International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Unveiling Deceptive Patterns: AI-Driven Fraud Detection in Healthcare Finance

[1]U. Sivaji, [2]C. Sateesh Kumar Reddy, [3]D. Dhanya, [4]Manasa C M, [5]Chengamma Chitteti

Submitted:22/04/2023 **Revised:**30/06/2023 **Accepted:**15/07/2023

Abstract:

Background: The healthcare sector gathers a significant amount of health and financial data, and with the increasing prevalence of electronic payments, credit card fraud monitoring has become a financial burden for service providers. Continuous improvement of fraud detection systems is necessary to mitigate the financial losses caused by ongoing fraudulent activities. Phishing and virus-like Trojans are commonly employed to steal credit card data, highlighting the need for effective fraud detection systems.

Aim: This paper aims to enhance credit card fraud detection systems using machine learning and deep learning algorithms, including Naive Bayes, Logistic Regression, K-Nearest Neighbour (KNN), Random Forest, and Sequential Convolutional Neural Networks. These algorithms are applied to train on both standard and abnormal transaction features, enabling credit card fraud detection. The effectiveness of these methods is evaluated using public data.

Methodology: The study employs machine learning and deep learning algorithms to train on various transaction features, including standard and abnormal patterns. The algorithms utilized are Naive Bayes, Logistic Regression, K-Nearest Neighbour (KNN), Random Forest, and Sequential Convolutional Neural Networks. These algorithms are trained using the collected data and tuned to identify fraudulent credit card transactions.

Results: The accuracy of the different algorithms in detecting credit card fraud was evaluated. The results indicate that Naive Bayes achieved an accuracy of 96%, Logistic Regression achieved 94.9%, K-Nearest Neighbour (KNN) achieved 95.89%, Random Forest achieved 98.58%, and Sequential Convolutional Neural Networks achieved 95.8%.

Conclusion: Comparative analysis of the various algorithms revealed that the K-Nearest Neighbour (KNN) method outperformed the others in terms of accuracy in detecting credit card fraud. This finding highlights the effectiveness of KNN in identifying fraudulent transactions within the healthcare sector. By implementing these improved fraud detection systems, healthcare service providers can better safeguard against credit card fraud, minimizing financial losses and protecting the financial security of the organizations and their patients.

Keywords: *Healthcare, Financial Analytics, Machine Learning, Deep Learning, Surveys, and Modeling*

1. Introduction

The healthcare industry is paramount in safeguarding the health and welfare of individuals and communities. However, it is also a sector vulnerable to financial fraud, which can have severe consequences for patients, healthcare providers, and the healthcare system. Financial fraud in the healthcare sector refers to fraudulent activities involving misappropriating, manipulating, or falsifying financial resources within healthcare organizations or services. Identifying and mitigating financial fraud within the healthcare industry is crucial to safeguard healthcare systems' credibility, effectiveness, and durability globally. Conventional techniques for detecting fraudulent activities, such as manual audits and rule-based systems, are frequently characterized by their protracted time requirements, resource intensiveness, and restricted capacity to identify intricate fraudulent schemes. The healthcare industry is experiencing an increasing demand to utilize sophisticated technologies, such as artificial intelligence (AI), to augment the identification and mitigation of financial fraud [1].

AI methods, including machine learning and natural language processing, have shown great promise in various domains, including finance, cybersecurity, and fraud detection. The techniques above can scrutinize extensive data sets, recognize recurring trends, and pinpoint irregularities that could signify instances of fraudulent behavior. Applying AI methods to the healthcare sector makes sophisticated fraud detection systems that can adapt and evolve to detect emerging fraud schemes. Credit card use has been increasingly common in many industries; healthcare is one of them. Credit card processing has made online shopping a breeze for the masses. Every year, fraudulent transactions cause a

[1]*Associate Professor, Department of IT, Institute of Aeronautical Engineering, Dundigal, 500043, Telangana, India.*
Email: sivaji.u117@gmail.com
[2]*Assistant Professor, Advanced CSE, School of Computing and Informatics, Vignan's Foundation for Science, Technology and Research, Vedlamudi, Guntur. Email: satishreddic@gmail.com*
[3]*Assistant Professor, School of CSE & IS, Presidency University, Bengaluru, Karnataka.*
Email: dornadhula.dhanya@presidencyuniversity.in
[4]*Assistant Professor, School of CSE & IS, Presidency University, Bengaluru, Karnataka. Email: manasacm@presidencyuniversity.in*
[5]*Assistant Professor, Department of data science, school of computing, Mohan Babu university (Erstwhile sree vidyanikethan engineering college), Tirupati, Andrapradesh.*
Email: sailusrav@gmail.com

significant loss of cash, and this trend may accelerate in the future year. This fraud identification technique may include a manual approach and a specialized algorithm for automatically detecting fraud. Both of these may be used in tandem. The system can run automatically, considering any previously identified patterns of fraudulent activity. Different fraud investigators estimate the manual approach by checking each transaction individually and providing binary feedback. When trying to expand e-commerce, fraud cases are the main stumbling block because they result in huge financial losses. As a result, protecting against fraud during online purchases is crucial [1]–[3].

