JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Deep Learning Approaches for Detecting Fraudulent Claims in Medical Insurance

Lalit Udayraj Pal,Niraj Pramod Jain, Santosh Singh, Mithilesh Vishwakarma
Post Graduation
Thakur College of Science and Commerce

Abstract

The growing complexity of medical insurance claims has resulted in a notable increase in fraudulent activities, which in turn has led to substantial financial losses for insurance providers and has adversely affected healthcare costs overall. In light of this pressing issue, this study presents **FraudNet**, an innovative deep learning framework designed specifically to leverage cutting-edge machine learning techniques for the detection and mitigation of fraudulent claims within the realm of medical insurance.

FraudNet utilizes a sophisticated combination of **supervised** and **unsupervised learning** methodologies, enabling it to effectively process a wide array of data sources. These sources include not only **structured claim data** but also **unstructured textual information** derived from claim narratives, which can often contain valuable insights. The model incorporates advanced algorithms such as K-Nearest Neighbors (KNN) and Support Vector Machines (SVMs) to uncover intricate patterns and anomalies that are typically associated with fraudulent behavior.

The experimental results obtained from this research indicate that **FraudNet** significantly surpasses the performance of traditional fraud detection methods. It achieves a remarkably high accuracy rate while simultaneously reducing the occurrence of false positives, which is a common challenge in the field.

This research makes a meaningful contribution to the expanding domain of artificial intelligence in healthcare, providing insurers with a powerful tool to enhance their fraud detection capabilities. Ultimately, this advancement holds the potential to foster a more sustainable healthcare system, benefiting both insurers and patients alike.

1. Introduction

Fraudulent claims in medical insurance pose a significant challenge, leading to billions of dollars in losses annually. As healthcare costs continue to rise, insurers face increasing pressure to implement effective fraud detection mechanisms. Traditional methods, often reliant on rule-based systems and manual reviews, struggle to keep pace with sophisticated fraudulent schemes and evolving data patterns.

Deep learning, a subset of machine learning, has shown great promise in various domains, including image recognition and natural language processing. Its ability to learn complex representations from large datasets positions it as a powerful tool for fraud detection in medical insurance.

This study introduces **FraudNet,** a deep learning framework designed to analyze both structured and unstructured data associated with insurance claims. By integrating advanced neural network architectures, FraudNet aims to identify anomalies and patterns indicative of fraudulent activities more effectively than conventional approaches. This paper discusses the development, implementation, and evaluation of **FraudNet,** highlighting its potential to enhance fraud detection capabilities and reduce financial losses in the medical insurance industry. The findings underscore the importance of leveraging artificial intelligence to promote integrity and efficiency in healthcare claims processing. The complexity of medical claims data presents unique challenges for fraud detection. Claims often contain a mix of numerical data, text descriptions, and procedural codes, making it difficult for traditional methods to effectively analyze and interpret. Moreover, the evolving tactics employed by fraudsters, such as creating fictitious patient records or exaggerating service costs, necessitate more adaptive and intelligent systems.

FraudNet addresses these challenges by utilizing a multi-faceted approach that combines both supervised and unsupervised learning techniques. By training on extensive historical claims data, the model learns to recognize legitimate patterns and flag anomalies that deviate from the norm. Additionally, the integration of natural language processing allows **FraudNet** to extract insights from unstructured text, enabling the detection of subtler forms of fraud that may go unnoticed by conventional systems. This research not only aims to improve the accuracy of fraud detection but also seeks to reduce operational costs for insurers by minimizing false positives and enhancing the efficiency of claims processing. By fostering a more robust fraud detection framework, we contribute to the ongoing efforts to safeguard the integrity of the healthcare system, ensuring that resources are allocated appropriately to those in genuine need.

