720	492.00	185.277
Security breaches in healthcare	8	270	720	492.00	185.277
Valid N (list wise)	8				

With the objective of analysing healthcare fraud prevention through implementing advanced AI technologies, descriptive statistics have been performed. The outcomes of descriptive statistics have been represented in Table.1. Here, the highest mean value is 67196.88 for the parameter representing fraudulent activities in the healthcare sector, and the lowest mean value is 492 for the parameter representing healthcare and security data breaches. The minimum value of fraudulent activities happened in the year 2015, and the maximum value of fraudulent activities happened in the year 2022. The inference proves the H11 hypothesis. In addition to, it is contradicted to null hypothesis.
Hence from test analysis, the assessed hypothesis is **$H_1$1**: There has been a significant increase in fraudulent cases reported in the healthcare system of USA over the past eight years.

Hypothesis 2

$H_1$2: The implementation of advanced technologies such as artificial intelligence and machine learning can significantly reduce fraudulent billing in the USA healthcare system.
$H_0$2: The implementation of advanced technologies such as artificial intelligence and machine learning will not reduce fraudulent billing in the USA healthcare system.

Regression analysis

Regression analysis is a statistical method which is utilized to analyse associations between two or more variables. It aids to estimate how deviations in one (independent) variable are related to the modification occurred at another variable.

Table.2 Model Summary

Model	R	R Sq.	Adjusted R Sq.	Std. Error of the Estimate
1	.893[a]	.798	.764	90.037

a. Predictors: (Constant), Implementing AI in healthcare (in a million dollars)

Table.3 ANOVA[a]

Model		Sum of Sqs	df	Mean Sq.	F	Sig.
1	**Regression**	191651.585	1	191651.585	23.641	.003[b]
	Residual	48640.415	6	8106.736		
	Total	240292.000	7			

a. Dependent Variable: Security breaches in healthcare

b. Predictors: (Constant), Implementing AI in healthcare (in a million dollars)

Table.4 Coeff[a]						
Model		Unstandardized Coeff		Standardized Coeff	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	330.414	46.019		7.180	.000
	Implementing AI in healthcare (in a million dollars)	.037	.008	.893	4.862	.003
a. Dependent Variable: Security breaches in healthcare						

Table.2 and Table.3 represent the statistical reports of model summary and ANOVA statistics from the conductance of Regression analysis. Table.4 demonstrated the coefficient table that indicated with significant values of 0.000 and 0.003, which is comparatively less than the default value (0.05). And the measured R Sq. value is .798. The attained values have shown the statistical significance. Hereby, the assumed dependent variable is security breaches, and independent variable is AI implementation at healthcare sector. The outcome has denoted the impact of AI implementation in healthcare sector based on security breaches will optimize security threats in healthcare system. The undergone association between these cases have proven $H_1$2 hypothesis. As well, it contradicts null hypothesis. Hence from test analysis, the assessed hypothesis is **$H_1$2**: The implementation of advanced technologies such as artificial intelligence and machine learning can significantly reduce fraudulent billing in the USA healthcare system.

Hypothesis 3
H13: Implementing AI in the health care system will improve the beneficiaries and challenges.
$H_0$3: Implementing AI in the health care system will not improve the beneficiaries and challenges.

Paired Sample T Test
Paired Sample T Test is tested for grouping the two variables in the same means, if not the paired variables are correlated then each may position to differ paths. Hereby, two pairs are undertaken for the test computation.

Table.5 Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Healthcare data breaches	492.00	8	185.277	65.505
	Healthcare Providers	367.38	8	131.775	46.589
Pair 2	Security breaches in healthcare	492.00	8	185.277	65.505
	Theft/loss incidents	59.75	8	24.034	8.497

Table.6 Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Healthcare Data Breaches & Healthcare Providers	8	.992	.000
Pair 2	Security breaches in healthcare & Theft/loss incidents	8	-.872	.005

Table.7 Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Healthcare data breaches - Healthcare Providers	124.625	57.184	20.218	76.818	172.432	6.164	7	.000
Pair 2	Security breaches in healthcare - Theft/loss incidents	432.250	206.574	73.035	259.550	604.950	5.918	7	.001

Table.5 and Table.6 illustrate the obtained statistical report on Paired Test Sample testing regarding Statistics and correlations calculation. Then, paired sample test analysis on means and sig. value representation is noted in Table.7, which specifically the significant values 0.000 and 0.001, which is comparatively less than the default value (0.05). The paired sample test variables comprised two pairs, namely, healthcare data breaches and healthcare providers, as well as security breaches in healthcare and theft/loss incidents. This instructs the statistical model as significant. The highest mean value consists of healthcare and security data breaches. This shows the differences among healthcare data breaches and healthcare providers and the security breaches as well as theft/loss incidents in the healthcare sector. Henceforth, the implications contradict null hypothesis.