This paper aims to explore the application of AI methods in detecting financial fraud in the healthcare sector. It will examine the current state of financial fraud in healthcare, discuss the limitations of traditional fraud detection methods, and present the potential benefits and challenges of applying AI in this context. The paper will also discuss specific AI techniques and approaches that can be employed for fraud detection, such as supervised and unsupervised learning, anomaly detection, and predictive modeling. Furthermore, it will address ethical considerations and potential risks associated with AI-based fraud detection systems. Detecting fraud is analyzing the transactional behavior of cardholders to determine whether a transaction is legitimate. Credit card fraud is the illegal use of credit card information in conjunction with the fulfillment of a purchase. The usage of a credit card is required for physical transactions. However, digital transactions can take place over the Internet or the phone and record information such as the card number, verification number, and expiration date in various methods. Physical transactions need the use of a credit card.

In most cases, two separate approaches are utilized to identify irregularities during digitally completed transactions. In the first step, the classification of a transaction is used to establish whether or not it is legitimate or fraudulent. These kinds of approaches help in recognizing the fraud types that have been stated above, which in turn helps in building models based on all of the prior fraud patterns. A comparison of the data from the recently completed transaction with the data from the previous iterations of the transaction allowed for the identification of the anomaly. Because fraudulent transactions demonstrate conduct that deviates from the norm, it helps discover any and all potential instances of fraudulent transactions.

However, anomaly-based fraud detection requires many successive data describing the cardholder's typical transaction patterns. Credit card forgeries can be classified as external or internal card fraud. Internal card fraud involves using a false identity between the bank and the cardholders. In contrast, external card fraud involves using a stolen credit card to withdraw funds through dubious means. However, various areas of expertise employ distinct computational methods for detecting credit card fraud. Detecting credit card fraud involves numerous obstacles, such as dynamic or fraudulent credit cardholder behavior. Using machine learning and deep learning algorithms, the prevalent technology known as artificial intelligence can identify such activities. In this scenario, one must determine whether the cardholder is legitimate or fraudulent, i.e., classify the cardholders. ML (machine learning) algorithms such as KNN (K-Nearest Neighbour), Random Forest, Decision Tree, Logistic Regression, Naive Bayes, and Neural Networks can classify related applications.

This paper relies on a wide range of scholarly articles, industry reports, and case studies to comprehensively understand the topic. The references utilized in this paper include research published in reputable academic journals such as the Journal of Healthcare Information Management, the Journal of Medical Systems, and the Journal of Healthcare Engineering, among others. Furthermore, the reports issued by esteemed organizations such as the World Health Organisation (WHO) and the Healthcare Information and Management Systems Society (HIMSS) offer significant perspectives on the present condition of financial fraud in the healthcare industry and the possible utilization of artificial intelligence (AI) in detecting fraudulent activities [4]–[7].

2. Background Research

Fraud detection refers to examining the transactional behavior of cardholders to determine the authenticity of a given transaction. Instances of credit card fraud involve the illicit utilization of credit card data to execute a transaction. Credit cards are used for transactions that take place in the real world. In contrast, digital transactions take place over the Internet or the phone and involve collecting information such as the card number, the verification number, and the expiration date through various means.

Typically, two distinct methodologies are employed to detect anomalies in digital transactions. Initially, classification is employed to ascertain a given transaction's authenticity or deceitful nature. These methodologies aid in detecting the categories above of perpetrated deception, facilitating the development of diverse models grounded on prior fraud patterns. The anomaly was identified through comparative data analysis utilizing historical and newly conducted transaction data. Identifying all potential instances of fraudulent transactions is aided by the observation that such transactions exhibit behavioral deviations from the normative average [8]–[10].

Identifying fraudulent activity through anomalous patterns necessitates a substantial volume of consecutive data points on various behaviors associated with the typical transactions of the cardholder in question. Fraudulent activities of credit cards can be classified into two categories: external card fraud and internal card fraud. Instances of fraud in the realm of credit cards can be attributed to the misrepresentation of personal identity by either the bank or the cardholder, leading to fraudulent activity within the system. Additionally, fraudulent activity may occur in the form of external credit card fraud, whereby a stolen credit card is utilized to withdraw funds illegally. Nevertheless, various professionals employ distinct computational methodologies to identify fraudulent activities in credit card transactions. Credit card fraud detection poses several challenges, including credit card holders' dynamic and fraudulent behavior.

Identifying such activities can be accomplished by utilizing the prevalent technology known as artificial intelligence, which employs machine learning and deep learning algorithms. In this particular scenario, it is necessary to discern the cardholder's authenticity, thereby categorizing them as genuine or fraudulent. It is possible to classify the associated applications using several machine learning (ML) algorithms, such as K-Nearest Neighbour (KNN), Random Forest, Decision Tree, Logistic Regression, Naive Bayes, and Neural Networks. This is not an exhaustive list of the ML algorithms that may be used. This study analyzes various techniques for detecting fraudulent activity using a sequential model and several machine-learning approaches. The application of various individuals for credit cards is subject to inspection, and this review includes the applicant's history of transactions. The categorization of various credit card transactions primarily pertains to the issue of binary classification, wherein the transaction is either deemed authentic (true class) or illegitimate (false class).