1.1 Types of Frauds in Medical Insurance Claims

A variety of fraudulent practices can take place within the realm of medical insurance claims, significantly contributing to the growing financial strain experienced by both insurers and consumers. These fraudulent activities encompass several distinct categories, including the following:

1. **Billing Fraud:** In this scenario, dishonest healthcare providers may submit claims for services or procedures that were never actually performed. This leads to unjustified reimbursements that can have a detrimental impact on the overall healthcare system.
2. **Upcoding:** Providers may engage in a practice known as upcoding, where they bill for a service or procedure that is priced higher than what was actually provided. This results in inflated claim amounts that can further exacerbate the financial burden on insurers.
3. **Phantom Billing:** In this type of fraud, individuals create fictitious claims for patients or services that do not exist, with the aim of receiving illegitimate payments. This deceptive practice can be particularly challenging to identify and address.
4. **Unbundling:** Unbundling occurs when providers bill for multiple separate services that should ideally be billed together as a single package. This tactic is employed to increase reimbursement amounts, which can lead to inflated costs for insurers.
5. **Patient Fraud:** In some cases, insured individuals may engage in fraudulent activities themselves. This can include misrepresenting their medical conditions or seeking unnecessary treatments in order to exploit their insurance benefits.
6. **Identity Theft:** Fraudsters may also utilize stolen identities to obtain medical services and submit claims, which complicates the detection of fraudulent activities and poses significant challenges for insurers
7. **Collusion:** The phenomenon of collusion, which involves cooperative behavior between healthcare providers and insured individuals, can result in the submission of fraudulent claims that are fabricated. In such cases, the proceeds obtained through these deceitful activities are often divided between the parties involved, leading to significant financial losses for insurance companies and undermining the integrity of the healthcare system.

1.2 The Need for Advanced Fraud Detection

Traditional systems that rely on rule-based methods for detecting fraud have demonstrated certain limitations, particularly in their ability to keep up with the constantly evolving tactics employed by fraudsters.

These static rule sets frequently fall short in capturing the intricate patterns and subtle anomalies that can be found within large and complex datasets of medical claims.

Consequently, there is an increasing demand for more advanced and adaptive approaches that can effectively identify **fraudulent claims** and alleviate the financial burden that insurance fraud imposes on the industry.

1.3 Adopting K-means Cluster Machine Learning Approach

In this paper, we put forth a sophisticated machine-learning approach aimed at enhancing fraud detection within the realm of medical claim insurance, specifically to tackle the various challenges presented by different types of fraud. We have chosen to implement the K-means clustering algorithm, which is a widely recognized unsupervised machine-learning technique celebrated for its ability to identify patterns and group similar data points within extensive

datasets [14].

The application of K-means clustering will enable us to uncover natural clusters of medical claims based on their distinctive features, thereby assisting in the differentiation between legitimate claims and those that may be fraudulent. By iteratively partitioning the claims into separate clusters, our proposed methodology can proficiently detect outliers that display suspicious behaviors indicative of potential fraud. Furthermore, the unsupervised nature of K-means clustering provides a level of flexibility and adaptability, making it particularly well-suited for handling large-scale medical claim datasets that exhibit varying patterns of fraudulent activities [

16Through the adoption of this intelligent machine-learning approach and the utilization of K-means clustering, we aspire to make a meaningful contribution to the advancement of fraud detection in the field of medical claim insurance. Ultimately, our goal is to promote the sustainability and reliability of the insurance industry as a whole.

Keywords: Medical Claim Insurance, Fraud Detection, Machine Learning, Data Preprocessing, Comparative Analysis

Challenges:

1. Data Quality and Availability: Deep learning models require large,
high-quality datasets for effective training. In the context of medical insurance, data may be incomplete, biased, or contain inaccuracies, which can hinder model performance.

2. Imbalanced Datasets: Fraudulent claims are typically rare compared to legitimate claims, leading to imbalanced datasets. This imbalance can result in models that are biased toward the majority class, reducing their ability to accurately identify fraudulent cases.

3. Complexity of Claims Data: Medical claims involve various data types, including structured data (e.g., numerical codes) and unstructured data (e.g., clinical notes). Integrating and processing these diverse formats can complicate model development.

4. Evolving Fraud Tactics: Fraudsters continually adapt their strategies, making it challenging for static models to keep pace. Continuous learning and model updates are necessary to maintain effectiveness, which can be
resource-intensive.

5. Interpretability: Deep learning models, particularly neural networks, often function as "black boxes," making it difficult to understand their
decision-making processes. This lack of transparency can be problematic in regulatory environments where justification for claims decisions is required.

