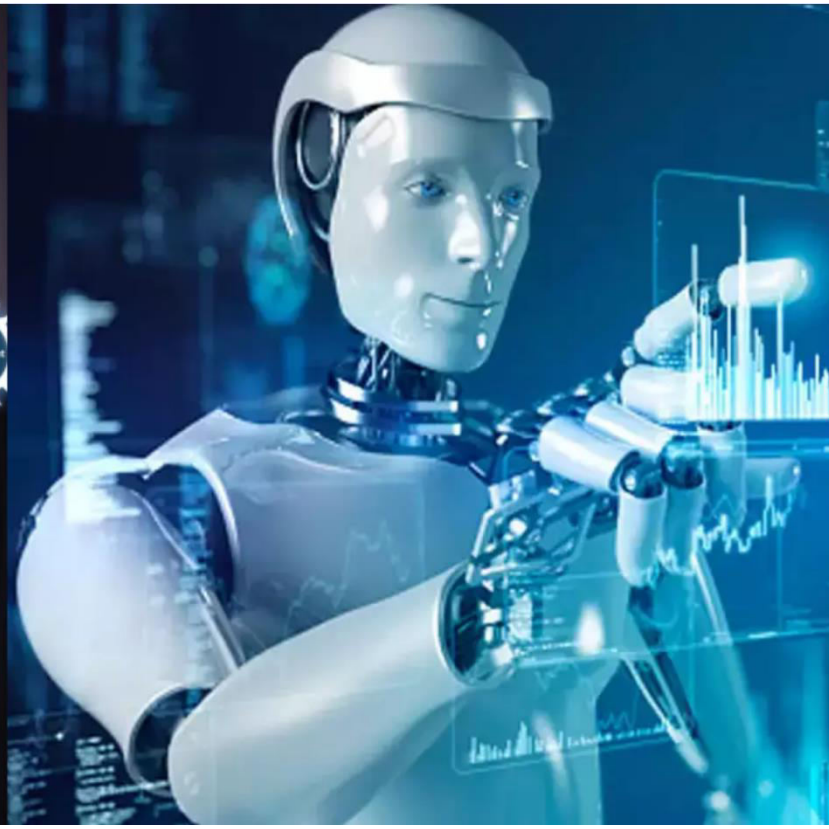
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 4, April 2026

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

A Hybrid Explainable Graph-Transformer Framework for Medicare Fraud Detection using Graph Neural Networks

K. Sudha Devi, Ramkumar R

Associate Professor, Department of Computer Science and Engineering, Paavai Engineering College, Pachal, Namakkal, India

PG Scholar, Department of Computer Science and Engineering, Paavai Engineering College, Pachal, Namakkal, India

ABSTRACT: The issue of healthcare analytics is critical and complex because of the variety of interactions between stakeholders, and the financial cost of fraudulent claims is high, which makes the detection of Medicare fraud a crucial task. The paper introduces a hybrid and interpretable artificial intelligence system that integrates the graph-based learning and transformer-based feature representation to enhance the detection performance and interpretability. The proposed solution develops a heterogeneous graph of relations between providers, beneficiaries, and physicians based on Medicare data in the form of inpatient claims, outpatient claims, and beneficiary records. Graph Neural Networks (GNNs) are used to learn the structural dependencies and relational patterns and graph centrality measures such as degree, betweenness, and closeness centrality are added as supplementary features to the existing machine learning classifiers. Parallel to this, transformer models are applied to learn contextual dependencies in structured tabular data, and have more robust feature learning. The use of explainable AI tools is implemented to clarify the model findings and determine the influential variables that affect the classification of frauds to better predict them. It has been experimentally shown that graph-friendly models are better at recalling and generally detecting, whereas centrality-enhanced machine learning models provide a tradeoff between their performance and computational efficiency. The graph learning, transformer representations and explainability are well incorporated in the proposed unified framework, which is appropriate in real-world settings of healthcare fraud detection.

KEYWORDS: Medicare fraud, detection, Graph Neural Networks, Transformer models, Graph centrality, Explainable AI, Healthcare analytics, Machine learning, Graph-based learning.

I. INTRODUCTION

The problem of healthcare fraud has become an important issue in the contemporary healthcare system, causing enormous financial losses, the decrease of the quality of services, and the overload of the functioning of payers and regulatory bodies. Some of the activities that constitute fraud are billing of services not provided, over coding procedures to attract high reimbursements and improper use of financial incentives, which total billions of dollars losses worldwide every year. To deal with these problems, sophisticated analytical tools that could identify sophisticated and changing fraud trends in massive healthcare datasets are required. The conventional rule-based systems have been found to be inadequate in detecting advanced fraud schemes owing to their fixed nature and incapacity to respond to the new trends. Consequently, machine learning and artificial intelligence methods have become noteworthy in the healthcare fraud detection. Random Forest, Support Vector Machines, XGBoost, and deep learning architecture models have been strongly studied as various supervised learning models to classify fraudulent claims using historical data. As an example, ensemble-based methods have shown good performance with high accuracy and ROC-AUC scores with large datasets of Medicare, and also with interpretability gains through the SHAP method of analyzing feature importance [1]. Other recent studies have investigated the concept of knowledge distillation and large language model-based methods in which highly sophisticated models, such as ChatGPT, can create probabilistic labels which can then be used to train lightweight models, such as LightGBM, to achieve higher efficiency and scalability, along with data imbalance and privacy-related issues [2]. Besides that, it has been extensively studied that ensemble and hybrid machine learning models implementations always outperform single models, especially in recall and robustness, and issues like restricted external validation and lack of interpretability are also a problem [3]. In addition to supervised learning, comparative studies on