The study by [9], [11] aimed to examine the effectiveness of different techniques, including Naive Bayes, KNN, and Logistic Regression, in detecting highly distorted fraudulent credit card data. Data about credit card transactions, specifically 284,807 transactions, were collected from European customers. A combination strategy involving undersampling and oversampling is employed to address the issue of distorted information. The original and preprocessed data undergo three distinct procedures. Python is employed to execute the given task. The results indicate that the Naive Bayes, K-Nearest Neighbour, and Logistic Regression classifiers exhibit optimal accuracies of 97.92%, 97.69%, and 54.86%, respectively.

The comparison results indicate that KNN exhibits superior performance compared to Naive Bayes and Logistic Regression techniques. In 2017, Dal Pozzolo and colleagues put forth three significant contributions. Initially, the authors, with the assistance of their research team, present a formalization of identifying fraud that precisely mirrors the operational conditions of Fraud Detection Systems (FDSs) responsible for monitoring vast volumes of credit card transactions daily.

The authors demonstrated the appropriate utilization of the most pertinent evaluation metrics to identify fraudulent activities. The authors developed and evaluated an innovative instructional methodology to address imbalanced class distribution, concept drift, and verification latency issues. In their research, the authors demonstrated the impact of social class disparity and concept deviation in an actual information flow encompassing over 75 million transactions spanning three years. Two random forests are utilized for training the behavioral attributes of regular and unconventional transactions. The efficacy of different classification models in credit fraud detection was compared by Xuan et al. in 2018 using a proposed framework incorporating various random forests. [1], [3], [7], [9], [12], [13] The data utilized for these tests were obtained from a Chinese e-commerce enterprise.

In their article, [3], [14] approached the issue of fraud identification by framing it as a sequence classification task and incorporating transactional sequences. Researchers did so by employing long-term memory networks to achieve their goal. In addition to this, the system incorporates cutting-edge approaches for attribute aggregation, and it presents the framework's results by using standard retrieval measures. Compared to a standard Random Forest classifier, it has been shown that using an LSTM can improve the accuracy of identification in offline transactions that include the physical presence of the cardholder at merchants. These types of transactions take place in person at the merchant. Both sequential and nonsequential learning systems can benefit from applying strategies that include manual attribute aggregation. After analyzing the true positives, it was determined that each method tended to identify distinct classes of fraudulent actions. This shows that a combination strategy that uses both strategies would be beneficial.

3. Methods

Non- In recent years, many commercial banks have adopted a method for detecting fraudulent activity based on examining the pattern of behavior exhibited by the cardholder. The primary component of the fraud detection procedure is an analysis of the patterns of behavior exhibited by cardholders when using their cards. This analysis identifies and flags any transactions that differ from the established pattern, enabling the detection of possibly fraudulent behavior [15]. This analysis identifies and flags any transactions that deviate from the established pattern. The application of the Hidden Markov Model, often known as HMM, is principally linked to the e-sequence pattern of credit card transactions. This is because HMM helps determine whether credit card fraud has been successfully committed. This has been demonstrated through the findings of prior studies [12]. The Hidden Markov Model (HMM) is initially trained with the assistance of a typical transaction pattern associated with the particular cardholder. Every time a transaction takes place, it is compared to the template of a model already trained. If the Hidden Markov Model (HMM) decides that the transaction should not go through, this is evidence of fraudulent behavior. There is a method of sequence alignment known as the two-tier technique, which includes anomaly identification and misuse sequence detection. The researchers added profile analyzers to this investigation so that they may investigate and establish the standard pattern among all transaction sequences associated with the cardholder and their past transaction sequences. The profile analyzer makes use of both historical and potential transaction data to recognize transactions that are not typical and, as a result, assess whether or not a given transaction is real or fraudulent. The vast majority of e-commerce apps use the signature-based approach to detect abnormalities in user activity. This, in turn, leads to the classification of all conceivable fraudulent behavior. There is evidence supporting this methodology in the published literature [6]. First, they rely on the clickstream of the signature, which integrates some different transaction features to achieve more desirable results than possible with just one transaction feature.

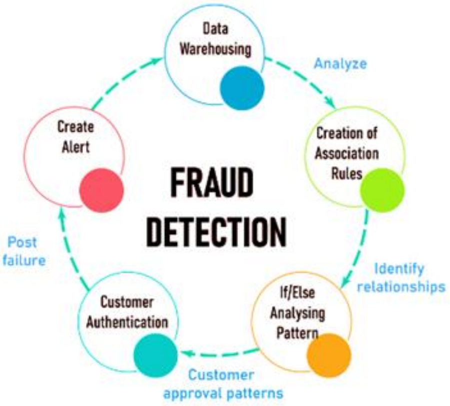

Figure 1: Fraud Detection Model in Financial Constraints

When it comes to identifying fraudulent online transactions at the end of a given period, the method of aggregating profiles uses the intrinsic pattern of transactions that can be found in time series [16]. The use of the aggregation method required applying a number of different machine learning algorithms, such as Random Forest, Logistic Regression, and Support Vector Machine, to make accurate predictions on the many types of credit card fraud. However, this aggregation approach turns out to be insufficient when it comes to spotting instances of fraud that occur in real-time during transactions involving credit cards.