Pearson correlation (PC)

The Pearson correlation coefficient is also known as the product-moment correlation coefficient or Pearson's r. It is a statistical form for calculating direction and strength of the connectivity among two quantitative variables. If the value holds +1, it is perfect positive linear correlation. If it shows -1, it is conveyed as a perfect negative linear correlation, and if it pointed in 0, it considered as a no-linear correlation.

Table.8 Correlations

		Healthcare Providers	Business Associate	Hacking/ IT incidents	Theft/loss incidents	Implementing AI in healthcare (in a million dollars)
Healthcare Providers	PC	1	.898**	.981**	.862**	.838**
	Sig. (2-tailed)		.002	.000	.006	.009
	N	8	8	8	8	8

Business Associate	PC	.898**	1	.959**	.871**	.903**
	Sig. (2-tailed)	.002		.000	.005	.002
	N	8	8	8	8	8
Hacking/ IT incidents	PC	.981**	.959**	1	.880**	.913**
	Sig. (2-tailed)	.000	.000		.004	.002
	N	8	8	8	8	8
Theft/loss incidents	PC	.862**	.871**	.880**	1	.811*
	Sig. (2-tailed)	.006	.005	.004		.015
	N	8	8	8	8	8
Implementing AI in healthcare (in million dollars)	PC	.838**	.903**	.913**	.811*	1
	Sig. (2-tailed)	.009	.002	.002	.015	
	N	8	8	8	8	8
**. Correlation is significant at the 0.01 level (2-tailed).						
*. Correlation is significant at the 0.05 level (2-tailed).						

The results of the correlation analysis have been defined in Table 8. The test considered variables that describe healthcare providers, business associates, hacking/ IT incidents, theft/loss incidents and AI implementation in the healthcare sector. The Pearson coefficient value of 1.000 denotes that a positive association is identified among every variable considered. The values' outcome shows a positive association between healthcare providers, business associates, hacking/IT incidents, and theft/loss incidents in healthcare. It is analysed from the Pearson coefficient's attained value of 1.000. This means that when AI technologies are implemented in the healthcare system, data breaches can be somewhat avoided.

Similarly, the Pearson coefficient value of 1.000 signifies the positive correlation between the variables of determined advantages through AI implementation in the healthcare sector and existing fraudulent activities. The significant value of this connection is remarked with the value of 0.15, which indicates the statistical value significance. In addition, it has been evidenced that variables encompassed to AI usage can enable it to work optimistically in the prevalence of fraudulent activities, exemplifying that AI implementation in the healthcare sector is a beneficiary practice. As the implication has been verified, the H13 and null hypotheses were rejected.

Hence, from the test analysis, the assessed hypothesis is H13: Implementing AI in the health care system will improve the beneficiaries and challenges.

Hypothesis 4

$H_1$4: Machine learning technologies will impact the identification of healthcare frauds and the elevation of security.

$H_0$4: Machine learning technologies will not impact the identification of healthcare frauds and elevation of security.

PC

Table.9 Correlations

		Fiscal year	Implementing AI in healthcare (in million dollars)	Healthcare data breaches	Security breaches in healthcare
Fiscal year	PC	1	.877**	.966**	.966**

	Sig. (2-tailed)		.004	.000	.000
	N	8	8	8	8
Implementing AI in healthcare (in million dollars)	PC	.877**	1	.893**	.893**
	Sig. (2-tailed)	.004		.003	.003
	N	8	8	8	8
Health care data breaches	PC	.966**	.893**	1	1.000**
	Sig. (2-tailed)	.000	.003		.000
	N	8	8	8	8
Security breaches in healthcare	PC	.966**	.893**	1.000**	1
	Sig. (2-tailed)	.000	.003	.000	
	N	8	8	8	8
**. Correlation is significant at the 0.01 level (2-tailed).					