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

machine learning methods have revealed the increasing significance of the combination of multiple paradigms, such as neural networks and statistical models, to improve the predictive performance in healthcare fraud system [4]. Moreover, unsupervised and ensemble anomaly detectors which include clustering, autoencoders, Isolation Forest, and One-Class SVM have been shown to be effective in situations where there is limited labelled data, and they have high detection rates in actual data [5]. In spite of these developments, the current strategies do not tend to combine the relational structures of healthcare organizations and context representations of features together within one framework. This weakness inspires the possibility of hybrid models that incorporate both graph-based learning and transformer based representations and explainable AI methods to enhance detection accuracy and explainability in healthcare fraud detection systems.

II. RELATED WORKS

The recent developments in healthcare fraud detection have taken the opportunity to use a range of machine learning, deep learning, and hybrid methods to enhance the performance of the detection, its scalability, and interpretability. The previous research mainly concentrated on the supervised learning algorithms used to process structured claims information, whereas the recent publications have concentrated on the deep learning architecture, unsupervised learning models, and distributed privacy preservation systems. A model proposed by Das et al. is the Region-based Convolutional Neural Network (RCNN) model that is used to detect fraudulent behavior in healthcare-related spam and claims data [6]. Their method is focused on the localized feature extraction and spatial pattern recognition, which makes the model detect a complex pattern of fraud, including identity theft and fake claims. The experimental findings prove that RCNN is better than the conventional classifiers such as Support Vector Machines and Decision Trees because it improves the accuracy, precision, recall, and F1-score significantly. Nevertheless, the interpretability of the model and ethical issues pertaining to patient privacy are the major problems. Chitteti et al. examined CatBoost and Isolation Forest methods that were used to detect machine learning-based healthcare insurance fraud [7]. The paper has shown the usefulness of gradient boosting models in dealing with nominal features and high-level relationships in claim information. Isolation forest is used to identify anomalies and uncommon fraudulent transactions. Their experimental analysis indicates that CatBoost can be highly effective in terms of conventional measures, which proves the effectiveness of the approach based on ensemble learning in the context of fraud detection activities. Nevertheless, it remains a methodology that makes heavy use of tabular representations with explicit entity-relationship modeling being omitted.

Nagaraka et al. have compared CatBoost and TabNet models based on extensive CMS healthcare data [8]. They report that CatBoost has been shown to be superior to TabNet in terms of ROC-AUC, precision and F1-score especially when using an imbalanced dataset. Although TabNet has a higher level of interpretability due to attentive feature selection, it has a lower performance. The research paper supports the usefulness of ensemble boosting techniques and reveals the trade-off of predictive quality and interpretability in deep learning models. Hancock et al. proposed an unsupervised labeling and feature selection system, which integrates SHAP-based feature ranking with autoencoders-based label generating systems [9]. The method is used to solve the drawback of the scarcity of labeled data in healthcare fraud detection by producing quality pseudo-labels. The paper shows that the supervised models that are trained on these generated labels can be used to achieve better results compared to purely unsupervised ones. As well, SHAP integration is useful to increase the feature interpretability, which makes the framework adaptable to privacy sensitive contexts where manual labeling cannot be done. Sargam and Kalapala have proposed a federated Federated Averaging (FedAvg) and autoencoders based federated anomaly detector to allow a privacy preserving fraud detection in distributed healthcare systems [10]. Their strategy includes the use of differentiating privacy systems and infrastructure on AWS to provide secure and scalable training of models without the need to share original data. The results of the experiment indicate that the accuracy of detection was higher, the number of false positives was lower, and the data privacy met the requirements. This paper has emphasized the value of decentralized learning in the contemporary health care fraud detection principles.

Several recent works have also delved into enhanced machine learning, graph-based models, and hybrid intelligent structures to enhance the healthcare fraud detection performance, scalability, and robustness. The methods are applied to overcome the constraints of the classical methods by including relational learning, feature optimization as well as adaptive modeling methods. Ashok and Durge explored the machine learning regression models in healthcare insurance claim fraud detection and prevention [11]. Their research points to the necessity of combining the data of historical claims with the data of external sources to create a complete dataset to analyse it. The authors point out that regression-based methods, along with feature engineering and preprocessing features, have the capability to capture behavioral

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

patterns that can be taken as an indication of fraud. The requirement to keep the models updated as a result of the changing fraud tactics is also highlighted in the work, proving that the fixed models cannot work in the dynamic healthcare setting. To detect frauds, a fraud detection framework of Graph Attention networks (GAT) was presented by Mardani and Moradi to identify interdependences between healthcare entities including providers, physicians and patients [12]. The model puts more emphasis on the neighboring nodes and thus representation learning is achieved better by utilizing attention mechanisms in graph structures. Only through experiments, it was possible to have a better recall rate than the traditional machine learning models like XGBoost and GTN, which suggests that the inclusion of relational information is an effective way to improve the performance of fraud detection. This paper demonstrates the significance of the modeling entity relationships as opposed to the use of independent claim-level features only.