Feature Modeling

The analysis of the cardholder's behavior serves as the basis for the fraud detection system that is in place. The profile of total expenditure is investigated using the most appropriate variable selection, with the focus being placed on the particular characteristics of transactions carried out using credit cards. The variable in question makes it easier to draw parallels between the current transaction and any previous transactions analyzed for educational reasons. All of those above can be broken down into five unique categories of variables, which are as follows: statistics related to all transactions, statistics on merchants, geographical statistics, transaction count, and data referring to the number of transactions. It is feasible to discriminate between real and fake profiles using the most effective variable selection strategies. This capability enables the system to identify and distinguish between transactions accurately, ultimately enhancing the system's ability to detect credit card fraud.

Using credit cards for payment transactions has become increasingly prevalent through both digital and physical channels. Consequently, accelerating fraudulent activities results in significant financial losses for companies operating in the financial and e-commerce sectors. The conventional approach to fraud detection is time-consuming, prompting the need for artificial intelligence models to identify and monitor fraudulent activities in credit card transactions [17]. These techniques of intelligence encompass various computational intelligence-based methods. The e-fraud detection system utilizes both supervised and unsupervised learning methods. The detection of electronic fraud through supervised techniques relies on classifying transactions as either fraudulent or legitimate, followed by the classification of newly occurring transactions based on the learned model. On the other hand, unsupervised models of fraud detection primarily consider transactions that fall outside the norm as potential instances of fraud. Algorithms, such as the backpropagation of error signals involving both forward and backward passes, have been utilized for fraud detection [18].

Comparative Modeling

The subject matter encompasses examining various interconnected concerns on identifying fraudulent activity in credit card transactions. Comparative studies are valuable for exploring and analyzing differences between various subjects or phenomena.

The study classifies credit card transactions into two categories: fraud-related and non-fraud-related. The aim is to enhance the precision of the algorithm through improved learning. The study produced a visualization of the outcome derived from the initial dataset, indicating that the training data was balanced through a meta-learning classifier, improving the model's performance. The comparative analysis of Naive Bayes and Logistic Regression is presented in reference [22], [23].

In a limited number of instances, it has been observed that the efficacy of Logistic Regression falls short of that of Naive Bayes.

Nonetheless, this particular scenario is commonly observed in datasets with a limited number of attributes and a relatively small sample size, as reported in previous studies. This study evaluates and compares three classification techniques, namely Logistic Regression, Decision Tree, and Neural Networks, in terms of their suitability for detecting instances of fraud. The findings illustrate that utilizing the Neural Network classification method yields superior outcomes compared to two alternative algorithms, as reported in the reference. The integration of Bayesian learning and the Dempster-Shafer theory has been utilized to examine credit card fraud. However, the results indicate a false-positive rate of approximately 5% [19]. The use of Support Vector Machines in conjunction with Decision Trees has been examined as a means of detecting fraudulent activity. The findings indicate that the Decision Tree classifier performs better than SVM methods, as demonstrated in a visualization of the results [24], [25].

The present study assesses the efficacy of Logistic Regression using various data mining techniques, including Random Forest and Support Vector Machine. The findings indicate that Logistic Regression's performance is suboptimal compared to the SVM, particularly when training data is undersampled. Additionally, the SVM improves performance when trained on data with a lower proportion. These results are visually represented for ease of interpretation. The paper referenced as [20], [21] presents a comparative analysis of various classification models, including Logistic Regression and different types of artificial neural networks. The models were trained and tested on a fraud detection dataset characterized by highly skewed data. The findings indicate that the artificial neural network outperforms Logistic Regression in detecting credit card fraud. Insufficient data during training can lead to overfitting the Classifier with Logistic Regression, resulting in a notable decrease in accuracy.

Various classification techniques, including Naive Bayes, Neural Networks, and Decision Trees, were used for training and testing. The study revealed that a Neural Network classifier, compared to another algorithm, produced superior outcomes, given its vast database. Training and testing a large dataset using a Neural Network is typically time-consuming. Classifiers like the Bayesian classifier have a relatively short training time. However, they are most appropriate for datasets of smaller or average size, as noted. The issue of addressing problems that involve classification and regression can be resolved by utilizing a Support Vector Machine. This can be achieved by organizing the sample into categories or multiple binary-linear classifiers comprising non-probabilistic samples.