Results from correlation analysis are demonstrated in Table 9 illustration. The test considers variables representing the fiscal year, AI implementation in the healthcare system (in a million dollars), healthcare data breaches, and security breaches in healthcare procedures. The Pearson coefficient value of 1.000 signifies a positive association between every considered variable. Hence, it reveals the positive association between fiscal year, AI implementation in the healthcare system (in a million dollars), healthcare data breaches, and security breaches in healthcare is identified. The value of the Pearson coefficient of 1.000 states this. Thus, it concludes that when AI technologies are held in the healthcare field, the problem of data breaches could be nullified.

Similarly, the Pearson coefficient value of 1.000 signifies the positive correlation between the variables determining the advantages of implementing AI techniques in the healthcare sector to prevent fraudulent activities. The significant value of this relationship is 0.00, which indicates that the model is statistically significant. Additionally, the variables demonstrating AI usage in optimizing data breaches have a Pearson coefficient of 1.000, which exemplifies that when AI is implemented in the healthcare sector.

T- Test

The most general hypothesis tests in the statistics. The t-test defines the differences between the sample mean and the population mean. The T-test comprises three types: independent, paired, and one sample t-test. The one-sample t-test relates a group against the standardized values.

Table.10 One-Sample Statistics

	N	**M**	**Std. Dev**	**Std. Error Mean**
Fiscal year	8	2018.50	2.449	.866
Fraudulent cases	8	67196.88	5626.144	1989.142
Health care data breaches	8	492.00	185.277	65.505
Security breaches in healthcare	8	492.00	185.277	65.505

Table.11 One-Sample Test

	Test Value = 0

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Fiscal year	2330.763	7	.000	2018.500	2016.45	2020.55
Fraudulent cases	33.782	7	.000	67196.875	62493.30	71900.45
Health care data breaches	7.511	7	.000	492.000	337.10	646.90
Security breaches in healthcare	7.511	7	.000	492.000	337.10	646.90

The one-sample test analysis depicted in Table.10 and Table.11 exemplifies the assumed cases are attained their significant value of 0.000, which is less than the default value (0.05). The one-sample test encompasses the variables such as fiscal year, fraudulent cases, healthcare data breaches, and security breaches in the healthcare sector. The specified significant value represents the statistical value significance. This shown differential impacting cause at fraudulent cases and fraud activities. While screening the results, fraudulent cases resulted with highest mean value of 67196.88 rather than fraud activities. Therefore, inferences declared null hypothesis rejection in evaluating hypothesis 4.

Regression Analysis

Table.12 Model Summary

Model	R	R. Sq.	Adjusted R. Sq.	Std. Error of the Estimate
1	.998[a]	.995	.994	13.825

a. Predictors: (Constant), Hacking/ IT incidents

Table.13 ANOVA[a]

Model		Sum of Sqs	df	M. Sq	F	Sig.
1	Regression	239145.212	1	239145.212	1251.209	.000[b]
	Residual	1146.788	6	191.131		
	Total	240292.000	7			

a. Dependent Variable: Security breaches in healthcare

b. Predictors: (Constant), Hacking/ IT incidents

Table.14 Coeff[a]

Model		Unstandardized Coeff		Standardized Coeff	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	224.639	9.001		24.957	.000
	Hacking/ IT incidents	.901	.025	.998	35.372	.000

a. Dependent Variable: Security breaches in healthcare

Table.12 and Table.13 demonstrate the statistical Tabulation report from the conductance of Regression analysis. The coefficient table of regression analysis illustrated in Table.14 wherein, the attained significant values are 0.000 and 0.000 have moderately sets lower to default value (0.05), therefore it conquers statistical significance. These results depicted the impact of hacking and IT applications can intervene into secured progress in the healthcare sector. Also, here the R Sq. value is .995, which has exposed the positive impact of dependent upon the independent variables, as here dependent variable is security breaches at healthcare, and independent variables is settled with hacking/ IT incidents. Henceforth, the conclusion contradicted null hypothesis.