6. Regulatory Compliance: The healthcare sector is heavily regulated, and any automated decision-making system must comply with legal and ethical standards. Ensuring that deep learning models meet these requirements can complicate deployment.

7. Computational Resources: Training and deploying deep learning models can require significant computational power and infrastructure, which may not be feasible for all

organizations, especially smaller insurers.
This research not only aims to improve the accuracy of fraud detection but also seeks to reduce operational costs for insurers by minimizing false positives and enhancing the efficiency of claims processing.

2. Literature Review

The topic of fraud detection within the domain of medical claim insurance has attracted significant attention and extensive research efforts, primarily due to its substantial implications for both the insurance sector and the wider healthcare system. In this section, we intend to provide a comprehensive and detailed review of the existing literature and research findings, highlighting the various approaches and methodologies that have been utilized to effectively combat fraudulent activities associated with medical claims.

In the initial phases of research dedicated to fraud detection, there was a predominant reliance on manual review processes and rule-based systems.. These rule-based systems functioned by employing predefined thresholds and heuristics to identify claims that seemed suspicious, based on specific patterns or characteristics. While these methods demonstrated some level of effectiveness, they encountered considerable challenges in adapting to the constantly evolving nature of fraudulent tactics. This often resulted in high rates of false positives and limited accuracy in detection.

As the field advanced, researchers began to investigate data-driven approaches, particularly the application of supervised machine learning algorithms, to enhance the capabilities of fraud detection. The early techniques in machine learning included support vector machines, decision trees, and logistic regression. These supervised learning models utilized historical claim data that had been labeled as either fraudulent or non-fraudulent, enabling them to learn patterns and make predictions regarding new claims. Although these approaches exhibited promising results, they were heavily reliant on the availability of labeled datasets, which can often be scarce and costly to obtain.

With the increasing volume and complexity of medical claim data, researchers shifted their focus towards unsupervised learning techniques to address the challenges associated with unlabeled data. By employing algorithms such as
k-means and hierarchical clustering, similar claims were grouped together in an effort to identify potential abnormalities. However, these methods frequently faced difficulties in distinguishing between legitimate anomalies and fraudulent claims, which ultimately led to suboptimal performance in detection.

In more recent developments, advanced machine learning algorithms, including ensemble methods and deep learning models such as neural networks, have been explored for their potential in fraud detection within medical claim insurance. Ensemble methods operate by combining multiple models to enhance predictive accuracy, while deep learning models leverage complex neural architectures to automatically learn intricate patterns from raw claim data. These innovative approaches have shown promise in achieving higher detection rates and reducing the occurrence of false positives; however, they often necessitate substantial computational resources and a significant amount of labeled data for effective training.

Despite the valuable contributions made by the aforementioned research, several challenges continue to persist in the field. The dynamic nature of fraudulent activities requires ongoing updates and adaptations of models to keep pace with new fraud schemes. Additionally, the

imbalanced nature of medical claim datasets, where instances of fraud are frequently outnumbered by legitimate claims, presents a significant challenge for traditional machine learning algorithms.

In light of these challenges, we propose an intelligent machine learning approach that employs the K-means clustering algorithm for the purpose of fraud detection in medical claim insurance. By embracing unsupervised learning and integrating domain knowledge, our approach aims to enhance detection accuracy and adaptability while effectively addressing the limitations encountered in previous techniques.

4. Methodology

In this study, we developed **FraudNet**, a deep learning framework aimed at detecting fraudulent medical insurance claims by leveraging advanced machine learning techniques. The model combines both **supervised** and **unsupervised learning** methods to effectively process and analyze large, diverse datasets. The key components of the methodology are outlined below:

4.1 Data Preprocessing

The input data consists of **structured claim information** (e.g., patient demographics, treatment codes, payment details) and **unstructured textual data** (e.g., claim narratives, medical notes). The structured data undergoes standard preprocessing steps, including **data cleaning**, **normalization**, and **feature selection**, while the unstructured textual data is preprocessed using **Natural Language Processing (NLP)** techniques such as **tokenization**, **stemming**, and **TF-IDF vectorization**.

4.2 Feature Engineering

To capture meaningful patterns from the data, both **domain-specific** and **generic features** are extracted. These include attributes such as **claim frequency**, **medical procedure types**, and **cost outliers**. In addition, **N-gram models** are applied to the unstructured text to identify common linguistic patterns related to fraudulent claims.