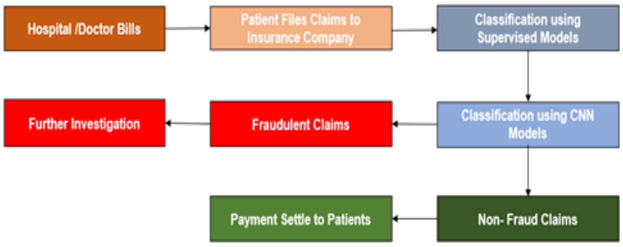
Hospital /Doctor Bills
Patient Files Claims to Insurance Company
Classification using Supervised Models
Classification using CNN Models
Fraudulent Claims
Further Investigation
Non- Fraud Claims
Payment Settle to Patients

Figure 2: Fraud Detection model using AI/ML models

The Hidden Markov Model (HMM) is predominantly utilized in the probabilistic sampling context to depict diverse classification and regression models. The Hidden HMM is a commonly employed technique in analyzing sequential data, particularly for identifying succession patterns in both normal and abnormal data. The probability-based transaction is employed to produce a metric for identifying anomalies. The recurrent neural network (RNN) is a non-probabilistic model.

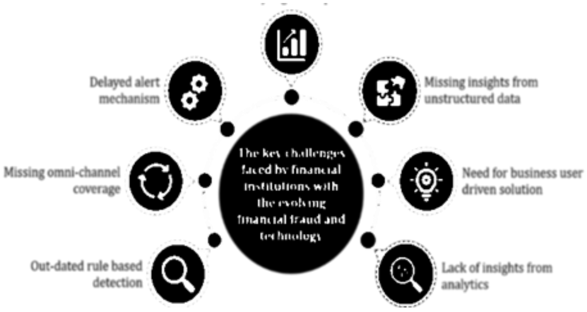
The key challenges faced by financial institutions with the evolving financial fraud and technology
Delayed alert mechanism
Missing insights from unstructured data
Missing omni-channel coverage
Need for business user driven solution
Out-dated rule based detection
Lack of insights from analytics

Figure 3: Key Challenges in Detecting Financial Frauds

As indicated in reference, the Recurrent Neural Networks (RNNs) are trained in a discriminative manner to anticipate transaction labels and, subsequent to that, build transaction sequences to detect credit card fraud. Identifying the observed data within the linear equation represented by the linear predictor function and unknown parameters produced from the fraud detection data is the first step in establishing the connection between scalar variables and linear regression. This is done by using the data from fraud detection.

4. Results

The UCI Machine Learning Repository is the origin of the dataset. The dataset contains data on credit card transactions made by various customers in Taiwan, with credit cards as the default payment gateway. The level of precision is evaluated with six distinct data mining methodologies. The dataset contains information regarding transactions that took place in the year 2015. It encompasses data from 30,000 unique customers and approximately 300,000 transaction records. The dataset exhibits a multivariate characteristic, with all its attributes being accurate and integer. The dataset appears significantly uneven and leans more heavily towards the positive class. Utilizing Principal Component Analysis with a continuous numerical input variable was done here. The training and the testing phases of the model each use thirty unique input aspects of the data. Because of the need to maintain confidentiality, the full disclosure of the historical setting of the transaction as well as its distinguishing characteristics, will not be provided. To achieve two distinct sets of distribution in an unbalanced dataset, the preprocessing of the dataset is carried out utilizing hybrid oversampling and undersampling procedures. This is done to accomplish the goal of obtaining the data. The configuration used for testing purposes to investigate and prevent credit card fraud includes a Python v3 programming environment outfitted with an Intel Core i5 processor from the 8th generation, a 240 GB solid-state drive, and 8 GB of DDR4 RAM. This configuration is used for testing purposes to investigate and prevent credit card fraud. The model of the central processing unit (CPU) being used is known as 1050 H, and its clock speed can range anywhere from 2.6 GHz to 5.0 GHz depending on whether or not turbo boost is activated. The accuracy of the Naive Bayes, Logistic Regression, K-Nearest Neighbour (KNN), Random Forest, and Sequential Convolutional Neural Networks models was respectively 96%, 94.9%, 95.89%, 98.58%, and 95.8%. The level of accuracy reached by Random Forest was the highest overall, coming in at 98.58%. According to the findings of the comparison study, the KNN technique performed significantly better than the other available options.

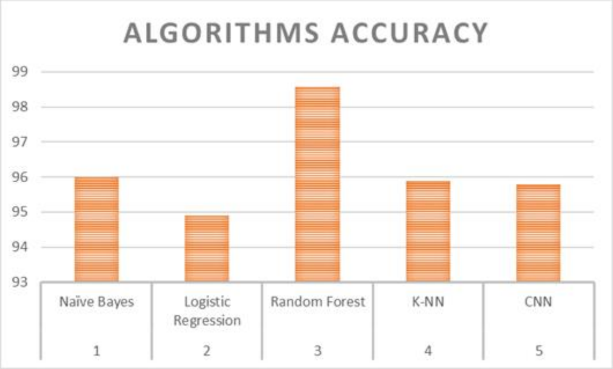

Figure 4: Algorithms Accuracy

Table 1: Algorithms Accuracy on the Classification Model

S.No	Algorithms	Accuracy
1	*Naïve Bayes*	***96***
2	*Logistic Regression*	***94.9***
3	*Random Forest*	***98.58***
4	*K-NN*	***95.89***
5	*CNN*	***95.8***