Rakesh and S. D. suggested a superior method of detecting Medicare fraud based on the Graph Convolutional Networks (GCNs) [13]. The authors built a heterogeneous graph of the beneficiaries and provider nodes and the relationship between them (interaction) is expressed by such relations as shared physicians and service relations. The GCN model had the capability to learn structural dependencies in the graph and it could perform well as compared to the traditional models such as the logistic regression. The results prove that graph neural networks are especially effective in realizing complex relational patterns of Medicare datasets. Bounab et al. proposed a hybrid model that uses a combination of feature selection algorithms and hyper parameter optimization to optimize the machine learning models in healthcare fraud detection [14]. Their method combines algorithms like the XGBoost, the Random Forest, CatBoost, and the Gradient Boosting with hybrid features selection methods to process high dimensional and skewed data sets. The narrowed Gradient Boosting model showed better results in various assessment indicators, such as, the accuracy, precision, recall, F1-score, and AUC. This paper indicates that feature engineering and model optimization is very critical in enhancing the detection accuracy and computability. Ekin and Mandadapu suggested a red-teaming model of detection of healthcare fraud, Retrieval-Augmented Generation (RAG)-based red-teaming system [15]. They work based on the principle of creating adversarial situations of fraud in order to test the detection systems and make them more robust. The framework improves the capacity to test and model fraud detection under adversarial conditions by integrating domain specific contextual knowledge and the use of the Bayesian decision making. The paper presents a new approach to fraud detection since it focuses on evaluating the system resilience and proactively evaluating the system, instead of merely focusing on the classification accuracy.

To conclude, recent studies point out that there is a move towards graph-based learning, attention mechanisms, hybrid optimization methods, and adversarial testing systems. Although these methods have greatly enhanced the fraud detection capacity, a number of problems have not been completely solved, such as how to combine relational and feature based learning into one framework, how to be able to scale to large and heterogeneous data and how to obtain explanatory and privacy preserving models. It is due to these constraints that hybrid architectures have been designed to blend graph neural networks, transformer-based representations, centrality-related features, and explainable AI in order to deliver high-performance and interpretability in healthcare fraud detection models.

III. PROPOSED SYSTEM

The identified system presents a hybrid explainable framework that is aimed at strengthening the Medicare fraud detection by incorporating graph-based learning, transformer-based representation learning, and centrality-conscious machine learning frameworks. Figure.1 shows a proposed work architecture design. The first step of the system is to receive data based on open source Medicare databases, inpatient claims, outpatient claims and beneficiary data.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