4.3 K-Nearest Neighbors (KNN)

The **K-Nearest Neighbors (KNN)** algorithm is employed as a **baseline** supervised learning model to classify claims as fraudulent or non-fraudulent based on their proximity to known instances in the training data. KNN evaluates the **Euclidean distance** between new claims and historical examples, identifying **similarity patterns** that might indicate fraud. KNN is useful for its simplicity and efficiency in cases where labeled data is available.

4.4 Support Vector Machines (SVM)

For more complex decision boundaries, **Support Vector Machines (SVMs)** are used. **SVM** with a **kernel trick** allows the model to transform the input data into a higher-dimensional space, enabling the identification of more complex patterns and relationships. By creating an optimal **hyperplane** that maximizes the margin between fraudulent and legitimate claims, SVM provides robust classification capabilities, especially in cases of non-linear data.

4.5 Ensemble Learning

To further enhance predictive performance, we implemented an **ensemble learning** approach that combines the strengths of both KNN and SVM models. This involves training multiple models and aggregating their predictions to minimize error and reduce overfitting. In particular, **bagging** and **boosting** techniques are applied to create a more stable and accurate model.

4.6 Unsupervised Learning for Anomaly Detection

In addition to the supervised models, **unsupervised learning** techniques are utilized to detect anomalies in the dataset. **Autoencoders** and **clustering algorithms** such as **K-means** are employed to identify abnormal patterns in the claims data that may not have been labeled as fraudulent but exhibit suspicious characteristics. This approach allows for the detection of **emerging fraud schemes** that may not have been captured by historical data.

4.7 Model Training and Evaluation

The model is trained using a dataset comprising both fraudulent and legitimate claims, with a **70/30 split** for training and testing purposes. Performance is evaluated using standard metrics, including **accuracy**, **precision**, **recall**, and **F1 score**. Additionally, **cross-validation** is conducted to ensure the model's generalizability.

4.8 Results and Performance Metrics

Experimental results show that FraudNet consistently outperforms traditional fraud detection models. **SVM**, in particular, provides excellent results in terms of **high accuracy** and **low false positives**, while KNN helps in handling cases where similar claim patterns exist. The **ensemble model** further improves performance by combining the best features of each algorithm.

Data Collection Process for Fraud Detection

Detecting fraudulent claims in medical insurance can also be approached using classical machine learning techniques like K-Nearest Neighbors (KNN) and Support Vector Machines (SVM). These models are more interpretable and computationally efficient for certain types of datasets compared to deep learning models like CNNs and RNNs. Below, I outline how KNN and SVM can be applied to this task, as well as the data collection and preprocessing steps necessary for effective fraud detection. Data Sources To train and evaluate KNN and SVM models for detecting fraudulent claims in medical insurance, you need access to comprehensive, labeled datasets. Potential data sources include:

- **Public Datasets:**
- Medicare Claims Data: Publicly available data can be used to model healthcare fraud.
- Health Insurance Fraud Detection Datasets: Kaggle and other repositories sometimes provide synthetic or real-world anonymized datasets.
- ICD (International Classification of Diseases) Codes: Medical data classified with ICD codes can help identify the legitimacy of claims based on diagnosis and procedures.
- Private Insurance Data:
- Insurance Claims Records: Collaborating with private insurance companies to access claims, demographics, and payment histories.

● Government and Research Data:
● Regulatory agencies or academic institutions sometimes release fraud-related healthcare data, either aggregated or anonymized.

Data Features

To build effective KNN or SVM models, your dataset should include the following types of features:

● Claim Details:
● Claim amount, submission date, claim type (inpatient, outpatient), service duration, and frequency.
● Medical History:
● Diagnosis codes, procedures performed, types of treatment, prescription drugs, and whether they match the claimed condition.
● Patient Information:
● Age, gender, location, insurance type, and medical history.
● Provider Details .Provider ID, hospital or clinic location, and any past involvement in fraudulent activities.
● Behavioral Features:
● Patterns of claim submissions, repeated claims for the same diagnosis, and deviation from typical patterns seen in
non-fraudulent claims.
.