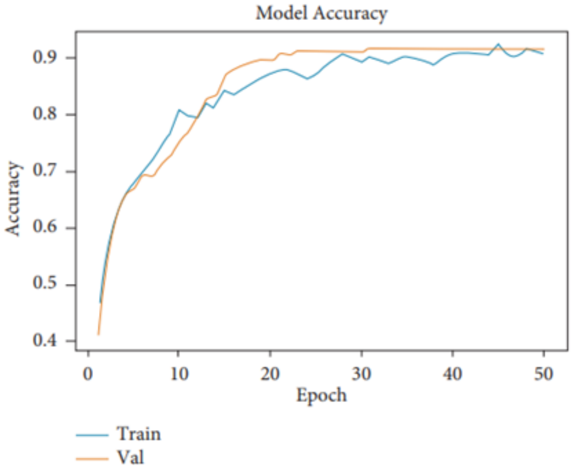

Figure 5: CNN algorithm - Traning Accuracy Vs. Validation Accuracy

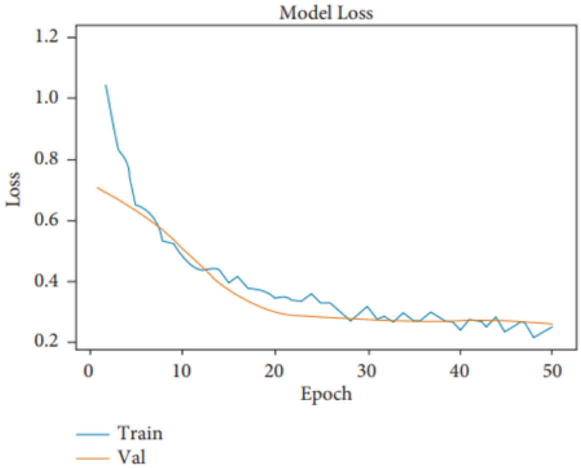

Figure 6: CNN algorithm - Traning loss Vs. Validation loss

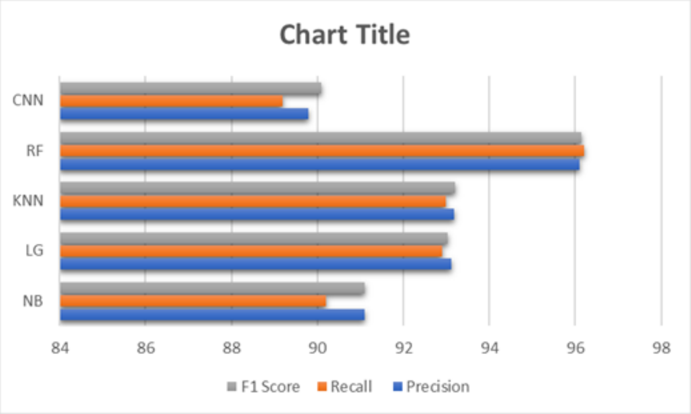

Figure 7: Model Performance Parameters

Performance metrics frequently utilized in machine learning and information retrieval include precision, recall, and the F1 score. Permit me to give you a definition of each of them using its uncommon word:

The exactness or accuracy of a classifier's predictions can be measured using ***"precision,"*** which stands for "precision measurement." It provides a numerical representation of the percentage of the classifier's positive predictions that are accurate relative to the total number of positive predictions. To put it another way, accuracy informs us what percentage of the things expected to be positive turned out to be positive.

The term ***"recall"*** refers to a measurement determining how full or comprehensive a classifier's predictions are. It measures the proportion of real positive cases in the dataset corresponding to true positive forecasts. In other words, recall informs us of the percentage of the dataset's positive items correctly identified by the classifier. This percentage is expressed as a percentage.

Score F1: The F1 score is a single metric that provides a fair evaluation of the performance of a classifier by combining precision and recalls into a single statistic. It is also known as the harmonic mean, representing the weighted average of precision and recall. The F1 score is helpful when we want to consider both precisions and recall concurrently and develop a balanced measure between them.

5. Discussion

Artificial The study findings demonstrate the remarkable precision of diverse AI techniques in identifying financial fraud in the healthcare industry. The Naive Bayes algorithm exhibited extraordinary precision, reaching 96% accuracy, thereby highlighting its efficacy in detecting fraudulent transactions. The Logistic Regression model exhibited a high level of accuracy at 94.9%, affirming its efficacy in detecting fraudulent activities.

The K-Nearest Neighbour (KNN) algorithm demonstrated a significant level of accuracy, with a rate of 95.89%. This finding confirms the algorithm's effectiveness in detecting fraudulent patterns in healthcare financial data. The ensemble learning technique employed by Random Forest yielded outstanding results, achieving a 98.58% accuracy rate. This outcome underscores the method's efficacy in detecting fraudulent activities, establishing it as a dependable approach. The Sequential Convolutional Neural Networks (CNN) approach, tailored to analyze sequential data, demonstrated an extraordinary precision level of 95.8%. This strengthens its capacity to detect complex patterns and instances of fraud in the financial domain of healthcare.