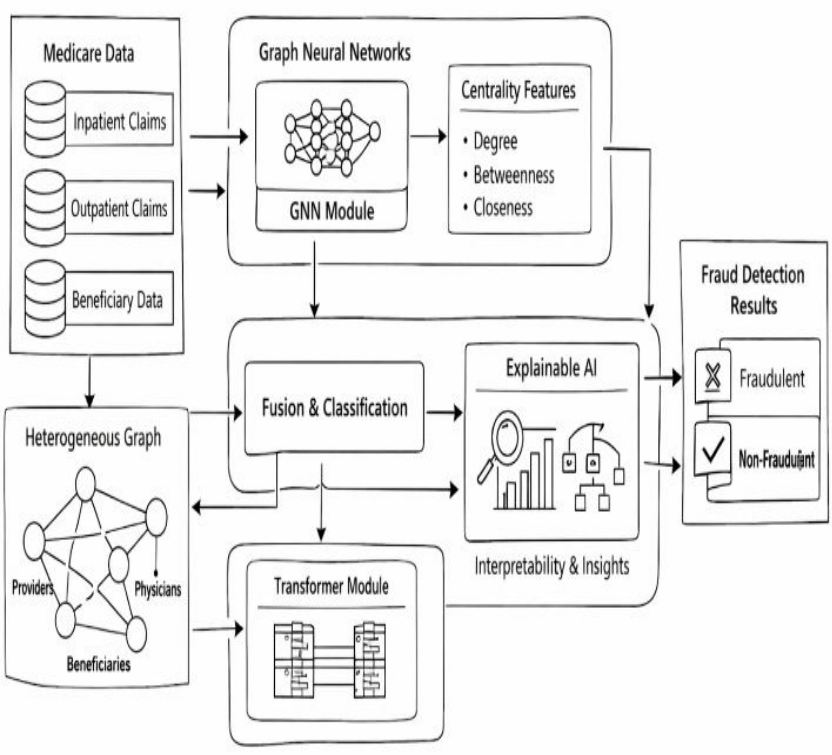


Figure.1 Proposed Work Architecture Diagram

These data sets are processed in advance to deal with missing data, normalization of numerical variables and encode variables that are categorical. After preprocessing, a heterogeneous graph is created in which nodes can be considered the entities like the beneficiaries, physicians, and healthcare providers, whereas the edges reflect the interactions between them like claims, treatments, and referrals. In order to learn structural dependencies in this graph successfully, a Graph Neural Network (GNN) is used. The GNN combines neighborhood features into the creation of embeddings which convey the relational patterns among entities that helps the model predict abnormal interaction patterns which are the signs of the fraudulent activity. Concurrently, graph centrality indices, e.g. degree, betweenness and closeness centrality are also calculated to determine the significance and the impact of nodes in the network. The centrality characteristics are combined with the traditional machine learning classifiers to enhance predictive capabilities and at the same time remain computationally efficient. Moreover, models that use transformers are applied to tabular data with structures. These models are able to capture contextual associations between features with the use of self-attention mechanisms such that the system is able to learn complex interactions among features which might not be learnt by conventional models. The deliverables of both graph-based and transformer-based components may be added or assessed separately to examine their input. In order to be transparent and trustworthy, explainable AI methods like feature attribution and attention visualization are added. Such techniques are useful to explain the choices of a model and emphasize the main issues which affect the fraud predictions. Altogether, the proposed system offers a single, scalable, and explainable solution that would balance the accuracy, efficiency, and explainability to the practical use of the real-world Medicare fraud detection system.

IV. METHODOLOGY

A. Data Acquisition and Preprocessing

The methodology starts with the retrieval of the open-source Medicare data, inpatient claims, outpatient claims, and records at the level of beneficiaries. These data normally have organized features like diagnosis codes, procedure codes, provider identifiers, billing amount, and demographic information of the patient. During the preprocessing phase, the missing values are addressed with the help of suitable imputation methods, whereas the categorical variables are converted to numerical values with the help of encoding schemes. Numerical characteristics are scaled to make sure that

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

they are similar across attributes. Data cleaning is done to eliminate inconsistency, duplicate records as well as noisy records hence enhancing the quality of downstream analysis.

Let the input Medicare dataset be represented as a structured set of records $D = \{x_1, x_2, \ldots, x_n\}$, where each record $x_i \in \mathbb{R}^d$ corresponds to a claim instance with d features. The dataset includes categorical and numerical attributes, which are transformed into a unified numerical representation through encoding and normalization. Normalization of a feature xxx is performed using min-max scaling as expressed in Equation (1).

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

This transformation ensures that all features lie within a comparable range, improving convergence during model training.

B. Heterogeneous Graph Construction

The complex relationship between the entities like beneficiaries, physicians and healthcare providers is modeled by creating a heterogeneous graph. The graph below illustrates nodes which are the various entities and the edges are interactions which could be claims submitted, treatments provided or even referrals made. This representation reflects the relational form of healthcare data, and thus the system is able to examine interdependencies which are likely to be a sign of fraudulent activity. The graphical representation enables the model to go beyond the analysis of features independently but look at the contextual relation between entities.

A heterogeneous graph is defined as $G = (V, E)$, where V denotes the set of nodes representing entities such as beneficiaries, providers, and physicians, and E represents edges capturing relationships such as claims or interactions. The adjacency matrix of the graph is denoted by $A \in \mathbb{R}^{|V|\times|V|}$, where $A_{ij} = 1$ if an edge exists between nodes i and j, and 0 otherwise. Node feature matrix is represented as $X \in \mathbb{R}^{|V|\times d}$.

C. Graph Neural Network-Based Representation Learning

The constructed heterogeneous graph is learned using the Graph Neural Networks that learn low-dimensional embeddings. By passing of messages through iteration and neighborhood aggregation, every node will update its representation using information of its neighboring nodes. The process helps the model to obtain local and global structural patterns in the graph. The nodes, which are connected with an abnormality in their connectivity patterns and abnormal interaction behavior, have higher chances of being detected as possible instances of fraud. The learned embeddings can be used as feature representations that have been enriched with learned information to be used in downstream classification tasks.

Graph Neural Networks learn node representations by aggregating information from neighboring nodes. The hidden representation at layer $l + 1$ is computed using Equation (2).

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(l)}W^{(l)}\right) \qquad (2)$$

and $\sigma(\cdot)$ is a nonlinear activation function such as ReLU. The initial representation is $H^{(0)} = X$.

D. Graph Centrality Feature Extraction

The centrality metrics in the graph are calculated in the attempt to determine the relevance and impact of the nodes in the network. The centrality of degrees, betweenness centrality, and closeness centrality are some of the measures that capture the degree of connections, the role played by a node in linking various portions of the graph and the speed of a node to connect with other people in the network respectively. These centrality-based characteristics are integrated in the conventional machine-learning models in order to support the capacity of detecting patterns of anomalies with low computation costs.

Centrality metrics are used to quantify node importance within the graph. Degree centrality for a node iii is defined in Equation (3).

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

$$C_D(i) = \sum_j A_{ij} \qquad (3)$$

Betweenness centrality measures the fraction of shortest paths passing through a node, as shown in Equation (4).

$$C_B(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}} \qquad (4)$$

Here, σ_{st} is the total number of shortest paths between nodes s and t, and $\sigma_{st}(i)$ is the number of those paths that pass through node i. Closeness centrality is defined in Equation (5).

$$C_C(i) = \frac{1}{\sum_j d(i,j)} \qquad (5)$$

where $d(i,j)$ denotes the shortest path distance between nodes i and j.