Labeling Data

For supervised learning, the dataset must be labeled as either fraudulent or non-fraudulent. Labels can be obtained through:

● Historical Data: Known fraudulent and legitimate claims from past records.
● Expert Validation: Claims labeled by fraud experts based on anomalies or suspicious behavior.
● Anomaly Detection: Unsupervised learning methods could first detect anomalies, which can later be labeled as fraud or not.

Preprocessing Steps

Before feeding data to KNN or SVM, preprocessing is crucial to ensure that the models perform effectively.

● Handling Missing Data: Impute or remove missing values to avoid introducing bias or errors into the model.

2. Normalization/Standardization:

● KNN: Distance-based models like KNN require feature scaling (normalization or

standardization) to ensure that no feature dominates due to its scale.

● SVM: Standardization improves the performance of SVM models by ensuring that all features contribute equally.

● Encoding Categorical Data: Convert categorical variables like gender, diagnosis codes, or insurance type into numerical representations (e.g., using one-hot encoding, label encoding, or embeddings).

● Feature Selection: Select the most relevant features for classification to reduce dimensionality, especially for KNN, as it suffers in
high-dimensional spaces. Techniques like Principal Component Analysis (PCA) or recursive feature elimination can be used.

●Handling Imbalanced Data: Fraudulent claims are often rare, so balancing the dataset is important.Oversampling: Use techniques like SMOTE (Synthetic Minority Over-sampling Technique) to generate synthetic fraudulent examples.

Undersampling: Reduce the number of legitimate claims to balance the dataset.

Class-weighting: In SVM, you can assign higher weights to the minority (fraudulent) class to reduce bias toward the majority class.

3. Model Implementation Using KNN and SVM

K-Nearest Neighbors (KNN)

KNN classifies claims based on the closest "k" neighboring claims in the feature space. The basic workflow involves:

● Input Data: A feature matrix where each claim is represented as a point in n-dimensional space.

● Distance Metric: Commonly, Euclidean distance is used to measure similarity between claims, but other metrics like Manhattan distance can also be tested.

● Prediction: For a given test claim, the algorithm looks at the "k" closest claims (neighbors) and assigns the label that is most common among those neighbors.

Example Code (KNN using Scikit-learn):

Data Collection Process for Fraud Detection

Detecting fraudulent claims in medical insurance can also be approached using classical machine learning techniques like K-Nearest Neighbors (KNN) and Support Vector Machines (SVM). These models are more interpretable and computationally efficient for certain types of datasets compared to deep learning models like CNNs and RNNs. Below, I outline how KNN and SVM can be applied to this task, as well as the data collection and preprocessing steps necessary for effective fraud detection.

Results and Analysis

The models—**K-Nearest Neighbors (KNN)** and **Support Vector Machine (SVM)**—achieved an impressive **90% accuracy** in detecting fraudulent medical insurance claims.

1. **KNN Performance**:
○ KNN's accuracy stems from its ability to classify claims based on similarity to known fraudulent or legitimate cases. However, it can be computationally expensive on larger datasets due to distance calculations.
2. **SVM Performance**:
○ SVM showed strong classification performance, particularly in separating borderline cases with its optimal hyperplane.

Comparative Analysis:

● **Accuracy**: Both models provided high accuracy, with KNN slightly faster in training but SVM handling complex boundaries more effectively.
● **Precision & Recall**: SVM exhibited higher precision for identifying fraud, while KNN had balanced recall across classes.

Python Code:

Output :

Output :

7. Conclusions and Future Work

This research paper presents a novel, practical, intelligent machine-learning approach for fraud detection in medical claim insurance using K-means clustering. We have demonstrated the approach's superior performance, interpretability, and potential for real-world deployment through a comprehensive experimental evaluation.

Key Findings: The key findings from our research are as follows:

1. The proposed intelligent machine learning approach achieved an impressive accuracy of 0.88, indicating a high percentage of correctly classified instances in detecting fraudulent and legitimate claims.

2. With a precision of 0.92, the approach demonstrated the ability to minimize false positives, reducing the risk of incorrectly flagging legitimate claims as fraudulent.

3. The recall value of 0.85 showcases the approach's capacity to capture a substantial proportion of actual fraudulent claims, minimizing the possibility of undetected fraud.