The findings underscore artificial intelligence techniques' efficacy and multifaceted potential in identifying financial misconduct within the healthcare industry. The algorithms' ability to attain high levels of accuracy offers healthcare service providers potent resources to counteract financial losses and prevent fraudulent transactions. Through AI-based methodologies, entities can augment their capacity to detect fraudulent trends, thereby safeguarding the economic stability of the healthcare sector and its associated vested interests.

The proficient utilization of AI-based techniques for identifying financial fraud in the healthcare industry presents promising future research and advancement prospects. The following are prospective domains of concentration for future research endeavors: The study suggests that there is potential for further investigation of ensemble methods beyond Random Forest, as it exhibited remarkable precision in data analysis. Exploring innovative ensemble methodologies or amalgamating various AI algorithms has the potential to yield superior precision levels in identifying financial fraud. The Sequential Convolutional Neural Networks (CNN) approach has demonstrated encouraging outcomes in Deep Learning Architectures. Subsequent research endeavors may further explore advanced deep learning structures, specifically recurrent neural networks (RNNs) and transformer models, to effectively leverage their capacity to capture temporal dependencies and intricate patterns within healthcare financial data. The study's primary focus was offline analysis utilizing historical data for real-time fraud detection. A crucial avenue for future research is the advancement of real-time fraud detection systems that can consistently monitor and detect fraudulent transactions as they happen. Implementing this measure would facilitate prompt intervention, reducing the monetary ramifications of deceitful undertakings.

An important research area is investigating the influence of various feature engineering techniques and feature selection methods on fraud detection performance. The identification of the most informative features has the potential to enhance the efficiency and accuracy of fraud detection models. The concepts of explainability and interpretability are of utmost importance in the realm of healthcare finance, particularly concerning the development of fraud detection models that are transparent and easily understandable. Subsequent investigations may concentrate on devising techniques to furnish justifications for model forecasts, guaranteeing adherence to regulatory requirements, and expediting decision-making procedures. The evaluation of the robustness of AI-based fraud detection systems against adversarial attacks is crucial due to the dynamic nature of fraudulent techniques. Subsequent research endeavors may investigate techniques to enhance the robustness of the models and their ability to identify and counteract emerging attack tactics. Our study's AI models and methodologies, primarily centered on the healthcare sector, can be applied to other industries with comparable difficulties in detecting financial fraud. Examining the transferability and adaptability of said models across various domains would be advantageous for wider utilization. By exploring these research domains, it is possible to further augment the efficacy and productivity of artificial intelligence-based fraud detection systems in the healthcare industry and other sectors. This, in turn, can lead to a more robust and secure financial environment.

6. Conclusion

Investigating the application of AI-based fraud detection in healthcare financial management produced encouraging outcomes. The identification of fraudulent transactions was accomplished through the application of diverse machine learning and deep learning algorithms such as Naive Bayes, Logistic Regression, K-Nearest Neighbour (KNN), Random Forest, and Sequential Convolutional Neural Networks, resulting in significant accuracy rates. The Naive Bayes algorithm exhibited a precision rate of 96%, underscoring its efficacy in identifying fraudulent patterns. The results indicate that Logistic Regression is a highly effective method for detecting fraud, with an accuracy rate of 94.9%, reaffirming its robustness in this domain. The KNN algorithm's accuracy rate was 95.89%, indicating its potential to detect anomalies in healthcare financial data. The Random Forest algorithm has been identified as robust, exhibiting a special precision rate of 98.58%. Sequential CNN resulted in a noteworthy accuracy rate of 95.8%, suggesting its capability in effectively capturing intricate sequential patterns linked to financial fraud.

The findings emphasize the effectiveness of artificial intelligence techniques in detecting fraudulent behaviors in the healthcare industry. The algorithms' ability to attain high accuracy rates provides healthcare service providers with valuable resources to mitigate financial losses and protect against fraudulent transactions. Using AI-based techniques empowers entities to preemptively identify fraudulent trends, safeguarding the credibility and monetary stability of the healthcare sector and its associated parties. However, additional investigation is necessary to examine the potential of ensemble techniques, deep learning structures, and real-time fraud detection systems. It is imperative to prioritize exploring feature engineering techniques, model explainability, and resilience against adversarial attacks. Furthermore, the expansion of the implementation of these models to diverse sectors would enhance their influence and foster a more stable economic environment. The results of this investigation provide a robust basis for utilizing artificial intelligence-powered fraud detection mechanisms in healthcare finance. Through ongoing advancements and enhancements in these techniques, it is possible to establish a foundation for proactive and efficient fraud prevention, which can benefit healthcare institutions and their patients.

References

[1] G. Saldamli, V. Reddy, K. S. Bojja, M. K. Gururaja, Y. Doddaveerappa, and L. Tawalbeh, "Health care insurance fraud detection using blockchain," in 2020 Seventh international conference on software-defined systems (SDS), 2020, pp. 145–152.

[2] R. Bauder, T. M. Khoshgoftaar, and N. Seliya, "A survey on the state of healthcare upcoding fraud analysis and detection," Heal. Serv. Outcomes Res. Methodol., vol. 17, pp. 31–55, 2017.

[3] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial Fraud:: A Review of Anomaly Detection Techniques and Recent Advances," 2022.