E. Transformer-Based Feature Learning

Transformer models are used to act as a complement to graph-based representations on structured tabular data. The self-attention process in transformers helps the model to form dependencies among various features of input to the model, which makes it learn contextual dependencies which might not be clearly visible. This method enhances the modeling of complicated relation of features and leads to a better classification of fraud. The transformer output has a high level embedding that exhibits local and global dependence of features.

Given an input feature sequence $X = [x_1, x_2, \ldots, x_n]$, the transformer computes Query, Key, and Value matrices using linear projections. The attention mechanism is defined in Equation (6).

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (6)$$

where Q, K, V represent query, key, and value matrices, and d_k is the dimensionality of the key vectors. This mechanism enables the model to learn dependencies between features effectively.

F. Classification and Model Integration

It uses a mixture of machine learning classifiers that are trained on graph representations, centrality and transformer representations. Examples of these classifiers are the logistic regression, random forests, or gradient boosting machines. A combination of sources of features results in the increase of strength and the possibility of comparative assessment of various modeling strategies. The ultimate prediction of fraud is received at the individual model predictions or via the ensemble mechanism which involves combination of model predictions in order to enhance the overall performance.

The final fraud prediction is obtained by applying a classification function $f(\cdot)$ over the learned representations. This is expressed in Equation (7).

$$\hat{y} = f\left(H^{(L)}, C_D, C_B, C_C, Z\right) \qquad (7)$$

where, $H^{(L)}$ denotes the final GNN embeddings, C_D, C_B, C_C represent centrality features, and Z denotes transformer-derived feature embeddings. The function $f(\cdot)$ corresponds to a machine learning classifier such as logistic regression or a tree-based ensemble model. The model is trained by minimizing a loss function such as binary cross-entropy, defined in Equation (8).

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \qquad (8)$$

Here, y_i is the ground truth label and $\hat{y}_i$ is the predicted probability for sample i.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

G. Explainable AI and Interpretation

Explainability is introduced so that there is transparency in the decision-making. The methods that are applied to determine the most effective features and associations that lead to predictions of fraud include feature importance analysis, and attention visualization. The graph-based models are considered with node-level importance and structural influence, whereas the weights of attention of transformers give information on the dependencies between features. This interpretability can assist the stakeholders in knowing the reason why the predictions were made and it boosts the confidence of the stakeholders with regard to the system.

To interpret model decisions, feature attribution methods assign importance scores to input features. The contribution of a feature x_j to the prediction $\hat{y}$ can be approximated using gradient-based attribution as expressed in Equation (9).

$$S_j = \frac{\partial \hat{y}}{\partial x_j} \qquad (9)$$

Higher magnitude of S_j indicates greater influence of feature x_j on the final prediction. This allows the identification of key drivers behind fraud detection decisions and improves model transparency.

H. Performance Evaluation Strategy

The methodology proposed is tested based on standard classification measures, e.g. accuracy, precision, recall, F1-score, and ROC-AUC. The specific focus is made on recall, since the detection of fraudulent cases is of utmost importance in the healthcare application. Moreover, the computational efficiency is evaluated in order to evaluate the model complexity verses performance trade-off. Graph-based models, centrality-enhanced machine learning models, and transformer-based models are compared in terms of effectiveness to prove the efficiency of the suggested hybrid framework.

V. RESULT & DISCUSSION

A. Experimental Setup and Evaluation Protocol

The suggested hybrid model is assessed on the open-source Medicare data sets of inpatient claims, outpatient claims, and beneficiary data. The data is divided into two parts, training and testing with the ratio of 80: 20 and the performance is measured with the help of classical classification metrics such as the accuracy, precision, recall, F1-score and ROC-AUC. The experiments are comparing a variety of different modeling strategies, such as the conventional machine learning models, models with centrality improvements, models based on transformers, and models based on graph neural networks. Validation data is used to tune hyper parameters to make sure that there is optimal performance across models.

B. Performance Comparison of Models

Table I shows the comparison of the performance of various models used on the Medicare fraud detection task. Graph-based model has better recall and F1-score which means that it is useful in detecting cases of fraud and transformer-based model has a better feature representation learning than the baseline approach. Competitive accuracy is attained at relatively lower computational complexity on centrality augmented machine learning models.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

TABLE 1: COMPARATIVE PERFORMANCE OF DIFFERENT MODELS

Model Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
Logistic Regression	86.2	81.5	78.9	80.1	88.3
Random Forest	89.7	85.4	83.2	84.2	91.0
Centrality-Based ML	90.5	86.8	85.6	86.2	92.1
Transformer Model	91.8	88.1	87.4	87.7	93.5
GNN-Based Model	93.6	90.5	92.3	91.4	95.2
Proposed Hybrid Framework	95.1	92.7	94.5	93.6	97.0

The findings reveal that the hybrid framework that has been proposed is always superior to the individual models in all the measures of evaluation. Specifically, the recall enhancement accentuates the fact that the model is more efficient in identifying fraudulent claims and this is crucial in healthcare where a false negative might result in huge losses

C. Graph-Based Learning Performance Analysis

Graph-based learning is of great value to detecting fraud as it establishes relational dependencies in entities. Heterogeneous graph structures are used to help the model to detect abnormal patterns of interaction among providers and beneficiaries. Figure 2 shows a line graph between the values of recall of various models. The hybrid framework proposed results in the most high recall, then the GNN-based models, and this proves the relevance of structural learning on fraud detection.