Paired Sample T test

Table.15 Paired Samples Correlations				
		N	**Correlation**	**Sig.**
Pair 1	**Health care data breaches & Health plan**	8	.852	.007
Pair 2	**Security breaches in healthcare & Hacking/ IT incidents**	8	.998	.000

Table.16 Paired Samples Test									
		Paired Differences					**t**	**df**	**Sig. (2-tailed)**
		M	**Std. Dev**	**Std. Error Mean**	**95% Confidence Interval of the Difference**				
					Lower	**Upper**			
Pair 1	**Health care data breaches - Health plan**	424.625	169.244	59.837	283.133	566.117	7.096	7	.000
Pair 2	**Security breaches in healthcare - Hacking/ IT incidents**	195.375	23.940	8.464	175.361	215.389	23.083	7	.000

The paired sample test analysis is depicted in Table. Sixteen specified significant values, 0.000 and 0.000, which is comparatively less than the default value (0.05), so it stipulates statistical significance. The set of paired samples comprised two typical pairs: healthcare data breaches to health plans and the other pair characterized on condition with security breaches in healthcare and hacking/ IT incidents. While Pair 1 accomplished a higher mean value than the other pair. This shows that data breach and security uphold two distinct extensions assigned to the healthcare sector. Also, the acquired inferences validate the existence of the H14 hypothesis and mutually reject the null hypothesis.

Hence, from test analysis, the assessed Hypothesis Statement is H14: Machine learning technologies will impact the identification of healthcare frauds and the elevation of security.

6.2 Comparison with Previous Studies

The existing study (Mohammed & Rahman, 2024) employed a quantitative analysis to quantify the congregated data and eventually to examine the opinion and perception of participants about AI's role in detecting fraudulent activity occurrence in private retail organizations in Saudi Arabia. The prior study undertaken survey enclosed eleven questions schemed to Likert scale options. The core objective of the study investigation is to gather knowledge about AI technologies and insights into AI influences in the detection of fraudulent activities in organization procedures. The total number of participants was 373 out of 5,000 individuals from seven private retail organizations. Remarkably, the responses included 84% appropriated insights, and 78% fortified the implementation of AI notion to the private retail organization. Also, the correlation exposes a modest positive association among AI techniques in improvising fraud detection on various grounds. According to (Lekkala, 2023), the healthcare sector is a gradually evolving field in the contemporary world that encourages technological advances on a factual basis, as the study insisted ML beneficiaries to the healthcare system. This signified a typical framework that delivers a significant enhancement with better efficiency and accuracy in the healthcare sector needing fraud detection. The conventional research (Mazumder et al., 2024) employed secondary data sources from different healthcare sectors that influenced significant improvements in perceiving fraudulent activities, including applying modernized fraud detection mechanisms. Noticeably, the study deliberated that implementing Predictive modelling, data analytics, Natural language processing and machine learning had extensively improved the healthcare sector's ability to detect diverse fraud patterns. The collaborative importation of data analytics also impacted healthcare providers to identify hidden anomalies and patterns in billing systems that the conventional techniques have failed to identify. The findings of the study revealed the capability of the adaptive and dynamic nature of machine learning techniques in the healthcare system.

Similarly, (Jambukar, 2021) also attempted to recognize the fraudulent activities of healthcare providers with the use of MI algorithms. Here, the research proposed a KDD-based method for MI implication that enhanced the performance in the fraudulent reduction in the healthcare system's security assurance on the extensive range of data collection. The research utilized technical statistical methods such as ANOVA and MANOVA sampling techniques to predict better efficiency regardless of the extensive datasets. This is generated through the supervised machine learning frameworks such as SVM, random forest, LogisticGAM, and XGBoosting in the code logic to improve enhanced results in fraud detection. From evaluating and comparing the proposed four models, LogisticGAM was determined to have the highest fraud detection with a recall value of 88%. The development in the healthcare sector has increased extensively to intensify the quality assurance of safety and security. The prevailing study (Mary & Claret, 2023) explored another framework model that enhances the classification and pre-processing phase for the prediction of fraudulent provider identification. Initially, the congregated datasets are pre-processed through a relative risk-based Map Reduce framework, which constructs a structured relationship set among patients, claiming, and disease variables. The classification phase enhances the Recurrent Neural Network (RNN). Henceforth, the experimental outcomes described that the framework provides better accuracy than other techniques. The accuracy, recall, and precision rates comprised 88.09%, 32.80%, and 14.15% with 92.30 seconds of computational time. It is determined that technologically advancing features can enable a better improvisation in the safety reliance of the healthcare system. As to denoting other technologies exceptions to AI, the prevailing research (Mackey et al., 2020) formulated a framework utilizing blockchain to monitor and document transaction claims and data in default format, permitting authentication access, as here, patient validating node. A Blockchain prototype and framework were developed for healthcare fraud prevention utilizing application layers comprised of smart contracts, tokens, governance, consensus algorithms, and critical blockchain tools derived from the Ethereum platform. As a result, the framework designated has enhanced the healthcare claims summaries with blockchain technology for consensus mechanisms in the healthcare system and secured data storage in preventing fraudulent activities; therefore, the operation of the arbitration process is more patient-centric in preventing and detecting fraud access.