4. The F1-score of 0.88 represents a balanced trade-off between precision and recall, indicating the approach's effectiveness in detecting fraudulent and legitimate claims.
➢**Significance and Contributions:** The proposed approach offers several significant contributions to the field of fraud detection in medical claim insurance:

➢**Unsupervised Fraud Detection:** The approach does not require labeled training data for fraud detection by leveraging unsupervised learning through K-means clustering. It makes it

adaptable to scenarios where labeled instances of fraudulent claims are limited or costly.

➢**Interpretability and Explainability:** The approach's interpretability, facilitated by the generated clusters, provides valuable insights into potential fraud patterns, aiding insurance investigators in understanding the underlying characteristics of claims flagged as potentially fraudulent

➢**Scalability and Adaptability:** The scalability of the approach enables efficient handling of large-scale medical claim datasets. In contrast, its adaptability allows it to adjust to dynamic changes in fraud patterns over time.

8. References

1. Yaqoob, I., Salah, K., Jayaraman, R., & Al- Hammadi, Y. (2021). Blockchain for healthcare data management: opportunities, challenges, and future recommendations. Neural Computing and Applications, 1-16.

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

3. Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement. IEEE Access, 8, 58546-58558.

4. Doan, R. (2011). The False Claims Act and the Eroding Scienter in Healthcare Fraud litigation. Annals Health L., 20, 49.

5. Drabiak, K., & Wolfson, J. (2020). What should healthcare organizations do to reduce billing fraud and abuse?. AMA Journal of Ethics, 22(3), 221-231.

6. 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.

7. Ashfaq, T., Khalid, R., Yahaya, A. S., Aslam, S., Azar, A. T., Alsafari, S., & Hameed, I. A. (2022). A machine learning and blockchain-based efficient fraud detection mechanism. Sensors, 22(19), 7162.

8. Ajemunigbohun, S. S., Isimoya, O. A., & Ipigansi, P. M. (2019). Insurance claims fraud in homeowner's insurance: Empirical evidence from the Nigerian insurance industry. Facta Universitatis, Series: Economics and Organization, 103-116.

9. Berdik, D., Otoum, S., Schmidt, N., Porter, D., & Jararweh, Y. (2021). A survey on

blockchain for information systems management and security. Information Processing & Management, 58(1), 102397.

11. Brodersen, C., Kalis, B., Leong, C., Mitchell, E., Pupo, E., Truscott, A., & Accenture, L. (2016). Blockchain: securing a new health interoperability experience. Accenture LLP, 1-11.

12. Said, A. M., Yahyaoui, A., & Abdellatif, T. (2021). Efficient anomaly detection for smart hospital IoT systems. Sensors, 21(4), 1026.

13. Kumar, S., Bharti, A. K., & Amin, R. (2021). Decentralized secure storage of medical records using Blockchain and IPFS: A comparative analysis with future directions. Security and Privacy, 4(5), e162.

14. Bhardwaj, M., & Agarwal, S. (2022). Decision-making optimization in insurance market using big data analytics survey. In Big Data Analytics in the Insurance Market (pp. 57-80). Emerald Publishing Limited.

15. Jamil, F., Ahmad, S., Iqbal, N., & Kim, D. H. (2020). Towards remote monitoring of patient vital signs based on IoT-based blockchain integrity management platforms in smart hospitals. Sensors, 20(8), 2195.

16. Houtan, B., Hafid, A. S., & Makrakis, D. (2020). A survey on
blockchain-based self-sovereign patient identity in healthcare. IEEE Access, 8, 90478-90494.

17. 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.

18. Bauder, R. A., & Khoshgoftaar, T. M. (2017, December). Medicare fraud detection using machine learning methods. In 2017 16th IEEE international conference on machine learning and Applications (ICMLA) (pp. 858-865). IEEE.

19. Verma, J. (2022). Application of Machine Learning for Fraud Detection–A Decision Support System in the Insurance Sector. In Big Data Analytics in the Insurance Market (pp. 251–262). Emerald Publishing Limited.
-0
20. Kose, I., Gokturk, M., & Kilic, K. (2015). An interactive
machine-learning-based electronic fraud and abuse detection system in healthcare insurance. Applied Soft Computing, 36, 283-299