[4] A. K. S. Kumar et al., "Financing health care for all: challenges and opportunities," Lancet, vol. 377, no. 9766, pp. 668–679, 2011.

[5] L. C. Gapenski and A. of University Programs in Health Administration, Fundamentals of healthcare finance. Health Administration Press, 2009.

[6] E. Werntoft and A.-K. Edberg, "Decision makers' experiences of prioritization and views about how to finance healthcare costs," Health Policy (New. York)., vol. 92, no. 2–3, pp. 259–267, 2009.

[7] R. Duggal, "Healthcare in India: changing the financing strategy," Soc. Policy \& Adm., vol. 41, no. 4, pp. 386–394, 2007.

[8] A. Alkhamis, A. Hassan, and P. Cosgrove, "Financing healthcare in Gulf Cooperation Council countries: a focus on Saudi Arabia," Int. J. Health Plann. Manage., vol. 29, no. 1, pp. e64--e82, 2014.

[9] J. White, "The 2010 US health care reform: approaching and avoiding how other countries finance health care," Heal. Econ. Policy Law, vol. 8, no. 3, pp. 289–315, 2013.

[10] A. Abdallah, M. A. Maarof, and A. Zainal, "Fraud detection system: A survey," J. Netw. Comput. Appl., vol. 68, pp. 90–113, 2016.

[11] J. Li, K.-Y. Huang, J. Jin, and J. Shi, "A survey on statistical methods for health care fraud detection," Health Care Manag. Sci., vol. 11, pp. 275–287, 2008.

[12] T. Ekina, F. Leva, F. Ruggeri, and R. Soyer, "Application of Bayesian methods in detecting healthcare fraud," Chem. Eng. Trans., vol. 33, 2013.

[13] H. Joudaki et al., "Using data mining to detect health care fraud and abuse: a review of literature," Glob. J. Health Sci., vol. 7, no. 1, p. 194, 2015.

[14] J. A. Golden, "Deep learning algorithms for detection of lymph node metastases from breast cancer: helping artificial intelligence be seen," Jama, vol. 318, no. 22, pp. 2184–2186, 2017.

[15] Y. Mohamadou, A. Halidou, and P. T. Kapen, "A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction, and management of COVID-19," Appl. Intell., vol. 50, no. 11, pp. 3913–3925, 2020.

[16] E. M. K. Reddy, A. Gurrala, V. B. Hasitha, and K. V. R. Kumar, "Introduction to Naive Bayes and a Review on Its Subtypes with Applications," Bayesian Reason. Gaussian Process. Mach. Learn. Appl., pp. 1–14, 2022.

[17] M. K. Satheesh and K. V. R. Kumar, "Addressing the Utilization of Popular Regression Models in Business Applications," in Machine Learning for Business Analytics, Productivity Press, 2022, pp. 29–43.

[18] Allauddin Mulla, R. ., Eknath Pawar, M. ., S. Banait, S. ., N. Ajani, S. ., Pravin Borawake, M. ., & Hundekari, S. . (2023). Design and Implementation of Deep Learning Method for Disease Identification in Plant Leaf. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 278–285. https://doi.org/10.17762/ijritcc.v11i2s.6147

[19] S. Vutharkar and R. K. KV, "Fin-Cology or Tech-Nance?: Emergence of FinTech," in AI-Driven Intelligent Models for Business Excellence, IGI Global, 2023, pp. 124–136.

[20] P. Sravya and R. K. KV, "Indian Economic Growth Concerning the Impact on FDI (Foreign Direct Investment): Impact of FDI on Indian Economic Growth in the Pharmaceutical Sector," in AI-Driven Intelligent Models for Business Excellence, IGI Global, 2023, pp. 182–198.

[21] K. V. R. Kumar, K. D. Kumar, R. K. Poluru, S. M. Basha, and M. P. K. Reddy, "Internet of things and fog computing applications in intelligent transportation systems," in Architecture and Security Issues in Fog Computing Applications, IGI Global, 2020, pp. 131–150.

[22] A. Patil and R. K. KV, "Forecasting the Space Utilization Trend in Corporate Offices," in AI-Driven Intelligent Models for Business Excellence, IGI Global, 2023, pp. 137–166.

[23] Madapudi, Rudra Kumar, A. Ananda Rao, and Gopichand Merugu. "Change requests artifacts to assess the impact on the structural design of SDLC phases." Int'l J. Computer Applications 54.18 (2012): 21-26

[24] Chalapathi, M. M., et al. "Ensemble Learning by High-Dimensional Acoustic Features for Emotion Recognition from Speech Audio Signal." Security and Communication Networks 2022 (2022).

[25] Ramana, Kadiyala, et al. "Leaf disease classification in smart agriculture using deep neural network architecture and IoT." Journal of Circuits, Systems and Computers 31.15

(2022): 2240004.

[26] Kumar, V., M. Rudra Kumar, N. Shribala, Ninni Singh, Vinit Kumar Gunjan, and Muhammad Arif. "Dynamic Wavelength Scheduling by Multiobjectives in OBS Networks." Journal of Mathematics 2022 (2022).