6.3 Practical Implications

The present research investigated the augmentation of AI's qualified implication in avoiding fraudulent cases in the healthcare system. Fundamentally, to prevent fraudulent access, assessing upcoding is an essential validation. Further, the conclusions have resulted in employing machine learning and artificial intelligence approaches to foster improved system practices in the healthcare sector. In order to avoid fraudulent activities, the research is concerned with a prospective evolution of AI implementation in the healthcare field that can reliably actuate a safety threat by the action of fraudulent detection and hacking technology. In addition, with reduced healthcare costs and fraudulent billings, the challenging factors like data availability and requirements for continued research will be reduced. Therefore, collaborative efforts of AI in healthcare practices could reinforce fraud prevention in the sector.

6.4 Challenges and Solutions

Healthcare fraud is increasing due to factors such as population development and constantly evolving improvements made researchers demand an updated attitude with modernized techniques and trends uphold fraudulent activities are chiefly engaging in the current era of healthcare structure (Bauder et al., 2017). Healthcare fraud could influence various adverse effects, such as individuals receiving unnecessary medications or treatment, which may lead to significant effects on life protection. To evade these challenges, the present research considerately deliberates on the utilization of technology. Hence, the study emphasizes the operation of AI to prevent fraudulent activity cause in healthcare. It is determined that ML technologies can examine fraud patterns in classifying deceptive practices like duplicate claim submissions and billings that are not legally approved. Also, integrating AI technology into the healthcare sector could ensign endangered claims through prediction that orders in providing preventive control measures. Remarking to broad sense, coordinative communication among payers, law enforcement agencies, and healthcare providers can control preventive fraudulent activities in the past. Even though AI contributes to delivering innovative solutions, maintenance and implementation procedures may require costs and technical knowledge. Hence, the healthcare sector must oversee quality, cost-effectiveness, and safety, along with modernized technologies like ML and AI applications. From the observation of research investigation, it is evidenced that ML and AI could hold effective preventive measures to reduce fraudulent activities, and the healthcare sector can directly implement these challenges prudently to enhance fraud prevention.

6.5 Limitations

Every study has its limitations, and so does the present research. The first limitation of the study is that research has only focused on and declared the solution in the regard of the United States. Thus, the research is geographically constrained. However, AI models depend on high-quality data for precise extrapolations. Biased, incomplete, or outdated data could lead to erroneous results. Fraudsters could falsify AI and ML systems by injecting misleading data, conceding the model's model's integrity. Maintaining, developing, and deploying AI and ML involves resources, and the smaller healthcare sector might need help to deploy these resources. Addressing inequalities, making impartial decisions, and maintaining patient privacy are precarious when utilizing AI for fraud prevention. In addition, congregating primary data for analysis instead of collecting secondary data would have aided in fetching consistent and reliable data for quantitative research.

CONCLUSION

7.1 Summary of Findings

The secondary data analysis exposes that machine learning and Artificial Intelligence technologies embrace the enormous potential for improvising fraud prevention in the healthcare sector. By employing advanced analytics competencies, AI mechanisms could effectively handle large sets of data to detect patterns and anomalies of fraudulent activities compared to conventional techniques such as manual audits and rule-based systems. Integrating ML and AI technology into the healthcare system will allow fraudulent activities without requiring systematic infrastructure service, facilitating capital resources for advanced technologies. Moreover, AI-automated claims could accelerate the verification of claim details, reduce human errors, and quickly flag anomalies or discrepancies. Therefore, the present research positively impacts employing advanced technologies in the healthcare sector to detect fraudulent activities.