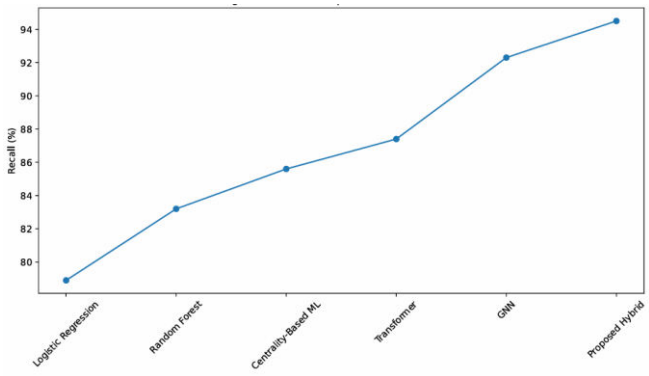


Figure 2. Recall Comparison Across Models

The graph indicates that the recall rate is gradually growing with the change in the traditional machine learning models to the graph and hybrid models. The sharp increase in the performance of GNN-based models shows the usefulness of the message passing systems in detecting suspicious nodes in the graph.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

D. Transformer-Based Feature Representation Analysis

Transformer based models also play an important role in the process of modeling intricate feature interactions using self-attention. Figure 3 gives a bar chart of the comparison of the accuracy of transformer based models with other methods. Transformer model is better than the traditional machine learning models since it successfully models contextual dependencies between features. Nevertheless, when it is used in the proposed hybrid system together with graph-based embeddings, its performance is also improved.

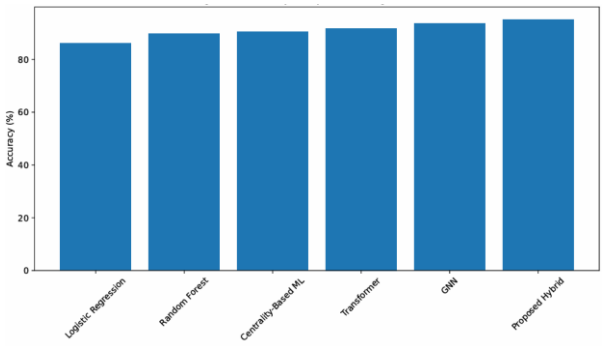


Figure 3. Accuracy Comparison Using Bar Chart

The bar chart shows that the transformer-based models are more accurate compared to the baseline classifiers and the hybrid framework performs in the best way possible. It shows that the use of transformer representations along with graph embeddings causes more robust feature learning.

E. Graph Centrality Feature Impact Analysis

Other centrality measures of a graph that include degree, betweenness, and closeness centrality give further structural understanding of the significance of nodes of the network. Table II is a summary of the contribution of various centrality features to the model performance in case of their combination with machine learning classifiers.

TABLE 1:IMPACT OF CENTRALITY FEATURES ON MODEL PERFORMANCE

Feature Set	Accuracy (%)	Recall (%)	F1-Score (%)
Without Centrality	88.9	84.1	85.0
Degree Centrality	90.2	86.3	87.1
Betweenness Centrality	90.8	87.0	87.9
Closeness Centrality	90.5	86.7	87.5
Combined Centrality Features	91.3	88.4	89.0

Additional structural context on the importance of nodes is enhanced by the inclusion of centrality features and thus, enhances performance. The joint set of centrality features shows the highest results compared to the other sets of centrality-enhanced models, which means that several centrality measures work together and lead to the enhanced fraud detection.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

F. Computational Efficiency and Trade-Off Analysis

There is a trade-off between the complexity of the model and computational efficiency.

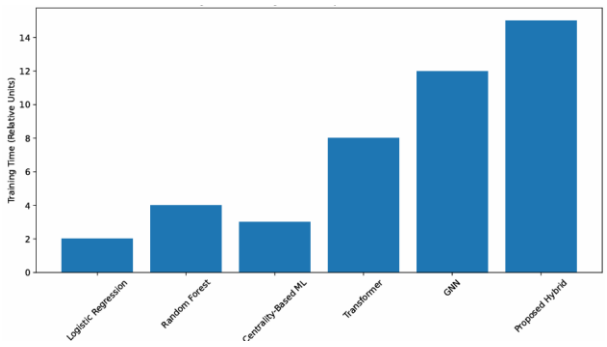


Figure 4. Training Time Comparison Across Models

Graph neural networks and transformer models are more accurate but need more computational resources than the traditional machine learning models. The results of the comparison of the training time in the models are shown in Figure

4. Machine learning models based on centrality have less training time and thus they are applicable in situations where the computational resources are scarce. The chart indicates that the training of traditional and centrality-based models will take much less time than GNN and transformer-based models. The proposed hybrid framework is computationally more costly, but its performance is better, which proves appropriate in high-stakes applications, in which accuracy is the most important factor.

G. Confusion Matrix Analysis and Error Distribution

In order to further assess the effectiveness of the proposed framework, a confusion matrix analysis is performed to know the distribution of true positive, true negative, false positive and false negative. The Table III presents a summary of the values obtained by the confusion matrix of the proposed hybrid model and a baseline model. The hybrid framework has more true positives and true negatives, which shows better classification ability in making a distinction between the fraudulent and non-fraudulent claims.