7.2 Recommendations

- ✓ ML/AI algorithms could scrutinize large datasets of the healthcare sector to identify fraudulent activities like fraudulent billings for the services not provided.
- ✓ Facilitating collaboration among healthcare providers, law enforcement agencies, and payers to employ AI/ML techniques could improve fraud prevention in the healthcare sector.
- ✓ The evolving development in the healthcare sector should automatically identify fraud prevention using modernized technologies like Blockchain, the Internet of Things, deep learning, and Robotic process automation. These technologies could be leveraged and contributed to optimal accuracy.

7.3 Future Research Directions

Implementing advanced technologies like AI and ML approaches could efficiently identify fraudulent occurrences in the healthcare sector. Comprehensively, this systematic bibliographical review and discussion will aid future research in using AI and ML technologies in various sectors. In addition, the research outcomes will be helpful in the field community in delivering detailed insights to promote healthcare fraud prevention. The depiction of challenges endured is also indicated, which led to further considerations. Also, it directs diverse perspectives in proposing technological advances to healthcare Fraud Detection. Hence, demonstrating a detailed bibliographical review may help researchers convey heightened social wellness possibilities among a broad community.

REFERENCES

1. Ahmed, A., Xi, R., Hou, M., Shah, S. A., & Hameed, S. (2023). Harnessing big data analytics for healthcare: A comprehensive review of frameworks, implications, applications, and impacts. IEEE Access.
2. Amponsah, A. A., Adekoya, A. F., & Weyori, B. A. (2022). A novel fraud detection and prevention method for healthcare claim processing using machine learning and blockchain technology. Decision Analytics Journal, 4, 100122. doi:https://doi.org/10.1016/j.dajour.2022.100122
3. Bauder, R., Khoshgoftaar, T. M., & Seliya, N. (2017). A survey on the state of healthcare upcoding fraud analysis and detection. Health Services and Outcomes Research Methodology, 17, 31-55.
4. Baur, N. Linearity vs. Circularity? On Some Common Misconceptions on the Differences in the Research Process in Qualitative and Quantitative Research.
5. Beyer, B., Draeger, M., & Rapley, E. T. (2024). Audit process ineffectiveness: evidence from audit report errors. Journal of Accounting Literature.
6. David, M., & Bommu, R. (2024). Navigating Cost Overruns in Civil Engineering Projects: AI-Powered Root Cause Analysis. Unique Endeavor in Business & Social Sciences, 3(1), 85-98.
7. Ezeji, C. L. (2024). Artificial Intelligence for detecting and preventing procurement fraud. International Journal of Business Ecosystem & Strategy (2687-2293), 6(1), 63-73.
8. Filippello, S. (2022). Healthcare Fraud Investigations: Overview of Overbroad Investigative Regime and Recommendations for a More Targeted Approach. Annals Health L., 31, 141.
9. Ismail, L., & Zeadally, S. (2021). Healthcare Insurance Frauds: Taxonomy and Blockchain-Based Detection Framework (Block-HI). IT Professional, 23(4), 36-43. doi:10.1109/MITP.2021.3071534
10. Jambukar, A. (2021). Fraudulent Healthcare Providers detection using Machine Learning Algorithms.

Dublin, National College of Ireland,

11. Johnson, J. M., & Khoshgoftaar, T. M. (2023). Data-Centric AI for Healthcare Fraud Detection. SN Comput Sci, 4(4), 389. doi:10.1007/s42979-023-01809-x

12. Kapadiya, K., Patel, U., Gupta, R., Alshehri, M. D., Tanwar, S., Sharma, G., & Bokoro, P. N. (2022). Blockchain and AI-empowered healthcare insurance fraud detection: an analysis, architecture, and future prospects. IEEE Access, 10, 79606-79627.

13. Kumaraswamy, N., Ekin, T., Park, C., Markey, M. K., Barner, J. C., & Rascati, K. (2024). Using a Bayesian Belief Network to detect healthcare fraud. Expert Systems with Applications, 238, 122241. doi:https://doi.org/10.1016/j.eswa.2023.122241

14. Lekkala, L. R. (2023). Importance of Machine Learning Models in Healthcare Fraud Detection. Voice of the Publisher, 9(4), 207-215.

15. Li, J., Lan, Q., Zhu, E., Xu, Y., & Zhu, D. (2022). A study of health insurance fraud in China and recommendations for fraud detection and prevention. Journal of Organizational and End User Computing (JOEUC), 34(4), 1-19.

16. Mackey, T. K., Miyachi, K., Fung, D., Qian, S., & Short, J. (2020). Combating Health Care Fraud and Abuse: Conceptualization and Prototyping Study of a Blockchain Antifraud Framework. J Med Internet Res, 22(9), e18623. doi:10.2196/18623

17. Mary, A. J., & Claret, S. P. A. (2023). Design and development of big data-based model for detecting fraud in healthcare insurance industry. Soft Computing, 27(12), 8357-8369. doi:10.1007/s00500-023-08296-5

18. Matloob, I., Khan, S. A., & Rahman, H. U. (2020). Sequence Mining and Prediction-Based Healthcare Fraud Detection Methodology. IEEE Access, 8, 143256-143273. doi:10.1109/ACCESS.2020.3013962

19. Mazhar, S. A., Anjum, R., Anwar, A. I., & Khan, A. A. (2021). Methods of data collection: A fundamental tool of research. Journal of Integrated Community Health (ISSN 2319-9113), 10(1), 6-10.

20. Mazumder, M. S. A., Rahman, M. A., & Chakraborty, D. (2024). PATIENT CARE AND FINANCIAL INTEGRITY IN HEALTHCARE BILLING THROUGH ADVANCED FRAUD DETECTION SYSTEMS. Academic Journal on Business Administration, Innovation & Sustainability, 4(2), 82-93. doi:10.69593/ajbais.v4i2.74

21. Mehbodniya, A., Alam, I., Pande, S., Neware, R., Rane, K. P., Shabaz, M., & Madhavan, M. V. (2021). [Retracted] Financial Fraud Detection in Healthcare Using Machine Learning and Deep Learning Techniques. Security and Communication Networks, 2021(1), 9293877.

22. Mohammed, A. F. A., & Rahman, H. M. A.-A. (2024). The Role of Artificial Intelligence (AI) on the Fraud Detection in the Private Sector in Saudi Arabia. (100), 472-506.

23. Najjar, R. (2024). Digital Frontiers in Healthcare: Integrating mHealth, AI, and Radiology for Future Medical Diagnostics.

24. Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2024). Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. Communication Methods and Measures, 18(2), 124-141.

25. Pilcher, N., & Cortazzi, M. (2024). 'Qualitative'and'quantitative'methods and approaches across subject fields: implications for research values, assumptions, and practices. Quality & Quantity, 58(3), 2357-2387.

26. Research, P. (2024). Healthcare Fraud Detection Market in the U.S. 2024 to 2033. Retrieved from https://www.precedenceresearch.com/healthcare-fraud-detection-market

27. Settipalli, L., & Gangadharan, G. (2021). Healthcare fraud detection using primitive sub peer group analysis. Concurrency and Computation: Practice and Experience, 33(23), e6275.

28. Singh, K. (2024). HEALTHCARE FRAUDULENCE: LEVERAGING ADVANCED ARTIFICIAL INTELLIGENCE TECHNIQUES FOR DETECTION.

29. Stowell, N. F., Pacini, C., Wadlinger, N., Crain, J. M., & Schmidt, M. (2020). Investigating healthcare fraud: Its scope, applicable laws, and regulations. William & Mary Business Law Review, 11(2), 479.

30. Zanke, P. (2023). AI-Driven Fraud Detection Systems: A Comparative Study across Banking, Insurance, and Healthcare. Advances in Deep Learning Techniques, 3(2), 1-22. Retrieved from https://thesciencebrigade.com/adlt/article/view/182

31. Zhang, C., Xiao, X., & Wu, C. (2020). Medical Fraud and Abuse Detection System Based on Machine Learning. International Journal of Environmental Research and Public Health, 17(19), 7265. Retrieved from https://www.mdpi.com/1660-4601/17/19/7265