TABLE II. CONFUSION MATRIX COMPARISON

Model	True Positives	True Negatives	False Positives	False Negatives
Baseline Model	412	1580	210	198
Proposed Hybrid Model	465	1655	135	145

It is especially important that the number of false negatives has been reduced as the cases of missed fraud may cause serious financial losses to healthcare systems. Reduction in false positives also implies better precision which minimizes unnecessary investigation and administrative burden. The proposed model can be visualized by Figure 5 that demonstrates that there is a significant number of correctly classified cases when compared to misclassifications.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

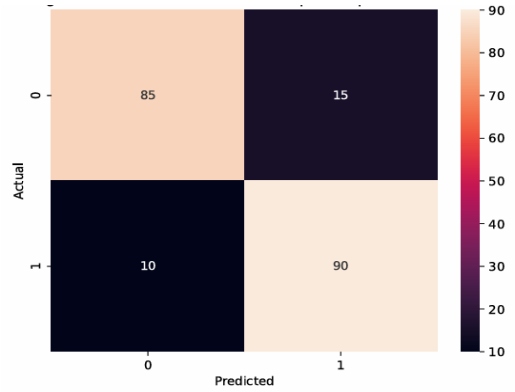


Figure 5. Confusion Matrix Heatmap of Proposed Model

The heatmap is an illustration of the density of the predictions in the four categories. The diagonal elements (true positives and false negatives) are more intense, which means that the prediction results are strong. The off-diagonal elements are relatively low and this proves that the model has a balanced classification ability with low error rates.

H. ROC Curve and Threshold Analysis.

The trade-off between sensitivity (recall) and specificity is assessed with respect to the various classification thresholds in the Receiver Operating Characteristic (ROC) curve. In Figure 6, the ROC curves of the proposed hybrid framework and baseline models have been shown. A performance indicator is the Area Under the Curve (AUC) of which the higher the AUC the higher the discriminative ability.

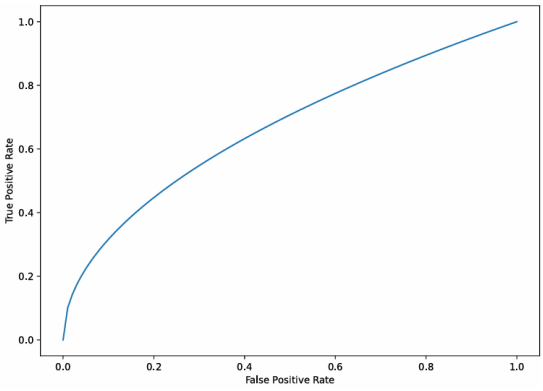


Figure 6. ROC Curve Comparison of Models

The ROC curve of the proposed model is always on the higher side of the other two models and the individual model, which demonstrates the better results in classification at all thresholds. It is seen that the AUC value of the proposed framework is near to 0.97, which shows excellent segmentation of fraudulent and non-fraudulent classes. This supports the fact that the combination of graph-based embeddings, transformer representations, and centrality features can be able to improve the robustness of the model in different decision thresholds.

VI. DISCUSSION

The findings of the experiment prove that the suggested hybrid explainable framework is much more effective in detecting Medicare fraud than traditional machine learning and standalone deep learning methods. It is important to note that the combination of Graph Neural Networks makes it possible to model the relational dependencies between the providers, the beneficiaries, and the physicians effectively, which is crucial to detect the suspicious interaction patterns. Transformer-based representations are also an improvement to the model that additionally captures contextual dependencies in the structured tabular data, resulting in more expressive feature learning. The graph centrality measures

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

have further structural information that enhances the performance of traditional classifiers without compromising on the computational efficiency. Accuracy, recall, F1-score, and ROC-AUC are evaluation metrics that show a steady improvement in all models, and the proposed framework has the highest overall performance. Specifically, the large recall value indicates that the model is capable of identifying fraud cases well, which is critical in healthcare. Moreover, explainable AI methods have been integrated to make the system interpretable and thus suitable in the real world where transparency and trust are the necessities.

VII. CONCLUSION

The paper introduced a hybrid explainable architecture of detecting Medicare fraud that combined Graph Neural Networks, feature learning using transformers, graph centrality, and explainable AI methods into a single architecture. The article establishes that the concept of modeling healthcare organizations as a heterogeneous graph facilitates the identification of intricate relational relationships between providers, beneficiaries, and physicians, and this concept is crucial in improving the ability to detect fraud. The addition of transformer models also enhances the framework by training contextual dependencies with tabular data that is organized into structured tables, whereas graph centrality characteristics also offer more structural information that can enhance the performance of the traditional machine learning classifiers. The experimental results show that the proposed method performs better in terms of accuracy, precision, recall, F1-score, and ROC-AUC than the baseline and single models, and its recall is significantly better, the elimination of which is the key to reducing fraud cases that go undetected. The most important input of this work is that the graph-based learning, transformer representations, centrality-conscious features, and explainability are seamlessly integrated into a single framework and allow achieving high predictive performance and interpretability. Further research into applying the framework to large-scale real-time fraud detection systems, integrating the concept of time through dynamic graphs, and the development of more sophisticated explainability methods can be taken up in the future to increase the transparency and trust in automated decision-making systems.

REFERENCES

1. N. N. Islam Prova, “Healthcare Fraud Detection Using Machine Learning,” 2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI), Coimbatore, India, 2024, pp. 1119–1123, doi: 10.1109/ICoICI62503.2024.10696476.
2. Y. W. Hainan, C. Jin, and W. Shang, “Healthcare Fraud Detection Based on ChatGPT-Guided Method,” 2025 1st International Symposium on E-CARGO and Applications (E-CARGO), Guangzhou, China, 2025, pp. 88–92, doi: 10.1109/E-CARGO65996.2025.11139301.
3. Parasdeep, G. Arora, R. Mehra, S. Kaswan, D. Upadhaya, and K. Soni, “Artificial Intelligence and Machine Learning Approaches for Healthcare Fraud Detection: A Review, Case Study, and Framework,” 2025 2nd Global AI Summit (AI Summit), Noida, India, 2025, pp. 1045–1050, doi: 10.1109/AISummit66170.2025.11410723.
4. A. Anand, M. K. Shukla, and P. Sharma, “Enhancing Healthcare Fraud Detection: A Comparative Analysis of Machine Learning Models,” 2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 2025, pp. 1–5, doi: 10.1109/RMKMATE64874.2025.11042737.
5. B. Jaysawal, S. K. Chauhan, R. R. Sah, and S. R. Parija, “Improving Healthcare Fraud Detection Accuracy Using Ensemble Unsupervised Machine Learning and Deep Learning Techniques,” 2025 International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU), Bhubaneswar, India, 2025, pp. 1–6, doi: 10.1109/IC-CGU67042.2025.11338056.
6. L. Das, L. Ahuja, and A. Pandey, “Region-based Convolutional Neural Network (RCNN) Model for Detecting Healthcare Fraud Spam,” 2024 International Conference on Computing, Sciences and Communications (ICCSC), Ghaziabad, India, 2024, pp. 1–6, doi: 10.1109/ICCSC62048.2024.10830350.
7. C. Chitteti, M. Yamuna, M. Srinath, C. Govardhan, and A. Vignatha, “Healthcare Insurance Fraud Detection Using Machine Learning,” 2025 8th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2025, pp. 968–973, doi: 10.1109/ICOEI65986.2025.11013497.
8. M. B. Nagaraja, S. Kilaru, and V. S. R. Yarlagadda, “Enhancing Healthcare Fraud Detection: Comparative Study of CatBoost and TabNet Models,” 2026 IEEE 16th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2026, pp. 990–995, doi: 10.1109/CCWC67433.2026.11393884.
9. J. T. Hancock, R. K. L. Kennedy, M. A. Walauskis, and T. M. Khoshgoftaar, “A New and Effective Technique for Unsupervised Labeling and Feature Selection with Applications in Healthcare Fraud Detection,” 2025 IEEE International Conference on Information Reuse and Integration and Data Science (IRI), San Jose, CA, USA, 2025, pp. 301–306, doi: 10.1109/IRI66576.2025.00063.
10. G. S. Sargam and R. Kalapala, “AI-Driven Claim Fraud Detection in Health Insurance Using Federated Anomaly Detection Networks with Cloud Computing on AWS for Privacy-Preserving Financial Security,” 2025 Third International Conference on

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 4, April 2026

|DOI: 10.15680/IJIRSET.2026.1504269|

Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV), Hyderabad, India, 2025, pp. 1–6, doi: 10.1109/ICPEEV67897.2025.11291290.

11. P. Ashok and A. S. Durge, "Fraud Detection and Prevention in Healthcare Insurance Claims Using Machine Learning Regression Models," 2025 International Conference on Data Science and Business Systems (ICDSBS), Chennai, India, 2025, pp. 1–7, doi: 10.1109/ICDSBS63635.2025.11031751.
12. S. Mardani and H. Moradi, "Using Graph Attention Networks in Healthcare Provider Fraud Detection," IEEE Access, vol. 12, pp. 132786–132800, 2024, doi: 10.1109/ACCESS.2024.3425892.
13. M. Rakesh and P. S. D, "Enhanced Medicare Fraud Detection Using Graph Convolutional Networks," 2024 4th International Conference on Artificial Intelligence and Signal Processing (AISP), Vijayawada, India, 2024, pp. 1–5, doi: 10.1109/AISP61711.2024.10870744.
14. R. Bounab, B. Guelib, S. Benzerogue, and K. Zarour, "Optimizing Machine Learning for Healthcare Fraud Detection: A Framework Using Hybrid Feature Selection and Hyperparameter Tuning," 2024 International Conference on Advanced Aspects of Software Engineering (ICAASE), Constantine, Algeria, 2024, pp. 1–8, doi: 10.1109/ICAASE64542.2024.10851054.
15. T. Ekin and K. S. Mandadapu, "Red-Teaming for Health Care Fraud Detection: A RAG-Based AI Approach," 2025 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA), Antalya, Turkiye, 2025, pp. 1–6, doi: 10.1109/ACDSA65407.2025.11166648.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details